# Personalized Information Recommendation for Chinese Paper Cutting Culture Heritage

Weibo Huang<sup>1</sup>, Bairu Xie and Xiaoyue Liao Guangdong University of Foreign Studies, Guangzhou, China

Abstract. Globalization and information technology are developing rapidly. Chaozhou Chinese paper cutting exists physically. Chaozhou Chinese paper cutting is facing the difficulties of preservation and dissemination. The purpose of this study is to further realize the preservation and personalized recommended dissemination of intangible cultural heritage represented by Chaozhou Chinese paper cutting. In this paper, information technologies such as knowledge graphs, multimedia applications, Bi-RNN neural networks, and probabilistic decomposition models are used. This paper delves into the visual communication and personalized recommendation of intangible culture to preserve, pass on, and spread it better.

Keywords. Chaozhou Chinese paper cutting, Personalized Information Recommendation, Knowledge Graph.

### 1. Introduction

The trend of economic globalization and world multi-polarization is irreversible, and international competition is a competition of comprehensive national power based on economic strength and scientific and technological strength. In the twenty-first century, to promote Chinese culture and spread Chinese voice, Chaozhou Chinese paper cutting maybe a piece of high ground that cannot be ignored. However, the network platform of Chinese paper cutting mainly displays information about Chinese paper cutting, without a personalized information customization service. How to spread the Chinese paper cutting in Chaozhou vigorously and radiate more and more people is the challenge of Chaozhou Chinese paper cutting heritage.

Yu Fuye and Zhang Xizhong have proposed the importance of establishing a knowledge map and building a database on the inheritance and development of Manchu Chinese paper cutting. In terms of knowledge graph construction, knowledge graphs are built around the domain aspects of various industries, while the knowledge graph construction of Chinese paper cutting information and application services are few and far between. [1]

In terms of personalized recommendation, scholars have mainly divided traditional recommendation algorithms into three main categories. For example, collaborative filtering recommendation algorithms, content-based recommendation algorithms, and hybrid recommendation algorithms are proposed with the shortcomings and defects of the combined recommendation algorithms. Also, personalized recommendation

<sup>&</sup>lt;sup>1</sup>Corresponding Author: Huang Weibo, Network and Information Center, Experimental Teaching Center, Guangdong University of Foreign Studies, China; Email: hwb444@163.com

systems with multiple sources of big data, and knowledge graphs that fuse contextual information [2]. However, the existing algorithms are not always able to maintain a high accuracy rate due to the dynamic changes of user interests and the long-term instability of user interests and other elements.

This study takes the relationship between information and attributes of Chaozhou paper cuttings as a starting point. This paper composes the relationship between Chaozhou Chinese paper cutting and its attributes. This paper constructs a knowledge graph on this basis and incorporates the knowledge graph into specific algorithms. This paper is constructed in a way to have a better sorting of Chaozhou Chinese paper cutting attributes. This paper is based on the construction of the knowledge graph of Chaozhou Chinese paper cutting and the integration of the personalized recommendation algorithm. This paper will be better and more accurate to make personalized information recommendations for users. [3]

#### 2. Knowledge graph construction

#### 2.1. Definition of Chaozhou Chinese paper cutting attributes

The construction of the knowledge map of Chaozhou Chinese paper cutting starts with defining the attributes of Chaozhou Chinese paper cutting. In other words, the process of defining the attributes of Chaozhou Chinese paper cutting is also the process of extracting and re-understanding the information of Chaozhou Chinese paper cutting.

To identify the attributes of Chaozhou Chinese paper cutting that can be used as entities for the knowledge graph construction of Chaozhou Chinese paper cutting. Knowledge and integration of offline and online information related to Chaozhou Chinese paper cutting are required. By understanding and researching offline and online, it is possible to determine information about the attributes common to Chaozhou paper cutting. Information on the attributes of Chaozhou Chinese paper cutting is Chinese paper cutting types, Chinese paper cutting knives, subject matter, Chinese paper cutting forms, pattern styles, emotional connotations, life uses, and belief customs. [4] By defining the attributes of Chaozhou paper cuttings, entities that can be used for knowledge graph construction are abstracted. These entities will be applied to the subsequent process of knowledge graph construction.

The information from the online platform was analyzed and studied to identify several important attributes of Chaozhou Chinese paper cutting as entities for the knowledge mapping of Chaozhou Chinese paper cutting. The common attributes of Chaozhou Chinese paper cutting information are Chinese paper cutting types, Chinese paper cutting knives, subject matter, Chinese paper cutting forms, pattern styles, emotional connotations, life uses, and belief customs.

#### 2.2. The construction of Chaozhou Chinese paper cutting knowledge graph

To ensure the quality of the Chaozhou Chinese paper cutting knowledge graph, this paper adopts a top-down construction approach. The top-down approach firstly determines the schema layer for the knowledge graph and uses the schema layer for the upper management of the knowledge graph. By upper-level management of the schema layer, valid data and entities can be added to the knowledge base. Furthermore, in the

process of adding data to the knowledge base, invalid data is excluded at the same time. In the process of constructing the knowledge graph, the knowledge graph sick cycle and some characteristics in terms of the Chaozhou Chinese paper cutting domain will also be considered. Because of this, the Chaozhou Chinese paper cutting knowledge graph collusion process is in order of original data acquisition, ontology construction, knowledge fusion, and knowledge storage. [5] It is worth mentioning that each step of this construction method develops a suitable scheme according to the actual situation.

# 3. Chaozhou Chinese paper cutting knowledge graph recommendation algorithm

In recent years, recommendation algorithms, including collaborative filtering algorithms, have suffered from shortcomings in cold-start, overfitting, and data sparsity [6]. These shortcomings lead to inefficient use of data and the inability to effectively match their users' individual needs.

Many researchers use different auxiliary information including Social Networks, the user or item attributes, and Knowledge Graphs to complement the algorithm to improve the recognition of user characteristics. Among them, recommendation algorithms using knowledge graphs are undoubtedly the hot research topic nowadays. The research results of recommendation algorithms for knowledge graphs are also being expanded and applied in industries such as music, film and television, and business. [7] Elena Hernández-Nieves et al. investigate how to use personalized recommendation algorithms (PER) for financial product recommendations. [8] To enhance the information linkage in the recommendation model for the implicit layer, Depeng Dang et al. proposed to use a deep perceptual neural network model (DKN) [9] Wang YueQun et al., on the other hand, tried to use RippleNet and a multi-task recommendation model (MKR) oriented to knowledge graphs. This model enhances the recommendation model's ability to mine user information and interactively combine the knowledge graph with the user's historical data. Thus, this approach improves the performance of the model for movie, music, and book recommendations. [10]

However, the above studies still have certain shortcomings. Subject to the problems of meta-road design and merging of model and auxiliary information, it has some degree of limitation in application scenarios. Moreover, among the existing recommendation algorithms for knowledge graphs, there is a lack of models that can efficiently identify and extract low-level features from user information. To address this situation, this paper proposes the use of collaborative filtering algorithms based on Bi-RNN neural networks to solve the data sparsity problem. In this paper, we improve the cold start problem by measuring the similarity of user features and finally integrate it with the knowledge graph embedding technique (KGE). Thus, this paper makes full use of the contextual relationship between user information and auxiliary information. This approach enhances the mining, identification, and extraction of users' features for the paper-cut recommendation algorithm, and thus improves its recommendation accuracy for the paper-cut algorithm.

#### 3.1 Model design

This model first collects and measures the similarity feature indexes of users, and cross-fuses the knowledge graph feature information with the user information using the knowledge embedding technique.

This model uses a collaborative filtering algorithm based on a Bi-RNN neural network to identify multi-level features in the feature vector of user preferences about Chinese paper cutting mapping and contextual relationships. The model recommends suitable Chinese paper cutting knowledge graphs for different users. The overall framework of the model is shown in Figure 1.



Fig. 1. Recommended model of the paper-cut knowledge graph.

### 3.1.1 User preference similarity index construction

For users with similar age and educational backgrounds, their preferences for Chinese paper cutting charts are perhaps more similar. Therefore, it makes sense to recommend similar or the same type of clippings to users with similar attribute characteristics. This also improves the cold start problem of the clippings recommendation model very well. For the construction indicators of user attribute characteristics, mainly including age indicators as well as education indicators, the relevant formulas are as follows.

### (1) User age similarity measurement

There is a high probability that users of similar age have similar preferences for paper-cut mapping. This model sets the age gap threshold between users to 3. Users with an age gap less than or equal to 3 will be considered as users with the same feature attribute and their similarity is marked as 1. The formula is as follows.

$$SIM(N,U)_{age} = \begin{cases} \frac{3}{|a_N - a_U|}, |a_N - a_U| > 3\\ 1, |a_N - a_U| \le 3 \end{cases}$$
(1)

N, U: User N as well as user U.

 $a_N, a_U$ : The ages of user N as well as user U. SIM $(N, U)_{age}$ : The age similarity between user N as well as user U.

### (2) Calculation of similarity of users' education

The preferences between users with similar academic qualifications are also likely to be similar, and the similarity for users' academic qualifications is calculated as follows.

$$SIM(N,U)_{Edu} = \begin{cases} 1, \ E_N = E_u \\ 0, \ E_N \neq E_u \end{cases}$$
(2)

N, U: User N as well as User U.

 $E_N, E_u$ : User N and the educational status of User U.

 $SIM(N,U)_{Edu}$ : The similarity of education between user N as well as user U is recorded as 1 if they are the same, and 0 if they are not.

Based on the 2 similarity features mentioned above, the formula used to represent the similarity of the overall user preference features is proposed as follows.

$$SIM(N,U)_{Preference} = \alpha * SIM(N,U)_{age} + \beta * SIM(N,U)_{Edu},$$
  

$$\alpha + \beta = 1 \ \alpha, \beta \in [0,1]$$
(3)

 $SIM(N, U)_{Preference}$ : The similarity of the overall preference characteristics of user N as well as user U.

 $\alpha, \beta$ : The weights of different similarity factors.

#### 3.1.2 Chinese paper cutting knowledge graph embedding

The MLP model is used to process the deep-level feature relationship between the entity vector of the paper-cut mapping and the feature vector of the paper-cut mapping. Thus, the deep-level feature representations of the relationship vectors between them can be obtained. Further, the potential feature representations of the data are concatenated to obtain the final paper-cut mapping feature vectors. The formula is as follows.

$$P_T = \phi_{C \sim E(H)}[M^T(C, P)[P]]$$
(4)

$$R_T = S^T(S^{T-1}(...S(R)))$$
(5)

$$\hat{f} = \sigma(F_{MLP}^Z(P_T, R_T)) \tag{6}$$

$$S^T = \sigma(W_R R + B_R) \tag{7}$$

$$F_{MLP} = \sigma(W_{KGE}(U_T, C_T) + B_{KGE})$$
(8)

 $P_T$ : Deep representation of paper-cut mapping entity feature vectors.

 $M^T$ : The cross-compression unit of the Lth layer.

 $M^{T}(C, P)$ : Paper-cut mapping entity features and deep feature representation of papercut mapping obtained by cross-compression cells of T-layer.

 $R_T$ : Deep representation of the relation T-vector.

[P] : Distinguishing marks for paper cutting entity vectors.

S(R): The network layer in the fully connected state.

 $F_{MLP}^Z$ :Z-layer MLP network.

 $\sigma$ : Nonlinear activation function.

 $W_{KGE}$ ,  $W_R$ : Weights.

 $B_{KGE}$ : Deviation.

 $\hat{f}$ : Prediction vectors for paper-cut mapping.

## 3.1.3

The user preference feature vector, as well as the mapping feature vector are crosscompressed with the following equation.

(1) For the paper-cut mapping prediction vector as well as the user similarity feature vector, crossover processing is performed.

$$M_{Z} = \hat{f}_{Z} S_{Z}^{T} = \begin{bmatrix} f_{Z}^{(1)} S_{Z}^{(1)} & \cdots & f_{Z}^{(1)} S_{Z}^{(E)} \\ \vdots & \ddots & \vdots \\ f_{Z}^{(E)} S_{Z}^{(1)} & \cdots & f_{Z}^{(E)} S_{Z}^{(E)} \end{bmatrix}$$
(9)

*Z*: The embedding dimension of the feature vector.

 $M_Z$ : The feature matrix obtained by intersecting the user similarity vector and the paper-cut mapping prediction vector.

(2) Compression of the feature vectors after cross-processing.

$$C_{z+1} = M_Z W_Z^{CC} + M_Z^T W_Z^{CP} + B_Z^C = C_Z P_Z^T W_Z^{CC} + P_Z C_Z^T W_Z^{CP} + B_Z^C$$
(10)

$$P_{Z+1} = M_Z W_Z^{CP} + M_Z^T W_Z^{PP} + B_Z^p = C_Z P_Z^T W_Z^{CP} + P_Z C_Z^T W_Z^{CP} + B_Z^C$$
(11)

$$[C_{z+1}, P_{z+1}] = M(C_{z+1}, P_{z+1})$$
(12)

 $M(C_{z+1}, P_{z+1})$ : The feature vector representation of the paper-cut mapping obtained after cross-compression.

[C], [P]: The output of user feature vectors and the output of paper-cut mapping feature vectors.

3.1.4

The cross-compressed processed knowledge mapping feature vectors are fed into the Bi-RNN neural network. Thus, its deep multi-feature representation can be obtained, and the structure of the Bi-RNN neural network is shown in Figure 2.

135



Figure 2. Bi-RNN neural network.

The formula is as follows:

$$\overrightarrow{H_Z} = gru(M(C_N, P_N), \overrightarrow{H_{Z-1}})$$
(13)

$$\overleftarrow{H_Z} = gru(M(C_N, P_N), \overleftarrow{H_{Z-1}})$$
(14)

$$H_Z = W^T \overleftarrow{H_Z} + W^C \overrightarrow{H_Z} \tag{15}$$

$$R = Softmax(W^T M(C_N, P_N))$$
<sup>(16)</sup>

R: The final value of the weight matrix.

 $\overrightarrow{H_{Z}}$ : Forward hidden layer state vector.

 $\overleftarrow{H_Z}$ : Backward hidden layer state vector.

 $W^{T}$ ,  $W^{C}$ : The weight coefficients of the preceding and following state vectors. And the expressions of users' prediction scores for different knowledge graphs are as follows.

$$\widehat{\mathbf{G}} = \mathbf{C}_{\mathbf{N}} \mathbf{P}_{\mathbf{N}}^{\mathrm{T}} = \sum_{a=1}^{Z} \mathbf{C}_{\mathbf{N}a} \mathbf{P}_{\mathbf{N}a} \tag{17}$$

C<sub>N</sub>: The number of rows of the matrix where user C is located.

P<sub>N</sub>: The number of rows of the matrix where the features P of the map are located.

 $C_{Na}$ : The relationship between user C and the a<sup>th</sup> mapping feature.

 $P_{Na}$ : The relationship between the graph feature P and the a<sup>th</sup> user.

After optimization, the loss function is:

$$Loss = \sum_{(N \in TRAIN)} \left( G_N - \sum_{a=1}^{Z} C_{Na} P_{Na} + \theta(||C_N||^2 + ||P_N||^2) \right)$$
(18)

# 4. Experimental results

The experimental environment is shown in Table 1-1:

Category	Parameter		
The operating system	10 the 32-bit Windows		
The processor	Intel(R) Core(TM) i5-3337U @ 1.80ghz quad Core		
GPU	NVIDIA GeForce GT 720M		
Memory	6G		

Table 1-1. MKFM recommendation algorithm training process.

## 4.1 Data set

In this paper, two datasets, the publicly available Chinese paper cutting website 1M and some survey questionnaires 2M made by ourselves, are selected.

The rating file in the dataset includes 3040 users, 22 types of Chaozhou paper cutting, and user rating scores for different paper cutting types. The cumulative number of user ratings in this data set has reached 1327, with ratings ranging from 1 to 5. The higher the rating, the more the user likes paper cutting. The user attributes file contains information about user attributes. For example, age, education, etc. The paper cutting attribute information. For example contains emotions, categories, etc.

#### 4.2 Data pre-processing

The display data in the two obtained datasets 1M, 2M are converted into implicit data. If a user rates a paper cutting, the entry is marked as 1. Then an additional set of unrated paper cutting is matched and marked as 0. The score threshold in dataset 1M is 4. The entry is marked as 1 only when the user rates the paper cutting as 4. This means the user has rated a type of paper cutting positively. The opposite is true for negative ratings. The score and count for all interactions are uniformly marked as 1.

The data set	The number of users	Value of the user feature domain	Number of paper cutting types	Number of interactions	Number of tuples of a knowledge graph
1M	2040	3.1	1,347	754,23	2,873
2M	1000	1	334	34,234	1,231

Table 1-2. Data set statistics.

Building the Knowledge Graph FG database uses Microsoft Satori to build the FG database for each dataset. Further, the names of the paper-cut types are matched with the tail features of the knowledge graph triples. Thereby, the set of IDs in the FG subset is obtained. Then, the IDs of paper-cut types with duplicate matches or missing matches are removed, and the remaining IDs are matched with the head and tail features of the knowledge graph triad. In turn, all the triads with successful matches and good matches are filtered out, as shown in Table 1-2.

## 4.3 Experimental Results

This algorithm uses both ACC and AUC as the measurement criteria for CTR prediction for 1M as well as 2M. The results are shown in Tables 1-3.

Model	1M		2M		
	AUC	ACC	AUC	ACC	
Bi-RNN	0.834	0.656	0.810	0.687	
PER	0.801	0.648	0.802	0.657	
CKE	0.768	0.637	0.789	0.643	
DKN	0.754	0.571	0.764	0.632	

Table 1-3. CTR prediction tasks on the data set.

## 4.4 Analysis of experimental results

The results of the experiments show that the AUC, as well as AAU metrics of this model are higher than the rest of the models, both on 1M and 2M. Even in the case of sparse data, the model can perform the recommendation of a Chinese paper cutting knowledge graph with high accuracy.

## 5. Conclusion

The art of Chinese paper cutting in Chaozhou is unique in its expression and has certain artistic value. Chaozhou Chinese paper cutting has been included in the list of China's non-material cultural heritage. There is a certain necessity to establish a knowledge graph and personalized recommendation for Chaozhou Chinese paper cutting to realize its modernized communication. In this paper, information technologies such as knowledge graph, multimedia application, Bi-RNN neural network, and collaborative filtering model are used. This paper delves into the visual communication and personalized recommendation of intangible culture to preserve and transmit it better. In addition, the research contents and methods of this project are well-grounded in theory. The project is innovative and highly experimental, with high feasibility of the scheme and a certain prior research foundation. The research of this paper will provide scientific and intellectual support for Chaozhou Chinese paper cutting and the preservation and dissemination of intangible cultural heritage.

## Acknowledgement

This study was financially supported by the Undergraduate Innovation Training Project of Guangdong University of Foreign Studies in 2021.

# References

 Hussein E., Daoud S. and Alrabaiah H. et al. Exploring Undergraduate Students' Attitudes towards Online Learning during COVID-19: A Case from the UAE. Children and Youth Services Review, 2020, 119: 105699.

- [2] Chen Wenyan, Wang Xin, Fu Xueqing, Li Renjie and Zhang Junhai. Multidimensional visualization method of regional intangible cultural heritage landscape. Science of surveying and mapping, 2013, 38(01): 87-89+95.
- [3] Depeng Dang, Chuangxia Chen, Haochen Li, Rongen Yan, Zixian Guo and Xingjian Wang. Deep knowledge-aware framework for web service recommendation. The Journal of Supercomputing, 2021, May 11 published online. DOI:10.1007/s11227-021-03832-2.
- [4] Elena Hernández-Nieves, Guillermo Hernández, Ana-Belén Gil-González, Sara Rodríguez-González and Juan M. Corchado. Fog computing architecture for personalized recommendation of banking products. Expert Systems with Applications, 2019, 140: 112900.
- [5] Jiang Xinjing, Qi Xiaogang and Liu Lifang. Journal of Intelligent Systems, 2018, 13 (02):189-195.
- [6] Scatter tiger. Without suitable soil, local art will eventually decline -- shaanxi paper cutting Inheritance dilemma. Art Observation, 2014(10):21-22.
- [7] Rendle S. Factorization Machines, The 10th IEEE International Conference on Data Mining, Sydney, Australia, 14-17 December, 2010, IEEE, 995–1000.
- [8] Wang Q., Mao Z. and Wang B et al. Knowledge graph embedding: a survey of approaches and applications. IEEE Transactions on Knowledge and Data Engineering, 2017, 29(12): 2724-2743.
- [9] Wang YueQun, Dong LiYan, Li YongLi and Zhang Hao. Multitask feature learning approach for knowledge graph enhanced recommendations with RippleNet. PLOS ONE, Public Library of Science, 2021, 16(5): 1-21.
- [10] Yang Yuji, Xu Bin and Hu Jiawei et al., An efficient method for constructing domain knowledge graph, Journal of Software, 2018, 29(10): 2931-2947.