Eng Int Syst (2026) 5: 641–646
© 2026 CRL Publishing Ltd

Engineering
Intelligent Systems

Precise Recommendation of Sichuan Cultural and Tourism Information in Short-Video Apps Based on User Profiles

Yangjian Hu*

Digital Economy Teaching and Research Office, School of Digital Economy and Management, Sichuan Technology and Business University, Meishan, Sichuan 620010, China

This paper briefly introduces the construction of user profiles and a user profile-based short video recommendation algorithm. In the algorithm, the user profiles are first classified using the clustering algorithm improved by particle swarm optimization (PSO), and then short videos are recommended to users based on the profiles. Simulation experiments were then conducted on the Douyin short video platform. The results showed that the PSO-improved K-means clustering algorithm effectively classified user profiles, with the best classification effect when the K value was 4. The proposed recommendation algorithm could provide users with more accurate short videos on Sichuan culture and tourism.

Keywords: User profile, short video, recommendation algorithm, clustering algorithm

1. INTRODUCTION

Given the deep integration of digital media and mobile Internet, short-video platforms have become the core channels through which the public obtain cultural and tourism information, stimulate travel intentions, and make destination decisions (Shang et al., 2017). Generation Z users learn about travel destinations on platforms such as Douyin, Kuaishou, and Xiaohongshu, and the different forms of content on these platforms can significantly influence users' travel choices. However, with the explosive growth of content, the traditional means of broadly disseminating information is facing disadvantages such as information overload, user distraction, and decreasing conversion efficiency (Yang et al., 2019). On the one hand, there are differences in the demand for cultural and tourism content among different user groups. On the other hand, even if the regions promoting cultural and tourism development have rich cultural and tourism resources, failing to achieve precise matching of person, content, and scene can easily lead to the submerging of high-quality resources or loss of interest by users. By collecting user profiles, such as behavioral data, demographic attributes, and interest tags of users on short-video platforms, personalized recommendations of cultural and tourism resource content can be derived to improve the efficiency of content reach, user dwell time, and tourism conversion rate (Qi & Li, 2021).

Several relevant studies have been conducted including a deep learning-based short video content analysis and recommendation algorithm proposed by Yang et al. (2025) exhibited enhanced feature extraction and personalization functions. Yuan et al. (2024) proposed a THACIL algorithm that optimized the attention mechanism of the original model and verified it through experiments. Liu et al. (2023) proposed adding modules for integrating time and user context with the traditional collaborative filtering

*Corresponding address: No. 9, Mindong Section, Meizhou Avenue, Meishan City, Sichuan 620010, China. Email: hyangjian@hotmail.com

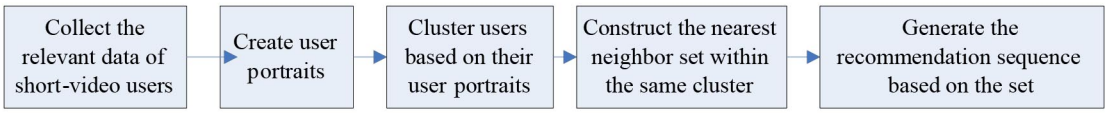


Figure 1 Process of the short video recommendation algorithm.

algorithm to focus on mining user interests. This current study constructs user profiles and proposes a short video recommendation algorithm based on these profiles. In the recommendation algorithm, the user profiles are first classified using the particle swarm optimization (PSO)-improved clustering algorithm, and then short videos are recommended to users based on the classified user profiles. Moreover, simulation experiments were conducted on the Douyin short video platform.

2. RECOMMENDATION ALGORITHM FOR SHORT VIDEOS ON CULTURAL TOURISM INFORMATION BASED ON USER PROFILES

In recent years, Sichuan has sparked a "Sichuan fever" across the Internet with its intellectual properties such as "Anyi Panda", "Woxingchuan" series, and "Ba Xiaohu" and high-frequency, highly creative short video content. Its official account matrix posts more than 60 videos per day, with cumulative views exceeding 10 billion times. However, with the explosive growth of content supply, the traditional mode of casting a wide net to collect and disseminate information is facing bottlenecks such as information overload, scattered user attention, and decreasing conversion efficiency. On the one hand, young tourists prefer "food + street photography + trendy check-in", the silver-haired group focuses on "health + slow pace + historical culture", and families with children focus on "study tours + safety + interactive experiences". On the other hand, Sichuan is rich in cultural and tourism resources, including the natural wonders of Jiuzhaigou, the mysterious bronze of Sanxingdui, the bustling market of Chengdu, the snow-capped plateau of Ganzi, and the sunny health care of Panzhihua. If the precise matching of people, content, and tourist site cannot be achieved, it is very easy for high-quality resources to be submerged or users' interest to be lost.

To enhance promotion efficiency and save costs (Ma et al., 2021), precise recommendations for short videos containing cultural and tourism information are required. The short video recommendation algorithm based on user profiles is shown in Figure 1 (Mou et al., 2021), and the specific steps are as follows.

① Relevant information of users are collected from the database and log files of a short video platform, including basic, interest, and social attributes of users (van Deen et al., 2022).

② User profiles are constructed based on the collected user-related information. For the basic user attributes, one-heat encoding is used to construct feature vectors. The dimension of a feature vector constructed by the one-heat code depends on the number of possible states of the feature. For example, gender has two states: males and females. Then, the one-heat code of this feature is a two-dimensional vector, the one-heat code for males is $\{1, 0\}$, and the code for females is $\{0, 1\}$. Interest attributes are expressed as:

$$\begin{cases} V_{user} = (T_1 \cdot w_1, T_2 \cdot w_2, \cdots T_i \cdot w_i) \\ T_i = \frac{\sum_{e_j \in E_i} e_j \cdot I_j}{n_i} \\ I_j = P \cdot K_p + D \cdot K_d + Z \cdot K_z \\ K_p + K_d + K_z = 1 \\ w_i = \frac{c_i}{\sum_{i=1}^{n} c_i} \end{cases} \tag{1}$$

where V_{user} is the feature representation of the user's interest attribute, T_i is the feature vector of type i videos, E_i is the collection of type i videos that the user has viewed, e_j is the feature vector of the j-th video belonging to E_i, obtained by the term frequency-inverse document frequency (TF-IDF) method, I_j denotes the user interest degree of e_j, n_i is the number of type i videos that the user has viewed (Wang & Li, 2022) P, D, Z are the number of comments, likes, and reposts for e_j, K_p, K_d, K_z are weights corresponding to P, D, Z behaviors, w_i is the weight of type i videos, and c_i is the number of category videos. The social attribute part is represented by a binary relation matrix:

$$R = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1j} \\ r_{21} & r_{22} & \cdots & r_{2j} \\ \cdots & \cdots & \cdots & \cdots \\ r_{i1} & r_{i2} & \cdots & r_{ij} \end{bmatrix} \tag{2}$$

where r_{ij} is the index indicating user i's social behaviors toward user j, which values 1 if user i follows user j and 0 if it does not.

③ The feature vectors of a user's basic, interest, and social attributes are concatenated to form the feature vectors of the user profile. This paper uses an improved K-means clustering algorithm based on PSO (Jicheng, 2021) to cluster users. The steps are as follows. First, randomly generate several groups of candidate solutions, each of which represents a group of initial clustering centers and also represents the position of a particle in the PSO population in the search space. The fitness value of the population particles is then calculated using the following formula:

$$J = \sum_{i=1}^{N} \sum_{j=1}^{c} \left\| x_i - c_j \right\|^2 \tag{3}$$

Table 1 Some user data.

User ID	1000212	1001023	1002341	...
User gender	Male	Female	Female	...
User age	Under 18 years old	19–30 years old	31–60 years old	...
User educational background	Below primary school	Undergraduate	High school	...
User region	Eastern region	Central region	Western region	...
User occupation	Student	Teacher	Merchant	...
Video viewing record	[101,3,2,3]; [203, 2, 3, 4]; [213,2,3,2]; ...	[111,2,2,3]; [204,1,3,4]; [211,3,2,2]; ...	[103,3,1,3]; [233,2,2,4]; [263,2,0,2]; ...	...
Subscribes	1002311; 1003969; 1002789; ...	1002784; 1004156; 1002589; ...	1006911; 1004669; 1003289; ...	...
Followers	1012361; 1004869; 1003219; ...	1002987; 1001458; 1006936; ...	1008524; 1004651; 1005478; ...	...

where J is the fitness values of the particles, x_i is the feature vector of user i's profile, c_j is the j-th cluster center within a candidate solution, N is the number of users, and c is the number of cluster centers in the candidate solution (Wang & Zhao, 2025). Then, the population is iterated:

$$\begin{cases} V_{t+1} = V_t + a_1 \cdot c_1 \cdot (pbest_t - P_t) \\ \quad +a_2 \cdot c_2 \cdot (gbest_t - P_t) \\ P_{t+1} = P_t + V_{t+1} \end{cases} \tag{4}$$

where P_t and V_t are the current position and velocity of particle t, $pbest_t$ and $gbest_t$ are the historical optimal position and the current global optimal position respectively, a_1 and a_2 are learning factors, c_1 and c_2 are random numbers that are evenly distributed between 0 and 1. The above calculation of the population particle fitness and the iteration of population particles are repeated until the fitness value converges to stability. Then, the candidate solution represented by the optimal particle within the population is used as the initial clustering center, and the K-means clustering algorithm is used to divide the users (Yan et al., 2019).

④ For target user u, the nearest neighbor user set is constructed by calculating the comprehensive similarity between the target user and the other users in the same cluster. The formula for calculating the comprehensive similarity is:

$$\begin{cases} sim_{user,cf}(u,v) = \alpha \cdot sim_{cf}(u,v) \\ \quad +(1-\alpha) \cdot sim_{up}(u,v) \\ sim_{cf}(u,v) = \frac{\sum_{a \in A} I_{u,a} I_{v,a}}{\sqrt{\sum_{a \in A} I_{u,a}^2} \sqrt{\sum_{a \in A} I_{v,a}^2}} \\ sim_{up}(u,v) = \frac{V_u \cdot V_v}{\|V_u\| \|V_v\|} \end{cases} \tag{5}$$

where $sim_{user,cf}(u,v)$ is the comprehensive similarity between users u and v, $sim_{cf}(u,v)$ is the similarity between users' evaluations of short videos, $sim_{up}(u,v)$ is the similarity of user profiles between the two users, $I_{u,a}$ and $I_{v,a}$ are users' ratings for video a, i.e., user interest degree, and $V_u$andV_v are user profiles feature vectors. After sorting the similarities in descending order, the top few users are selected to construct the set of user$u$'s nearest neighbors.

⑤ Based on the ratings of the unviewed videos in the user u's nearest neighbor set (Zhang & Liu, 2021; Liu & Chen, 2018), these videos are recommended in descending order based on the predicted ratings. The formula for predicting the ratings of unviewed video is:

$$P_{u,b} = \bar{I}_u + \frac{\sum_{v \in kNN(u)} sim_{user,cf}(u,v) \times (I_{v,b} - \bar{I}_v)}{\sum_{v \in kNN(u)} sim_{user,cf}(u,v)} \tag{6}$$

where $P_{u,b}$ is the user's predicted rating for the video that $u$has not viewed, $\bar{I}_u$, $\bar{I}_v$ are the average rating given by users u and v to the video, and $I_{v,b}$ is user v's rating for the video.

3. SIMULATION EXPERIMENT

3.1 Experimental Data

The simulation experiment used data from the Douyin short video platform, including the historical records and basic information of users. Due to space limitations, only partial user information is presented here (Table 1).

Considering that the main purpose of this article is to achieve precise push of Sichuan culture and tourism information, only videos related to Sichuan culture and tourism were retained in the collected user video viewing records.

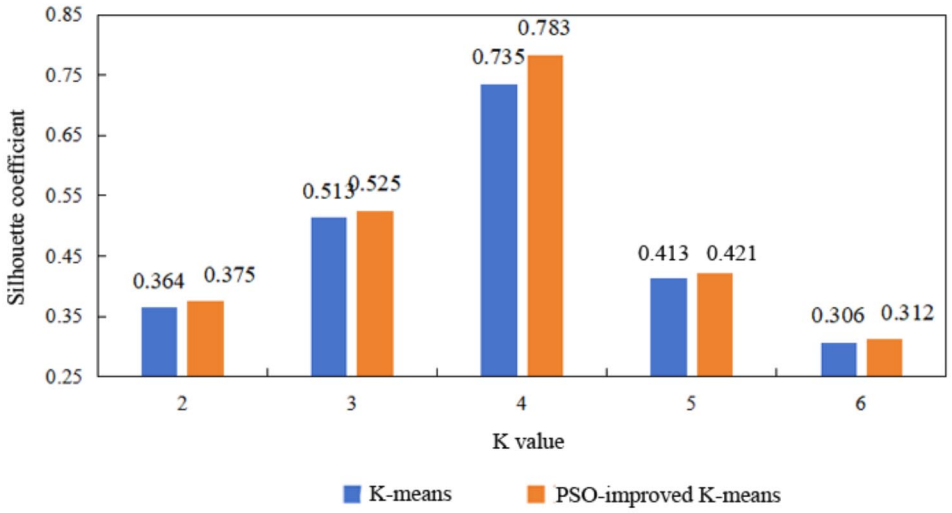


Figure 2 Clustering effect.

3.2 Methods

(1) Clustering effect analysis

The relevant parameters for the PSO section were set as follows: a population size of 10, $c_1 = 1.5$, $c_2 = 1.0$, an inertia weight of 0.6, and a maximum of 100 iterations. The PSO-improved algorithm was compared with the traditional K-means clustering algorithm, and the K value was set to 2, 3, 4, 5, 6.

The clustering effect was measured using the silhouette coefficient:

$$S = \frac{1}{n}\sum_{i=1}^{n}\frac{b_i - a_i}{\max(a_i, b_i)}, \tag{7}$$

where S is the silhouette coefficient, a_i is the average distance between sample i and the other samples within the cluster, b_i is the average distance between the sample and samples outside the cluster, and n is the number of samples.

(2) Analysis of the effectiveness of short video recommendations based on user profiles

To further verify the performance of the proposed recommendation algorithm, it was compared with the algorithms based on association rules and collaborative filtering (Ma &Sawadee, 2026; Liu et al., 2026). The formula for measuring the algorithm's performance is:

$$\begin{cases} P = \frac{\sum_{u\in U}|R(u)\bigcap T(u)|}{\sum_{u\in U}|R(u)|} \\ R = \frac{\sum_{u\in U}|R(u)\bigcap T(u)|}{\sum_{u\in U}|T(u)|} \\ F = \frac{2\cdot P\cdot R}{P+R} \end{cases}, \tag{8}$$

where P denotes the accuracy, R denotes the recall rate, F is the combined value of accuracy and recall rate, $R(u)$ is a collection of recommended videos generated by the algorithm, and $T(u)$ is a collection of short videos that users actually like.

3.3 Experimental Results

Silhouette coefficients were used to measure the clustering performance of the traditional and PSO-improved k-means algorithms at different K values, as shown in Figure 2. It can be seen that as the K value increased, the silhouette coefficients of both clustering algorithms rose first and then fell. When the K value was 4, the silhouette coefficients of both clustering algorithms were the largest, and the clustering effect was the best at this time. Under the same K value, the PSO-improved algorithm had a larger silhouette coefficient and better clustering effect.

The recommendation performance of the three recommendation algorithms is shown in Figure 3. It can be seen that the algorithm based on user profiles exhibited the best recommendation performance. Some of the recommendation results are shown in Table 2. The recommendation results generated by the proposed algorithm were more in line with the user profiles and personalized compared to the other two algorithms.

4. CONCLUSIONS

This paper briefly introduces the construction of user profiles and a short video recommendation algorithm based on user profiles. In the recommendation algorithm, the user profiles are first classified using the PSO-improved clustering algorithm, and then short videos are recommended to users based on the classified user profiles. Simulation experiments were conducted on the Douyin Short Video platform. It was found from the results that the K value increased, the silhouette coefficients of both clustering algorithms rose first and then fell. When the K value was 4, the silhouette coefficients of both clustering algorithms were at their maximum, and the PSO-improved algorithm performed better than the traditional algorithm. The recommendation algorithm based on user profiles demonstrated the best recommendation performance. The recommendation results

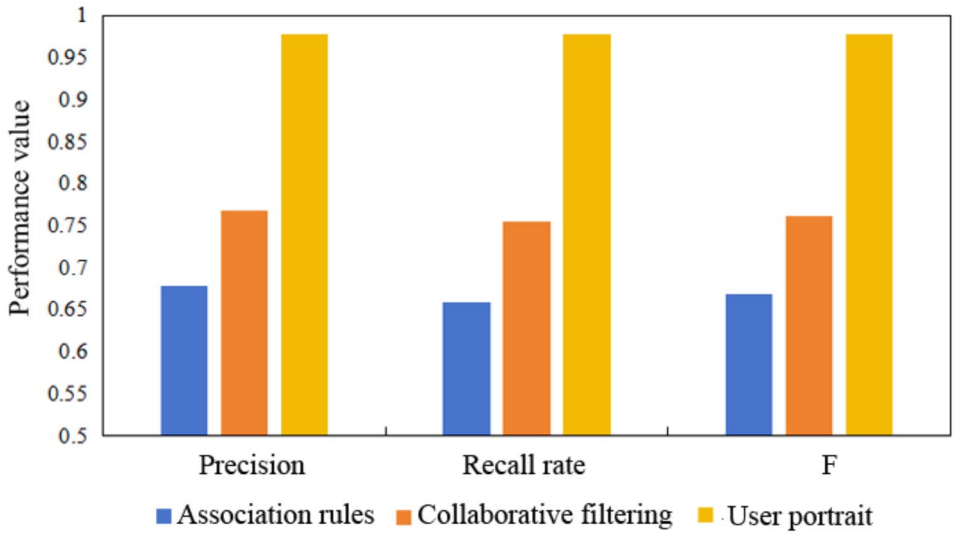


Figure 3 Recommendation performance.

Table 2 Part of recommendations.

User ID	User profile (preference)	Results recommended by the association rules-based algorithm	Results recommended by the collaborative filtering-based algorithm	Results recommended by the proposed algorithm
100231	Scenery	"Sixteen 5A Scenic Spots Only in Sichuan", "Top 10 Intangible Cultural Heritages of Sichuan", "Mysterious Quadrangle Courtyards Discovered in the Deep Mountains of Sichuan"	"One Day You will Come to Western Sichuan for any reason", "Bipeng Valley is Now So Beautiful", "Already Looking Forward to the Autumn in Jiuzhaigou"	"A Riot of Colors, This is the Autumn Color of Bipeng Valley", "One Day You will Come to Western Sichuan for any reason", "Bipeng Valley is Now So Beautiful"
100211	Culture	"Top 10 Intangible Cultural Heritages of Sichuan", "Lazy People's Guide to Eating in Chengdu during National Day Holidays", "How Valuable is the Pension in Leshan?"	"Smiling Pottery Figurines from the Han Dynasty at Sichuan Museum", "Hanjiang Ancient City", "Jinsha Site Museum in Chengdu"	"Exploring the Luojiaba Site Museum of the Ba Culture", "Hanjiang Ancient City", "Searching for Temple Remains in the Deep Mountains"
100215	Food	"Top 10 Must-Eat Dishes in Sichuan", "Mysterious Quadrangle Courtyards Discovered in the Deep Mountains of Sichuan", "Smiling Pottery Figurines from the Han Dynasty at Sichuan Museum"	"Lazy People's Guide to Eating in Chengdu during National Day Holidays", "36 Authentic Delicacies of Chengdu", "Top 10 Special Delicacies of Chengdu, Sichuan"	"36 Authentic Delicacies of Chengdu, "Sichuan Babayan", "One Day Tour of Neijiang, Sichuan, Eating at 15 Restaurants"
100451	City exploration	"Mysterious Quadrangle Courtyards Discovered in the Deep Mountains of Sichuan", "Smiling Pottery Figurines from the Han Dynasty at Sichuan Museum", "Hanjiang Ancient City"	"How Valuable is the Pension in Leshan?", "Read the 2,300-year-old History of Chengdu in One Go", "Welcome to Chengdu, China"	"Read the 2,300-year-old History of Chengdu in One Go", "Splendid Tianfu, Comfortable Sichuan", "200 Yuan for a Rich Tour of Deyang, Sichuan"

generated by the proposed algorithm were more in line with the user profiles and personalized compared to the other two algorithms.

REFERENCES

1. Jicheng, Z. (2021). Analysis of Short Video Production and Dissemination from the Perspective of Mobile Multimedia. *Journal of Physics: Conference Series*, *1915*(4), 1–6.
2. Liu, C. L., &Chen, Y. C. (2018). Background music recommendation based on latent factors and moods. *Knowledge-Based Systems*, *159*(NOV.1), 158–170.
3. Liu, J., Tu, Y., & Yu, W. (2026). Optimization of Course Sequence in Private Universities Based on Random Forest and Association Rules. *Engineering Intelligent Systems*, *34*(1), 119–127.
4. Liu, W., Wan, H., &Yan, B. (2023). Short Video Recommendation Algorithm Incorporating Temporal Contextual Information and User Context. *Computer Modeling in Engineering & Sciences*, *135*(1), 239–258.
5. Ma, D., Luo, L., &Fang, Y. (2021). Short video recommendations based on analytic hierarchy process and collaborative filtering algorithm. *Journal of Physics: Conference Series*, *1774*(1), 1–7.
6. Ma, T., &Sawadee, Y. (2026). Art Health Product Design and User Experience Optimization Based on Fuzzy Algorithm. *Engineering Intelligent Systems*, *34*(1), 97–108.
7. Mou, X., Xu, F., &Du, J. T. (2021). Examining the factors influencing college students' continuance intention to use short-form video APP. *Aslib Journal of Information Management*, *73*(6), 992–1013.
8. Qi, W., &Li, D. (2021). A User Experience Study on Short Video Social Apps Based on Content Recommendation Algorithm of Artificial Intelligence. *International Journal of Pattern Recognition and Artificial Intelligence*, *35*(2), 2159008.1–2159008.13.
9. Shang, S., Shang, W., Shi, M., Feng, S., &Hong, Z. (2017). A Video Recommendation Algorithm Based on Hyperlink-Graph Model. *International Journal of Software Innovation*, *5*(3), 49–63.
10. van Deen, W. K., Simpson, M., Dupuy, T. P., Khalil, C., Bonthala, N. N., &Spiegel, B. (2022). Social Media for the Dissemination of Educational Videos About Inflammatory Bowel Diseases. *The American Journal of Gastroenterology*, *117*(8), 1320–1323.
11. Wang, L., &Zhao, X. (2025). Analyse of physical exercise short video production and transmission mechanism based on big data analysis. *International Journal of Medicine and Science of Physical Activity and Sport*, *25*(99), 59–74.
12. Wang, Y., &Li, P. (2022). Development and Strategy Analysis of Short Video News Dissemination under the Background of Artificial Intelligence. *Mobile Information Systems*, *2022*(Pt.24), 2750925.1–2750925.11.
13. Yan, W., Wang, D., Liu, J., Ma, L., &Li, Z. (2019). Multi-Channel and Fusion Encoding Strategy based Auto Encoder Model for Video Recommendation. *IEEE Access*, *7*, 86004–86017.
14. Yang, F., Xie, H., &Li, H. (2019). Video associated cross-modal recommendation algorithm based on deep learning. *Applied Soft Computing*, *82*, 105597.
15. Yang, Y. (2025). Short video content and recommendation algorithm based on deep learning. *Journal of Computational Methods in Sciences and Engineering*, *25*(5), 4116–4128.
16. Yuan, D., Wang, S., Liu, Y., Zhang, M., Wei, H., &Jiang, H. (2024). Research on Short Video Personalized Recommendation Algorithm based on Attention Mechanism. *2024 International Conference on Computers, Information Processing and Advanced Education (CIPAE)*, 619–625.
17. Zhang, M., &Liu, Y. (2021). A commentary of TikTok recommendation algorithms in MIT Technology Review 2021. *Basic Research of Natural Science*, *1*(6), 846–847.