Personalized Recommendation Algorithm Based on Trajectory Mining Model in Intelligent Travel Route Planning

Jingya Shi, Qianyao Sun*
Department of Management Engineering,
Inner Mongolia Vocational and Technical College of Communications, Chifeng, 024005, China

***Abstract*—With the increasing demand for personalized travel, traditional travel route planning methods are no longer able to meet the diverse needs of users. In view of this, on the ground of the analysis of user trajectory data at the temporal and spatial levels, a new scenic spot recommendation model is proposed by combining personalized recommendation algorithms. Meanwhile, improved genetic algorithm and minimum spanning tree algorithm were introduced to adjust the structure of the personalized recommendation model. After matching the visit sequence of scenic spots, the final new personalized tourism route recommendation model was proposed. The experiment demonstrates that the optimal pause time for the personalized scenic spot recommendation model is 45 minutes, the pause distance is 15 meters, and the clustering radius is 500 meters. And the model has the highest accuracy in the Tok-10 testing environment, with a maximum value of 90%. In addition, the new personalized tourism route recommendation model has the highest accuracy of 85.6%, the highest recall rate of 88.7%, the highest F1 value of 92.4%, and an average convergence rate of 88.9%. In summary, the new scenic spot and route recommendation model proposed in the study can achieve more intelligent and personalized travel route planning, providing new guidance for the intelligent development of travel route recommendation.**

Keywords—Trajectory mining; personalized recommendations; travel routes; genetic algorithm; visiting sequence of scenic spots

I. Introduction

As the boost of intelligent technology and people's increasing pursuit of personalized experiences, personalized recommendations have become an indispensable part of smart travel route planning to enhance the travel experience of travelers [1]. Many domestic and foreign researchers have conducted varying degrees of exploration to address the problems in this field. And relevant researchers have successively developed route planning techniques using global positioning systems and geographic information systems, and proposed personalized recommendation models for travel route planning [2]. These technical models can to some extent meet the line planning requirements of users. But with the diversification of demand, these technologies have also exposed issues such as slow real-time performance, poor accuracy, and low interactivity [3]. With the continuous development of trajectory mining technology, it is widely used in location services, logistics management, and other areas due to its superior real-time data monitoring and efficient data-driven characteristics [4]. In view of this, the study attempts to innovatively introduce user data trajectory mining technology on the basis of existing personalized recommendation algorithms. By analyzing the trajectory changes of users in time and space and adjusting the structure of recommendation algorithms, a new intelligent travel route planning model can be achieved. The rationale for this initiative lies in the lack of solutions in the current market that can provide personalized travel advice by taking into account users' historical behavioral data and real-time location information. Its significance is reflected in its ability to greatly enhance the user's travel experience and plan more personalized and reasonable travel routes for travelers through intelligent data analysis. The core research questions and objectives are closely centered on how to effectively use trajectory mining techniques to achieve personalized recommendations, as well as to improve algorithms to optimize the structure of the recommendation model to improve the accuracy and applicability of real-world travel planning. This directly addresses the core challenges in the field of intelligent travel planning, such as dealing with large spatio-temporal datasets and providing personalized travel recommendations that match user needs. The study first outlines the progress and limitations of related research and clarifies the research objectives. Then, the process of constructing a personalized recommendation model based on spatio-temporal trajectories is introduced. The validity of the model is verified through experiments. Finally, the research results are summarized, its application in the field of smart travel planning is discussed, and future research directions are proposed.

II. Related Works

With the continuous development of technology, intelligence has penetrated into every aspect of people's lives, and the travel industry is no exception. Yao Z et al. found that existing tourism route planning techniques have lower planning accuracy when facing complex environments. In view of this, the research team proposed a new travel route map matching method by combining mobile phone trajectory switching data under 5G networks. The experiment demonstrates that this method has high accuracy in planning user travel routes, and can switch to view parallel roads with smaller spacing at any time through mobile phones [5]. Huang F et al. found that existing travel route planning methods mainly focus on single planning problems for specific tasks, but cannot be applied to other tasks. In view of this, the research team proposed a multi task deep travel route planning framework by combining interest attributes, user preferences, and historical route data. The experimental results show that the framework is more effective in general path planning compared to similar planning methods and better meets user needs [6]. Khamsing N et al. proposed a novel optimal decision model for family tourism route planning by combining adaptive large neighborhood search method to explore the optimal solution in daily family tourism route planning problems. The experiment demonstrates that the average total travel cost of the optimal route under this decision model is relatively low, and the average travel satisfaction is 89% [7]. Zhu S proposed a multi-objective mixed linear programming model for circular tourism to maximize the utilization of tourist attractions by cyclists and minimize the total travel time by combining multi-objective algorithms. The experiment demonstrates that the model can continuously update the optimal path plan in actual bicycle tourism path planning, greatly increasing the service level of bicycle tourism path planning [8].

With the development of position sensing technology and the popularization of smart devices, the acquisition of trajectory data has become easier. The application fields of trajectory mining technology are also becoming increasingly widespread. To achieve accurate prediction of flight delays, Shao W et al. proposed a flight prediction model combining trajectory mining technology by utilizing various vehicle trajectories and related sensor data on the airport apron. The experimental results show that the error rate of the test results of the model in the simulation environment is only 2.56% [9]. Jiang L et al. found that when trajectory data shows low quality, the map matching effect cannot achieve satisfactory results. In view of this, the research team proposed a trajectory data augmentation technique that combines deep learning. The experiment demonstrates that this technology, with its superior migration mode and high-quality trajectory data expression, performs far better than other data augmentation models of the same type. With the rapid development of the Internet and the explosive growth of information, users often feel confused and anxious when facing massive amounts of information. The emergence of personalized recommendation algorithms provides an effective solution to this problem [10]. Chen et al. found that users find it difficult to find resources of interest in large capacity interactive calligraphy experience devices. In view of this, the research team proposed a hybrid personalized recommendation algorithm that combines content and coordinated filtering. The experiment demonstrates that the algorithm can accurately predict user selection, demonstrating certain effectiveness and superiority [11]. Zou F et al. found that traditional recommendation systems only ensure the accuracy of recommendations and lose the diversity of recommendations. In view of this, the research team proposed a two-stage recommendation algorithm that combines collaborative filtering (CF) and multi-objective teaching decomposition. The experiment demonstrates that this method is highly effective and efficient on the Movielens dataset [12].

In summary, although the previous studies have made progress in the field of smart travel planning, they mainly focus on static user preferences and do not sufficiently consider the complexity of spatio-temporal data, resulting in the inability to accurately capture users' real-time behaviors. In addition, traditional recommendation algorithms suffer from inefficiency when dealing with large-scale spatio-temporal trajectory data. And the study proposes a personalized recommendation algorithm using trajectory mining aims to address these limitations. By deeply analyzing users' spatio-temporal trajectory data, the algorithm can dynamically capture changes in user preferences. Combining the improved genetic algorithm and the minimum spanning tree algorithm, the study optimizes the recommendation structure, enhances the planning efficiency and accuracy, and achieves intelligent and personalized travel route planning, overcoming the key gaps in existing research.

III. Construction of a Smart Travel Route Planning Model Combining Trajectory Model and Personalized Recommendation Algorithm

To improve the overall performance of the final smart travel route planning model, this study first mined user trajectory data and obtained a recommendation model for the user's target attractions through personalized algorithm data analysis. Secondly, on the ground of the personalized attraction recommendation model, improvements were made and the final personalized intelligent travel route planning model was proposed.

A. Construction of Personalized Recommendation Model on the Ground of Spatiotemporal Trajectory

In this era of information explosion, personalized recommendations have become an important way for the public to obtain information and enjoy services [13]. Data mining is nothing but the best personalized recommendation method, among which web scraping technology is the most classic. This technology simulates browser behavior by writing programs to automatically obtain information on the Internet. The working steps are shown in Fig. 1.

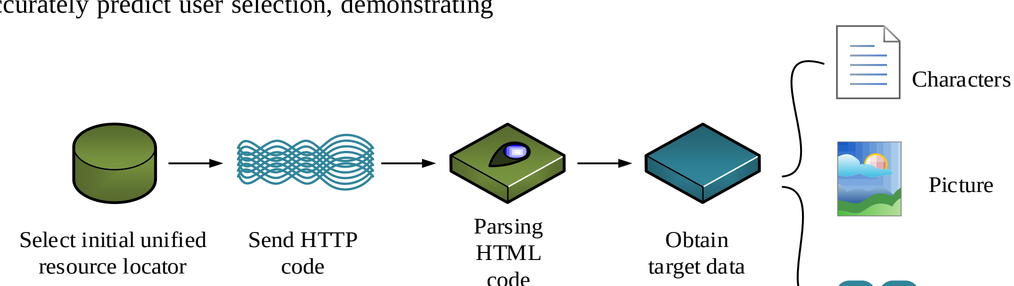


Fig. 1. Crawler technology workflow.

As shown in Fig. 1, the process of web crawling technology can be roughly divided into four steps. That is, sending HTTP code requests, parsing HTML code, extracting target data and storing it in the database. After completing the data crawler, the spatiotemporal trajectory analysis algorithm can analyze the user's temporal and spatial data, thereby strengthening the preference judgment of the user's historical data. The general spatiotemporal trajectory analysis uses Euclidean distance as a measurement unit to determine two similar data objects [14]. For different time nodes on two trajectories, it calculates the corresponding Euclidean distance meanwhile. The calculation formula for this process is shown in Eq. (1).

$$Dist(P,Q)=\sum_{i=1}^{n} dist(p_i,q_i) \tag{1}$$

In equation (1), p_i and q_i represent the nodes on the p and q trajectories at time point i, respectively, and n represents the total time point. The specific calculation of $dist(p_i,q_i)$ is shown in Eq. (2).

$$dist(p_i,q_i)=\sqrt{(p_{ix}-q_{ix})^2+(p_{iy}-q_{iy})^2} \tag{2}$$

In equation (2), (p_{ix},p_{iy}) represents the two-dimensional coordinates of node p_i, and (q_{ix},q_{iy}) represents the two-dimensional coordinates of node q_i. For the convenience of analysis, the Euclidean distance is converted into similarity calculation, as shown in Eq. (3).

$$sim(P,Q)=1-\frac{Dist(P,Q)}{\min(m,n)} \tag{3}$$

In equation (3), m and n represent the length values of trajectories p and q. Therefore, the similarity value interval after conversion can be determined as $(0,1)$, and the larger the value, the greater the similarity between the two trajectories. However, when faced with relatively large computational data, spatiotemporal trajectory analysis algorithms still face certain challenges. Therefore, the study focuses on scenic spots as recommendation objects and introduces the Mean Shift clustering algorithm to construct the user core access matrix [15]. It takes the location and time of each user's stay as a pause point, and after connecting all the pause points, it generates the user's travel path. The reasoning process of the pause point is shown in Eq. (4).

$$\begin{cases} L_i=\{l_1,l_2,l_3,\cdots,\mathrm{l_n}\} \\ \mathrm{l_n}=\{long_n,latt_n\} \end{cases} \tag{4}$$

In equation (4), L_i represents all pause records of the user, and $\mathrm{l_n}$ represents the pause location at time n. $long_n$ and $latt_n$ represent the longitude and latitude of the pause point, respectively. To avoid the calculation of the maximum number of nearest pause points, the study first excludes pause points for users with shorter time and distance, as shown in Eq. (5).

$$\begin{cases} \min long_s=\dfrac{1}{s}\sum_{i=1}^{s} long_i \\ \min latt_s=\dfrac{1}{s}\sum_{i=1}^{s} latt_i \end{cases} \tag{5}$$

In equation (5), s represents all pause points of shorter time and shorter distance. After excluding these pause points, the corresponding clustering labels are established using the Mean Shift clustering algorithm, and a user core access matrix is constructed [16]. The process of this matrix is shown in the Fig. 2.

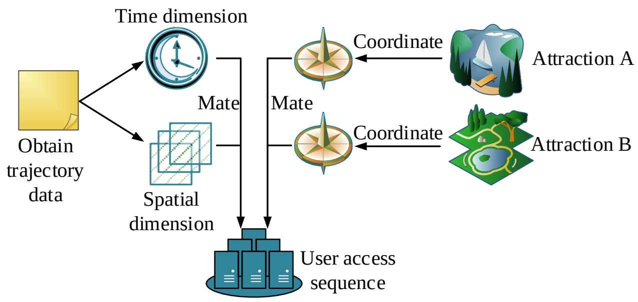


Fig. 2. User core access matrix flow.

As shown in Fig. 2, it first collects user access data and performs pause point analysis. It then establishes clustering labels for pause data in both temporal and spatial dimensions. In addition, it establishes a dataset of urban tourist attractions and classifies them through longitude and latitude coordinates [17]. Finally, it matches the pause data points after the label with tourist attractions, and the user access point at this point is the target point. It assumes that the vector of any point in the d-dimensional space after Mean Shift offset is shown in Eq. (6).

$$\begin{cases} M_h(x)=\dfrac{1}{k}\sum_{x_i\in S_h}(x_i-x) \\ S_h(x)=\{y\,|\,(y-x)(y-x)^T\le h^2\} \end{cases} \tag{6}$$

In equation (6), x_i is the i-th point, x represents any point, S_h is a low latitude sphere with a radius of h, T represents the number of iterations, and k represents the convergence center value. The entire algorithm obtains the final position that stabilizes the sphere by continuously approximating the offset vector towards any point x. The vector representation of the optimal point position at this time is shown in Eq. (7).

$$M_h(x)=\frac{\sum_{i=1}^{n} G(\frac{x_i-h}{h_i})w(x_i)(x_i-x)}{\sum_{i=1}^{n}\sum_{i=1}^{n} G(\frac{x_i-h}{h_i})w(x_i)} \tag{7}$$

In Eq. (7), h represents an element in a positive definite diagonal matrix, and h_i represents an element. $w(x_i)$ represents sample weight, $G(x)$ represents unit kernel function. To match the format of label data for the established visit sequence of scenic spots, the study abstracts any scenic spot, and the expression for this process is shown in Eq. (8).

$$\begin{cases} latt_{x`} \le latt_l \le latt_{y`} \\ long_{y`} \le long_l \le long_{x`} \end{cases} \tag{8}$$

In Eq. (8), l represents any pause point, and $(x`, y`)$ represents the coordinates of any scenic spot S_i. If the coordinate interval of the pause point is exactly within this range, it indicates that the pause point has visited scenic spot S_i. After similar frequent mining, a large number of tourist attraction visit sequences can be established, as shown in Eq. (9).

$$\varepsilon \le \frac{sum(l_{ck}^i \in vk)}{sum(ck)} \tag{9}$$

In equation (9), ε represents the critical threshold, ck represents the clustering sequence with many pause points, vk represents the clustering sequence with many pause points visiting scenic spot S_i, and l_{ck}^i represents the clustering pause point with time i. In summary, the personalized recommendation model combining user spatiotemporal trajectory data is shown in Fig. 3.

As shown in Fig. 3, the model structure can be roughly divided into five parts. The first part is the target users, who have unique ideas about tourist attractions and path planning and prefer intelligent recommendations. Secondly, through spatiotemporal trajectory analysis, the second part can be obtained, which is the user pause data sequence, which records the user's spatiotemporal historical browsing data. After analyzing these pause point data through clustering algorithms, the third part, namely the user access matrix, was obtained. Meanwhile, it establishes visit sequences for frequently followed attractions, and finally personalized recommendations are made by matching the similarity between the two.

B. Construction of Travel Route Planning Model Combining Personalized Recommendation Algorithm and Improved Genetic Algorithm

In practical life, to improve the functionality of personalized recommendation algorithms, this study not only needs to construct a recommendation model for scenic spots, but also needs to substantially propose route planning methods for these scenic spots [18]. To avoid user resistance caused by the large number of personalized recommended attractions and the scattered distribution of attractions, the study introduced the Minimum Spanning Tree (MST) clustering method to prioritize the segmentation of attractions. The schematic diagram of MST is shown in Fig. 4.

Fig. 4 shows that after MST segmentation, the distance between recommended tourist attractions within the established range is significantly reduced, enabling users to visit multiple tourist attractions within a specific time range, improving the quality of travel and saving time. In addition, unlike recommending tourist attractions, the problem of recommending tourist routes is complex and variable, that is, there are multiple possibilities for planning a route to a certain location [19]. Therefore, the study introduced an improved genetic algorithm (GA) on the ground of personalized recommendation algorithms. The traditional GA is shown in Fig. 1.

Fig. 5 shows that the traditional GA in tourism route planning can be roughly divided into nine steps, including recommending a set of tourist attractions, encoding recommended attractions, determining the population, calculating individual fitness, cross mutation, selecting the best individual, updating the population, condition judgment, and outputting the best route. Traditional GAs tend to construct initial populations randomly. If a smaller initial population appears, it will affect the convergence of subsequent algorithms, leading to lower individual fitness. Therefore, the study introduced the Greedy Algorithm to improve the initialization process of the population. The improved initialization population calculation formula is shown in Eq. (10).

$$t_i = random(T), T - S \ne \varnothing \tag{10}$$

In Eq. (10), t_i represents a randomly selected attraction, T represents a set of all attractions, and S represents the initial population. At this point, the selection of nearby attractions is shown in Eq. (11).

$$\begin{cases} returnS = S.insert(t_j) \\ t_j = \min\_dis\tan ce(T - S, t_i) \end{cases} \tag{11}$$

In Eq. (11), t_j represents other attractions that are closer to the first attraction, and $S.insert(t_j)$ represents all the best individual attractions. At this point, to stabilize the individual's fitness value, the study introduced a fitness function with multiple constraints. This function includes two parts: the shortest route and the optimal access time. The fitness function calculation for the shortest route problem is shown in Eq. (12).

$$f_1(C) = \sum_{i=1}^{n`-1} d(c_i, c_{i+1}) + d(c_{n`}, c_1) \tag{12}$$

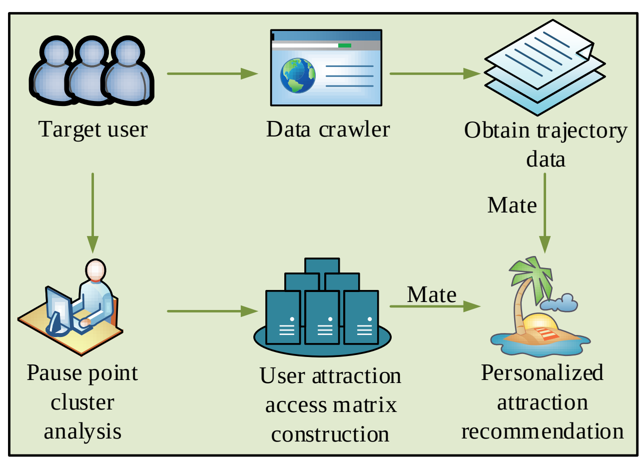


Fig. 3. Personalized recommendation model structure diagram.

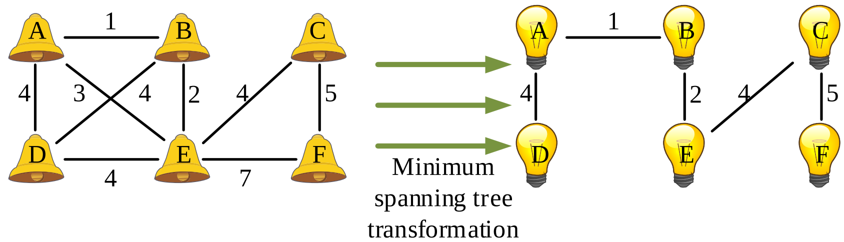


Fig. 4. MST schematic diagram.

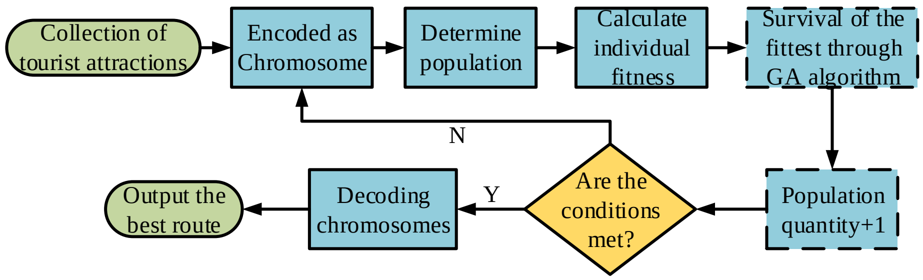


Fig. 5. The process of traditional GA.

In Eq. (12), C represents the collection of scenic spots, n\` represents the number of paths that satisfy user preferences, and $d(c_i, c_{i+1})$ represents the sum of distances between c_i and c_{i+1} paths. The optimal visit time needs to be planned in conjunction with the designated opening hours of the scenic spots. Due to the different opening hours of each scenic spot, the study divided the optimal visit time into three different visit times as dividing points [20]. The expression is shown in Eq. (13).

$$R(ci,t) = \begin{cases} \dfrac{2t-2t_{io}}{t_{ic}-t_{io}}, t_{io} \le t \le \dfrac{t_{io}+t_{ic}}{2} \\ \dfrac{2t_{ic}-2t}{t_{ic}-t_{io}}, \dfrac{t_{io}+t_{ic}}{2} \le t \le t_{ic} \\ 0, t < t_{io} \;\&\; t > t_{ic} \end{cases} \tag{13}$$

In Eq. (13), t_{io} represents the opening time of the attraction, t_{ic} represents the closing time of the attraction, c_i represents the attraction, and the user's visit time is t_i. Although the user's stay time at each attraction cannot be estimated, the visit time can be calculated on the ground of the popularity of the attraction. The calculation formula for conversion estimation is shown in Eq. (14).

$$t_{is} = T_N + \frac{count(c_i)}{\max(count(c_1), count(c_2), \cdots count(c_n))} \tag{14}$$

In equation (14), $count(c_i)$ represents the number of popular searches for attraction c_i, and T_N represents the time constant. By using this formula, the visit time and stay time of each attraction in a set of attractions C can be calculated. The optimal fitness function of the GA at this time is shown in Eq. (15).

$$\begin{cases} f_2(C) = \dfrac{1}{n}\sum_{i=1}^{n} R(c_i, t) \\ f(C) = \alpha f_1(C) + \beta f_2(C) \end{cases} \tag{15}$$

In Eq. (15), α and β represent the limiting weights of the fitness function for the shortest route and optimal access time, respectively. The best individual selected, crossed, mutated, and determined through GA is the optimal travel planning route for user needs. In summary, a new intelligent travel route recommendation model has been proposed by combining personalized recommendation algorithms and improved GAs. The structure of the model is shown in Fig. 6.

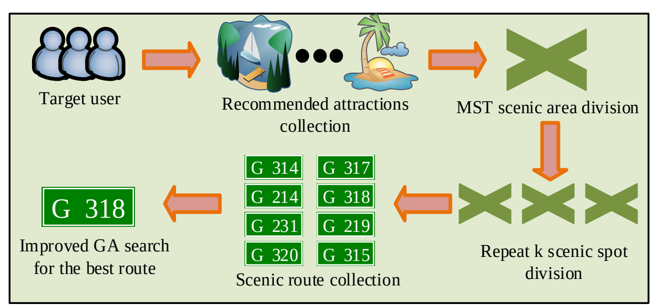


Fig. 6. Smart travel route recommendation model structure diagram

Fig. 6 shows that the entire smart travel route planning model consists of six parts. Firstly, it obtains the trajectory data information of the target user and their preferences. Secondly, it filters the target node set that best meets the user's needs through a personalized recommendation model. Then, the MST algorithm is used to divide the collection of scenic spots into regions, to reduce the recommended target. After k repeated reductions, multiple path planning schemes were finally obtained to determine the distance between scenic spots. Finally, an improved GA algorithm was used to select the optimal path set and output the best route.

IV. RESULTS

To verify the performance of the proposed new intelligent personalized recommendation model, this study first tested the scenic spot recommendation model and determined the optimal operating parameters of the algorithm. Then it was compared with similar recommendation algorithms. In addition, the new tourism route recommendation model was tested to determine its optimal iteration times and fitness function values. It was also compared with similar recommendation models.

A. Test Results of Scenic Spot Recommendation Model

To verify the performance of the personalized scenic spot recommendation model proposed in the study, which combines trajectory data mining, the Windows 10 operating system was used, with an Intel Core 2.5Hz dual core CPU and 16GB of memory. To ensure the authenticity of the test, this study used a global dataset of tourist attractions and routes. This dataset contains information on various tourist attractions and related routes from around the world, totaling approximately 100000 pieces. It divides the dataset into training and testing sets in an 8:2 ratio, and the training set sample data is used to train personalized recommendation models. The study introduced two variables, dwell time and dwell distance, to analyze the pause points of users. Meanwhile, to prevent model training caused by too large or too small variables, the study sets the dwell time to within 1 hour and the dwell distance to within 50 meters. In addition, the Mean Shift clustering radius is used as a variable to detect changes in the number of clusters. The specific test results are shown in Fig. 7.

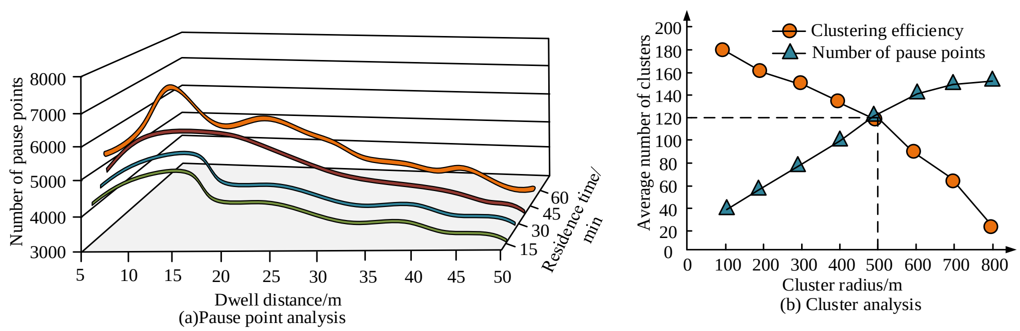


Fig. 7. Pause point analysis and cluster analysis parameter testing.

Fig. 7(a) shows the changes in the pause time and pause distance parameters of the pause point, and Fig. 7(b) shows the changes in the clustering radius parameters of the clustering analysis. As shown in Fig. 7, the pause change curve is most stable when the pause time is 45 minutes, and the maximum number of pauses at this time is 6000 when the pause distance is 18 meters. In addition, as the clustering radius increases, the average number of clusters gradually decreases, while the number of pause points gradually increases. When the clustering radius is 500m, that is, at the intersection, the number of the two performs best. Therefore, in subsequent research, the set parameters include a pause time of 45 minutes, a pause distance of 15 meters, and a clustering radius of 500 meters. To verify the performance difference between the personalized recommendation model proposed in the study and existing models of the same type, Tok-k accuracy was used as a reference indicator. This indicator represents the proportion of the top k results with the highest probability in the prediction results that contain correct labels, for example, Tok-5 is a test environment with 5 recommended sets. Meanwhile, the content attribute personalized recommendation model (Item based), rating personalized recommendation model (Mark based), and image personalized recommendation model (Graphics based) were introduced. The test results are shown in Fig. 8.

Fig. 8 shows that the personalized recommendation method proposed in the study generally has high accuracy in three testing environments. The highest accuracy rate in Tok-5 is 88% for the study of the proposed model, 90% for the study of the proposed model in Tok-10, and 84% for the study of the proposed model in Tok-15. In Tok-10, the accuracy of the proposed model is 33% higher than that of the content attribute personalized recommendation model and 31% higher than that of the rating personalized recommendation model. In summary, it can be concluded that the personalized scenic spot recommendation model proposed in the study, which combines temporal and spatial data trajectory mining, has the best performance. In addition, with accuracy, recall, F1 value, and recommendation similarity as reference indicators, comparative tests were continued on the four models, and the test results are shown in Table I.

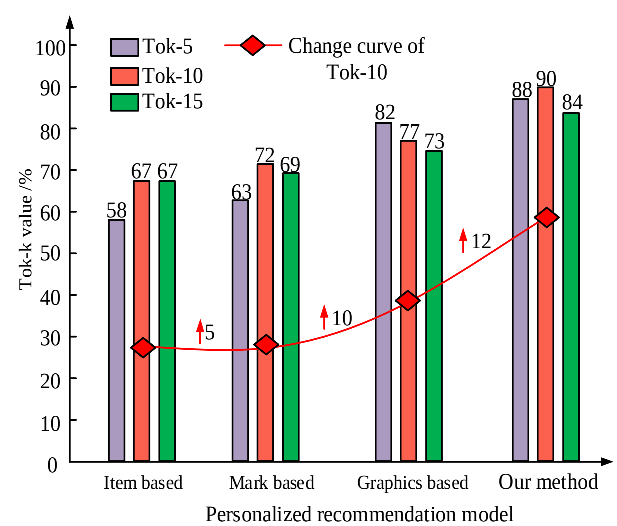


Fig. 8. Comparison results of Tok-k accuracy of different recommendation models.

TABLE I. COMPARATIVE TEST RESULTS OF PERSONALIZED RECOMMENDATION MODELS

Model	Precision/%	Recall/%	F1/%	Recommended similarity/%
Item based	64.8	65.3	67.1	68.5
Mark based	68.4	71.6	74.8	77.9
Graphics based	74.2	75.3	76.7	79.5
Our method	82.4	84.6	85.9	89.3

As can be seen from Table I, the Item based model has generally low indicators in various categories, with its highest P value of 64.8%, highest R value of 65.3%, highest F1 value of 67.1%, and highest recommendation similarity of 68.5%. In contrast, the study proposes that the recommendation model performs the best, with the highest accuracy of 82.4%, the highest recall of 84.6%, the highest F1 value of 85.9%, and the highest recommendation similarity of 89.3%. The values rose 17.6%, 19.3%, 18.8% and 20.8% respectively compared to the lowest Item based model. In summary, it once again proves that the model proposed in the study is more in line with user choices, with higher recommendation and usage rates.

B. Test Results of Route Recommendation Model

Due to the introduction of an improved GA on the ground of personalized recommendation algorithms, this study first tested the iterative performance of the improved GA. It determines the optimal number of iterations and fitness function values to facilitate subsequent model testing. In addition, to enhance the reliability of testing, traditional GAs, CF, and Particle Swarm Optimization (PSO) were introduced in the study. These algorithms were tested separately on the training and testing sets, and the test results are shown in Fig. 9.

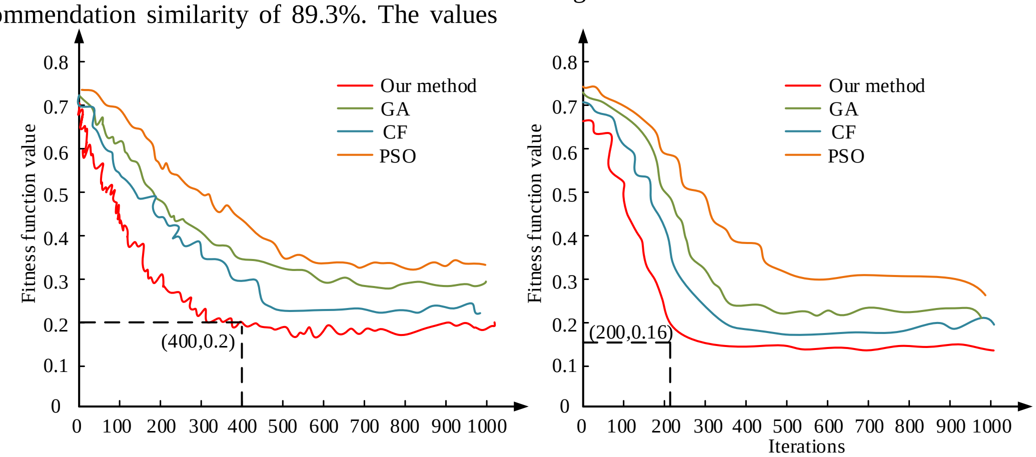


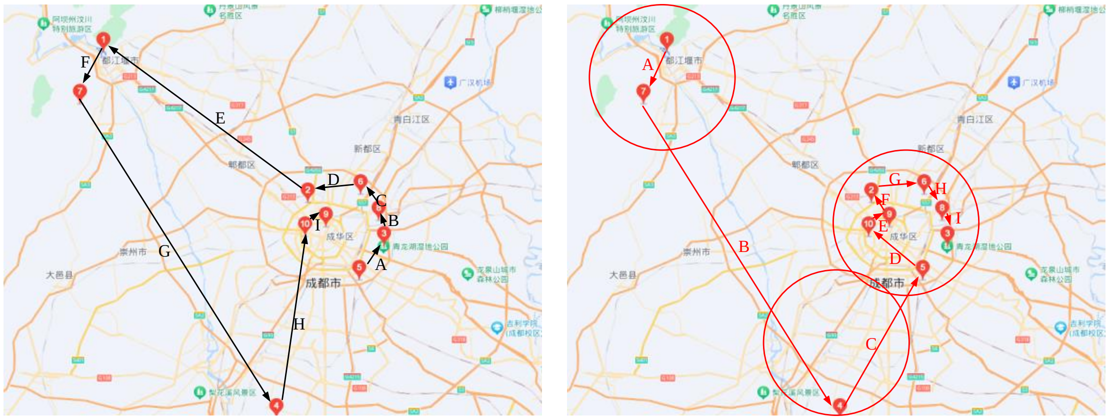
Fig. 9. Iterative performance testing of different optimization algorithms.

Fig. 9(a) shows the iterative performance test results of four optimization algorithms in the training set. Fig. 9(b) shows the iterative performance test results of four optimization algorithms in the test set. Fig. 9 shows that as the number of iterations increases, the fitness function values of all four algorithms decrease, but then tend to stabilize. The results of the training set show that the lowest fitness function value of the proposed model is 0.2 when the number of iterations is 400. In the test set, when the number of iterations is 200, the fitness function value at this point is as low as 0.16. Therefore, it is necessary for subsequent research to use the number of iterations and fitness function values as benchmarks for performance testing of similar algorithms. It conducts 50 repeated experiments on four algorithms and takes the average of the ratio of the optimal solutions obtained each time as the average convergence degree of the algorithm. Meanwhile, it continues to introduce three reference indicators: accuracy, recall, and F1 value. The test results are shown in Table II.

As can be seen in Table II, the GA algorithm has the worst performance in the metrics test, with the highest accuracy of 54.2%, the highest recall of 57.6%, the highest F1 value of 64.5%, and an average convergence of 58.8%. This is followed by PSO algorithm and CF algorithm, while the personalized recommendation algorithm proposed in the study has the highest accuracy of 85.6%, the highest recall of 88.7%, the highest F1 value of 92.4%, and the average convergence of 88.9%. To more accurately reflect the performance of the personalized tourism route recommendation model proposed in the study, the top 10 popular tourist attractions in Chengdu were selected as the target locations. They are 1) Dujiangyan Irrigation Project Water Conservancy Project, 2) Chengdu Happy Valley, 3) Qinglong Lake Park, 4) Huanglongxi Ancient Town, 5) Sansheng Flower Town Scenic Spot, 6) Giant Panda Base, 7) Qingcheng Mountain, 8) Eastern Suburb Memory, 9) Wenshu Academy and 10) Du Fu Thatched Cottage. The actual results of comparing the traditional personalized recommendation route and the new personalized recommendation route are shown in Fig. 10.

TABLE II. PERFORMANCE TEST RESULTS OF DIFFERENT OPTIMIZATION ALGORITHMS

Model	Precision/%	Recall/%	F1/%	Mean Convergence %
GA	54.2	57.6	64.5	58.8
PSO	68.4	73.2	79.5	73.7
CF	74.2	77.7	83.8	78.6
Our method	85.6	88.7	92.4	88.9



(a)Traditional personalized recommendation routes (b)New personalized recommendation route

Fig. 10. Comparison results of personalized recommended routes.

Fig. 10(a) shows the traditional personalized recommendation route, and Fig. 10(b) shows the new personalized recommendation route. As shown in Fig. 10, the traditional personalized recommendation route connects 10 scenic spots in pairs, resulting in a total of nine moving paths, represented by black arrows. The red arrow represents the new recommended path, while the red circle tends to indicate the clustering range of scenic spots. Considering the actual situation, the maximum daily travel itinerary for users is three scenic spots. Therefore, under traditional methods, the longest crossing path for the same day's itinerary takes more time and is not conducive to users choosing fixed accommodation. The new personalized recommendation path can select at least one and at most seven stopping points within the established stopping range, and set accommodation at the center of each circle for the most convenient. In summary, the new personalized tourism path recommendation model proposed in the study performs better.

V. Conclusion

Traditional travel planning methods are usually static, only considering the user's departure and destination, without considering the user's actual behavior and real-time location. In view of this, the study established a personalized recommendation model through mean shift clustering and trajectory analysis after decomposing user trajectory data in time and space. After introducing an improved GA, a new intelligent personalized tourism route recommendation model was proposed. The experimental results show that when the pause time is 45 minutes, the pause distance is 15 meters, and the clustering radius is 500 meters, the performance of the personalized scenic spot recommendation model is the best. Compared to personalized recommendation models of the same type, the proposed model has the highest accuracy in the Tok-10 testing environment, with a maximum value of 90%. The highest model accuracy is 82.4%, the highest recall is 84.6%, the highest F1 value is 85.9%, and the highest recommended similarity is 89.3%. In addition, testing the personalized travel recommendation route model found that compared to other models, the new route recommendation model proposed in this study has the lowest iteration number of 200 and a fitness function value of 0.16. The highest accuracy of model recommendation is 85.6%, the highest recall is 88.7%, the highest F1 value is 92.4%, and the average convergence is 88.9%. Simulation tests have shown that the new model can plan more reasonable and suitable routes for the public's actual tourism, save time on route expenses, and facilitate accommodation arrangements. In summary, the new personalized attraction and route recommendation model proposed in the study can improve the effectiveness and experience of travel planning and meet user needs. However, this study only analyzed user trajectories from time and space sequences. Further research can add more user characteristic information analysis, such as subjective requirements and preferences, to enhance the completeness of the study.

VI. Discussion

The study successfully constructed a novel attraction recommendation model and travel route recommendation model by introducing improved genetic algorithm and minimum spanning tree algorithm. These two personalized recommendation models perform well in several performance metrics, highlighting their potential in the field of intelligent travel route planning. First, the optimal stopping time of the personalized attraction recommendation model is set to 45 minutes, the stopping distance is 15 meters, and the clustering radius is 500 meters when these parameters are optimized to ensure that the model can accurately capture the actual user behaviors and deviations. In particular, the model achieves 90% accuracy in the Tok-10 test environment, which is much higher than traditional personalized recommendation models, such as content attribute personalized recommendation model, rating personalized recommendation model, and image personalized recommendation model. This result emphasizes the importance of spatio-temporal trajectory data analysis in improving recommendation accuracy, and also demonstrates that the performance of recommender systems can be significantly improved by fine-grained user behavior analysis. In addition, the personalized travel route recommendation model has the highest accuracy of 85.6%, recall of 88.7%, F1 value of 92.4%, and average convergence of 88.9%. These metrics not only reflect the model's efficiency and accuracy in the field of travel route recommendation, but also show the effectiveness of the improved genetic algorithm in dealing with complex route planning problems. By introducing the greedy algorithm to optimize the initial population and adopting the fitness function with multiple constraints, the study successfully improves the convergence speed of the algorithm and the quality of recommendations. However, there are still some potential challenges and limitations of these models. First, although the studies have achieved significant results in specific datasets and environments, their generalization capabilities still need to be validated in a wider range of scenarios and complex user behavior patterns. In addition, the performance of the model relies heavily on high-quality spatio-temporal trajectory data, and thus, it may face challenges of privacy protection and data security in data collection and processing. In response to the above discussion, future research could further explore the application of the model in different cultural and geographic contexts to validate its generalization ability. Second, considering the importance of data privacy and security, future work should focus more on the anonymization of user data and the application of encryption techniques. Finally, given the diversity and dynamics of user preferences, the development of more flexible and adaptive models will be the key to improving the accuracy and user satisfaction of recommender systems.

Acknowledgment

The research is supported by: Special ideological and political project of scientific research projects in Colleges and universities of the autonomous region in 2021, Research on the teaching method of integrating the red tour into the ideological and political course of Tourism Specialty in Higher Vocational Colleges, (NO. NJSZZX2174).

References

[1] Sardianos C, Varlamis I, Chronis C, Dimitrakopoulos G. The emergence of explainability of intelligent systems: Delivering explainable and personalized recommendations for energy efficiency. International Journal of Intelligent Systems, 2021, 36(2): 656-680.

[2] Nitu P, Coelho J, Madiraju P. Improvising personalized travel recommendation system with recency effects. Big Data Mining and Analytics, 2021, 4(3): 139-154.

[3] Mou N, Jiang Q, Zhang L. Personalized tourist route recommendation model with a trajectory understanding via neural networks. International Journal of Digital Earth, 2022, 15(1): 1738-1759.

[4] Koc E, Yazici Ayyildiz A. An overview of tourism and hospitality scales: Discussion and recommendations. Journal of Hospitality and Tourism Insights, 2022, 5(5): 927-949.

[5] Yao Z, Wang Y, Yang F, Chen Y, Ran B. Map matching for travel route identification based on Earth Mover's Distance algorithm using wireless cell trajectory data. Journal of Intelligent Transportation Systems, 2021, 25(1):644-656.

[6] Huang F, Xu J, Weng J. Multi-task travel route planning with a flexible deep learning framework. IEEE Transactions on Intelligent Transportation Systems, 2020, 22(7): 3907-3918.

[7] Khamsing N, Chindaprasert K, Pitakaso R, Sirirak W, Theeraviriya C. Modified ALNS algorithm for a processing application of family tourist route planning: A case study of Buriram in Thailand. Computation, 2021, 9(2): 23-24.

[8] Zhu S. Multi-objective route planning problem for cycle-tourists. Transportation Letters, 2022, 14(3): 298-306.

[9] Shao W, Prabowo A, Zhao S, Koniusz P, Salim F D. Predicting flight delay with spatio-temporal trajectory convolutional network and airport situational awareness map. Neurocomputing, 2022, 472(2):4280-293.

[10] Jiang L, Chen C, Chen C. With deep models for low-quality gps trajectory data. ACM Transactions on Knowledge Discovery from Data, 2023, 17(3): 1-25.

[11] Chen S, Huang L, Lei Z, Wang S. Research on personalized recommendation hybrid algorithm for interactive experience equipment. Computational Intelligence, 2020, 36(3):1348-1373.

[12] Zou F, Chen D, Xu Q, Jiang Z, Kang J. A two-stage personalized recommendation based on multi-objective teaching–learning-based optimization with decomposition. Neurocomputing, 2021, 452(9):716-727.

[13] Fang Z, Pan L, Chen L, Du Y, Gao Y. MDTP: A multi-source deep traffic prediction framework over spatio-temporal trajectory data. Proceedings of the VLDB Endowment, 2021, 14(8): 1289-1297.

[14] Chen J, Yang J, Huang J. Robust Truth Discovery Scheme Based on Mean Shift Clustering Algorithm. Journal of Internet Technology, 2021, 22(4): 835-842.

[15] Mehdi G, Hooman H, Liu Y, Peyman S, Arif R. Data Mining Techniques for Web Mining: A Survey. Artificial Intelligence and Applications, 2022, 1(1):3-10.

[16] Zhou X, Su M, Liu Z. Smart tour route planning algorithm based on naïve Bayes interest data mining machine learning. ISPRS International Journal of Geo-Information, 2020, 9(2): 112-113.

[17] Ntakolia C, Iakovidis D K. A route planning framework for smart wearable assistive navigation systems. SN Applied Sciences, 2021, 3(1): 104-105.

[18] Huang F, Xu J, Weng J. Multi-task travel route planning with a flexible deep learning framework. IEEE Transactions on Intelligent Transportation Systems, 2020, 22(7): 3907-3918.

[19] Ho R C, Amin M, Ryu K. Integrative model for the adoption of tour itineraries from smart travel apps. Journal of Hospitality and Tourism Technology, 2021, 12(2): 372-388.

[20] Hu W C, Wu H T, Cho H H. Optimal route planning system for logistics vehicles based on artificial intelligence. Journal of Internet Technology, 2020, 21(3): 757-764.