引文格式：卫晨,杨丽琼,徐浩然. 基于自适应遗传算法的个性化旅游景点智能推荐方法[J]. 地理空间信息, 2025, 23(11): 65-69. doi:10. 3969/j. issn. 1672-4623.2025.11.013

基于自适应遗传算法的个性化旅游景点智能推荐方法

卫　晨[1,2]，杨丽琼[2]，徐浩然[1]
（1. 巢湖学院，安徽 合肥 238024；2. 中国旅游研究院，北京 100005）

摘　要：针对目前用户旅游景点推荐存在考虑因素不全面、推荐性能低的问题，设计了数据收集、构建用户旅行历史和基于自适应遗传算法的景点智能推荐框架；建立了综合考虑强制性POI数量、POI总数、用户访问持续时间、POI受欢迎程度、整体评分、成本等因素的个性化旅游景点推荐目标函数；提出了具有权重调整机制的自适应遗传算法。实验结果表明，与MaxN、MaxP方法相比，该方法在强制性POI数量、POI总数、POI受欢迎程度、总时间、成本和整体评分等指标上的综合性能更优，验证了方法的有效性，可为改善用户旅行体验和个性化服务订制提供助力。

关键词：行程推荐；智能化旅游；机器学习；遗传算法；自适应；目标函数
中图分类号：P208　　**文献标志码：**B　　**文章编号：**1672-4623（2025）11-0065-05

Intelligent Recommendation Method of Personalized Tourist Attractions Based on Adaptive Genetic Algorithm

WEI Chen[1,2], YANG Liqiong[2], XU Haoran[1]
(1. Chaohu University, Hefei 238024, China;
2. China Tourism Academy, Beijing 100005, China)

Abstract: Aiming at the problems of incomplete consideration and low recommendation performance in the current user recommendation of tourist attractions, we designed an intelligent recommendation framework of tourist attraction based on data collection, user travel history construction and adaptive genetic algorithm, established a objective function of personalized tourist attraction recommendation comprehensively considering the mandatory POIs number, the total number of POIs, the duration of user access, the popularity of POIs, the overall score, cost and other factors, and proposed an adaptive genetic algorithm with weight adjustment mechanism. The experimental results show that compared with MaxN and MaxP methods, the proposed tourist attraction recommendation method has better comprehensive performance in mandatory POIs number, the total number and popularity of POIs, total time, cost and overall score, verifying the effectiveness of proposed method, which can provide help for improving user travel experience and personalized service customization.

Key words: itinerary recommendation, intelligent tourism, machine learning, genetic algorithm, adaptive, objective function

随着网络、社交媒体技术的飞速发展[1-2]，越来越多的旅行用户开始分享其兴趣点（POI）并对旅行体验进行评论。通过研究这些社交媒体数据，可了解用户的个性化旅游需求，进而推荐旅游行程。近年来，旅游行程推荐系统是计算机科学、运筹学、图论等相关专业领域的一个热门话题[3-4]。部分学者[5-7]从成本预算、时间限制和起点终点景点等角度考虑并解决了用户旅游景点推荐问题；但个性化旅游景点智能推荐还应考虑游客感兴趣的相关因素，如强制性景点、景点之间的距离和游览时长等。为提高旅游推荐服务质量，本文提出了一种考虑用户个性化需求的旅游景点智能推荐方法，综合考虑了强制POI数量、POI受欢迎程度、POI评分、起点/终点POI、各POI之间的距离、成本预算、时间限制等用户个性化需求，为各因素分配了不同的权重，并基于自适应遗传算法（AGA）生成个性化的行程推荐，从而更好地满足用户的多种偏好。

1　模型框架

个性化旅游景点智能推荐模型框架包括数据收集、用户旅行历史构建和基于AGA的景点智能推荐3个部分。数据收集主要从Flickr、微博、小红书、知

项目来源：巢湖学院2023年博士科研启动基金资助项目（KYQD-2023069）；贵州省高校哲学社会科学实验室资助项目（黔教哲[2023]08号）；宿州学院2021年博士科研启动基金资助项目（2021BSK027）；2025年全省青少年成长与发展理论研究重点课题资助项目（2025AHQSNCZFZ038）。
第一作者简介：卫晨（1990—），博士，工程师，研究方向为文化旅游、乡村旅游、文化旅游，E-mail：weicheng369852@163.com。

乎、旅游论坛等社交媒体或网络中抓取带有地理标记的照片，确定每个城市中各POI的访问量，并从百度百科中获取POI的相关信息，包括但不限于POI的评分、POI之间的旅行时间和距离、访问成本、用户访问持续时间、POI受欢迎程度。用户旅行历史构建是在收集了足够的POI数据后，整合多个用户的数据进而统计每个POI的访问频率、平均评分、平均访问持续时间等信息，用于后续的个性化旅游推荐。基于AGA的景点智能推荐专注于满足用户的个性化旅游需求，是本文研究的重点内容。

2 景点智能推荐方法设计

2.1 相关概念

本文将用户的旅行轨迹定义为一个有序序列[8] $T_{\text{raj}}=((q_1,a_{q_1},d_{q_1}),\cdots,(q_k,a_{q_k},d_{q_k}))$，其中 q_i 为访问的POI，a_{q_i}、d_{q_i} 为到达和离开 q_i 的时间。此外，每个 q_i 都有两个属性 $l_{q_i}^{\text{at}}$ 和 $l_{q_i}^{\text{ng}}$，即 q_i 的经纬度坐标。将用户的旅行记录定义为集合 $R_{\text{eco}}(u_i)=((q_1,f_{q_1}^{p},r_{q_1}),\cdots,(q_k,f_{q_k}^{p},r_{q_k}))$，其中 $f_{q_i}^{p}$ 为用户 u_i 拍摄的照片数量，r_{q_i} 为用户 u_i 给出的评分。将两个POI之间的地理空间距离定义为 $D_{\text{ist}}(q_i,q_j)$，通过其地理坐标 $l_{q_i}^{\text{at}}$、$l_{q_i}^{\text{ng}}$、$l_{q_j}^{\text{at}}$ 和 $l_{q_j}^{\text{ng}}$ 计算。需注意，为简化模型本文只取绝对值，即 $D_{\text{ist}}(q_i,q_j)=D_{\text{ist}}(q_j,q_i)$。POI的访问持续时间是所有访问过该POI的用户的平均访问持续时间。令 $U=u_1,u_2,\cdots,u_m$ 为 q_i 的用户集合（m 为用户数量），则 q_i 的访问持续时间可定义为：$V_{\text{dur}}(q_i)=\frac{1}{m}\sum_{j=1}^{m}(d_{q_i}(u_j)-a_{q_i}(u_j))$，其中 $d_{q_i}(u_j)$、$a_{q_i}(u_j)$ 分别为用户 u_j 离开、到达的时间。将两个POI的旅行时间定义为 $D_{\text{ist}}(q_i,q_j)$ 与旅行速度 p_s 的比值。给定用户旅行记录的数量 m，则POI照片频率可表示为 $F_{\text{req}}(q_i)=\sum_{j=1}^{m}f_{q_i}^{p}(u_j)$。照片频率可直接反映POI的受欢迎程度[9]，即照片频率越大，该POI越受欢迎。为统一量纲，本文采用照片频率的对数衡量POI受欢迎程度，即 $P_{\text{opu}}(q_i)=\log(1+F_{\text{req}}(q_i))$。$q_i$ 的成本 $C_{\text{ost}}(q_i)$ 为进入 q_i 所需支付的费用。POI总体评分是所有用户评分的平均值。

本文将个性化旅游景点智能推荐问题定义为多目标优化[10-11]问题，即将行程定义为一条路径，该路径由指定的起始POI和结束POI连接，并包含至少一个其他POI，且行程中的所有POI只能访问一次。此外，本文还考虑了个性化旅游中存在的强制性POI。令POI集合为 $C=\{q_1,q_2,\cdots,q_L\}$，强制性POI集合为 $M=\{m_1,m_2,\cdots,m_K\}$，$K<L$。在理想情况下，行程可表示为 $I=\{q_s,\cdots,m_1,\cdots,m_K,\cdots,q_d\}$，其中 q_s、q_d 为起始、结束POI，$q_s,q_d\notin M$。个性化旅游景点推荐[12]的目标是推荐一个包含尽可能多的强制性POI和其他POI的旅游路径，使用户的访问持续时间、POI受欢迎程度、整体评分在时间预算内最大化，并使用户在POI的成本尽可能低。

2.2 AGA算法

GA算法[13-15]是一种典型的群智能优化算法，无法较好地解决多目标问题。为解决个性化旅游景点推荐问题，本文提出了具有交叉和变异概率的AGA算法。令种群集合 $P=\{p_1,p_2,\cdots,p_N\}$，N 为种群数量，$p_i=\{q_s,\cdots,q_d\}$ 为POI的行程序列，即每个POI为种群中一个基因。由于各个体的基因序列长度可能不同，因此本文采用归一化方法将所有值限制在[0,1]内。AGA算法的适应度函数可表示为：

$$f(p_i)=\omega_1T_m(p_i)+\omega_2T_n(p_i)+\omega_3T_p(p_i)+\omega_4T_v(p_i)+\omega_5T_c(p_i)+\omega_6T_r(p_i) \tag{1}$$

式中，ω_i 为各因素的权重，可根据用户偏好进行调整；$T_m(p_i)$、$T_n(p_i)$、$T_p(p_i)$、$T_v(p_i)$、$T_c(p_i)$、$T_r(p_i)$ 分别为与强制POI数量、POI总数、POI受欢迎程度、总旅行时间、成本、整体评分相关的子适应度函数。

$$T_m(p_i)=\frac{N_m(p_i)}{N_M} \tag{2}$$

$$T_n(p_i)=\frac{N_m(p_i)}{N_P} \tag{3}$$

$$T_p(p_i)=\frac{T_{\text{otp}}(p_i)}{\max P_{\text{opu}}} \tag{4}$$

$$T_{\text{otp}}(p_i)=P_{\text{opu}}(q_s)+\cdots+P_{\text{opu}}(q_d) \tag{5}$$

$$T_v(p_i)=\frac{T_{\text{otv}}(p_i)}{\max T_{\text{rav}}} \tag{6}$$

$$T_{\text{otv}}(p_i)=V_{\text{dur}}(q_s)+\cdots+V_{\text{dur}}(q_d) \tag{7}$$

$$T_c(p_i)=\frac{T_{\text{otc}}(p_i)}{\max C_{\text{ost}}} \tag{8}$$

$$T_{\text{otc}}(p_i)=T_{\text{otc}}(q_s)+\cdots+T_{\text{otc}}(q_d) \tag{9}$$

$$T_r(p_i)=\frac{T_{\text{otr}}(p_i)}{\max R_{\text{ati}}} \tag{10}$$

$$T_{\text{otr}}(p_i)=R_{\text{ati}}(q_s)+\cdots+R_{\text{ati}}(q_d) \tag{11}$$

式中，$N_m(p_i)$、N_M 为行程 p_i 和总行程中包含的强制POI数量；N_P 为总行程中包含的POI数量；$\max P_{\text{opu}}$ 为行程中POI受欢迎程度最大值；$\max T_{\text{rav}}$ 为最长旅

行时间；$\max C_{ost}$ 为最大成本；$\max R_{ati}$ 为最大评分。

AGA算法中交叉和变异概率是动态的，可防止程序陷入局部最佳解。交叉、变异概率的定义为：

$$\xi_c=\begin{cases}p_{c1}-\dfrac{(p_{c1}-p_{c2})(f'-f_{avg})}{f_{max}-f_{avg}} & f'\geqslant f_{avg}\\ p_{c1} & f'<f_{avg}\end{cases} \tag{12}$$

$$\xi_m=\begin{cases}p_{m1}-\dfrac{(p_{m1}-p_{m2})(f_{max}-f)}{f_{max}-f_{avg}} & f\geqslant f_{avg}\\ p_{m1} & f<f_{avg}\end{cases} \tag{13}$$

式中，p_{c1}、p_{c2} 为[0,1]之间固定参数；f' 为两个个体的较大适应度得分，这两个个体将通过交叉产生下一代；f_{max} 为适应度最高分；f_{avg} 为适应度平均分；p_{m1}、p_{m2} 为[0,1]之间固定参数；f 为将要变异个体的适应度得分。

AGA算法迭代过程中会将一些未访问的POI通过自适应交叉和变异插入到下一代的每个个体中，从而提高解决方案的质量；但这些操作可能引起总旅行时间的增加，因此本文定义总旅行时间约束为：

$$T_{time}(p_i)=T_{rav}(q_s,q_{s+1})+\cdots+T_{rav}(q_{d-1},q_d)+T_v(p_i)\leqslant T_{con} \tag{14}$$

式中，T_{con} 为用户自定义的时间阈值，设定时需综合考虑用户需求、个人旅行习惯、可用时间窗口和目的地间实际距离。

通过多次试验，基于经验法确定AGA算法的参数取值，设置 p_{c1} 为0.9、p_{c2} 为0.6、p_{m1} 为0.1、p_{m2} 为0.001、ω_1 为10、ω_2 为5、ω_3 为2、ω_4 为2、ω_5 为1、ω_6 为1。AGA算法执行流程为：①基于决策空间降维策略删除部分劣质解，从而降低种群搜索难度，并获得降维搜索空间；②对于每个推荐行程，基于初始化策略生成初始种群；③自适应基因操作，通过交叉、变异操作生成新的种群；④计算种群中所有个体的适应度函数值，采用非支配排序和拥挤距离排序获得的新种群并替换初始种群；⑤判断算法是否满足退出条件，若满足则终止算法，并输出推荐结果，否则返回至自适应基因操作。

3　实验与结果分析

3.1　数据集与实验环境

本文采用的数据为从Flickr、微博、小红书等创意共享中爬取的图像，包含约10 000张照片和600个大小不等的视频。本文选取了上海、重庆、云南、武汉、深圳和北京6个具有代表性的旅游特色地区，每个地区随机生成200个行程（该值由每个地区的数据集大小和为确保统计结果的显著性确定）。针对云南等地域辽阔、景点分散的地区，设定较高的时间阈值，以确保推荐行程的实用性与可行性。所有区域的POI列表与评分从微博、TripAdvisor中爬取，POI之间的距离通过高德地图估算，旅行时间阈值通过综合用户历史信息估计。此外，实验中每个行程的起始和结束POI以及强制性POI均为随机生成并记录，且每个地区随机生成若干个行程。用户历史旅游数据统计见表1。

表1　用户历史旅游数据整理统计表

省市	POI数量	用户数量	POI访问数量	照片数量
上海	38	3 848	18 448	36 124
重庆	29	4 854	32 894	82 165
云南	28	4 048	39 449	15 757
武汉	25	1 598	11 425	29 025
深圳	21	424	3 655	18 355
北京	25	958	7 751	32 421

实验通过Java和Matlab混合语言对个性化旅游景点智能推荐模型进行编译，部分参数以xml配置文件读取。实验的硬件环境为Windows10操作系统的台式服务器，配置为Intel i7-8750H 2.20 GHz CPU、32.0 GB RAM、2 T硬盘。

3.2　基准算法与评估指标

将本文方法与不同基准方法进行比较，以评估其推荐性能。本文方法首先生成 N 个行程作为初始种群 P，对于每个 $p_i=p_{i1}\cup p_{i2}$，p_{i1} 在 p_{i2} 之前生成，p_{i1} 包括起始POI、结束POI和强制性POI；p_{i2} 包含其他POI；再采用AGA算法生成最佳解决方案（推荐行程），包含最具强制性的POI和其他POI，用户的访问持续时间在时间预算内最大化，总体POI受欢迎程度和整体评分最大化，用户进入POI的成本最低。该方法综合考虑了用户的个性化旅游需求，规划的行程通常包括游客经常想要访问的热门或特殊POI，还考虑了POI受欢迎程度和用户访问持续时间等其他需求。

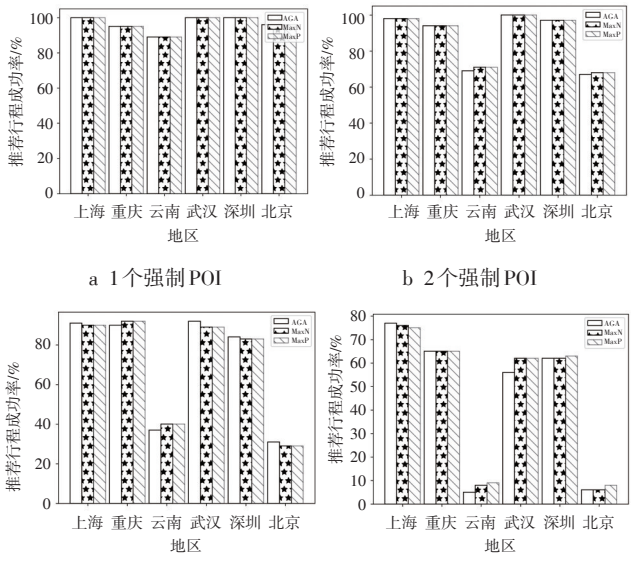
MaxN方法是生成包含最多POI的行程。首先根据访问持续时间基于贪婪策略添加强制POI，再对剩余POI采用相同的方法，只关注行程中包括的POI数量，因此在时间预算内具有最多数量的强制性POI和POI。该方法提供了强制性POI数量和POI数量的基准。MaxP方法是生成一个总体POI受欢迎程度较大的行程。首先通过贪婪策略根据POI受欢迎程度值添加强制POI，再对其他POI采用相同方法，只关注受欢迎程度，因此总行程具有较大的POI受欢迎程度。该

方法提供了POI受欢迎程度因素的基准。

为评估本文方法的性能，实验将强制POI数量、访问POI数量、旅行时间、访问持续时间、POI受欢迎程度、成本、评分等指标应用于个性化旅游行程推荐方法的评估。

3.3 结果分析

实验时还考虑了4个强制性POI集合，分别包括1、2、3、4个强制POI，且所有POI均从一个地区的整个POI集合中随机选择。将包含强制性集合中所有元素的行程定义为成功的推荐行程，不同方法生成的推荐行程成功率见图1，可以看出，随着强制性POI数量的增加，推荐成功率逐渐下降，MaxN和MaxP方法具有更高的推荐成功率。其原因在于，MaxN和MaxP方法都是先添加强制性POI，再通过考虑不同因素的贪婪策略添加其他POI来生成推荐行程；而本文方法是随机添加每个POI，仍然取得了较好的结果。

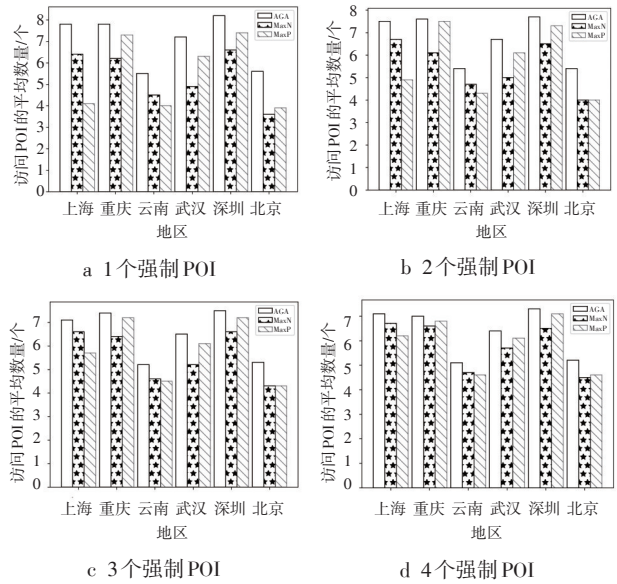


a 1个强制POI　b 2个强制POI

c 3个强制POI　d 4个强制POI

图1 不同方法生成的推荐行程成功率

不同方法生成的行程访问的POI平均数量见图2，可以看出，随着强制POI数量的增加，访问POI的平均数量逐渐下降；本文方法具有最好的性能指标。其原因在于，较多的强制POI限制了算法搜索更多其他POI的能力，尽管MaxN方法通过优先添加访问时间短的POI来生成行程，但POI之间的旅行时间超出了其控制范围，因此MaxN在所有地区的规划结果并不总是优于MaxP方法。此外，MaxN方法在实验中显示出较少POI数量，分析其原因：首先MaxN方法虽然旨在最大化POI数量，但在实际操作中可能会受到时间预算、成本和距离等其他约束的限制；其次在添加POI时采用了贪婪策略，可能导致其优先选择那些访问时间短的POI，从而在时间预算内无法包含更多POI；再次可能未考虑POI之间的距离和旅行时间，这在景点之间距离较远的地区会更加明显，因此实际能访问的POI数量可能会减少。

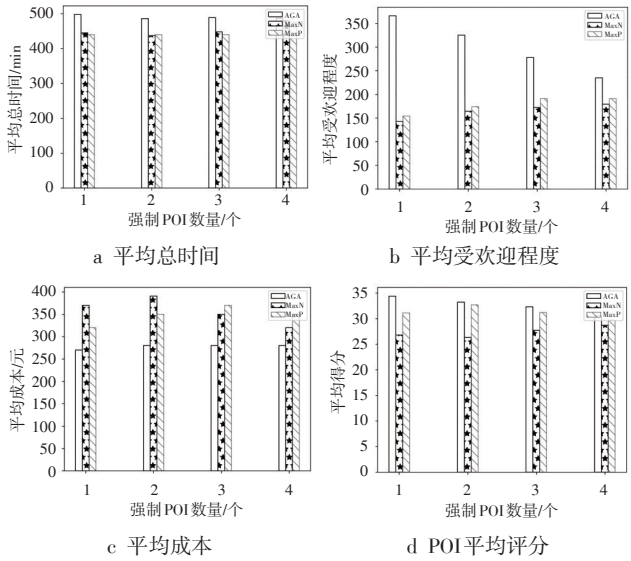


a 1个强制POI　b 2个强制POI

c 3个强制POI　d 4个强制POI

图2 不同方法生成的行程访问的POI平均数量

由于不同地区的POI分布、用户偏好和旅行条件可能存在很大差异，为评估算法的整体性能，本文分别从推荐行程平均总时间、平均受欢迎程度、平均成本、平均评分等方面评估算法的通用性和鲁棒性。结果表明（图3），本文方法有效利用了时间预算，总访问时间优于MaxN和MaxP方法；本文方法平均受欢迎程度最高、MaxP方法次之，主要原因是本文方法为POI受欢迎程度因素分配了较高的权重，MaxP方法通过优先添加受欢迎的POI生成行程，因此受欢迎程度指标较高；MaxN和MaxP方法的局部成本更低，其原因是本文方法规划的行程包括更多的POI，成本较



a 平均总时间　b 平均受欢迎程度

c 平均成本　d POI平均评分

图3 各方法通用性和鲁棒性比较

高；本文方法在评分指标总体上优于MaxN和MaxP方法，其原因是本文方法能生成更多POI的行程，因此局部具有较高的评分。

综上所述，尽管有些指标与MaxN和MaxP方法相比不占优势，但本文方法在强制性POI数量、POI访问量、POI受欢迎程度、总时间（包括旅行时间和访问时间）、成本和整体评分方面整体较优，说明本文方法可基于用户个性化需求规划旅游景点，为提高用户旅行体验和个性化订制服务提供了助力，具有广泛的应用前景。

4　结　论

随着社交媒体的兴起，用户生成的大量数据为旅游推荐系统提供了丰富的资源。本文提出了一种基于用户个性化需求的旅游景点智能推荐方法，通过整合社交媒体数据和POI相关信息构建了综合用户旅行历史数据库，并利用AGA方法生成满足用户多样化偏好的旅游行程推荐。与传统的旅游推荐系统相比，本文方法在考虑用户个性化需求方面取得了一定进步，特别是考虑了强制POI、受欢迎程度、评分、成本预算和时间限制等因素。本文仍存在一些局限性：首先数据来源的偏差可能影响推荐结果的准确性；其次算法的计算复杂度较高，耗时较长，这对于实时推荐场景可能是一个挑战；此外用户隐私和数据安全问题也需要得到更多关注。

未来的工作可以关注推荐系统的泛化能力，即如何使算法在不同的数据集和场景下都能保持良好的性能；提高推荐系统的可解释性，让用户理解推荐背后的逻辑，也是提升用户信任和满意度的关键；跨领域推荐和多语言支持有助于为全球用户提供更加个性化和多元化的旅游体验；考虑旅游活动对环境的影响，推荐系统可鼓励用户选择更环保的旅游选项，以促进旅游业的可持续发展。

参考文献

[1] 徐明华,代婉琦. 国际传播中社交机器人的转译机制与网络治理启示:基于行动者网络理论视角的分析[J]. 信息技术与管理应用,2023,2(5):52-63

[2] 乔禹涵,贾彩燕. 基于图自监督对比学习的社交媒体谣言检测[J]. 南京大学学报(自然科学),2023,59(5):823-832

[3] 兰美辉,高炜. 路径因子临界可避免图的参数条件[J]. 云南师范大学学报(自然科学版),2021,41(1):31-35

[4] 杜永潞,梅心蕊,周文月,等. 基于蚁群算法的乡村旅游行程定制[J]. 信息与电脑(理论版),2022,34(8):64-66

[5] 王维莉,邹雪儿. 疫情防控常态化下多目标旅游路线规划[J]. 重庆交通大学学报(自然科学版),2023,42(9):99-105

[6] 班航,王忠群. 基于情境聚类扩展用户画像的旅游景点推荐方法[J]. 安庆师范大学学报(自然科学版),2023,29(2):75-79

[7] 杨佳鹏,俎毓伟,纪佳琪,等. 基于Spark框架的瀑布型融合旅游推荐系统[J]. 智能计算机与应用,2023,13(4):142-146

[8] 韦鑫. 基于用户兴趣模型的个性化旅游线路生成建模研究[J]. 自动化技术与应用,2024,43(3):177-180

[9] 潘兰,魏嘉银,卢友军,等. 融合时空信息个性化旅游兴趣点推荐算法[J]. 智能计算机与应用,2023,13(12):120-123

[10] 李晓芳. 基于注意力机制和画像的旅游路线推荐研究[D]. 南京:南京邮电大学,2023

[11] 符常君. 融合用户偏好和情境感知的旅游路线推荐方法研究[D]. 长沙:中南大学,2022

[12] 高天贺. 考虑顾客需求和在线评论的旅游产品个性化设计方法研究[D]. 沈阳:辽宁大学,2022

[13] 吴雁,张韩. 基于混沌遗传算法和图论分析的BIM模型协作与决策优化[J]. 建筑机械,2024(8):55-60

[14] 周立俊,黄腾,王思捷,等. 基于GA-SVR的地铁隧道沉降预测[J]. 地理空间信息,2021,19(3):115-117

[15] 许文祥,邱晨,何惠婷,等. 基于遗传算法的多设施选址问题研究[J]. 物流科技,2024,47(15):5-9

（上接第11页）

[10] 王子斌,张全,谢聪,等. 基于知识图谱与BERT的安全领域汉字文本纠错模型[J]. 计算机应用,2023,43(增刊1):75-80

[11] 郭欣毅. 基于深度学习对中文拼写纠错模型的研究[D]. 兰州:兰州大学,2023

[12] 曹军航,黄瑞章,白瑞娜,等. DCsR:一种面向中文文本的集成式纠错框架[J]. 计算机工程与科学,2023,45(2):370-380

[13] 徐美仙,谢晓尧,郑欣. 基于BERT的流行病学调查信息文本的命名实体识别[J]. 贵州师范大学学报(自然科学版),2022,40(3):73-81

[14] 尚海怡,黄继风,陈海光. 基于Transformer融合词性特征的中文语法纠错模型[J]. 计算机应用,2022,42(增刊2):25-30

[15] 汪苏琪. 面向规范性文件的基于BERT的文本纠错模型[D]. 南昌:江西师范大学,2021

[16] 高印权. 基于深度学习的文本语法自动纠错模型研究与实现[D]. 成都:电子科技大学,2020

[17] 张大骞,郭昌达,陶旸,等. 省级自然资源监测数据库建设研究[J]. 地理空间信息,2024,22(2):109-112

[18] 杨梦梅,黄波,谢三五,等. 基础地理信息元数据批量制作方法研究[J]. 地理空间信息,2022,20(11):147-149