

基于通勤轨迹的潮汐交通拥堵路段识别与分析

饶磊¹, 刘艳芳^{1,2*}, 罗园园¹, 王楠楠¹

(1. 武汉大学资源与环境科学学院, 湖北 武汉 430079; 2. 武汉大学教育部地理信息系统
重点实验室, 湖北 武汉 430079)



摘要: 提出一种基于通勤轨迹数据, 结合地图匹配算法和潮汐交通特征计算的潮汐交通拥堵识别方法, 轨迹数据具有时间和空间上的连续性, 而道路流量受单条轨迹数据的缺失影响比较小。潮汐路段空间分布结果显示武汉市都市发展区的潮汐交通分布在环线之间的连接路、跨江通道的出入口路段以及重要交通枢纽附近。武汉市目前潮汐路段相对较多而可变车道数量较少, 本研究的识别结果有助于相关部门规划潮汐车道具体位置的决策, 以及对可能潮汐拥堵路段的预测。

关键词: 拥堵识别; 潮汐交通; 通勤轨迹; 地图匹配; 隐式马尔可夫链 (HMM)

中图分类号: P208

文献标志码: B

文章编号: 1672-4623(2021)04-0089-08

随着我国城市规模的扩大、经济水平的上升, 住宅区逐渐往城市外围延伸, 职住分离的现象越来越明显, 人群在早高峰期间向市中心汇聚, 在晚高峰期间则向城市外围扩散, 这种潮汐交通流造成了道路双向流量不均衡的现象, 形成潮汐拥堵^[1]。不同于道路的双向拥堵, 潮汐拥堵路段作为常发性拥堵^[2]可以通过设置可变车道、错峰上下班等方式来有效缓解, 从而减少出行成本、提高市民对城市的满意度等^[3]。

研究潮汐交通的核心是识别和分析潮汐路段。关于识别城市潮汐拥堵路段, 当前研究往往是利用道路拥堵指数计算时段性路段双向不平衡性和稳定性这两个潮汐交通的特征属性来综合判断^[4], 但拥堵指数实际上仅仅描述了研究对象的一个快照^[5], 缺少时间上的连续性。同时, 利用拥堵指数计算潮汐等级是将道路作为独立单元, 忽视了潮汐现象的产生直接结果, 而拥堵一旦发生便会迅速扩散^[6-8], 在交通路口附近产生明显的共生现象^[9]。另外, 潮汐路段特征指标的阈值设定受到路段拥堵等级数据缺失的影响很大。

百度通勤轨迹数据是出行轨迹链中一系列定位点的集合, 它具有时间连续性^[10]的特征, 单条轨迹数据的缺失或者匹配有误并不会对道路整体流量产生过大的影响^[11], 但是轨迹数据由于定位精度的原因, 使用前首先需要进行地图匹配。过往的研究中使用到的地图匹配算法可以概括为以下几个类型: 基于几何信息、基于路网拓扑、基于概率计算以及综合方法^[12]。基于几何信息的算法往往效率高, 但是精度不足; 基于路网拓扑的算法常见的有通过构建路网约束利用最短路径进行地图匹配^[13]; 在基于概率计算的算法方面,

杨易^[14]提出使用 D-S 证据推理进行轨迹点匹配, 综合方法主要涉及机器学习或者数据挖掘相关方法^[15]; 吴刚^[16]使用了一种基于隐式马尔可夫模型的算法, 该算法在使用了轨迹点几何信息的同时考虑了路网的拓扑关系, 匹配精度高, 但是时间成本也比较高。

针对上述问题, 本文提出了一种基于百度通勤轨迹数据, 构建隐式马尔可夫 (HMM) 模型进行位置匹配保证匹配精度, 采用二级空间索引结构大大提高匹配效率, 统计道路流量并使用流量时段性双向不平衡性和稳定性作为特征属性的潮汐路段识别方法。同时, 该方法使用了具有时段连续性的轨迹点数据, 在计算转移矩阵时考虑了路网的拓扑关系。因此该方法可以有效提取出潮汐路段, 识别出拥堵在交通路口的聚集现象^[17], 为相关部门的潮汐车道的规划以及拥堵预测提供有效参考^[18]。

1 数据与方法

1.1 数据与区域

1.1.1 区域选择

武汉是长江经济带的中心城市之一, 人口流动大, 就业人口多, 职住分离现象明显, 机动车拥有量超过 310 万辆, 由于武汉市的交通供给能力有限, 居民在通勤过程中不可避免地产生拥堵问题。武汉市发布的互联网行业报告显示, 超过 98% 的网民通过手机上网。而百度作为中国最大的搜索引擎提供商, 绝大多数手机浏览器都会内置百度搜索, 百度系手机应用在国内所占比例也非常大, 不论是主动还是被动大多数只能机用户都会产生定位信息。由于公共交通的供给在一天

收稿日期: 2020-09-25。

项目来源: 地理大数据统一表达、聚合与可视化挖掘 (017YFB0503601)。

(* 为通讯作者)

内保持稳定,因此道路流量主要受驾车数量的影响,本研究使用 835 650 条驾车轨迹数据,提取全天轨迹数据的起点(O)与终点(D),统计到公里格网分布^[19]

如图 1b 与图 1c 所示,武汉市的高出行区域集中分布在矩形框内的中心城区,远城区的出行量则比较少,出行终点相比于出行起点分布更加集中。

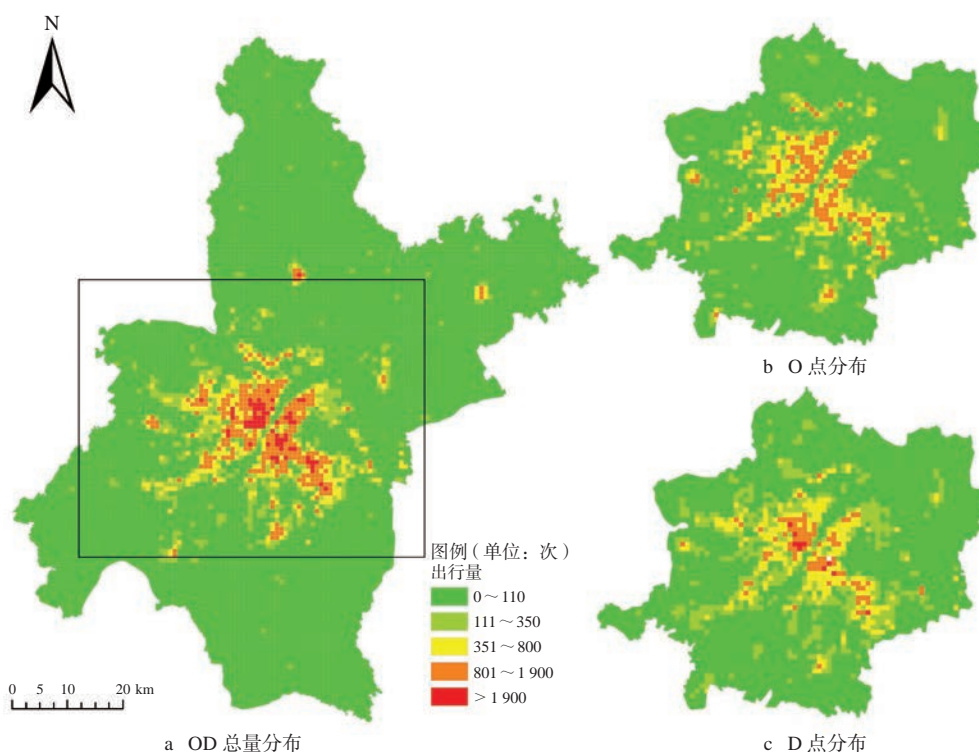


图1 全天轨迹 OD 分布公里格网统计值

本文选取 2019 年 6 月工作日在武汉市都市发展区内的基于百度轨迹数据的出行轨迹进行研究(矩形框中区域)。其中都市发展区指的是以 1h 通勤圈为基本范围的城市空间形态,是交通拥堵的常发区域。武汉市都市发展区面积约 3 100 km²,区域内的路段总数约为 1.5 万条,其中主干路占比为 16.2%、快速路占比为 5.7%、次干路的占比为 18.1%。选定区域后,需要对研究数据进行认知,而拥堵等级的识别本质是考虑流量相对值的差异,结果的受到数据的相对分布影响,因此本研究分别从 OD 的时序分布以及空间分布两个角度^[20-21]进行验证。

1.1.2 出行 OD 时序分布说明

OD 时序分布反映的是研究区域在不同时间的出行量,经统计百度轨迹还原数据的通勤出行次数从凌晨 1 点开始持续上升至早高峰(6:00~9:00)达到峰值,然后逐渐下降直到晚高峰(18:00~20:00)到达一个次峰值,时序分布规律显示,早高峰期间的百度出行量变化比较明显,而晚高峰由于错峰下班等因素的影响相比之下变化更平稳。武汉市居民通勤出行时间的洛伦兹曲线^[22]显示,不同时间段内的出行次数相对平均,基尼系数为 0.31,说明武汉市居民的上下班时间分布相对比较平均。另外,由于百度数据存在用户的

空间定位点不连续的情况,因此本研究采取百度用户长时间序列通勤打点与百度挖掘的通勤模式下的候选路径进行匹配后的轨迹还原结果。

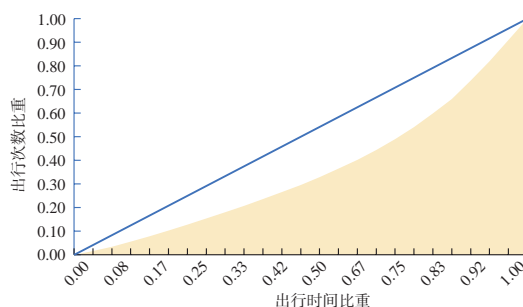


图2 武汉市居民通勤出行时间的洛伦兹曲线

1.1.3 出行 OD 空间分布验证

传统地理国情普查数据在数据总量上比较精确但是更新时间慢,大数据具有比较强的时效性而噪声数据比较多,综合两类数据的特点,为进一步验证出行相对量空间分布的合理性,本研究从 2 种数据出发验证数据的精度:①传统数据相对验证:以地理国情普查的武汉两实社区人口数据作为检验依据计算出空间分布的相关性;②多元大数据交叉验证^[23]:以联通职住 OD 数据作为检验依据计算相关性。将出行量统计到街道单元分别与 2 种数据计算斯皮尔曼相关系数,结果如表 1 所示。百度人口数据和两实人口的相关系

数为 0.819, 与联通街道人口的相关系数为 0.981, 说明百度通勤数据的出行量能比较准确的反映城市出行总量的相对信息。

表1 两实人口、联通手机信令和百度人口 Spearman 相关系数

		两实人口	联通街道
百度街道	相关系数	.819**	.981**
	显著性 (双尾)	.000	.000

1.2 方法

本研究使用的方法为首先通过构建地图匹配模型进行轨迹匹配并统计路段流量, 然后利用流量计算出路段的潮汐特征得分, 最后加权计算出道路潮汐等级, 并使用双变量 lisa 验证潮汐路段是否和双特征聚集区域分布一致。

构建地图模型包括 3 个部分: ① 计算发射概率, 即每个轨迹点到其候选路段的概率矩阵; ② 计算转移概率, 即候选点到其匹配道路转移到另一个候选道路的概率; ③ 利用维特比算法求解隐式马尔可夫模型。同时, 为了提高匹配效率, 在计算发射概率时引入二级空间索引结构, 即利用 rtree 索引筛选候选路段, 再利用 kdtree 筛选路段折点。

针对匹配结果统计研究区域内主干路、次干路、快速路的流量, 然后利用时段性双向不平衡性和稳定性的定义计算每个路段的双特征得分, 综合双特征得分识别出潮汐拥堵的空间分布和等级, 并与双变量 lisa 聚类地图对比验证空间分布的一致性, 研究方法流程如图 3 所示。

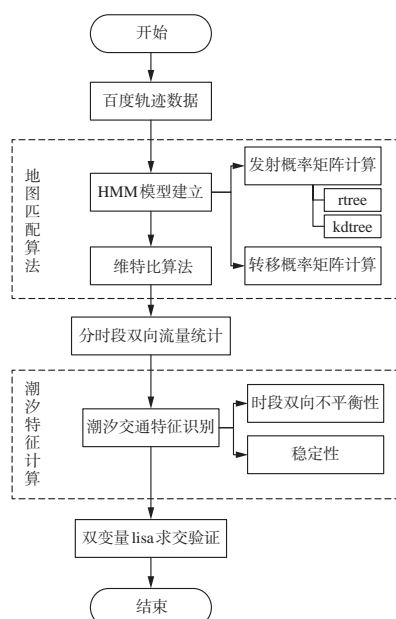


图3 研究方法图

1.2.1 基于路网约束的地图匹配模型算法

本研究将地图匹配的结果看成是一个隐式马尔科夫链 (HMM), 然后采用维特比算法求解隐式马尔科夫

模型预测问题^[24] (如图 4 所示)。HMM 模型中首先需要确定隐状态空间的集合, 对于地图匹配来说即所有轨迹点候选路段的集合 $\{\Phi_i\}$, 维特比算法的时间复杂度为 $O(N^2T)$, 其中 T 为序列长度, N 为隐状态空间的状态数量, 算法效率受隐状态影响很大, 因此需要尽可能减少候选路段数量, 本文采用 2 种方式实现该目标: ① 由于潮汐拥堵主要是由于机动车造成, 在数据预处理阶段通过对武汉市路网进行筛选, 保留主次支公路等机动车可以行驶的路段; ② 使用二级索引对每个轨迹点的 m 条候选道路快速筛选。

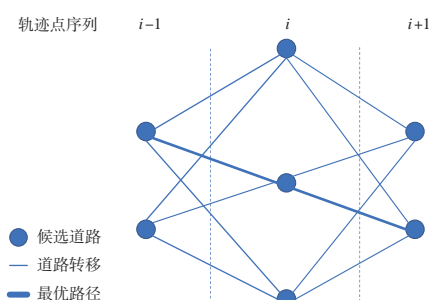


图4 HMM用于地图匹配算法

1.2.2 二级空间索引介绍

为了达到轨迹点数量尽可能少而拓扑连通性良好的目的, 本文引入了二级空间索引的概念, 对路段建立 R 树索引而对路段的折点建立 Kdtree 树索引, 具体介绍如图 5 所示。

1) rtree 索引。rtree 索引是 btree 索引在高维空间的扩展, 是一颗平衡树, 它对空间实体建立最小边界矩形 (MBR), 并将这些 MBR 当作叶子节点, 然后往上递归分割生成 MBR 分割空间, 结点越往上, MBR 框住的空间越大, 示意图如图 5a 所示。

本文通过对城市路网中所有的路段建立 MBR, 并作为叶子节点, 依次建立城市路网的 R 树索引, 能够快速查询轨迹点周围 n 邻近候选路段, 显著提高地图匹配算法效率。

2) kdtree。kdtree 是一种划分 k 维数据空间的数据结构, 本质上也是一棵二叉树, 但是每个节点的数据都是 k 维。其原理是对一系列维度为 k 的点中选取某个维度, 求这个维度值的中位数, 再将数据按中位数分为两个区域, 每个区域递归分割直到每个区域中的点数目为 1, 示意图如图 5b 所示。

由于道路往往是多条首位相连的折线相连组成, 不同折线与轨迹点方向向量形成的角度不同, 为了减少由于路段曲率过大导致的误差, 在获取邻近路段的基础上进一步确定轨迹点候选路段的折点 (即轨迹点属于道路的哪一段), kdtree 索引对空间的高分割性相比于 rtree 索引在检索折点的效率更高。使用二级索引

结构可以有效提升候选路段对应折点的查找速度,对每个轨迹点订单筛选候选路段集合 $\Phi_i | i=1, 2, 3, \dots, n$, 其

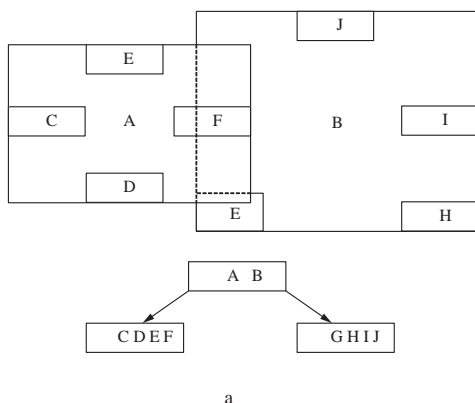


图5 空间索引示意图

1.2.3 发射概率

发射概率即隐含状态下观测值的概率,本文中的发射概率为轨迹点 Pt_j 到候选路段 Φ_i 的概率 $P(Pt_j | \Phi_i)$, 满足 $\sum_i P(Pt_j | \Phi_i) = 1$, 当 $P(Pt_j | \Phi_i)$ 的值越大时,说明轨迹点位于候选路段 Φ_i 的概率越大,根据数据特征,本研究使用距离相似性和方向相似性两个方面作为候选路段得分的计算因子,公式如下:

$$P(Pt_j | \Phi_i) = \frac{w_1 \times S_d(Pt_j | \Phi_i) + w_2 \times S_a(Pt_j | \Phi_i)}{\sum_{i=1}^n (w_1 \times S_d(Pt_j | \Phi_i) + w_2 \times S_a(Pt_j | \Phi_i))} \quad (1)$$

式中, $S_d(Pt_j | \Phi_i)$ 为距离相似度得分; $S_a(Pt_j | \Phi_i)$ 为方向相似度得分; w_1, w_2 分别为对应指标的权重。其中距离相似度得分使用的是轨迹点到候选路段的欧氏距离归一化后的结果:

$$S_d(Pt_j | \Phi_i) = \frac{d(Pt_j | \Phi_i) - d_{\min}(Pt_j | \Phi_i)}{d_{\max}(Pt_j | \Phi_i) - d_{\min}(Pt_j | \Phi_i)} \quad (2)$$

式中, $S_a(Pt_j | \Phi_i)$ 为方向相似度得分,由于百度轨迹还原数据本身不具有方向属性,因此使用当前轨迹点与下一个轨迹点所形成的向量 $\overrightarrow{\Delta Pt_i}$, 与候选路段的起点终点形成的向量 $\overrightarrow{\Delta \Phi_i}$ 的夹角,利用余弦公式计算夹角余弦值并归一化的结果:

$$\cos \theta_{ji} = \frac{|\overrightarrow{\Delta Pt_i} \times \overrightarrow{\Delta \Phi_i}|}{|\overrightarrow{\Delta Pt_i}| \times |\overrightarrow{\Delta \Phi_i}|} \quad (3)$$

$$S_a(Pt_j | \Phi_i) = \frac{\cos \theta_{ji \max} - \cos \theta_{ji}}{\cos \theta_{ji \max} - \cos \theta_{ji \min}} \quad (4)$$

从武汉市不同环线的个区域中随机抽出多条轨迹共500个轨迹点计算距离相似性和方向相似性占比分布(如图6所示),结果显示由于轨迹数据是通过工作日长时间序列下百度定位数据结合路径规划匹配后的结果,因此轨迹点距离路段比较近,相似性高的轨迹

中 n 为候选路段数量,在本研究中,结合效率和精确度设置 $n=5$ 。

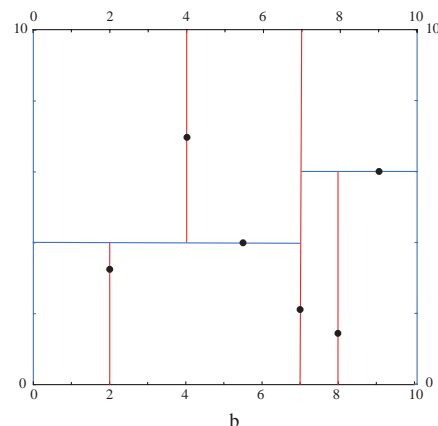


图6 相似性占比分布图

点数量占比很大,相似性低的轨迹点呈急剧下降的态势;方向相似性受到武汉市复杂的路网结构以及交通路口转弯现象的影响,所以相比于距离相似性的分布,方向相似性在低相似性区间仍存在比较高的占比。综上所述,本研究设置的距离相似性权重 $w_1=0.6$, 方向相似性权重 $w_2=0.4$ 。

1.2.4 转移概率

潮汐拥堵路段在交通路口具有比较强的相关关系,同时由于轨迹数据为路径规划后的结果,因此轨迹点分布均匀,前后轨迹点所在路段在路网中具有强拓扑相关,即前后轨迹点匹配的路段在拓扑相邻路段上概率更大,因此在计算转移概率矩阵时,给拓扑相邻路段更大权重,轨迹匹配以及进而流量统计的结果会在相邻路段相关性加强,转移概率计算公式:

$$P'(Pt'_{j+1} | Pt'_j) = \frac{w_i C_i}{\sum_{i=1}^m w_i C_i} \quad (5)$$

式中, $P'(Pt'_{j+1} | Pt'_j)$ 指的是轨迹点 Pt_j 经过道路 t 到轨迹点 Pt_{j+1} 的概率; m 为候选路段总数; w_i 是根据拓扑关系人为设置的参数; C_i 候选路段的出现频次。

1.2.5 维特比算法求解

维特比算法求解 HMM 模型本质上是利用动态规划(DP)求概率最大路径,根据原理,在一条连续轨迹记录中,从轨迹点 Pt_1 开始,递推计算该记录中其

余的轨迹点 Pt_j 到候选路段的概率最大值, 直到最后一个轨迹点 Pt_m 到其候选路段的概率最大值, 此概率即为 m 点的最优路径, 然后逐一反推得到各个轨迹点隐状态值 (候选路段), 最终得到该轨迹链的匹配路径 $I=(i_1, i_2, \dots, i_m)$ 。算法如下:

输入: 轨迹数据 $Pt=(Pt_j|j=1, 2, 3, \dots, m)$, HMM 模型 $\lambda=(E, T, S)$, 其中:

T 为转移概率矩阵, $T=[t_{ij}]_{n \times n}$

$$t_{ij}(k)=P(\Phi_j|\Phi_i), j=1, 2, \dots, n; i=1, 2, \dots, n \quad (6)$$

$t_{ij}(k)$ 是在轨迹点 Pt_j 处于候选路段 Φ_i 的条件下, 轨迹点 Pt_{j+1} 时转移到候选路段 Φ_j 的概率, n 是候选路段个数。

E 为发射概率矩阵 $E=[e_{ij}]_{m \times n}$

$$e_{ij}=P(Pt_j|\Phi_i), j=1, 2, \dots, n; i=1, 2, \dots, n \quad (7)$$

e_{ij} 是在轨迹点 Pt_j 处于候选路段 Φ_i 的条件下生成轨迹点坐标 Pt_j 的概率。

S 为初始概率向量 $S=(S_i)$

$$S_i=P(Pt_1|\Phi_i), i=1, 2, \dots, n \quad (8)$$

S_i 为第一个轨迹点到候选路段 i 的概率。

输出: 该轨迹链的匹配路径 $I=(i_1, i_2, \dots, i_n)$

流程: ①初始化:

$$\delta_1(i)=S_it_{i1}(Pt_1), i=1, 2, \dots, n \quad (9)$$

$$\Psi_1(i)=0, i=1, 2, \dots, n \quad (10)$$

式中, $\delta_1(i)$ 为轨迹点 Pt_1 在路段 i 处的匹配概率; $\Psi_1(i)$ 为最大节点。

②递推: 对于 $j=2, 3, \dots, m$

$$\delta_j(i)=\max_{1 \leq i \leq n} [\delta_{j-1}(i)e_{ij}]t_{ij}(k), i=1, 2, \dots, n \quad (11)$$

$$\Psi_j(i)=\arg \max_{1 \leq i \leq m} [\delta_{j-1}(i)e_{ij}], i=1, 2, \dots, n \quad (12)$$

式中, $\delta_j(i)$ 为轨迹点在路段 j 处的匹配概率; $\Psi_j(i)$ 为最大节点。

③终止与回溯: $P=\max_{1 \leq i \leq n} \delta_j(i)$, P 为最终路径匹配概率, 当其最大时, 对于 $j=m, m-1, \dots, 1$

$$I_j=\Psi_{j+1}(t+1) \quad (13)$$

求得最优匹配路径 $I=(i_1, i_2, \dots, i_n)$

1.2.6 潮汐交通特征计算

潮汐交通具有 2 个典型特征: ①时段性双向不平衡性: 同一路段的流量表现为路段某方向早高峰流量较大, 晚高峰交通流较小, 而另一方向流量状况则相反, 在工作日呈现明显的周期性; ②稳定性: 路段的交通状态较稳定, 即一个交通状态持续的时间较长。

因此在利用地图匹配算法将轨迹匹配到路段后,

首先需要统计出路段在各个时刻不同方向的流量值, 再根据轨迹数据的方向特征和时间特征得出流量值 $F_{i|dt}$, $i=1, 2, 3, \dots, n$; $d=d, d'$; $t=1, 2, \dots, 24$, 其中 n 为道路数量; d 为方向编号; t 为时间。

根据双向不平衡性主要表现为同一路段双向交通状态是否互异, 表现在交通流量上即为双向交通流量的差值是否大于阈值 a , 则 t 时刻双向不平衡得分 Pb_t :

$$Pb_t=\begin{cases} 1, & |F_{i|dt}-F_{i|d't}| \geq a \\ 0, & |F_{i|dt}-F_{i|d't}| < a \end{cases} \quad (14)$$

路段的时段不平衡性 SumPb 表示路段在早高峰时期 $T_m=[tm_1, tm_2, \dots]$ 、晚高峰时期 $T_e=[te_1, te_2, \dots]$ 的双向不平衡得分总和, 则有

$$\text{SumPb}=\frac{\sum_{t \in T_m} Pb_t + \sum_{t \in T_e} Pb_t}{m+e} \quad (15)$$

稳定性是通过将早晚高峰期间某方向路段前后时间片的差值与道路流量阈值 b 作比较, 则稳定性得分:

$$C_i=\begin{cases} 1, & |F_{i|dt+1}-F_{i|dt}| \leq b \\ 0, & |F_{i|dt+1}-F_{i|dt}| > b \end{cases} \quad (16)$$

判断路况的不平衡性以及稳定性, 首先需要确定阈值, 即 $a=\lambda \bar{F}_i$, $b=\mu \bar{F}_i$ 的系数 λ, μ (其中 \bar{F}_i 为路段的平均流量值), 根据阈值和路段数的统计量关系计算得当 $\lambda=0.625$, $\mu=0.375$ 时, 不平衡路段占比和稳定路段占比变化最大, 则有分级标准设置如表 2 所示。

表 2 潮汐交通特征等级与阈值关系

潮汐交通特征	系数	判断条件	等级
不平衡性	λ	< 0.625	平衡
		$> 0.625 \text{ and } < 0.75$	轻微不平衡
		$> 0.75 \text{ and } < 0.825$	中度不平衡
		> 0.825	显著不平衡
稳定性	μ	< 0.375	显著稳定
		$> 0.375 \text{ and } < 0.5$	中度稳定
		$> 0.5 \text{ and } < 0.625$	轻度稳定
		> 0.625	不稳定

2 结果

2.1 轨迹匹配精度校验

针对轨迹匹配结果, 从武汉市不同环线内随机抽取 4 条轨迹数据, 将原始轨迹点以及匹配的路网结果还原, 结果如图 7 所示。

案例 a 显示由于数据不完整导致路段拓扑不连续的情况下, 部分轨迹点匹配到其相邻轨迹点所属路段; 案例 b、案例 c 显示, 对于拓扑关系良好但是去除支路、乡道的路网中, 主干道、次干道、快速路的轨迹点能准确匹配到邻近道路上, 支路的轨迹点则匹配到

其相邻的主干道上；案例 d 显示轨迹点在交通路口处稀疏的情况下路网匹配结果存在少量的误差。手动匹配上述轨迹点的所属道路，对比得到这些区域的匹配精度分布图如下，结果显示抽样区域的平均匹配精度为 95.47%，最低匹配精度为 92.59%，地图匹配结果准确率高。

为了验证二级索引结构的效率提升，本研究另选取了通过建立轨迹点缓冲区的形式获取匹配点最邻近候选路段的方法，对比二级空间索引结构筛选候选路段的方法。从研究区域内分别随机选取了 500、1 000、5 000、10 000、50 000 条轨迹，分别统计 2 种筛选方法的轨迹匹配总耗时，如图 8 所示，二级空间索引方法由于建立空间索引的关系，在轨迹条数较少的情况下（如 500、1 000 条轨迹点）所用时间相对较长，而当轨迹点数量增加到 5 000 以后，二级空间索引结构的方法匹配时间明显缩短。当轨迹点数量达到 50 000 条时，轨迹匹配时间缩短了 41.42%，综上所述，二级空间索引结构可以有效提升轨迹匹配效率。

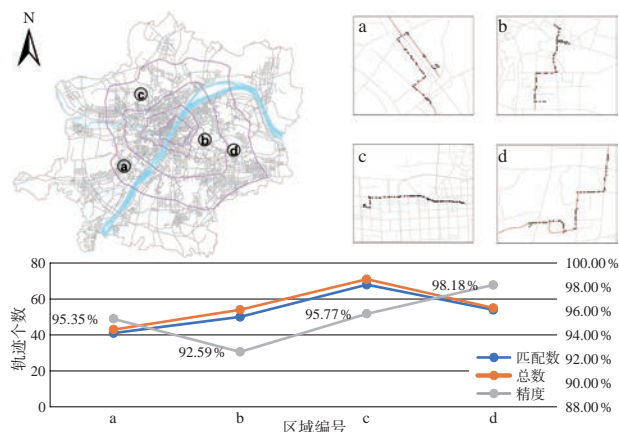


图7 抽样区域轨迹匹配结果验证

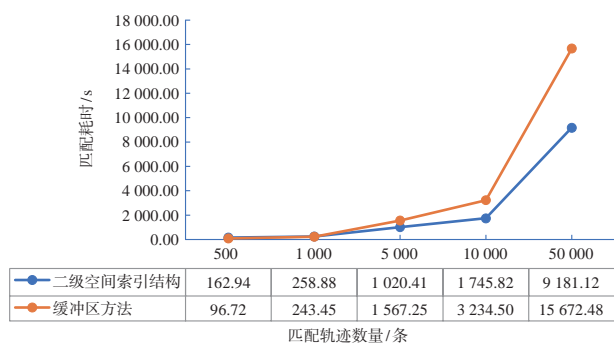


图8 轨迹匹配耗时对比图

2.2 轨迹匹配结果分析

统计轨迹匹配结果，然后计算工作日流量平均值以及不同类型道路流量总和。图 9a 显示，高流量道路主要分布在环线、快速路、跨江通道，三环以内区域流量显著高于三环以外；经统计，主干路全天流量总值最高约为 556 万，次干路约为 254 万，快速路约为 372 万。

整体来看武汉市道路流量分布呈现出内高外低的状态。

潮汐拥堵主要发生在工作日的早晚高峰时段，同时百度的轨迹数据存在一定的滞后性特征，选取早高峰 6:00~9:00，晚高峰 18:00~21:00 的流量数据进行研究。根据路段的时段性双向不平衡分级标准和道路稳定性分级标准显示了各级别时段性双向不平衡性的空间分布，其中显著不平衡（图 9b）和中度不平衡的路段（图 9c）主要分布在环线以及环线的连接路上，而在二环以内的中心城区则没有呈现明显的聚集现象，轻度不平衡的路段则在区域中分布比较广泛；稳定性道路主要分布在中心城区，跨江通道呈现出明显的流量稳定性，显著稳定路段在江北的分布多于江南。为进一步显示时段性双向不平衡性和稳定性的空间分异特征，使用双变量局部空间自相关分析，在 P 为 5%， Z 值检验基础之上（置信度 95%），绘制双变量 Lisa 聚集图，如图 9d 所示。高-高聚集即流量高稳定与高不平衡的路段主要分布在三环以内，尤其是过江通道附近的主次干道，环线以及附近的连接路，这些道路的早晚高峰持续时间长且不同方向的流量有显著差异；低-低聚集与高-低聚集的道路主要分布在三环以外，远城区道路的早晚高峰持续时间相对较短；低-高聚集即双向流量较平衡且比较稳定的道路主要分布在三环线以内二环线以外的支路。

2.3 潮汐交通空间分布

综合潮汐交通的两个特征，提取出时段性双向不平衡性大且稳定性比较高的路段，结合都市发展区流量分布特征，得到都市发展区内的潮汐交通空间分布如图 10 所示，潮汐路段的空间分布与双变量 Lisa 聚类图中的高高聚类路段空间分布之间存在高相关性，同时，潮汐路段主要分布在环线之间特别是二环以外的连接路上，主要跨江通道附近的主干路以及城市内部交通枢纽如光谷和洪山广场上也存在潮汐路段的聚集分布。计算潮汐路段的全局莫兰指数^[25]结果， z 得分为 9.4， p 值为 0，结果可信度高，全局莫兰指数为 0.11，潮汐拥堵路段的聚集特征显著，说明轨迹数据用于识别潮汐交通，可以考虑到交通路口处流量的扩散现象，对于识别路口处的交通拥堵聚集特征更有优势。

街道是城市的基本单元，是城市规划中的基础操作对象，因此本研究以街道为统计单元统计武汉都市发展区的潮汐路段以验证识别结果的正确性。如图 10 所示，潮汐路段显著分布在环线附近。三环线附近主要存在 3 个中心：位于东西湖区的径河街及吴家山街，四环线通往三环的连接路位于其中，是通勤的主要路段，因此潮汐拥堵路段聚集。由于光谷软件园等

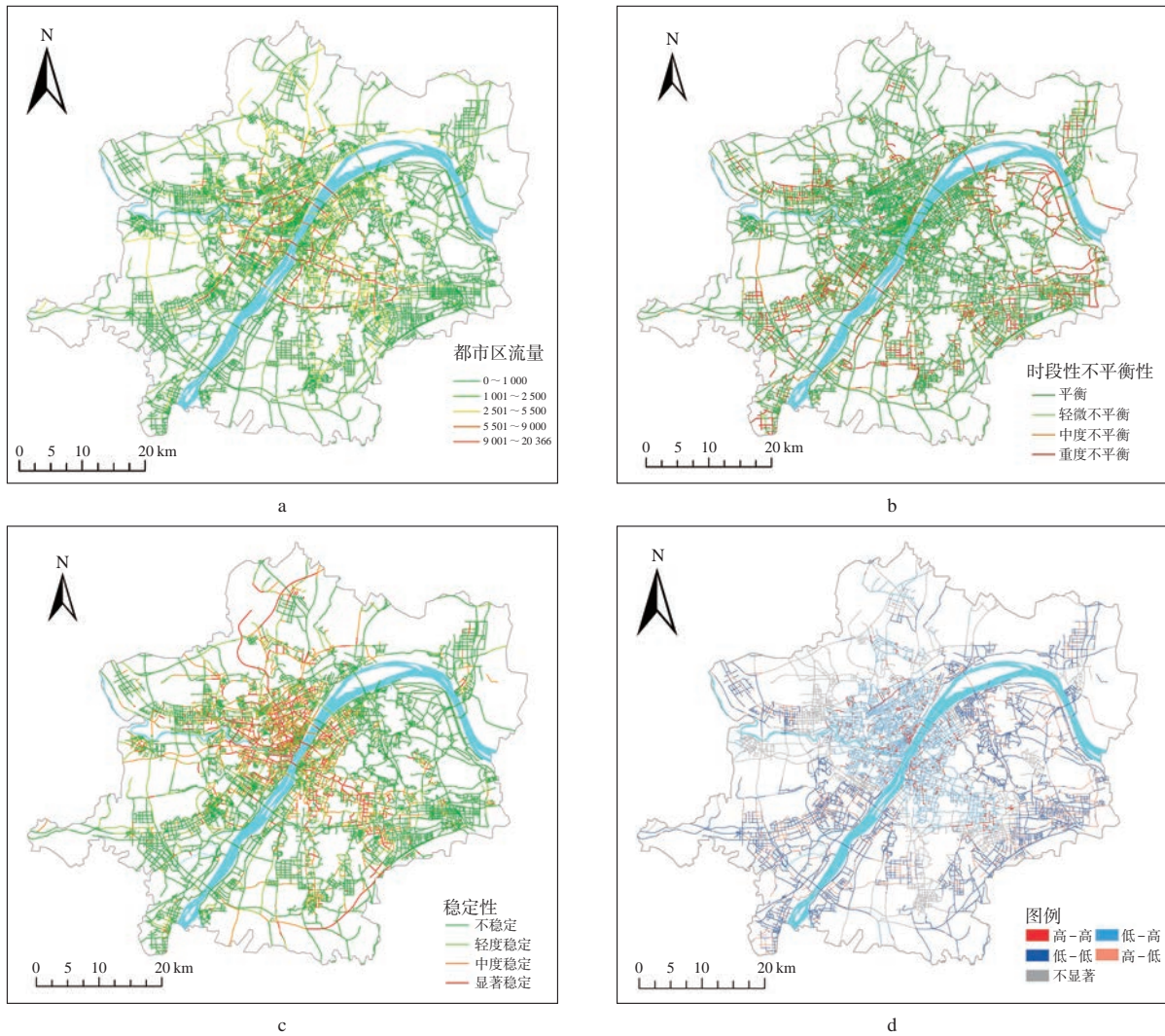


图9 轨迹匹配结果分析图

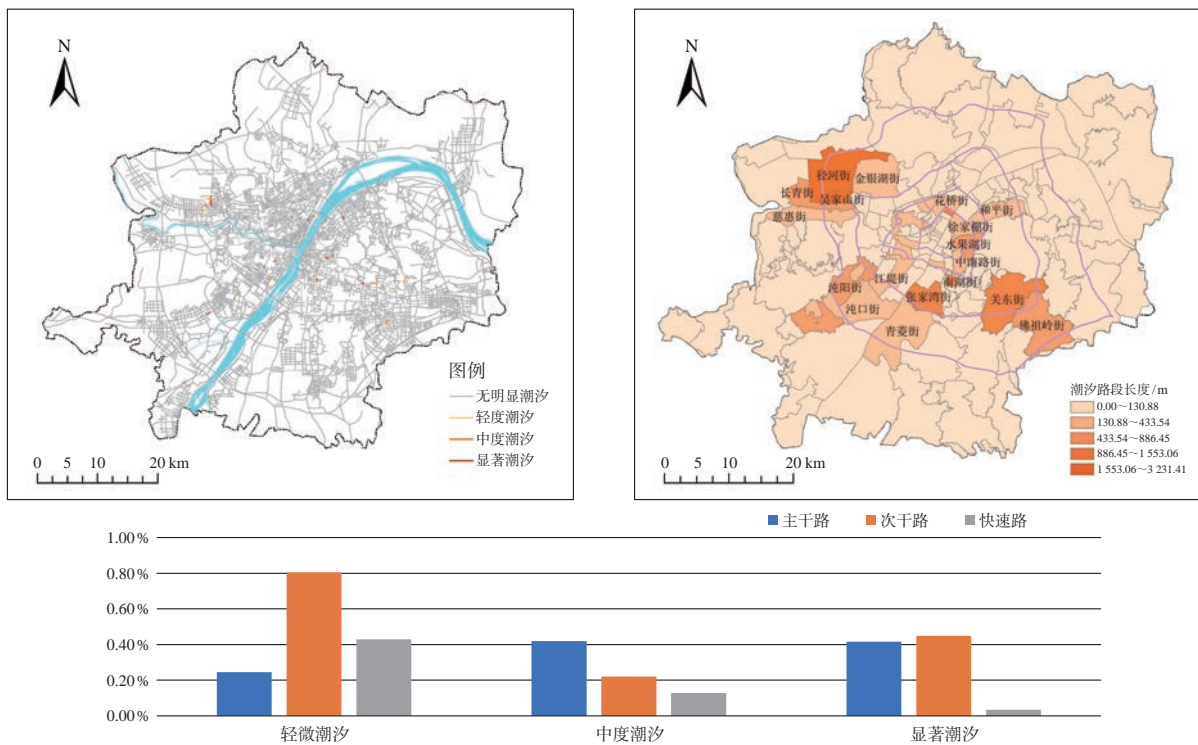


图10 潮汐交通空间识别结果

高新技术产业聚集,是武汉市重要的通勤目的地,因此在关东街附近潮汐拥堵现象显著;武汉经济开发区则是武汉汽车业、加工业的中心,同样也是通勤的重要目的地之一,因此沌口街、沌阳街潮汐现象显著。二环线附近的街道潮汐拥堵等级相对较弱而分布更加平均:洪山广场作为武汉重要的交通枢纽,影响着附近的中南路街和水果湖街,产生潮汐拥堵;位于江岸区的二七街及其附近街道与跨江通道相连,跨江通勤是武汉的交通特征之一,潮汐拥堵现象因而比较显著。根据武汉市交通发展蓝皮书,武汉市目前仅存在唯一一条潮汐可变车道设置在珞喻东路,位于关东街,可变车道的设置在了潮汐拥堵路段高值路段,比较合理。

筛选出主干路次干路快速路,得到武汉市潮汐路段总计 67 条,其中显著潮汐路段有 22 条,中度潮汐路段有 16 条,轻度潮汐路段有 27 条,统计各类型道路上的潮汐路段长度的占比如图 10 所示,对于轻微潮汐现象,次干路占比最高,其次是快速路,主干路占比最低;对于中度潮汐现象,主干路占比最高,次干路次之,快速路占比最低;对于显著潮汐现象,次干路占比最高,主干路次之,快速路最低。上述结果说明,对于武汉市都市发展区内的潮汐路段,大多数居民的通勤路线都会经过主干路,因此其更容易发生中度以及重度潮汐;次干路配合主干路组成城市路网,主要分布在环线与区域间的连接路,因此潮汐拥堵总占比最高;快速路作为通勤的另一个重要支柱,其主要分布在环线,不同区域的居民在通勤时倾向于将其作为主要路径,容易产生双向型拥堵而非潮汐型拥堵。

3 结 语

本文从轨迹数据出发,提出一种基于地图模型匹配方法与潮汐路段的道路流量不平衡性和稳定性双特征识别潮汐路段的方法。由于轨迹数据在一定时段内是连续的结果,通过路网的拓扑关系建立转移矩阵,使得该方法对于拥堵在路口的扩散导致的聚集现象能比较好的反映。识别结果显示,武汉市都市发展区的潮汐路段主要分布在环线之间的连接路、跨江通道的出入口路段以及重要交通枢纽附近。目前武汉市潮汐可变车道数量比较有限,可以考虑在这些位置设置可变车道缓解交通压力。

参考文献

- [1] 黄超. 杭州西湖“潮汐走廊”交通模型及评价研究[D]. 杭州: 杭州电子科技大学, 2016
- [2] 李霄. 基于 IC 卡数据的公交常发拥堵识别方法与特性分析[D].

- 北京: 北京交通大学, 2019
- [3] 丁心茹. 城市路网潮汐车道设置研究[D]. 吉林: 吉林大学, 2019
- [4] 曾丽燕, 张晶, 赵天天, 等. 基于实时路况数据的潮汐交通路段提取[J]. 地理空间信息, 2017, 15(3): 87-89
- [5] 韦伟, 毛保华, 陈绍宽, 等. 基于时空自相关的道路交通状态聚类方法[J]. 交通运输系统工程与信息, 2016, 16(2): 57-63
- [6] 叶晓飞, 陈峻. 交通拥堵信息对交通拥堵扩散的影响分析[C]. 中国大城市交通规划研讨会论文集, 苏州, 2010
- [7] Zheng J, Gao Z, Zhao X. Self-organized Diffusion of Congestion in Complex Networks[J]. Physica A: Statistical Mechanics and its Applications, 2010, 389(2): 342-348
- [8] 严淑惠, 白翰, 王亚龙. 城市交通拥堵形成机理及扩散演变规律研究[J]. 中外交流, 2016(4): 42
- [9] 赵宝新. 基于多源数据融合的城市交通拥堵扩散的研究及其应用[D]. 深圳: 中国科学院深圳先进研究院, 2020
- [10] Qianzhu W, Dechao J, Daqin P. Calculation Method of Congestion Metrics Based on Mobile Phone Signaling[J]. Procedia Computer Science, 2018(131): 1 081-1 088
- [11] Cui Z, Lin L, Pu Z. et al. Graph Markov Network for Traffic Forecasting with Missing Data[J]. Transportation Research Part C: Emerging Technologies, 2020(117): 102 671
- [12] Nikolić M, Jović J. Implementation of Generic Algorithm in Map-matching Model[J]. EXPERT SYST APPL, 2017(72): 283-292
- [13] Wu Z, Xie J, Wang Y. Map Matching Based on Multi-layer Road Index[J]. Transportation Research Part C: Emerging Technologies, 2020(118): 102 651
- [14] 杨易. 基于概率决策的车辆导航系统地图匹配算法[J]. 汽车工程, 2006, 28(10): 897-898
- [15] He M, Zheng L, Cao W. An Enhanced Weight-based Real-time Map Matching Algorithm for Complex Urban Networks[J]. Physica A: Statistical Mechanics and its Applications, 2019, 534: 122 318
- [16] 吴刚, 邱煜晶, 王国仁. 基于隐马尔可夫模型和遗传算法的地图匹配算法[J]. 东北大学学报(自然科学版), 2017, 38(4): 472-475
- [17] Kan Z, Tang L, Kwan M. Traffic Congestion Analysis at the Turn Level Using Taxis' GPS Trajectory Data[J]. Computers, Environment and Urban Systems, 2019(74): 229-243
- [18] 林树宽, 于伶姿, 乔建忠, 等. 基于 GPS 轨迹数据的拥堵路段预测[J]. 东北大学学报(自然科学版), 2015, 36(11): 1 530-1 534
- [19] Liu Y, Wang F, Xiao Y. Urban Land Uses and Traffic 'Source-sink Areas': Evidence from GPS-enabled Taxi Data in Shanghai[J]. Landscape Urban Plan, 2012, 106(1): 73-87
- [20] 张健钦, 仇培元, 杜明义. 基于时空轨迹数据的出行特征挖掘方法[J]. 交通运输系统工程与信息, 2014, 14(6): 72-78
- [21] 张振龙, 邱煜卿, 蒋灵德, 等. 基于实时路况的交通拥堵时空特征及其影响因素分析——以苏州古城区为例[J]. 现代城市研究, 2020(1): 104-112
- [22] 古杰, 周素红, 闫小培, 等. 居民日常出行时空集聚视角下的城市交通拥堵形成机制研究——以广州为例[J]. 地理科学, 2012, 32(8): 921-927
- [23] Muhammad T Z, Suhajito. Detection Traffic Congestion Based on Twtter Data Using Machine Learning. 4th International Conference on Computer Science and Computational Intelligence[C]. 2019
- [24] 刘旻, 李梅, 徐晓宇, 等. 一种基于 HMM 模型改进的地图匹配算法[J]. 北京大学学报(自然科学版), 2018, 54(6): 1 235-1 241
- [25] 刘康, 段滢滢, 陆锋. 基于拓扑与形态特征的城市道路交通状态空间自相关分析[J]. 地球信息科学学报, 2014, 16(3): 390-395

第一作者简介: 饶磊, 硕士, 主要从事地理大数据多元聚合研究。

Comparative Study on the Spatial Characteristics of Urban Agglomerations on the Mongolian Plateau Based on DMSP/OLS Data

by LI Yanyan

Abstract In this paper, we used DMSP/OLS nighttime lighting data after mutual correction, saturation correction and continuity correction to extract the urban built-up areas. Then, we measured the urban expansion of UD-UA and HBO-UA, and analyzed the spatial evolution characteristics. The results show that ①the built-up areas of each city extracted from the corrected nighttime light images all have a large increase. The extraction threshold of HBO-UA is larger than that of UD-UA. ②The expansion of HBO-UA is much more balanced than UD-UA. ③The urbanization processes of the two urban agglomerations are in a rapid development state. Many small towns around HBO-UA have been gradually developed, while the regional polarization of UD-UA is extremely significant.

Key words urban expansion, spatial characteristic, nighttime light data, Mongolian Plateau (Page: 84)

Identification and Analysis of Tidal Congested Section Based on Commuter Trajectory

by RAO Lei

Abstract In this paper, we proposed a tidal traffic congestion identification method based on commuter trajectory data, which combined with map matching algorithm and tidal traffic feature calculation. The trajectory data has the continuity in time and space, and the road flow is less affected by the lack of single track data. The results of spatial distribution of tidal sections show that the tidal traffic in Wuhan urban development zone is distributed in the connecting roads between loop roads, the entrance and exit sections of river crossing channels and the vicinity of important transportation hubs. At present, there are many tidal roads in Wuhan, but the number of variable lanes is relatively small. The identification results of this study are helpful for the relevant departments to plan the specific location of tidal lanes and predict the possible tidal congestion sections.

Key words congestion identification, tidal traffic, commuter trajectory, map matching, implicit Markov chain (Page: 89)

Application Analysis of the Land Survey Data Distribution Management System

by LIU Lingqi

Abstract The third national land survey is a major national survey on national conditions and power. Based on the results of the second national land survey, it comprehensively refines and perfects the basic data of the country's land, and grasps detailed and accurate national land use status and changes in land resources. The data results relate to national security and socio-economic development. Therefore, there are many special data items and thematic elements, large amount of data, and wide data sources in the course of this survey. Through the analysis and research of the data distribution management system of land survey, we realized the efficient application of multi-source and massive data, rational analysis, orderly distribution, accurate docking and rapid work, which could assist the rapid and accurate development of the third national land survey, and provide strong system support and technical support.

Key words management system, system construction, application analysis (Page: 97)

Application of MMS in the Key Steps of Rural Real Estate Cadastral Right Investigation

by LYU Yue

Abstract With the development of new technology and mapping technology equipment, we discussed how to solve the problems existing in the key steps of rural real estate cadastral right investigation by using mobile measurement system, intelligent data acquisition and processing, and quality control based on new surveying and mapping technology and equipment. The application was proved that this method was a new scheme to obtain reliable data results and improve work efficiency in rural real estate cadastral right investigation.

Key words rural real estate, cadastral right investigation, mobile measurement system, spatial data, attribute data (Page: 101)

Settlement Analysis of Xinjiayi Hotel in Quanzhou Based on PS-InSAR Technology

by WU Ping

Abstract Taking the Xinjiayi hotel which collapsed in March 2020 in Quanzhou City of Fujian Province for example, based on the multi-phase "Sentry-1" SAR image data about one year before the accident, we used the permanent scatterer technology to analyze the building settlement of the accident hotel. By monitoring the permanent scatterer points on the hotel and drawing the settlement curve, we found the phenomenon of the uneven settlement of the building before the collapse accident, which could provide data evidence for the accident cause analysis of related literature.

Key words PS-InSAR, hotel collapse, settlement analysis (Page: 105)

Transforming Method of the Third National Land Survey Field Data to Geographical Conditions Monitoring Field Data

by XU Weijie

Abstract In this paper, we demonstrated how to transform the third national land survey field data to geographical conditions monitoring field data by the idea of ETL pipeline processing and data filtering with the help of some tools like FME, Python and ArcMap. This method can finish the transform without any human interventions and prevent human errors except for the very beginning of the procedure, selecting photos.

Key words geographical conditions monitoring, the third national land survey, field check, ETL, FME, Python (Page: 109)

Contour Extraction of LiDAR Point Cloud Data Based on 3D Mapping Software

by DU Jinli

Abstract In this paper, based on Geoway 3D Mapping software, we filtered the DEM data of 2 m × 2 m interval generated by LiDAR point cloud processing, and then extracted the contour data of 1:10 000. Through experiments, we summarized the optimal method and steps of extracting contour and the common problems and solutions in production. The results of contour extraction have lifelike landform morphology, good contour continuity, beautiful shape and high precision, which can meet the requirements of specifications, and provide a good technical guarantee for improving the mapping quality and operation efficiency.

Key words LiDAR, DEM, contour, altitude point (Page: 113)

Representation Method of Ancient Chinese Map Symbols

by ZHANG Hanmei

Abstract Map content is expressed by map symbols. With the continuous enrichment of map content and the continuous improvement of map accuracy requirements, map symbols are constantly evolving. Based on the study of representative ancient Chinese maps from Warring States to Qing Dynasty, we summarized the three representation methods of ancient Chinese map symbols, such as landscape painting, abstract painting and symbolization, and expounded the changes of representation methods of ancient Chinese map symbols.

Key words ancient Chinese map, map symbol, landscape painting, abstract painting, symbolization (Page: 115)

Exploration and Reform of Carrying out the Ideological and Political Theories Education in the Cartography Course

by HUANG Jianyi

Abstract In this paper, we made an initial exploration of the discipline advantage and main direction of ideological and political theories education, based on the discipline characteristics of the *Cartography* course, in order to integrate the intangible value education and tangible professional knowledge, which could provide a reference for other professional courses.

Key words ideological and political theories education, course characteristic, Cartography (Page: 118)

Route Planning and Blind Area Identification of Illegal Land Dynamic Patrol

by ZHANG Yingyu

Abstract Based on the establishment of dynamic patrol track data and case business database of illegal land use, we used the buffer analysis and erasure analysis to calculate the blind area of patrol, and used the overlapping analysis and grid processing of case business data and patrol blind area to analyze the correlation between case patrol route and case business data, so as to obtain suggestions for optimizing the patrol route of illegal land use, and provide decision support for law enforcement of natural resources.

Key words buffer analysis, overlapping analysis, patrol track, GIS, natural resource law enforcement (Page: 121)

Research on the Technical Methods of Compiling Natural Resource Balance Sheet Based on Geographical Conditions Information

by SU Donghua

Abstract Taking Guang'an City as the research area, we researched the technical methods of compiling natural resource balance sheet from three aspects, such as quantity quality and value, and used the combination method of numerical statistics and spatial distribution patterns to compile the natural resource balance sheet of Guang'an City. The results show that ①as the basic data for making natural resource asset dataset, the geographical conditions vector data have great advantages in the area statistics and spatial distribution of natural resources assets, which can directly reflect the above two indexes. ②Multi-spectral remote sensing image data can be used to calculate some quality and value indexes. ③The compilation of natural resource balance sheet needs the support of multi-sectoral professional data.

Key words geographical conditions information, natural resource balance sheet, compilation, Guang'an City (Page: 126)