引文格式：王增利,雷科,朱宣霖,等. 在线地理编码平台精度指标的分析与评价[J]. 地理空间信息, 2025, 23(1): 110-114. doi:10. 3969/j.issn. 1672-4623.2025.01.024

在线地理编码平台精度指标的分析与评价

王增利[1]，雷　科[1*]，朱宣霖[2]，任亚利[2]
（1. 河海大学 地理与遥感学院，江苏 南京 211100；2. 河海大学 地球科学与工程学院，江苏 南京 211100）

摘　要：以某南方大型城市A区的2 394条地址数据为基础，引入混淆矩阵法对百度、高德、腾讯、天地图4个平台的精度指标进行定量分析，并利用聚类分析进行分类。实验结果表明，百度的总体编码数据质量最高；在各平台的指标中，高德的“门牌号”、百度的“80”、腾讯的“9”、天地图的“90”等指标表现稳定且具有较高精度，其中高德的“门牌号”数据质量最高，约70%的误差在50 m以内，超过90%的误差在500 m以内；其他指标的排序与网站介绍的等级顺序存在差异。研究结果可为网络地理编码平台用户选择数据提供依据。

关键词：地理编码；精度指标；混淆矩阵；聚类分析；网络地理编码平台；精度评价
中图分类号：P208　　**文献标志码：**B　　**文章编号：**1672-4623（2025）01-0110-05

Analysis and Evaluation of Accuracy Indexes of Online Geocoding Platforms

WANG Zengli[1], LEI Ke[1], ZHU Xuanlin[2], REN Yali[2]
(1. School of Geography and Remote Sensing, Hohai University, Nanjing 211100, China;
2. School of Earth Science and Engineering, Hohai University, Nanjing 211100, China)

Abstract: Based on 2 394 address data in District A of a large southern city, we used confusion matrix method to quantitatively analyze the accuracy indexes of Baidu, AutoNavi, Tencent and MapWorld, and used cluster analysis for classification. The experimental results show that Baidu has the highest overall coding data quality. Among the indexes of various platforms, “house number” of AutoNavi, “80” of Baidu, “9” of Tencent, and “90” of MapWorld perform have stable performance and high accuracy. Among them, “house number” of AutoNavi has the highest quality, with about 70% of the error below 50 m, and more than 90% of the errors within 500 m. The ranking of other indexes differs from the rank order described on the site. The results of this study can provide a basis for the users of network geocoding platform to select data.

Key words: geocoding, accuracy index, confusion matrix, cluster analysis, network geocoding platform, accuracy evaluation

移动互联网推动了导航与位置服务的发展，地理编码作为其关键技术，为坐标获取、路线规划和物流配送等应用提供了便利[1-6]。目前国内网络地图服务商开发了百度、高德、腾讯、天地图等网络地理编码平台，实时高效地将地址文本转化为空间坐标[7]；但由于各平台使用的参考地址数据和地址匹配算法不同，编码数据质量存在差异。为方便用户准确评估和选择使用数据，各平台均设置了相应指标来反映编码数据的质量，如百度地图将精度指标设置为0～100，数值越高精度越高；高德地图以匹配级代表编码数据质量，包括门牌号、道路等。精度指标作为评价平台服务质量的重要因素，直观展现了地理编码服务的质量。数据选取和质量评价时，现有研究往往假设平台提供的精度指标按顺序排列且稳定可靠，并基于这些指标对数据进行选取，如廖薇薇等[8]优化多源网络地理编码结果时，假设腾讯平台编码 “7”以下的数据为不准确；张弘弢等[9]使用多源网络地理编码服务时选取最高等级的数据作为准确数据。目前尚未对网络编码平台提供的指标可靠性进行评价；且各平台使用的精度指标不同，导致各平台之间、指标之间很难进行对比[10]。鉴于此，本文引入并改进了混淆矩阵方法，对各平台的指标进行了定量分析；再利用聚类优化方法对各指标进行排序和分类，进一步验证定量分析的结论；最后综合定量与定性分析结果，评价各平台精度指标的可靠性。

项目来源：教育部人文社科基金资助项目（22YJAZH108）；中央高校业务费资助项目（B220201021）。
第一作者简介：王增利（1984—），博士，副教授，主要研究方向为空间数据不确定性分析、面向公共安全大数据的空间模式分析及预测、地理空间数据挖掘与模式识别，E-mail：wzlgeo@hhu.edu.cn。
通信作者：雷科（2001—），本科，主要研究方向为地理编码、空间数据分析，E-mail：18893011287@163.com。

1　研究方法

1.1　基于混淆矩阵的指标精度分类评价

混淆矩阵也称误差矩阵，通过对比分类结果与正确分类结果来评价分类结果的精度。为使混淆矩阵可用于精度评价，本文对其进行改进：真实值采用距离阈值 D 代替，由一个数据的误差距离是否在 D 之内决定其归类于第一列还是第二列；预测值采用精度指标代替，由一个数据的编码结果指标是否属于该等级决定其归类于第一行还是第二行。一个指标是否能准确反映数据的质量等级代表了该指标的可靠性。

TP 类表示编码结果距离误差小于 D 且其指标为要筛选的指标名称；*FN* 类表示距离误差小于 D 但不为目标指标；*FP* 类表示指标名称为要筛选的指标但距离误差大于 D；*TN* 类表示既不是目标指标且距离误差大于 D。根据模型的分类结果，统计各类别的数量并计算精确率、召回率和F1_score等精度指标。其中，精确率反映一个指标对应的匹配结果接近真实位置（$\leqslant D$）的概率，值位于0~1之间，值越高说明该指标在距离 D 的分类效果越明显；召回率反映通过编码平台计算获取准确结果（$\leqslant D$）的概率，值位于0~1之间，值越高说明该指标分类的数据与其他指标的差异越大；F1_score反映指标精度和区别于其他指标的程度，值位于0~1之间，值越高说明该指标的精度和区分度越高。

1.2　基于系统聚类的定量分析结果优化

聚类是一种寻找数据间内在关系的方法。研究各平台指标时发现，有些指标代表的精度范围存在明显的相似性，可根据F1_score结果将相似的指标聚合为一类，不同指标分到不同的类别[8]。实验时每个指标为1个类，利用逐步聚类方法，以各指标F1_score的欧氏距离为特征，每次将距离最小的两个类合并为1个新类，直至所有指标合并为1类。观察聚类时生成的层次聚类树，完成聚类分析，并获取各指标所处的精度层级。各指标中包含的F1_score为聚类要素$\{X_1, X_2, \cdots, X_m\}$，则两个指标间欧氏距离的计算公式为[11]：

$$d_{ij}=\sqrt{\sum_{k=1}^{m}\left(X_{ik}-X_{jk}\right)^2}$$

式中，X_{ik}、X_{jk} 为样本 i、j 第 k 个聚类要素的值。

2　研究数据与分析

2.1　研究数据与处理

实验采用201*年某南方大型城市A区的地址数据，包含省、市、区县、道路名称+门牌号、小区名+楼号等信息。经人工检查，这些地址中没有出现错别字，且每条地址都可精确到建筑物所在的地块。该市A区占地面积约49.11 km^2，下设12个街道、112个社区，为实验提供了2 394条数据，主要内容为地址文本与对应的精确空间位置坐标数据。实验还采用从百度、高德、腾讯、天地图4个平台获取的对应的空间坐标数据和精度指标数据。

数据处理包括统一坐标系、获取真实位置和计算误差距离3个步骤。首先利用WanderGIS（2015）提供的坐标转换程序将所有实验数据投影至GCS_WGS_1984坐标系[12-13]；然后利用测绘部门的标准地籍数据库对实验数据进行配准，并对数据进行人工纠正，得到实验中用于比较分析的真实值；最后将纠正后的坐标与编码坐标导入ArcGIS中，计算编码坐标到真实坐标的距离，即编码结果的误差。

2.2　编码平台的精度指标

不同平台选用的精度指标规则存在较大差异。百度平台以误差距离为判定数据精度的标准，高德、腾讯、天地图则以地址文本为基础对编码数据精度进行标记；但这些平台指标都是按照一定顺序排列，且每个指标对应一定的精度等级。百度平台以坐标点误差距离为编码结果提供了11个指标，范围在0~100之间[14]，每个指标对应一定的误差范围（表1）。高德平台根据地址文本表征的含义为地理编码结果设置了15个精度等级，匹配级变化范围由国家逐步缩小到一个具体的站点[15]。匹配级所表示的区域大小由地址文本的具体内容决定，具体等级为：国家、省、市、区县、开发区、乡镇、村庄、热点商圈、兴趣点、门牌号、单元号、道路、道路交叉路口、公交站台和地铁站、未知。腾讯平台根据地址解析范围确定了11个精度等级[16]，从1~11级依次对应城市，区、县，乡镇、街道，村、社区，开发区，热点区域、商圈，道路，道路附属点，门址，小区、大厦和POI点。天地图平台根据输入的地址文本等级返回对应的指标[17]，指标范围为0~100，指标值越小，编码结果的质量越差。根据指标值大小，以整10的值为区间确定指标分类（表2）。

表1　百度编码数据指标及误差范围

可信度	坐标点误差范围/m	可信度	坐标点误差范围/m
100	<20	≥50	<1 000
≥90	<50	≥40	<2 000
≥80	<100	≥30	<5 000
≥75	<200	≥25	<8 000
≥70	<300	≥20	<10 000
≥60	<500	–	–

表2 天地图可靠性值与指标名对照表

可靠性值	指标分类名称	可靠性值	指标分类名称
100	天地图90	57	天地图50
96	天地图90	43	天地图0
89	天地图80	20	天地图0
71	天地图70	9	天地图0
60	天地图60	–	–

2.3 研究数据分析

1）空间分析。以100 m、500 m、1 000 m、5 000 m、5 000 m以上为区间对实验数据进行数量统计，结果表明，误差距离在100 m以内时，百度和高德的数据较多，腾讯次之，天地图最少；在100~500 m时，天地图的数据最多，腾讯次之，百度和高德较少；在500~1 000 m以及1 000~5 000 m两个区间时，天地图数据较多，其他平台数据较少；在5 000 m以上时，百度没有数据，高德、腾讯数据少于10个，天地图数据为39个。

2）统计分析。从中值、平均值、标准差来看（表3），百度平台的误差距离最大值、中值、平均值、标准差均最小；高德平台的误差距离中值、平均值、标准差仅大于百度；腾讯平台的中值、平均值、标准差仅小于天地图；天地图的中值、平均值、标准差均最大，最大误差次于腾讯，该平台的地理编码结果最差。

表3 点位误差统计分析表/m

平台	最大值	最小值	中值	平均值	标准差
百度	4 894.15	0	23.93	119.22	149.85
高德	6 837.63	0	26.27	141.03	186.50
腾讯	78 986.96	0	44.51	186.62	237.16
天地图	8 400.41	0	98.65	559.37	724.31

3 研究结果与分析

3.1 混淆矩阵精度分析结果与分析

1）精确率对比与分析。由各平台的精度指标和实验结果可知，大多数误差位于1 000 m以内，因此将混淆矩阵模型中的 D 设置为0~1 000 m。一个指标的精确率从左向右逐渐增长，若很快趋于稳定，则说明该指标的编码数据误差大多在某个距离内；若没有趋于稳定，则说明该指标的编码数据误差分布较分散，且没有明确地范围，分类效果不佳。

实验结果表明（图1），大多指标的曲线从50 m开始逐步上升，说明各指标均有50 m以内的误差；曲线在700 m左右到达0.8，并逐渐变缓，说明这些指标的误差大多（>80%）位于700 m以内。百度“50”、高德“门牌号”、腾讯“10”、腾讯“9”、腾讯“6”的精确率起始在0.7附近，说明这些指标的编码数据中约有70%的数据在50 m以内；这些曲线在500 m左右最接近1，并开始变缓，说明这些指标的大多（>90%）数据误差在500 m以内。天地图“50”和天地图“0”从50 m到1 000 m始终逐渐上升，说明这两个指标有多个不同大小的误差，且1 000 m以内的误差占比比其他指标低。其中部分精度指标的结果为0、0.5、1，且这些结果不随距离的增加而变化或变化不显著，说明这些指标的误差分布在某个特定的区间内，如百度“20”、高德“区县”、腾讯“2”的数据均超过最大距离阈值1 000 m；腾讯“6”对应数据的距离误差小于50 m。高德“热点商区”的数据结果为0.5和1，当距离阈值为350 m时开始变化，说明该指标数据的误差距离一半在50 m以下，另一半在350 m以下。腾讯“8”、腾讯“3”的数据结果为0和1，变化点分别为150 m和100 m，说明前者的误差距离在100~150 m，后者的误差距离在50~100 m。天地图“80”的数据结果为0和0.5，变化点为150 m，说明该指标数据的误差距离一半在100~150 m，另一半在大于1 000 m。

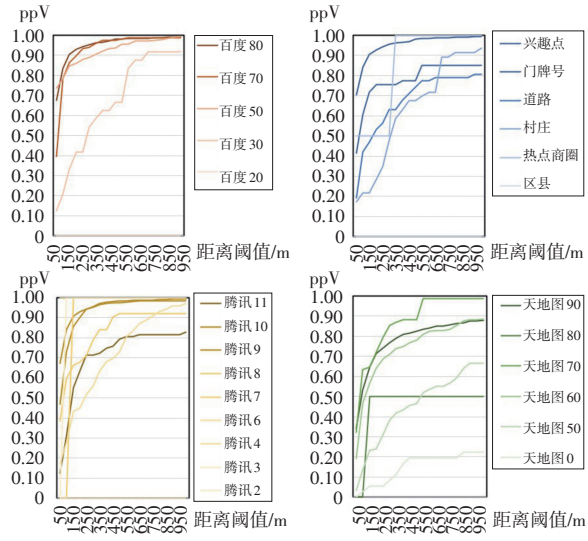


图1 各指标的精确率计算结果

2）召回率对比与分析。召回率的计算结果表明，大多数指标的召回率稳定在某个值附近，说明多数指标的误差分布较分散，没有准确的范围，这与精确率的分析结果一致。4个平台的所有指标中，高德“门牌号”的召回率最高，具有较高的区分度；其次依次是天地图“90”、百度“80”、腾讯“9”、腾讯“10”、百度“50”、天地图“60”；剩余指标的召回率均在0.1以下。从变化趋势来看，高德“门牌号”、天地图“90”、腾讯“9”、百度“50”等指标数据的召回率值在150 m以内有轻微下降，其中腾讯“9”下降较多，说明这些指标具有更多150 m以内的短误差；百度“80”、腾讯“10”、天地图“60”的召回率呈轻

微上升趋势，说明这些指标的代表性随距离的增加而轻微上升，但总体较平稳，这些与精确率的分析结果基本一致。

3）F1_score对比与分析。为反映各平台指标在不同距离阈值下的精度和代表性，综合各指标的精确率与召回率，计算了各指标在不同距离条件下的F1_score，结果表明高德“门牌号”、天地图“90”、百度“80”、腾讯“9”的F1_score较高，说明这些指标的误差比同平台内其他指标的区分度更高；腾讯“9”的误差集中在100 m以内，高德“门牌号”、百度“80”的误差大多集中在150 m以内，天地图“90”的误差集中在300 m以内。腾讯“10”、百度“50”的曲线大部分较平缓，说明这两个指标的误差在200 m以内，且与同平台其他指标具有一定的区分度。天地图“60”的F1_score在0.10~0.30区间上升，其余指标的F1_score均小于0.10，且分布密集。

3.2　聚类优化结果

利用聚类分析对F1_score进行聚类，进一步观察各指标的分类情况。高德“区县”、百度“20”、腾讯“2”的点位误差均大于最大距离阈值，召回率计算结果为0，导致F1_score结果为空，因此不参与聚类分析。聚类算法将所有指标分为5类：F1_score在0.85以上的为第一类；在0.60～0.85的为第二类；在0.30~0.60的为第三类；F1_score在0.30以下，但随着距离阈值的增加增长明显的，为第四类；F1_score在0.10以下，且变化不明显，为第五类（表4）。

表4　聚类分析

指标名	逐步聚类类别	指标名	逐步聚类类别
门牌号	1	村庄	5
百度80	2	热点商圈	5
腾讯9	2	腾讯11	5
天地图90	2	腾讯8	5
百度50	3	腾讯7	5
腾讯10	3	腾讯6	5
百度70	4	腾讯3	5
腾讯4	4	天地图80	5
天地图60	4	天地图70	5
百度30	5	天地图50	5
兴趣点	5	天地图0	5
道路	5	–	–

通过对F1_score结果进行分类，获取平台内部指标的精度大小排名以及具体分布情况（表5），从各平台来看，高德“门牌号”的精度最高，各距离阈值下F1_score均在0.80以上，其余指标结果均较差，对应的F1_score均小于0.10；百度、腾讯的指标结果分布类似，百度较好的两个指标的F1_score稳定于0.78、0.44，腾讯较好的两个指标的F1_score稳定于0.68、0.52，其余指标结果较差，均小于0.15；天地图“90”的结果不错，但其F1_score随距离阈值的增加变化较大，分布在0.50～0.85，除天地图“60”在0.20左右外，其余指标均小于0.10。各平台第一个指标以外的其他指标并不是按照顺序排列的。

表5　F1_score结果排序

平台	F1_score值结果排序					
百度	80	50	70	30	20	
高德	门牌号	兴趣点	道路	村庄	热点商圈	区县
腾讯	9	10	4	11	7	8、3、2
天地图	90	60	70	50	80	0

4　结　语

本文比较了百度、高德、腾讯、天地图4个网络地理编码平台精度指标的可靠性，并通过分析编码指标和综合排名，评价了它们的服务质量。

1）从平台内各指标排序来看，各平台除排序最高的指标外，其余指标的排序与平台介绍的排序有差异。各平台内排序最高的指标比其他指标精度高且更可靠。其中，高德“门牌号”、百度“80”、腾讯“9”、天地图“90”等指标在其平台内精度最高。

2）从平台内各指标误差分布来看，指标的误差大小分布不规律，将4个平台的指标分为3类：①误差大多（>80%）聚集在700 m以内，该类指标占据了大多数；②误差分散在1 000 m以内，主要为天地图“50”和天地图“0”；③误差超过1 000 m，主要包括百度“20”、高德“区县”、腾讯“2”以及一半的天地图“80”。

3）从平台间对比来看，各平台总体的编码质量存在较大差异，百度的总体编码数据质量最高，高德与腾讯质量接近，天地图的质量最低。从误差区分度来看，各平台的指标总体分为5个等级，高德“门牌号”的误差具有较高的区分度，其次依次为天地图“90”、百度“80”、腾讯“9”、腾讯“10”、百度“50”、天地图“60”。

根据上述结论，可为使用网络编码平台的用户提供以下建议：①在同时使用4个平台对同一批数据进行编码时，尽量选择带有高德“门牌号”、百度“80”、腾讯“9”、天地图“90”等标签的数据；此外，百度“50”、腾讯“10”、腾讯“6”等指标的误差也较小。②在只能选择一个平台作为编码平台的情况下，建议采用百度地理编码平台。

本文研究结果受限于实验区域和数据，对于其他区域和数据可能会有不同的结论。此外，日常使用的地址数据可能存在信息缺失和错误，与本文研究结果不一致。在后续研究中可以使用不同类型或质量的数据来验

证编码平台的精度和可靠性，进一步提升评价方法。本文提出的混淆矩阵改进方法可为其他地理编码数据质量评估提供参考，并对平台的精度指标进行评价。

参考文献

[1] 刘丹. 基于百度地图的多点导航系统实现[J]. 辽宁工业大学学报(自然科学版),2019,39(6):359-362

[2] 谭月,杨兴地. 基于“数字规划”平台的城市道路网络编码研究:以成都市中心城区路网为例[J]. 四川建筑,2016,36(5):10-13

[3] 边旭,沈树雄,王明兴,等. 通讯地址查询邮政编码系统设计[J]. 电子技术与软件工程,2016(6):193-194

[4] 肖圆秀, 钟先坤, 聂后德. “数字城市”地理信息公共平台建设研究[J]. 江西测绘,2013(1):11-13

[5] 彭明军,李宗华,刘辉,等. 基于中文分词的加权地理编码在COVID-19疫情防控空间定位中的应用 [J]. 武汉大学学报(信息科学版),2020,45(6):808-815

[6] 唐爽,向其权,邓仕虎,等. 基于GIS的市场主体分类监管系统设计与实现[J]. 地理空间信息,2020,18(8):47-49

[7] 田沁,巩玥,亢孟军,等. 国内主流在线地理编码服务质量评价[J]. 武汉大学学报(信息科学版),2016,41(10):1 351-1 358

[8] 廖薇薇,柳林,周素红,等. 多源在线地理编码服务分类优化模型[J]. 热带地理,2018,38(2):255-263

[9] 张弘弢,肖炼,周尧,等. 多源在线地理编码与地名检索服务聚合方法[J]. 地理与地理信息科学,2020,36(4):1-7

[10] 兀伟,邓国庆,黄琛,等. 地名地址地理空间位置与服务标准化探讨[J]. 地理空间信息,2018,16(11):10-13

[11] 徐建华. 现代地理学中的数学方法 [M]. 北京:高等教育出版社,1996

[12] 魏子卿. 2000 中国大地坐标系及其与 WGS84 的比较 [J]. 大地测量与地球动力学,2008,28(5):1-5

[13] 朱宣霖. 城市小区地址与道路地址的地理编码精度比较研究[J]. 江苏科技信息,2022,39(24):70-75

[14] 百度地图开放平台 [EB/OL]. (2019-06-18)[2023-05-20]. https://lbs.baidu.com/faq/api?title=webapi/guide/webservice-geocoding-base

[15] 高德开放平台 [EB/OL]. (2022-07-13)[2023-05-20]. https://developer. amap. com/api/webservice/guide/api/georegeo

[16] 腾讯位置服务 [EB/OL]. (2019-01-14)[2023-05-20]. https://lbs.qq.com//service/webService/webServiceGuide/webServiceGeocoder

[17] 国家地理信息公共服务平台 天地图 [EB/OL]. (2022-11-30) [2023-05-20]. http://lbs.tianditu.gov.cn/server/geocodinginterface. html

（上接第97页）

参考文献

[1] 穆屹,毛力增,钱汉涛,等. 北京市停车信息资源管理系统设计与实现[J]. 交通运输研究,2016,2(3):49-55

[2] 仝进,孙海瑞,周正. 北京市居住停车区域认证机制研究[J]. 城市交通,2019,17(5):66-70

[3] 王艳青,房燕. 公交都市建设地方性立法初步研究:以上海市公共汽电车行业为例[J]. 交通与港航,2022,9(2):46-50

[4] Liu J,Guan H Z,He Y L,et al. Searching for a Parking Spot in a Large City Center[J]. Journal of Highway and Transportation Research and Development (English Edition),2017,11(1):236-242

[5] Cai Y F,Chen J,Zhang C,et al. A Parking Space Allocation Method to Make a Shared Parking Strategy for Appertaining Parking Lots of Public Buildings[J]. Sustainability,2018,11(1):124-132

[6] Philip P, Selvaperumal S K, Lakshmanan R, et al. IoT Based Car Parking Management System with an Application Program Interface[J]. Journal of Physics: Conference Series,2019(1):36-45

[7] 史冲. 城市停车场信息管理系统设计[J]. 地理空间信息,2014,12(3):59-62

[8] Pitale A M,Parida M,Sadhukhan S. GIS-MCDM-based Approach to Determine the Potential Facility Locations for Park-and-Ride Facilities Along Transit Corridors[J]. Journal of Urban Planning and Development,2022(1):148

[9] 吴才锐,李铭,汪益纯. 基于停车普查的城市停车供需特征研究:以西宁市为例[C]//中国科学技术协会,交通运输部,中国工程院. 2018 世界交通运输大会论文集,2018:1 595-1 604

[10] 范东旭,刘威,赵云龙. 沈阳市停车普查实践与供需特征研究[J].《规划师》论丛,2020(1):266-272

[11] 周曦冰,王爽,谭龙,等. 西安市停车资源现状与空间分布研究[J]. 地理空间信息,2022,20(1):75-78

[12] 刘娈娈. 城市停车设施普查档案管理思考[J]. 黑龙江档案,2022(3):196-198

[13] 谢文瑄,张翼然,胡海鹤,等. 地理信息技术在停车设施信息普查中的应用[J]. 城市勘测,2022(6):41-44

[14] 王何忆. 现代建筑院落中的中国传统院落空间特质研究[D]. 杭州:浙江大学,2014

[15] 自然资源部办公厅. 关于开展2020年度全国国土变更调查工作的通知[EB/OL]. (2020-12-16)[2023-05-21]. http://gi.mnr. gov. cn/202012/t20201216_2595331. html

[16] Haklay M. How Good is Volunteered Geographical Information? A Comparative Study of OpenStreetMap and Ordnance Survey Datasets[J]. Environment and Planning B: Planning and Design,2010,37(4):36-49

[17] 杨芳. 基于ArcMap计算第三次国土调查线状地物宽度[J]. 测绘技术装备,2019,21(3):65-66