引文格式：宁夏薇,夏林元,祝宏宇,等. 基于增强指纹特征的WiFi室内定位方法[J]. 地理空间信息, 2024, 22(8): 102-106. doi:10. 3969/j.issn. 1672-4623.2024.08.023

基于增强指纹特征的WiFi室内定位方法

宁夏薇[1]，夏林元[1*]，祝宏宇[1]，刘俊求[1]
（1. 中山大学 地理科学与规划学院，广东 广州 510006）

摘　要：基于位置指纹的WiFi定位是室内常用的定位方法之一，但WiFi位置指纹库存在无线接入点（AP，access points）分布及稀疏样本等问题，严重影响室内局部区域的定位效果。为此，提出一种基于增强指纹特征的WiFi室内定位方法。在指纹库建立阶段，通过基于接收比率与方差的AP优选策略，筛选出更加稳定、质量更好的AP点；再对相应位置指纹库进行扩充，实现在减少前期建库工作量的同时提高指纹库质量；最后在匹配定位阶段，融合皮尔逊相关系数构建加权的k最邻近（k-Nearest Neighbor，kNN）定位算法。实际定位场景测试结果表明，与常规的位置指纹WiFi室内定位方法相比，该方法将定位精度提高了22.5%。所提出的基于增强指纹特征的WiFi室内定位可以在降低指纹库采集成本的同时，有效提升定位效益与精度。

关键词：指纹特征增强；稀疏样本；相关系数；WiFi室内定位
中图分类号：P228　　**文献标志码：**B　　**文章编号：**1672-4623（2024）08-0102-05

WiFi Indoor Location Method Based on Enhanced Fingerprint Feature

NING Xiawei[1], XIA Linyuan[1], ZHU Hongyu[1], LIU Junqiu[1]
(1. School of Geography and Planning, Sun Yat-Sen University, Guangzhou 510006, China)

Abstract: Fingerprint-based WiFi location is one of the commonly used indoor location methods, but the WiFi location fingerprint database has problems, such as wireless access point (AP) distribution and sparse samples, which seriously affect the location effect in indoor local areas. For this reason, we proposed a WiFi indoor location method based on enhanced fingerprint feature. Firstly, at the stage of building a fingerprint database, we used AP preferential selection strategy with joint reception percentage and variance to screen out more stable and better quality AP points. Then, we expanded the fingerprint database in the corresponding location, so as to reduce the workload of establishing the fingerprint database in the early stage and improve the quality of fingerprint database. Finally, at the stage of matching location, we fused the Pearson correlation coefficients to construct a weighted k Nearest Neighbor (kNN) location algorithm. Test results of actual location scenarios show that the location accuracy is improved by 22.5% using AP-preferred fingerprint database and the weighted kNN location algorithm incorporating Pearson correlation coefficients. The method can effectively improve the location accuracy while reducing the cost of capturing fingerprint databases.

Key words: fingerprint enhancement, sparse sample, correlation coefficient, WiFi indoor location

目前室内定位研究领域百花齐放，仍没有形成一个简单、有效的技术框架与方法[1-2]，主流的定位技术包括WiFi、蓝牙、超宽带、地磁、LED可见光、惯性传感器、RFID、ZigBee、伪卫星、计算机视觉等[3]。然而，上述技术仍存在一定的局限性及成本与便捷性问题，并不能满足所有场景的室内定位需求[4-5]。超宽带、计算机视觉等定位技术精度较高，可以达到厘米级，但设备成本过高、算法也较为复杂；蓝牙和ZigBee技术则需要提前布设大量设备，且传播距离较短；地磁定位技术虽无需额外部署设备，但室内区域电磁干扰较为严重，易降低匹配的准确率。随着5G技术发展，毫米波定位、窄带（NB）定位等技术也在发展之中。相比于其他室内定位技术，WiFi定位具有资源大众化、可操作性强、模型鲁棒性强和定位精度高等优点，近年来大量研究基于位置指纹的WiFi室内定位方法展开[6]。基于位置指纹的定位方法使用WiFi信号强度进行建模和测量，无需确定AP的确切位置。同时，因WiFi装置的广泛部署、使用以及智能移动设备的普及，位置指纹定位技术可在大多数室内场所使用，而且WiFi网络部署过程中不需要增加其他额外的硬件[7]。故相比于其他室内定位方法，基于位置指纹的WiFi定位技术具有成本低、覆盖范围广等优

收稿日期：2023-04-04。
项目来源：广东省重大专项资助项目（2015B010104003）；广州市产学研协同创新重大专项资助项目（201604046007）。
第一作者简介：宁夏薇（1996—），硕士研究生，研究方向为室内外导航与定位，E-mail：ningxw@mail3.sysu.edu.cn。
通信作者：夏林元（1964—），博士，教授，研究方向为卫星定位及其应用、室内外无缝定位、泛在定位与位置服务，E-mail：xialiny@mail.sysu.edu.cn。

点，且精度能够一定程度上满足室内定位需求。但在实际室内场景应用中，WiFi室内定位方法易受到环境影响造成信号波动，从而导致定位精度降低；另外，利用现有无线接入点意味着无法掌控所有AP的质量，导致WiFi信号的质量存在较大差距，这对于定位效果的影响非常严重；利用位置指纹定位的前期需要建立位置指纹库，这项工作需要消耗一定的人力物力成本。针对以上问题，本文提出一种基于增强指纹特征的WiFi室内定位方法。

1　位置指纹定位技术

1.1　基于位置指纹的室内定位方法

基于位置指纹的定位方法可以分为2个阶段，离线建库阶段和在线定位阶段：

1）离线阶段进行位置指纹数据的采集，并建立指纹库。首先利用指纹技术将所有参考点的RSSI特征值进行采集，然后对所采集数据进行预处理并建立指纹库。如表1，包含一个指纹库中应有的主要信息。

表1　指纹库信息

	MAC_1	MAC_2	MAC_3	…	MAC_n
$F(x_1,y_1)$	$RSSI^1_{(x_1,y_1)}$	$RSSI^2_{(x_1,y_1)}$	$RSSI^3_{(x_1,y_1)}$		$RSSI^n_{(x_1,y_1)}$
……					
$F(x_1,y_1)$	$RSSI^1_{(x_i,y_j)}$	$RSSI^2_{(x_i,y_j)}$	$RSSI^3_{(x_i,y_j)}$		$RSSI^n_{(x_i,y_j)}$

设指纹库 S，S 中所包含指纹用 $F(x,y)$ 来表示，其中 x 和 y 为该指纹对应参考点在物理空间中的坐标。$F(x,y)$ 中包含 n 个RSS数据，对应 n 个无线接入点（AP），每个AP有其对应的唯一的MAC地址。

2）在线阶段则是利用定位算法将待测点指纹信息与指纹库中信息进行匹配，从而确定待测点位置。在这一阶段，所使用的定位算法将直接影响定位的精度和效率。

1.2　加权kNN算法

k 最邻近（k-Nearest Neighbor，kNN）算法是机器学习中的一种监督学习方法[8]。利用 k 最近邻算法实现基于位置指纹的室内定位，针对一个待测点X，通过计算得到指纹库中所有参考点与待测点X之间的欧氏距离，可用RSSI值表示为：

$$L_i=\sqrt{(RSSI_1-RSSI_{i1})^2+\cdots+(RSSI_n-RSSI_{in})^2} \quad (1)$$

式中，L_i 为第 i 个参考点与待测点间的距离；$RSSI_1$，$RSSI_2$，…，$RSSI_n$ 为待测点所能接收到的 n 个AP的信号强度特征值；$RSSI_{i1}$，$RSSI_{i2}$，…，$RSSI_{in}$ 为第 i 个参考点所能接收到的 n 个AP的信号强度特征值。

按照欧氏距离排序得到最近的 k 个参考点，将这 k 个参考点坐标取平均值，便得到了待测点的位置坐标。

加权 k 最邻近（Weighted k-Nearest Neighbors）算法，是在kNN算法基础上进行改进得到的一种算法[9-10]。它对距离最近的 k 个参考点进行赋权，赋权的依据是各参考点与待测点之间的距离，距离越近，权值越大。待测点坐标则被表示为：

$$(x,y)=\frac{\sum_{k=1}^{K}w_k(x_k,y_k)}{\sum_{k=1}^{K}w_k} \quad (2)$$

式中，$w_k=1/L_k$，为第 k 个参考点的权重。

1.3　融合皮尔逊相关系数的加权kNN算法

基于欧氏距离的加权kNN算法使用欧氏距离为 k 个最近邻参考点赋权。但在复杂的室内环境中，信号的多径、散射、反射等现象均会影响RSS观测值，使其并非按照对数距离损耗模型变化。为了更好地表征样本间的相关性，引入皮尔逊相关系数与欧氏距离共同定权。

皮尔逊相关系数与加权kNN算法融合的指纹匹配步骤主要为：

1）算出待测点与各个参考点的欧氏距离 L_i，并将各参考点所得欧氏距离求和。

2）将待测点到各参考点的欧氏距离进行排序，选取欧氏距离最小的 k 个参考点指纹进行最后的定位。

3）计算 k 个参考点各自与待测点的皮尔逊相关系数 r_i。

4）计算 k 个参考点各自的权重。利用欧氏距离与相关系数融合加权，公式如下：

$$W=1/[(A-r_i)\times L_i] \quad (3)$$

式中，A为常数；r_i 为相关系数；L_i 为欧氏距离。

最后，计算待测点位置。

2　指纹特征增强

2.1　基于接收比率与方差的AP优选策略

目前已有的AP选择策略主要有随机AP选择算法、基于RSSI最大均值的AP选择算法（max mean，MM）[10]，以及基于RSSI均值最大累计差分（max cumulative difference，MCD）的AP筛选算法。这些方法没有顾及到接收信号强度过弱、AP信号的接收率低以及不同路由设备引起的信号差异性等问题对最终定位精度的影响。

综合考虑上述因素，本文提出一种基于接收比率与方差的AP优选策略，具体过程如图1所示。首先根

据WiFi信号特性对于所采集到的RSSI数据信息做出分析，根据其强度设立一个合适的阈值，仅保留在此阈值范围内的RSSI数据；然后按照AP分类，分析每个AP对于所有参考点（reference point，RP）,计算出能够接受到该AP的RP所占比例，即为接收比率。设立阈值，将接收比率低于此阈值的AP剔除，然后保留剩余AP的接收比率特征，以配合后续处理。经大量实验证明，当一个AP在全部参考点间接收比率小于60%时，将对定位效果产生可观的不利影响。利用长时间序列的观测数据，计算剩余每个AP的均值、极差和方差，对每一个AP节点的稳定性做出评价。将AP节点接收比率与稳定性相结合，最终筛选出信号强度较强、较稳定的若干个AP。

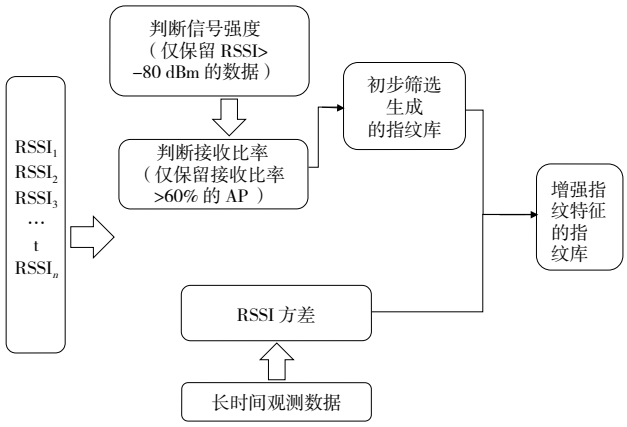


图1 AP筛选过程

2.2 指纹库扩充

对于基于WiFi指纹库的室内定位，离线阶段构建指纹库是一个特别费时费力的过程，而指纹库的构建又直接影响到定位精度。在真实的室内环境中，往往会有大量的位置点无法正常获取指纹数据。根据WiFi的传播特征我们可以得知，不同参考点所接收到的同一个AP点的WiFi信号强度，在空间上是呈现一定相关性的。那么在由于客观原因无法得到足量指纹数据而只能得到稀疏指纹数据样本的情况下，便可以根据这一特性扩充指纹库。即根据已采集到的指纹数据及其对应位置信息，在其他位置生成虚拟的指纹。同样，室内复杂场景中的定位薄弱区所产生的异常值，也可以通过此方法对指纹进行替换。

空间插值方法可以利用给定一组离散点或子区域形式的空间数据，找出最能代表整个曲面的函数，并预测其他点或其他子区域的值。克里金插值是一种在一定区域范围内对变量进行最优无偏估计的方法，它的理论基础主要包括变异函数理论和结构分析，属于地统计学的主要内容之一[11]。克里金插值方法符合地统计学的基本统计特征，因此使用其进行插值在计算得到预测结果的同时，还能得到预测误差,这种特性有利于对预测结果的不确定性进行一定的评估[12]。

克里金插值方法对周围已知点的特征值经过加权得到待测位置的预测值：

$$\hat{Z}_0=\sum_{i=1}^{N}\lambda_i Z_i \tag{4}$$

式中，$\hat{Z}$ 为待测点处的估计值；Z_i 为参考点 i 处的测量值；N 为参考点数；λ_i 为参考点 i 的权重，它是能够使得待测点处估计值与真实值的差值最小的一系列最优系数，即

$$\min_{\lambda_i}\mathrm{Var}(\hat{Z}_0-Z_0) \tag{5}$$

普通克里金插值方法假设条件为此空间均一，空间中各点其属性值是均一的[13]。因此，在空间内任意一点（x，y），都有相同的期望和方差，即

$$E[z(x,y)]=E[z]=c \tag{6}$$

$$\mathrm{Var}[z(x,y)]=\sigma^2 \tag{7}$$

同时，还需要满足无偏估计条件：

$$E(\hat{Z}_0-Z_0)=0 \tag{8}$$

变异函数是克里金法和模拟变化的关键输入，空间中各点的空间变异性通过变异函数进行量化处理[14]。考虑对于空间中的某一点，则变异函数可被定义为：

$$2\gamma(h)=E\{[Z(\mu)-Z(\mu+h)]^2\} \tag{9}$$

对于有限的观测量而言，半变异函数可被定义为：

$$\gamma(h)=\frac{1}{2N(h)}\sum_{i=1}^{N(h)}[Z(\mu_i)-Z(\mu_i+h)]^2 \tag{10}$$

半变异函数可通过经验模型进行描述，以量化不同观测值之间的空间自相关性。不同的经验半变异函数模型由变程、基台和块金共同构成。其中，变程是指该半变异函数模型的值的变化范围，当超过变程时，半变异函数达到极大值，即基台值；块金值是半变异函数在y轴上的截距，尽管该值理论上应为0，但对于无穷小的间距而言，由于存在测量误差、采样间隔大于0等原因，其变化通常要大于0。

3 实验及结果分析

3.1 实验环境

实验选在广东省广州市中山大学东校园工学院某办公室内进行，该办公室为长约20 m、宽约9 m的矩

形，面积约180 m²。办公室内分布着9组共计38个工位，能够更加真实地还原人们所需要定位的真实场景。将实验场所内可供行人活动区域分为1 m×1 m的方格，形成了8×18共计144个小方格。每个小方格的顶点作为参考点，点（0，6）和点（0，13）处为两根立柱，点（4，12）处难以采集WiFi信号，除去以上3点剩余共计168个参考点。

3.2　构建指纹库

RSSI数据的采集使用搭载Qualcomn Atheros AR956x Wireless Network Adapter无线网卡的笔记本电脑进行，用于采集的软件为“SMPex7”。

在实验区域内共进行了7次指纹库数据采集。每次采集均是在所涉及参考点处采集10 s，对采集数据进行处理后形成对应指纹库。采集过程中保留每个点能够接收到的所有WiFi的RSSI数据，如表2所示。

表2　所用指纹库的基本信息

指纹库编号	参考点个数	是否包括边界	采样比例/%
A	168	是	100
B	102	是	60
C	85	是	50
D	48	是	28.5
E	118	否	70.2
F	85	否	50
G	102	否	60

3.3　在线匹配定位及结果分析

为验证2.1中提出基于接收比率与方差的AP优选策略的有效性，利用该优选策略对指纹库 *A* 进行AP优选，生成新的指纹库 *H* 。分别使用指纹库 *A* 和指纹库 *H* 进行定位。 *K* 值选取3、4、5、6，AP优选前后的定位误差平均值如下表。经AP优选的指纹库定位精度，较未优选指纹库平均提升了约22.5%。如图2所示，未经优选指纹库定位误差小于4 m的概率为75.3%。而经过优选的指纹库，这一概率为88.8%，较优选前提高了13.5%。

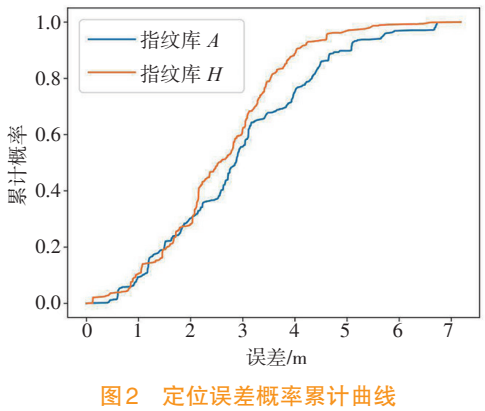


图2　定位误差概率累计曲线

为验证指纹库扩充的有效性，对经指纹库扩充后的指纹库进行定位误差分析。将指纹库 *B* 、 *C* 、 *D* 和指纹库 *E* 、 *F* 、 *G* 分别进行AP优选，经指纹库扩充后与指纹库H进行定位效果比较。指纹库 *B* 、 *C* 、 *D* 是包含边界的样本比例分别为60%、50%、28.5%的指纹库，定位误差及其累积概率分布如表3所示。可以看出样本比例为60%和50%的指纹库，经指纹库扩充处理后定位效果与全样本指纹库相近。定位误差小于3 m的概率分别为65.8%和64.5%，略高于全样本指纹库。即采用60%和50%的指纹库，经扩充后用于定位，其定位效果可与包含全部参考点信息的指纹库相当。那么在前期采集指纹数据阶段，便可减少一半的工作量。而样本比例28.5%的指纹库则定位效果相对较差，平均定位误差达到2.927 m，误差小于3 m的概率为54.2%。

表3　各指纹库定位误差平均值2

指纹库	*H*	*B*	*C*	*D*
定位误差平均值/m	2.572	2.605	2.617	2.927

指纹库 *E* 、 *F* 、 *G* 分别是不包含房屋边界的样本比例为70.2%、60%及50%的指纹库，定位效果如表4所示。可以看出3个指纹库与全样本指纹库的定位效果相近，甚至略好。定位误差小于3 m的概率分别为61.2%、71.5%及69.9%。样本比例同为60%，指纹库 *B* 的定位误差为2.605 m，而指纹库 *F* 则是2.508 m；样本比例同为50%，指纹库 *C* 的定位误差为2.617 m，而指纹库 *G* 则是2.543 m。这表明将位于房屋边界等薄弱区的参考点剔除后进行指纹库扩充，得到的指纹库定位效果更优。

表4　各指纹库定位误差平均值3

指纹库	*H*	*E*	*F*	*G*
定位误差平均值/m	2.572	2.508	2.493	2.543

4　结　语

针对目前室内定位研究领域仍没有形成一个简单、有效的技术框架，围绕WiFi室内定位展开研究，本文提出了一种基于增强指纹特征的WiFi室内定位方法。使用融合皮尔逊相关系数的加权kNN算法，并结合基于接收比率与方差的AP优选策略，能够有效提高WiFi指纹定位精度；利用克里金插值方法对样本稀疏的实测指纹库进行扩充，从而实现对指纹库进行增强的效果，且能达到与真实采集指纹库相近的定位精度。这将有利于降低复杂室内场景中薄弱点的异常值对定位效果的影响。

参考文献

[1] 宁津生,姚宜斌,张小红. 全球导航卫星系统发展综述[J]. 导航定位学报,2013,1(1):3-8

[2] Al-Ammar M A, Alhadhrami S, Al-Salman A, et al. Comparative Survey of Indoor Positioning Technologies Techniques and Algorithms[C]//International Conference on Cyberworlds (CW),2014:245-252

[3] 赵朋,何骞,程铭宇. 一种新的用于WiFi室内定位的最近邻搜索融合算法[J]. 测绘地理信息,2021,46(4):32-35

[4] Dinh T-M T,Duong N-S,Sandrasegaran K. Smartphone-based Indoor Positioning Using Ble iBeacon and Reliable Lightweight Fingerprint Map[J]. IEEE Sensors Journal, 2020, 20(17):10283-10294

[5] Wisanmongkol J, Taparugssanagorn A, Tran L C, et al. An Ensemble Approach to Deep-learning-based Wireless Indoor localization[J]. Iet Wireless Sensor Systems, 2022, 12(2):33-55

[6] Paul A S, Wan E A. WiFi Based Indoor Localization and Tracking Using Sigma-point Kalman Filtering Methods[C]//IEEE on Position Location and Navigation Symposium,2008:172-185

[7] 王威,陈正宇,张帅,等.多类智能终端基于灰色关联度的WKNN定位效果分析[J].地理空间信息,2021,19(3):64-67

[8] Zhang S, Li X, Zong M, et al. Efficient kNN Classification with Different Numbers of Nearest Neighbors[J]. IEEE Transactions on Neural Networks and Learning Systems, 2018,29(5):1 774-1 785

[9] 李新春,王欢. 基于稀疏指纹采集和改进WKNN的定位算法[J]. 重庆邮电大学学报(自然科学版),2019,31(4):451-459

[10] Deng Z, Ma L, Xu Y, et al. Intelligent AP Selection for Indoor Positioning in Wireless Local Area Network[C]//6th International ICST Conference on Communications and Networking in China (CHINACOM),2011:257-261

[11] Oliver M A, Webster R. Kriging: A Method of Interpolation for Geographical Information Systems[J]. International Journal of Geographical Information Systems, 1990, 4(3): 313-320

[12] 邓岳川,谷双喜. 克里金插值的多路径误差空间分布特征[J]. 测绘科学,2018,43(4):17-23

[13] Kleijnen J P C. Kriging Metamodeling in Simulation: A Review[J]. European Journal of Operational Research, 2009, 192(3):707-716

[14] Oliver M A, Webster R. A Tutorial Guide to Geostatistics: Computing and Modelling Variograms and Kriging[J]. Catena,2014,113:56-69

(上接第92页)

“数智化”“数治化”转型。

参考文献

[1] 自然资源部. 关于以“多规合一”为基础推进规划用地“多审合一、多证合一”改革的通知[EB/OL]. (2019-09-20)[2023-09-15]. http://gi. mnr. gov. cn/201909/t20190920_2468404. html

[2] 自然资源部国家税务总局中国银保监会. 关于协同推进“互联网+不动产登记”方便企业和群众办事的意见[EB/OL]. (2020-06-02)[2023-09-15]. https://www. gov. cn/zhengce/zhengceku/2020-06/02/content_5516803. htm

[3] 中华人民共和国中央人民政府. 国务院办公厅关于进一步优化营商环境更好服务市场主体的实施意见[EB/OL]. (2020-07-21)[2023-09-15]. https://www. gov. cn/zhengce/content/2020-07/21/content_5528615. htm

[4] 马万钟,詹长根,朱小杰. 基于人地关系的土地管理数据模型研究[J]. 中国土地科学,2016,30(6):3-13

[5] 龚勋,邹寒月,陆生强. 基于数据图谱的“一码管地”管理体系研究与应用[J]. 地理空间信息,2023,21(11):56-58

[6] 蔡先娈,李钢,尹鹏程. 土地管理业务全要素生命周期模型研究[J]. 中国土地科学,2013,27(12):46-51

[7] 郑建军,赵颖,冯文杰. 基于“自然资源码”的自然资源全生命周期管理研究[J]. 智能城市,2022,8(5):48-50

[8] 应荷香,赵骞,张朝忙,等. 土地全生命周期管理的知识图谱构建及应用 [J]. 测绘科学,2022,47(6):161-167

[9] 郑建军. 绍兴市上虞区创新实施“一码管地”构建自然资源管理“数字图景”[J]. 浙江国土资源,2021,36(5):42-44

[10] 周丹,徐德军,张欢,等. 面向全生命周期管理的“一码管地”建设思路与应用探讨[J]. 规划师,2022,38(8):106-110

[11] 刘冬荣,张伟娜,莫加伟,等. 土地全生命周期管理基本架构与探索实践[J]. 中国国土资源经济,2023,36(6):36-43

[12] 张丹,陈晓茜,牟紫微. 湖北省“一码管地”探索与实践[J]. 地理空间信息,2023,21(9):77-80

[13] 张帆. 创新土地“码”管理 助推工业项目审批制度改革[J]. 浙江国土资源,2021,36(2):42-43

[14] 周建平. 不动产“一码”全链式管理机制研究[J]. 浙江国土资源,2021,36(5):39-41

[15] 周文奇,乐恒,张伟伟. 基于“浙政钉”的土地供应多跨场景应用建设实践[J]. 自然资源信息化. https://kns. cnki. net/kcms2/detail/10. 1797. N. 20230621. 1722. 004. html

[16] 贾春霞,姚玉全,寻知锋,等. 面向土地全生命周期管理的“一码管地”建设探索与应用[J]. 山东国土资源,2023,39(7):66-71