



基于 MDT 重叠覆盖度数据的 KNN-DBSCAN 参数自适应调优研究

刘璐¹, 陈睿杰², 李嘉²

(1. 中国移动通信集团设计院有限公司重庆分公司, 重庆 401121;

2. 中国移动通信集团云南有限公司, 云南 昆明 650228)

摘要: 传统网络优化中路测工作存在难以全量测试道路及楼宇、测试工作量大、工作效率低、周期长、受人为因素影响等显性缺点, 无法动态关注每个区域网络质量情况, 且常规测量报告 (measurement report, MR) 数据不具备定位信息, 无法精确定位如重叠覆盖度问题发生位置。基于最小化路测 (minimization drive test, MDT) 精准定位系统通过采集底层基站 MDT 数据, 并根据重叠覆盖度算法输出高重叠覆盖度栅格, 再通过自适应 K 最近邻-具有噪声的基于密度的聚类方法 (K -nearest neighbor density-based spatial clustering of applications with noise, KNN-DBSCAN) 联合算法解决了 DBSCAN 算法对参数设置敏感性问题, 并对问题栅格进行非监督聚类, 收敛问题连片区域, 通过小区采样贡献度进行栅格区域映射, 最终达到精准调整全局最高优先级 (TOP) 小区, 降低小区高重叠覆盖度的目的。

关键词: KNN-DBSCAN 算法; MDT 数据; 重叠覆盖度; 小区贡献度

中图分类号: TN929.5

文献标志码: A

doi: 10.11959/j.issn.1000-0801.2022010

Research on adaptive optimization of KNN-DBSCAN parameters based on MDT overlapping coverage data

LIU Lu¹, CHEN Ruijie², LI Jia²

1. Chongqing Branch of China Mobile Communications Group Design Institute Co., Ltd., Chongqing 401121, China

2. Yunnan Branch of China Mobile Communications Group Co., Ltd., Kunming 650228, China

Abstract: In the traditional network optimization, the drive test (DT) work has obvious disadvantages, such as difficult to fully test roads and buildings, large test workload, low work efficiency, long cycle, affected by human factors, unable to dynamically pay attention to the network quality of each area, and the conventional measurement report (MR) data does not have positioning information, so it is impossible to accurately locate the location where the overlapping coverage problem occurred. Based on minimization drive test (MDT), the precision positioning system collected the MDT data of the underlying base station and outputted the grid with high overlapping coverage according to the overlapping coverage algorithm. Then, the sensitivity of DBSCAN algorithm to parameter setting was solved through the adaptive K -nearest neighbor density-based spatial clustering of applications with noise (KNN-DBSCAN)



joint algorithm. The problem grid was unsupervised clustered, the problem contiguous area was converged, and the grid area was mapped through the cell sampling contribution. Finally, the global top cell was accurately adjusted to optimize the high overlap coverage.

Key words: KNN-DBSCAN algorithm, MDT data, overlapping coverage, cell contribution

0 引言

随着近几年网络建设的大力发展, LTE 业务量不断增长、网优规模急剧扩大, 干扰问题日益严重, 网络容量、覆盖和质量检查矛盾日趋凸显。LTE 网络主要以 D/F 频段组网, 无法通过频率规划的方法精确有效降低同频干扰, 必须对网络结构进行调整与优化。目前常规采用手段如路测、测量报告 (measurement report, MR) 数据和道路扫频等数据进行优化, 但由于路测数据获取成本较高, MR 数据未携带定位信息, 扫频数据也仅局限于道路, 以及现网工参准确性较差等, 重叠覆盖分析误判较多, 难以准确定位和分析问题, 无法达到优化预期效果。

1 基于 MDT 重叠覆盖度优化思路

1.1 MDT 技术原理

最小化路测 (minimization of drive test, MDT) 技术功能主要通过扩展现无线资源管

理 (radio resource management, RRM) 测量功能和跟踪功能实现, MDT 数据测量及解析流程如图 1 所示。

基站根据网管配置的 MDT 测量任务下发相关测量配置至终端 UE, 终端 UE 在满足测量条件时测量并上报测量信息, 包含参考信号接收功率 (reference signal receiving power, RSRP)、参考信号接收质量 (reference signal receiving quality, RSRQ) 等字段, 最重要的是携带相关终端 UE 的定位信息 (GPS 经度、纬度)^[1], 基站将收到的终端测量结果和基站自身的测量结果按数据标准要求上报至网管或 MDT 数据存储处理服务器。

目前 MDT 支持的测量项主要包括 3 类, MDT 支持测量项见表 1。

1.2 MDT 重叠覆盖栅格定义

重叠覆盖问题主要指在网络拓扑空间内, 彼此存在深度交叠区域, 如 RSRP 较好但 SINR 较差, 或者多个小区间乒乓切换造成用户感知差问题, 重叠覆盖原因示意图如图 2 所示。

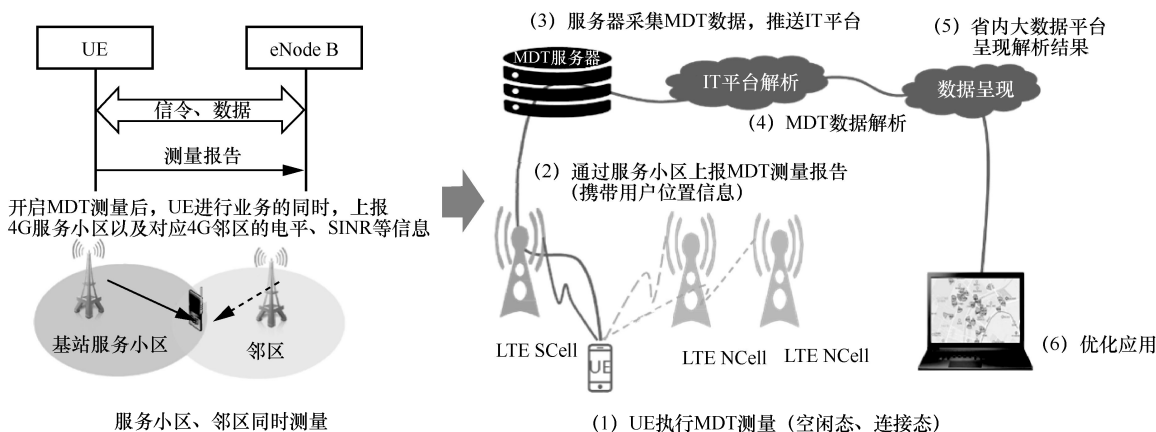


图 1 MDT 数据测量及解析流程

表1 MDT 支持测量项

MDT 测量项		测量上报信息	测量网元
即时 MDT 数据	M1	RSRP、RSRQ、位置信息等	UE
	M2	最大发射功率余量 (power headroom, PHR)	UE
	M3	接收干扰功率测量 (received interference power measurement, RIP)	eNode B
	M4	下行/上行数据吞吐量	eNode B
	M5	下行/上行调度 IP 吞吐率	eNode B
	M6	下行/上行数据包时延测量	eNode B
	M7	下行/上行数据丢包率测量	eNode B
已登录 MDT 数据		RSRP、RSRQ、位置信息等	UE
无线链路失败 (radio link failure, RLF)		发生 RLF 主服小区、重建小区、失败类型、位置信息等	UE/eNode B

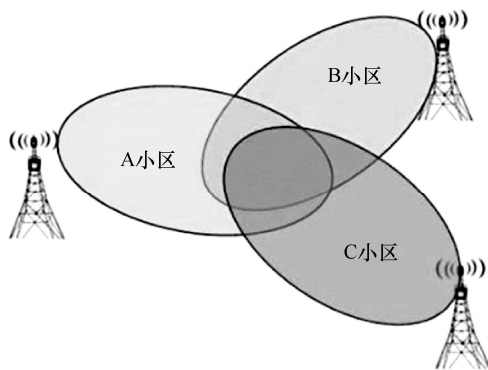


图2 重叠覆盖原因示意图

MDT 栅格:采用目前主流 GeoHash 位置服务技术,实现栅格定位普通精度为 $38\text{ m}\times 19\text{ m}$ 及 $5\text{ m}\times 5\text{ m}$ 高精度的 GIS 渲染。

基于上述 MDT 栅格进行高重叠覆盖度栅格的计算,涉及以下指标内容:

(1) 重叠覆盖采样点:在基站服务小区 $\text{RSRP}\geq -110\text{ dBm}$,同时邻区与主服小区 RSRP 的绝对差值在 6 dB 范围内的小区数量 $n^{[2-3]}$,即:
|主服小区 RSRP - 邻区 $n\text{RSRP}$ | $\leq 6\text{ dB}$ 。

(2) 栅格重叠覆盖度:栅格重叠覆盖度 ≥ 3 的采样点/栅格总采样点 $\times 100\%$ 。

(3) 高重叠覆盖度栅格:栅格重叠覆盖度 $\geq 15\%$,该阈值可根据实际情况进行动态调整。

重叠覆盖度栅格内小区采样贡献度:栅格内

总共有 N 个小区覆盖,其中, N 个小区重叠覆盖采样点共计 K 个,第 N_i 个小区的重叠覆盖采样点为 n_i 个,则第 N_i 小区重叠覆盖采样贡献度 $= n_i/K$ 。

1.3 传统高重叠覆盖问题优化思路

传统高重叠覆盖问题优化主要是基于 DT 路测数据以及 MR 数据进行渲染,呈现高重叠覆盖区域,在获取数据便捷性、成本性、时效性、精准性上均劣于 MDT 数据评估分析效果。

传统高重叠覆盖优化流程如图3所示,传统高重叠覆盖优化流程及判定分析方法均是侧重寻优单例问题小区,而本文主要是对传统高重叠覆盖问题评估阶段进行改进优化,采用较为精准的 MDT 数据,并依据 GeoHash 算法形成栅格级数据,通过目前较为成熟的 KNN 与 DBSCAN 联合非监督机器学习算法,进行高重叠覆盖度栅格的聚类,在此问题聚类区域进一步计算 TOP 小区贡献度,综合性、全局性地调整 TOP 小区以实现最终寻优目的。

2 基于 DBSCAN 算法研究分析

2.1 研究场景及采样策略

由于目前密集城区高重叠覆盖度问题相对凸显,且 DBSCAN 算法在样本集的密度不均匀、聚类间距差相差很大时聚类质量较差等原因,在分

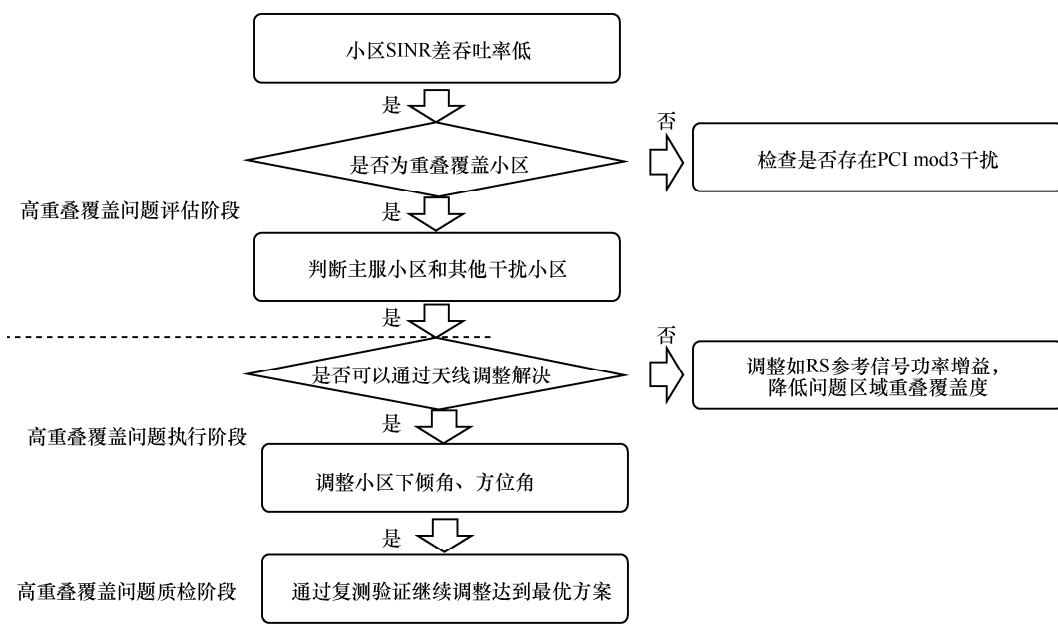


图3 传统高重叠覆盖优化流程

析场景上选取站间距较小、MDT 高重叠覆盖数据采样的稀疏程度相似、簇密度数据差异不明显的密集城区进行研究。

(1) MDT 测量采样策略

考虑全省 MR 以及 MDT 数据测量项同时开启、数据服务器负荷因素, 可以只开启以下基本测量项。

- 全省开启 IMM-MDT 的 M1 测量 (RSRP、RSRQ, 由 UE 测量)。
- M2 测量 (PHR 最大发射功率余量, 由 UE 测量)。
- M3 测量 (RIP 接收干扰功率测量, 由 eNB 测量)。
- 同时开启 LOGGED-MDT 以及 RLF-MDT 测量项。

(2) MDT 采样周期

全省每个月集中采样 10 天, 取最终全量并集数据。

(3) MDT 补采数据机制

通过步骤 (1) 和步骤 (2) 中数据采集策略及采集周期可能会出现因厂商 MDT 服务器 IP 地

址差异、服务器传输模式如 SFTP 但省端采集侧配置为 FTP 等问题出现数据采集不全的情况, 所以需要针对上述原因补采 MDT 数据。

(4) MDT 数据 ETL 策略

根据步骤 (1) ~ 步骤 (3) 获取初步 MDT 数据, 经 IT 支撑平台进行预处理, 可输出 MDT 采样数据, MDT 采样样例数据见表 2, 按优化经验对该数据做进一步数据分析处理。

- 栅格总采样点 > 5 采样点。
- 栅格重叠覆盖度 > 15%。

联合上述两个条件输出对应高重叠覆盖度栅格作为研究候选对象栅格。

2.2 经典 DBSCAN 算法分析

选取按第 2.1 节中条件过滤后的高重叠覆盖度栅格, 将栅格经度、纬度通过编程工具进行 GIS 渲染, 并使用 DBSCAN 密度聚类算法将多个邻近高重叠覆盖度栅格聚为一类。

(1) DBSCAN 密度聚类寻优流程^[4-5]如算法 1 所示。

算法 1 DBSCAN 密度聚类寻优流程

输入 D : 数据源集 (高重叠覆盖度栅格)

表2 MDT 采样样例数据

地市	栅格ID	栅格经度	栅格纬度	CGI	小区总采样点	小区重叠覆盖采样点	小区重叠覆盖率	小区在该栅格总采样点	小区在该栅格重叠覆盖采样点	小区在该栅格重叠覆盖率	该栅格总采样点	该栅格重叠覆盖采样点	该栅格重叠覆盖率
××	wk3n8ec6	102.680 1	25.069 86	830748-232	365	34	9.32%	7	1	14.29%	50	9	18.00%
××	wk3n8ec6	102.680 1	25.069 86	830841-230	254	19	7.48%	23	4	17.39%	50	9	18.00%
××	wk3n8ec6	102.680 1	25.069 86	830748-235	797	141	17.69%	12	4	33.33%	50	9	18.00%
××	wk3n8ec6	102.680 1	25.069 86	830746-230	424	33	7.78%	8	0	0	50	9	18.00%
××	wk3n0emz	102.686 3	24.980 08	818993-230	578	14	2.42%	1	0	0	1	0	0
××	wk3n98dk	102.725 4	25.052 35	569127-11	238	15	6.3%	1	0	0	9 293	1	0.01%
××	wk3n98dk	102.725 4	25.052 35	569307-55	580	41	7.07%	15	1	6.67%	9 293	1	0.01%
××	wk3n98dk	102.725 4	25.052 35	566634-27	1 109	197	17.76%	7	0	0	9 293	1	0.01%
××	wk3n98dk	102.725 4	25.052 35	566634-62	630	67	10.63%	8	0	0	9 293	1	0.01%
××	wk3n98dk	102.725 4	25.052 35	566634-8	312	18	5.77%	3	0	0	9 293	1	0.01%
××	wk3n98dk	102.725 4	25.052 35	566634-4	169	19	11.24%	3	0	0	9 293	1	0.01%
××	wk3n98dk	102.725 4	25.052 35	817975-232	82	5	6.1%	4	0	0	9 293	1	0.01%
××	wk3n8hks	102.662 6	25.072 95	815677-235	248	12	4.84%	2	0	0	9 293	1	0.01%
××	wk3n8hks	102.662 6	25.072 95	712463-143	136	8	5.88%	22	0	0	9 293	1	0.01%
××	wk3n8hks	102.662 6	25.072 95	815677-234	417	32	7.67%	12	0	0	9 293	1	0.01%
××	wk3n8hks	102.662 6	25.072 95	853851-145	3 492	159	4.55%	7	0	0	9 293	1	0.01%
××	wk3n8hks	102.662 6	25.072 95	830233-235	1 302	134	10.29%	1	0	0	9 293	1	0.01%



ε : 半径参数

MinPts: 邻域密度阈值

输出 基于密度的簇集合

标记集合 D 所有对象为 unvisited

do

集合 D 内随机选择 1 个 unvisited 对象 p

预设参数 ε 为以 p 为质心, ε 为指定半径的圆

if p 的 ε -邻域至少有 MinPts 个对象 (预设参数 MinPts)

创建一个新簇 C , 并把 p 添加至 C 中

令 N 为 p 的 ε -邻域中的对象集合

for N 中每个点 p'

if p' 为 unvisited

标记 p' 为 visited

if p' 的 ε -邻域至少有 MinPts 个对象,

将这些对象添加至 N 中

p' 还不是任何簇的对象点, 将 p' 添加至 C 中

end for

输出 C

else 标记 p 为 noise (噪点)

until 没有标记为 unvisited 的对象

(2) 实现经典 DBSCAN 主要过程

通过相关编程语句实现 DBSCAN 算法, 为便于寻优 DBSCAN 算法中超参数搜索半径 ε 、MinPts 与聚类数量、聚类区域范围及 noise 噪点数量的关系, 通过使用 for 循环判决函数、plot 函数等经典函数构建二维平面图形关系^[6]。

高重叠覆盖栅格簇群 GIS 渲染如下。

利用 PlotClusterinResult(X,IDX1)函数, 可得出在超参数搜索半径 ε 、MinPts 拟合模型函数, 设置不同超参数的聚类效果如图 4 所示。

由图 4 可知, DBSCAN 算法的搜索半径 ε 、MinPts 参数需要人为设定, 聚类结果敏感度、准确度受其超参数影响较大, 需要在高重叠覆盖干扰簇群范围大小与噪点包含度进行寻优平衡,

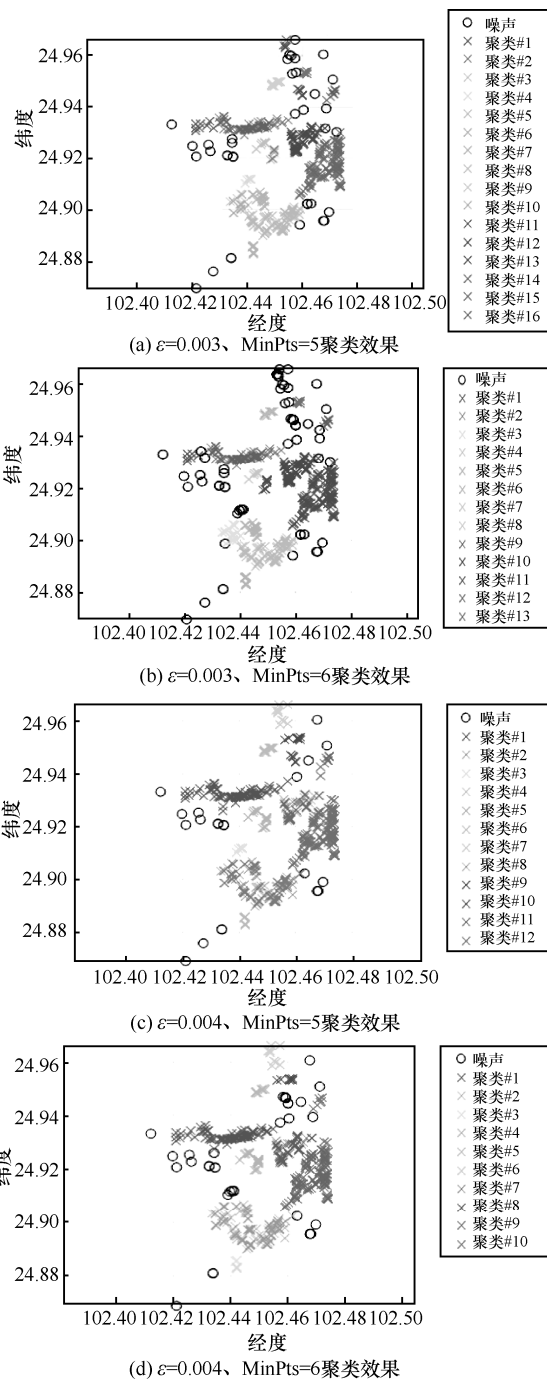


图 4 设置不同超参数的聚类效果

如干扰簇群搜索半径 ε 过大, 其涵盖小区过多会导致干扰簇群内小区贡献度指标被弱化, 无法凸显最严重干扰小区; 而干扰簇群搜索半径 ε 过小可能会导致无法包含相近样本, 从而误判为噪点情况, 所以为降低减少人为干预, 需要进行机器自适应的 DBSCAN 超参数调优的研究。

2.3 自适应 KNN-DBSCAN 联合算法调优的研究思路

基于 DBSCAN 算法的基本原理, 针对密集城区重叠覆盖度数据特征, 引入密度阈值 Density 参数, 定义密度阈值 Density 为以 ε 为搜索半径的圆内存在 MinPts 个数据点, 如式 (1) 所示^[7]。

$$D(\text{Density}) = \frac{\text{MinPts}}{\pi \cdot \varepsilon^2} \quad (1)$$

根据式 (2) 中密度阈值 D 过大, 可能会导致同簇群中内部被划分为多个集合; 如密度阈值 D 过小, 则可能引起不同簇群集合间的合并, 所以选取合适的密度阈值 D 非常重要。K-均值 DBSCAN 算法需要人为设定初始聚类中心, 不能很好地解决聚类结果随机性问题; AF-DBSCAN 算法需要选择种子代表对象的数量, 代表对象过多则无法发挥算法效率, 过少则容易造成对象丢失, 影响算法聚类质量。为尽可能减少人为干预, 确保聚类效果, 本文采用自适应参数 KNN-DBSCAN 算法构建最小距离矩阵, 并通过分析数据的 KNN 的数学统计规律, 寻优密度阈值列表中的超参数取值, 本文算法主要适用于低维且簇密度差异不明显的数据集。

(1) 确定搜索半径 ε 具体算法思路有以下 3 个步骤。

步骤 1 构建基于高重叠覆盖栅格数据的坐标位置矩阵 $L_{m \times 2}$, 该矩阵包含高重叠覆盖度栅格的位置信息 (long,lat), 并对位置信息维度数据进行升序排序。

步骤 2 采用 KNN 算法^[8]及 knnsearch 函数 [IDX,Dist]=knnsearch(a,b) 遍历计算每个栅格位置与其他栅格位置的距离, 并将最小距离值存储至 Dist 中, 记录其数据行号索引值存储至 IDX 中。

步骤 3 根据最小距离值 Dist 降序排列并绘制关系曲线, 寻找曲线拐点处的阶跃值, 即可确定搜索半径 ε 值。

寻优策略执行: 输入第 2.2 节的数据, 按照上述算法进行步骤 (1) ~ 步骤 (3) 调优, 并使用编程工具建模仿真, 搜索半径 ε 寻优过程 1 和 2 分别如图 5、图 6 所示。

由图 5 可知, 在第 4 个数据点处出现明显数据凸起, 说明在第 4~5 个点处出现数值阶跃现象。其在图 6 曲线上则反映 A、B 两点的拐点处纵坐标阶跃值最大, 说明该 B 点的纵坐标为搜索半径 $\varepsilon=0.00449$ 。

(2) 确定 MinPts 值具体算法思路有以下 3 个步骤。

步骤 1 基于算法思路 (1) 中坐标数据构建坐标间的距离分布矩阵 $D_{m \times m}$ ^[9], 即:

$$D_{m \times m} = \{\text{Dist}(p, q) | 1 \leq p \leq m, 1 \leq q \leq m\} \quad (2)$$

其中, $D_{m \times m}$ 是 m 维乘以 m 维的对称矩阵, m 为数据集 D 中包含的所有对象数量, $\text{Dist}(p, q)$ 表示在数据集 D 中第 p 个对象到第 q 个对象的空间距离。

步骤 2 将上述距离矩阵 $D_{m \times m}$ 的行向量按升序排列, 并令距离矩阵 $D_{m \times m}$ 中第 K 列元素为构成

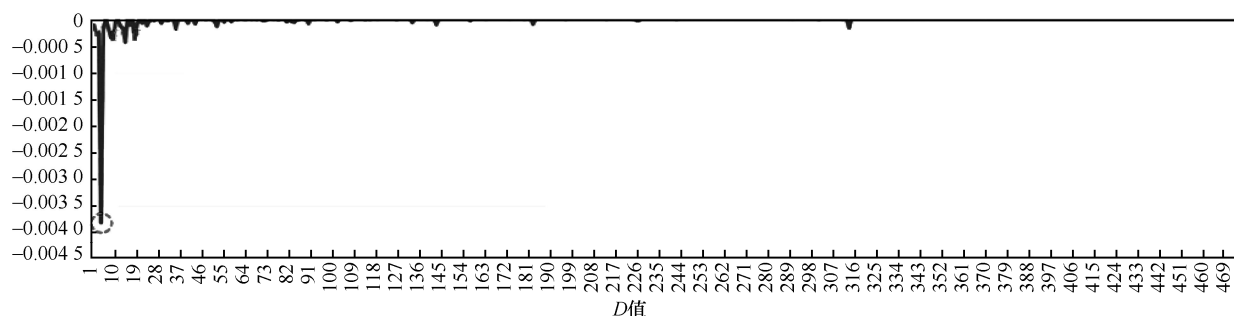
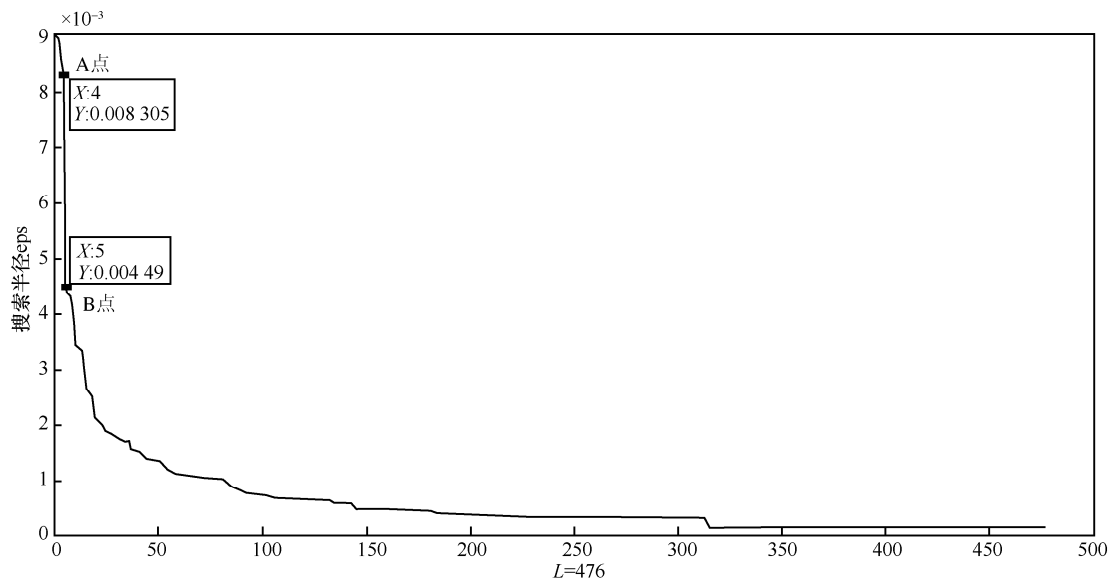


图 5 搜索半径 ε 寻优过程 1

图6 搜索半径 ε 寻优过程 2

所有数据点的 K 最近邻距离向量 D 。

步骤3 当 K 取不同值时绘制对应分布曲线, 以便于观察 K 取离散值 $1, 2, \dots, 15$, 该曲线从下至上依次为 K 值由小至大的变化图形, 结合已确定的搜索半径 ε 值, 与上述 K 值曲线相交处, 进行 MinPts 值的选取确定。

寻优策略执行: 按上述确定 MinPts 值具体算法思路及步骤采用编程实现, 对 K 不同取值曲线进行图形化渲染, MinPts 参数寻优过程如图 7 可

知, K 取离散值 $1, 2, \dots, 15$ 等 15 条曲线基本形似, 呈“前缓后陡”趋势。

根据上述自适应 KNN-DBSCAN 联合算法可知, 当 $\varepsilon=0.00449$ 时, 与 K 值曲线集合密度最大处相交 (图 7 圆圈处), 该 K 值曲线为 MinPts=8。

自适应 KNN-DBSCAN 联合算法调优聚类效果如图 8 所示, 当密度阈值 Density 参数列表 = $\{\varepsilon=0.00449, \text{MinPts}=8\}$ 时, 该数据集可取得最优解。

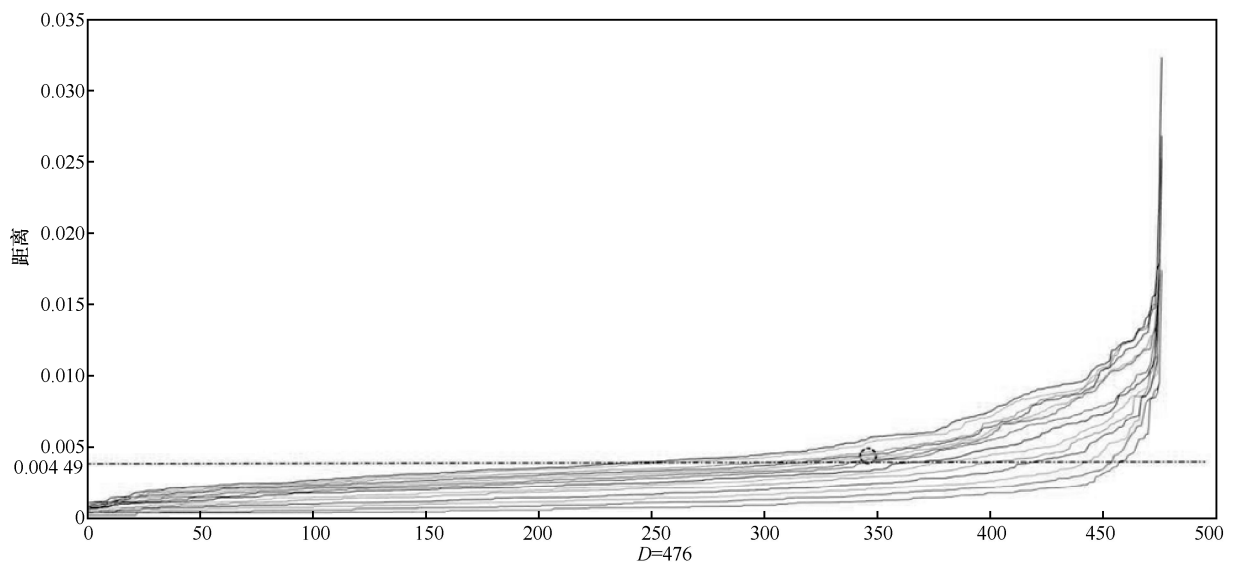


图7 MinPts 参数寻优过程

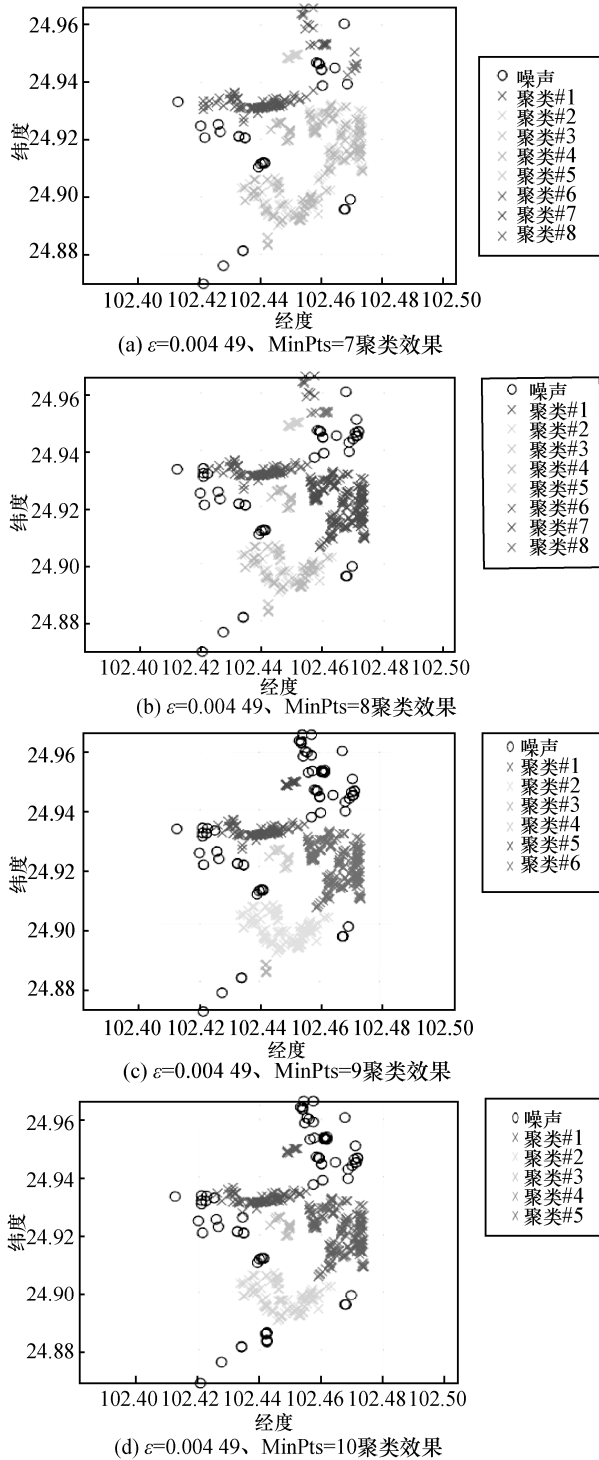


图8 自适应 KNN-DBSCAN 联合算法调优聚类效果

(3) 自适应 KNN-DBSCAN 联合算法模型的实效性评估

自适应 KNN-DBSCAN 联合算法模型主要是应用于网络高重叠覆盖度优化, 基于该算法特性

及实际网络优化工作需求, 除了考虑模型超参数搜索半径 ε 、MinPts 设置外, 还需要考虑模型输出的高干扰簇群范围以及其涵盖的小区数量, 平衡相关要素最终获取每个高干扰簇群中 TOP 小区贡献度, 为验证自适应 KNN-DBSCAN 联合算法模型有效性, 设置不同超参数进行相关指标对比验证, 不同超参数设置下算法模型 KPI 验证结果见表 3。其中, 2 个 KPI 评估验证指标如下。

- 噪点率=noise 噪点数 n /对象数 N 。
- 最大簇群内 TOP 问题小区贡献度占比=最大簇群中高重叠覆盖度栅格所辖 TOP5 小区的采样点/最大簇群中高重叠覆盖度栅格所有小区采样点。

通过表 3 可知, 在自适应 KNN-DBSCAN 联合算法调优层面, 当簇群 C_i 一定时, 噪点率越小越好; 在实际网络优化层面, 解决高重叠覆盖小区需要以优化最少小区为代价获得最好收效, 即最大簇群内 TOP 问题小区贡献度的占比越高, 该簇群的问题小区收敛性越好。

实验组别 9、10 的噪点率最小, 同为 4.83%, 而最大簇群内 TOP 问题小区贡献度仅为 38.97%; 实验组别 6、7、8 最大簇群内 TOP 问题小区贡献度同为 59.32%, 但噪点率依次为 9.03%、13.24%、15.34%, 均略高于实验组别 9、10。另实验组别 6 与实验组别 5 在同为 8 个簇群前提下, 虽然实验组别 6 比实验组别 5 噪点率稍劣于 2.73%, 但在最大簇内 TOP 问题小区贡献度占比方面实验组别 6 大幅优于实验组别 5, 领先幅度为 20.60%。

综合分析, 当密度阈值 Density 参数列表 = $\{\varepsilon=0.004\ 49, \text{MinPts}=8\}$ 时, 即实验组别 6 在噪点率及最大簇群内 TOP 问题小区贡献度占比两项 KPI 评估中平衡性最优, 从而验证自适应 KNN-DBSCAN 联合算法在实际网络优化中的有效性和实用性。



表3 不同超参数设置下算法模型 KPI 验证结果

实验组别	数据集	对象数 N	密度阈值 D (Density) 参数列表		噪点率	簇群数 C_i /个	最大簇群 $C_{\max}$ 覆盖近似面积/ km^2	最大簇群 $C_{\max}$ 包含问题小区数/个	最大簇内 TOP 问题小区贡献度占比
			搜索半径 ϵ ps	MinPts					
1	××地市密集城区MDT高重叠覆盖度栅格数据	476	0.003 00	7	15.76%	10	11.46	82	41.06%
2			0.003 00	8	18.07%	10	10.62	75	42.66%
3			0.003 00	9	20.38%	10	10.62	75	42.66%
4			0.003 00	10	25.00%	10	10.62	75	42.66%
5			0.004 49	7	6.30%	8	25.10	234	38.72%
6			0.004 49	8	9.03%	8	14.76	117	59.32%
7			0.004 49	9	13.24%	6	14.76	117	59.32%
8			0.004 49	10	15.34%	5	14.76	117	59.32%
9			0.005 00	7	4.83%	7	35.66	284	38.97%
10			0.005 00	8	4.83%	7	35.66	284	38.97%
11			0.005 00	9	9.87%	5	33.28	276	39.15%
12			0.005 00	10	13.66%	4	24.12	158	54.71%

3 结束语

对 MDT 高重叠覆盖度栅格基于 KNN-DBSCAN 的联合算法研究,利用编程工具对 DBSCAN 密度阈值 Density 参数中搜索半径 ϵ 、MinPts 两个超参数进行自适应寻优,从而解决了该算法对参数设置敏感性问题。结合实际网络优化工作特性,高效、快速构建优化区域内高重叠覆盖度簇群,精准识别其簇群中 TOP 小区,为一线生产提供有效的优化解决方案,确保移动用户的良好感知体验。

参考文献:

- [1] 谭钰山,周文金,何延. 基于MDT数据的重叠覆盖优化思路及实践[J]. 通讯世界, 2019, 26(12): 6-7.
TAN Y S, ZHOU W J, HE Y. Overlapping coverage optimization idea and practice based on MDT data[J]. Telecom World, 2019, 26(12): 6-7.
- [2] 张吉,赵凤,朱晓荣. 基于大数据挖掘的LTE网络重叠覆盖优化方法[J]. 南京邮电大学学报(自然科学版), 2020, 40(6): 92-99.
ZHANG J, ZHAO S, ZHU X R. Optimization method for overlapping coverage of LTE networks based on big data mining[J]. Journal of Nanjing University of Posts and Telecommunications (Natural Science Edition), 2020, 40(6): 92-99.
- [3] 谷欣杏. LTE 网络覆盖优化及无线定位优化算法的研究[D].

北京: 北京邮电大学, 2019.

- GU X X. Research on LTE network coverage optimization and wireless positioning algorithm[D]. Beijing: Beijing University of Posts and Telecommunications, 2019.
- [4] 宋金玉,郭一平,王斌. DBSCAN 聚类算法的参数配置方法研究[J]. 计算机技术与发展, 2019, 29(5): 44-48.
SONG J Y, GUO Y P, WANG B. Research on parameter configuration method of DBSCAN clustering algorithm[J]. Computer Technology and Development, 2019, 29(5): 44-48.
- [5] 赵文,夏桂书,苟智坚,等. 一种改进的 DBSCAN 算法[J]. 四川师范大学学报(自然科学版), 2013, 36(2): 312-316.
ZHAO W, XIA G S, GOU Z J, et al. An improved DBSCAN algorithm[J]. Journal of Sichuan Normal University (Natural Science), 2013, 36(2): 312-316.
- [6] 何正风. MATLAB 在数学方面的应用[M]. 北京: 清华大学出版社, 2012.
HE Z F. Application of MATLAB in mathematics[M]. Beijing: Tsinghua University Press, 2012.
- [7] 李文杰,闫世强,蒋莹,等. 自适应确定 DBSCAN 算法参数的算法研究[J]. 计算机工程与应用, 2019, 55(5): 1-7, 148.
LI W J, YAN S Q, JIANG Y, et al. Research on method of self-adaptive determination of DBSCAN algorithm parameters[J]. Computer Engineering and Applications, 2019, 55(5): 1-7, 148.
- [8] 王紫薇,徐凯,侯益明. 基于不同距离公式的 KNN 算法对鸢尾花的分类[J]. 无线互联科技, 2021, 18(13): 105-106.
WANG Z W, XU K, HOU Y M. Classification of iris based on KNN algorithm with different distance formulas[J]. Wireless Internet Technology, 2021, 18(13): 105-106.

- [9] 周红芳, 王鹏. DBSCAN 算法中参数自适应确定方法的研究[J]. 西安理工大学学报, 2012, 28(3): 289-292.

ZHOU H F, WANG P. Research on adaptive parameters determination in DBSCAN algorithm[J]. Journal of Xi'an University of Technology, 2012, 28(3): 289-292.

[作者简介]



刘璐 (1986-), 男, 中国移动通信集团设计院有限公司重庆分公司工程师、高级咨询设计师, 主要研究方向为无线网络智能优化。



陈睿杰 (1987-), 男, 中国移动通信集团云南有限公司工程师, 主要研究方向为大数据分析、AI、智慧运维。



李嘉 (1989-), 男, 现就职于中国移动通信集团云南有限公司, 主要研究方向为大数据分析、AI、智慧运维等。