

文章编号: 1001-0920(2006)02-0213-04

## 基于聚类分解的 WCDMA 基站布局规划算法

张宏远, 谷寒雨, 席裕庚

(上海交通大学 自动化研究所, 上海 200030)

**摘 要:** 针对大规模 WCDMA 无线网络基站布局规划问题, 提出一种基于聚类分解的分层算法。在聚类分解中, 以测试点信号增益矩阵构造聚类分解数据, 并给出了收敛判定函数和相似度计算方法。在分层算法中, 首先用  $K$  均值聚类将原问题分解为  $K$  个子规划问题, 然后对各子问题求解整数规划问题, 最后对各子问题的基站布局结果进行全局调整。仿真计算验证了该算法的有效性。

**关键词:** WCDMA;  $K$  均值聚类; 无线网络规划; 启发式算法

**中图分类号:** TN 929.5      **文献标识码:** A

## Planning Algorithm for WCDMA Base Station Location Problem Based on Cluster Decomposition

ZHANG Hong-yuan, GU Han-yu, XI Yu-geng

(Institute of Automation, Shanghai Jiaotong University, Shanghai 200030, China. Correspondent: ZHANG Hong-yuan, E-mail: irry-zhang@sjtu.edu.cn)

**Abstract** An algorithm based on cluster decomposition is presented for large scale WCDMA base station positioning problem. A data matrix based on the signal gain of cellular networks and some criterion functions are designed for  $K$  means clustering. In the algorithm, the original problem is firstly decomposed into  $K$  sub-problems by  $K$  means clustering. Each sub-problem is then solved by integer programming. Finally, the solutions of  $K$  sub-problems are coordinated to form an approximate solution of the global problem. Simulation result shows the validity of this algorithm.

**Key words:** WCDMA;  $K$  means clustering; Wireless network planning; Heuristic algorithm

### 1 引 言

在 WCDMA 网络中, 从降低网络运营成本的角度考虑, 运营商需要尽量减少网络设备(如基站数量)或造价; 从满足网络覆盖和网络质量要求的角度考虑, 需要增加设备的数量并合理布局, 以保证一定的通信覆盖质量。WCDMA 基站布局规划 W-BSP 就是为完成这一任务而提出的优化问题。该问题已被证明是 NP-hard 问题, 因此大量工作集中在近似算法的寻找上。Amaldi<sup>[1]</sup>提出了“加”和“减”两种贪婪算法及禁忌搜索算法; Wright<sup>[2]</sup>用 Nelder-Mead 直接搜索方法求解基站选址问题; Molina<sup>[3]</sup>讨论了遗传算法在该问题中的应用; Tutschku<sup>[4]</sup>介绍了以集

成算法为基础的 ICEPT 移动通信网络优化工具; Matthias<sup>[5]</sup>提出一种基于距离分割的启发式算法。

聚类技术广泛应用于模式识别、城市规划、医学、物流等领域<sup>[6,7]</sup>。利用空间分解, 可将大规模 W-BSP 问题分成多个小规模问题来求解, 但人为地按方块切割<sup>[5]</sup>过于简单。为此本文采用聚类对大规模问题进行分解, 在此基础上设计了 W-BSP 问题的优化算法。仿真计算验证了该算法的有效性。

### 2 规划模型

无线网络基站布局规划问题可以描述为: 在  $N$  个基站候选站址集合中, 选择适当的站址架设基站, 希望通信链路信干比 SIR 满足最小要求  $SIR_{min}$ , 并

收稿日期: 2004-12-14; 修回日期: 2005-04-01

作者简介: 张宏远(1971—), 男, 陕西眉县人, 博士生, 从事无线网络规划、优化算法的研究; 席裕庚(1946—), 男, 上海人, 教授, 博士生导师, 从事预测控制、智能机器人等研究。

保证  $M$  个测试点的业务量要求, 优化目标为减少网络造价. 为了清晰地描述 WCDMA 的网络基站布局规划问题, 首先给出以下符号定义:

$J$ : 基站候选站址集合,  $J = \{j = 1, 2, \dots\}$ ;

$I$ : 测试点集合,  $I = \{i = 1, 2, \dots\}$ ;

$O(J), C(J)$ : 分别为建设的基站集合和没有选中的基站集合,  $O(J) \cup C(J) = J$ ;

$J(r), I(r)$ : 分解得到的第  $r$  组基站候选站址集合和测试点集合;

$X_{ij}$ : 0-1 变量, 如果测试点  $i$  由基站  $j$  服务则  $X_{ij} = 1$ , 否则  $X_{ij} = 0$ ;

$Y_j$ : 0-1 变量, 如果在  $j$  处建基站则  $Y_j = 1$ , 否则  $Y_j = 0$ ;

$C_{ij}$ : 测试点  $i$  到基站  $j$  的无线信号传播损耗;

$C_j$ : 基站  $j$  的造价, 包括土地使用费、设备费、维护费等;

$U_i$ : 测试点  $i$  的通话量统计值, 单位为 Erl

对于上行链路无基站数量约束的 WCDMA 无线网络基站布局规划问题 W-BSP, 参照文献[1] 给出该问题的整数规划模型如下:

$$Z = \min \left[ \sum_{i \in I, j \in J} C_j Y_j + \lambda \sum_{i \in I, j \in J} U_i C_{ij} X_{ij} \right]; \quad (1)$$

$$\text{s.t.} \quad X_{ij} = 1, i \in I, j \in J, \quad (2)$$

$$X_{ij} \leq Y_j, i \in I, j \in J, \quad (3)$$

$$\frac{\sum_{i \in I} U_i X_{ij} C_{ij} / C_{ij} - 1}{j \in J} \geq \text{SR}_{\min} \times Y_j, \quad (4)$$

$$X_{ij}, Y_j \in \{0, 1\}, i \in I, j \in J. \quad (5)$$

式(1)第1项表示基站的建造总费用; 第二项表示所有移动用户的功率消耗, 其中  $\lambda$  用来平衡不同量纲的两个优化目标. 式(2)要求每个测试点必须唯一地接入一个基站. 式(3)表示一个测试点由一个基站覆盖的前提是该基站被选择建设. 式(4)要求每个建设基站上行链路的 SR 至少达到  $\text{SR}_{\min}$ .

该模型相对于复杂的 WCDMA 无线网络规划问题而言省略较多, 但仍是一个 NP-hard 问题, 并且约束(4)为整数非线性约束, 对求解造成了很大困难. 为此本文研究基于聚类分解的近似算法.

### 3 聚类分解

聚类是将物理或抽象的集合分成相似对象组成的多个类的过程,  $K$  均值聚类是一种划分的聚类方法<sup>[6]</sup>. 对于 W-BSP 问题, 基站和测试点都是分类对象. 为了便于聚类计算, 假设在每个候选基站位置存在一个测试点, 并以其他点到该虚拟测试点的信

号衰减增益为分类依据. 从测试点(包括虚拟测试点)到基站的衰减  $C_{ij}$  构成数据矩阵

$$M_c = \begin{bmatrix} C_{1,1} & C_{1,2} & \dots & C_{1,N} \\ C_{2,1} & C_{2,2} & \dots & C_{2,N} \\ \vdots & \vdots & \ddots & \vdots \\ C_{N+M,1} & C_{N+M,2} & \dots & C_{N+M,N} \end{bmatrix}$$

逐步进行中心化、极差正规化和标准化变换后,  $M_c$  被转换为列数据均值为 0, 方差为 1 的标准聚类分析数据  $C_{ij}$  构成的矩阵  $M_c^s$ , 行向量构成测试点的样本数据. 对测试点和基站候选站址的划分以  $M_c^s$  中的样本数据为依据.

收敛判定函数为

$$E = \sum_{i=1}^K P_j - m_i^2 \quad (6)$$

其中:  $P_j = (C_{j1}, C_{j2}, \dots, C_{jN})$  为被划分在第  $i$  类中的测试点  $j$  的样本数据,  $m_i = (m_{i1}, m_{i2}, \dots, m_{iN})$  为第  $i$  类中所有测试点的样本均值,  $CL_i$  为第  $i$  类中的测试点集合.

采用两个测试点  $i$  和  $j$  的样本数据的欧氏距离  $d_{ij}$  表示其相似度, 则

$$d_{ij}(2) = \left[ \sum_{k=1}^N |C_{ik} - C_{jk}|^2 \right]^{1/2}, \quad i, j = 1, 2, \dots, N + M. \quad (7)$$

根据式(6)和(7), 以数据矩阵  $M_c^s$  为分析对象, 对 W-CDMA 网络中的测试点和基站按标准的  $K$  均值聚类算法进行聚类分解.

下面用聚类分解结果设计分层算法, 以求解 W-BSP 近似解.

### 4 分层算法

$K$  均值聚类算法首先确定将要划分的聚类数目  $K$ , 本文利用经验给定  $K$  值. 经过分解, 原问题的  $M$  个测试点和  $N$  个基站候选站址被分解到  $K$  组集合  $I(r)$  和  $J(r)$  中,  $r = 1, 2, \dots, K$ . W-BSP 中的约束(4)为非线性约束, 为了利用现有的线性整数规划工具, 给定一个合理的极大值  $\alpha$  则其可转化为线性约束.

$$\sum_{i \in I(r), j \in J(r)} U_i X_{ij} G_{ij} / G_{ii} \leq \text{SF} / \text{SR} + \alpha(1 - Y_j), \quad i \in J(r). \quad (8)$$

针对基站候选站址集合  $J(r)$  和测试点集合  $I(r)$  的 W-BSP 问题记为 W-BSP( $r$ ), 称为原问题的子问题. W-BSP( $r$ ) 一般为小规模问题, 可以直接求解线性整数规划. 如果  $K$  值选得比较合理, 则子问题不会无解. 尽管如此, 各子问题的解仍然存在以下问题:

1) 一些候选基站处于子问题的边缘地带, 当子

问题独立求解时, 为了满足约束, 处于子问题边缘的基站将被选择, 并且只有该子问题中的测试点可以被其覆盖, 而其他子问题中的测试点即使离该基站很近也不会被覆盖

2) 子问题求解时, 没有考虑其他子问题中测试点产生的干扰, 因此最后的解不能保证原问题的干扰约束

为解决上述问题, 本文设计以下全局策略:

1) 以各子问题选中的基站构成原问题的基站集合  $O(J)$ , 则可构造以下 CP 问题:

$$Z = \min_{i \in I, j \in O(J)} U_i C_{ij} X_{ij}; \quad (9)$$

$$\begin{aligned} \text{s t } & X_{ij} = 1, i \in I, j \in O(J), \\ & X_{ij} \leq Y_j, i \in I, j \in O(J), \\ & X_{ij}, Y_j \in \{0, 1\}, i \in I, j \in O(J). \end{aligned} \quad (10)$$

CP 问题的目标函数与网络中用户的发射功率之和成正比, 其最优解将保证此时网络上功率之和最小. 在 WCDMA 网络中, 用户功率的降低将导致总体干扰水平的降低, 因此求解上述问题, 问题 1) 将得到纠正, 同时也会降低网络上链路的干扰水平. 问题 CP 的约束矩阵为全幺模阵, 松弛掉整数约束后其解仍具有整性, 因此可转化为等价的线性规划问题, 比较容易求解

2) 加算法: 参照文献[1], 本文的加算法在基站集合  $C(J)$  中选择性价比最高的基站加入基站集合  $O(J)$ , 然后重新求解 CP 问题. 以上过程迭代至各基站的上行链路 SIR 满足要求为止

3) 减算法: 参照文献[1], 本文的减算法在基站集合  $O(J)$  中选择性价比最小的基站加入集合  $C(J)$ , 然后重新求解 CP 问题. 以上过程迭代至各基站的 SIR 不再增加为止

本文的分层算法如下:

第 1 层: 分解求解

Step 1: 对原 W-BSP 问题利用  $K$  均值聚类算法分成  $K$  个小问题 W-BSP( $r$ ),  $r = 1, 2, \dots, K$ ;

Step 2: 对 W-BSP( $r$ ) 分别利用整数规划工具来求解;

Step 3: 以各子问题 W-BSP( $r$ ) 的解构成原 W-BSP 问题的近似解

第 2 层: 全局调整

Step 1: 构造问题 CP 并求解;

Step 2: 检验各选中基站的上行链路干扰水平 SIR, 如果干扰水平超标则执行加算法, 否则执行减算法

### 5 仿真分析

对以上算法进行仿真计算, 计算数据来源于意大利的 Dipartimento di Elettronica e Informazione (DEI). 程序设计中利用了 GLPK 规划软件包, 基站到测试点的传播损耗用 COST 231 模型估算. 目标信干比  $SIR_{min} = 6 \text{ dB}$ ,  $\lambda = 1$ . 计算中令各基站的造价相同, 因此网络总造价等价于基站建设数量

对于有 15 个基站候选地址, 100 个测试点的 W-BSP 问题, 用  $K$  均值聚类分解的结果如图 1 所示, 其中具有相同数字标识的测试点属于同一类. 经过相关性计算, 聚类分解得到的同一类测试点的相关性均值为 0.78, 相关性标准方差为 0.2. 直接分块划分结果的相关性均值为 0.73, 相关性标准方差为 0.25. 通过对 30 组数据的计算结果进行比较分析, 聚类分解效果优于简单的方块切割

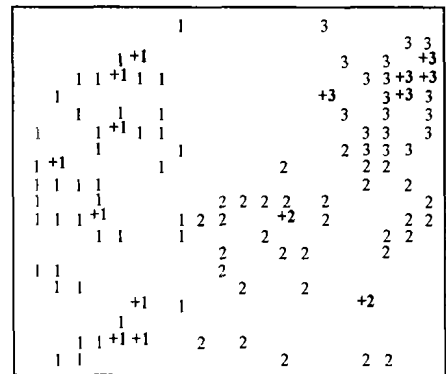


图 1  $K$  均值聚类分解结果

表 1 为 10 组数据的计算结果. 其中 SI 为网络中基站上行链路的最小 SIR, NB 为建设的基站数目, T 为计算时间. 启发式算法的设计参考了文献[1]中的加算法. 从计算结果可以看到, 虽然本文算法的计算时间大于启发式算法, 但仍然可以接受. 在算例中, 只有第 7 组数据分解计算后 SIR 达到规划要求, 其他问题的 SIR 严重超标, 这也说明了全局调整的必要性

进行第 1 步全局计算后, 所有 SIR 都得到很大改善. 对于部分算例如 3 和 4 等, 仍需要第 2 步全局调整以满足规划要求. 可以看到, 基于聚类分解的算法得到的基站数量明显少于文献[1]中的贪婪算法. 聚类分解方法较为充分地利用了 W-BSP 的数据信息, 而分块划分则难以利用这些信息. 当  $K = 6$  时, 聚类分解得到的子问题有解, 而此时分块划分已得不到有效的子问题, 因而分块划分时  $K = 4$ . 尽管如此, 仍有部分信干比没有达到目标值

表1 仿真计算结果

编号	聚类分解算法( $K=6$ )					$T/s$	启发式算法			分块划分( $K=4$ )	
	$SI^1$	$NB^1$	$SP^2$	$SP^3$	$NB^3$		$SI$	$NB$	$T/s$	$SI$	$T/s$
1	- 1.38	18	6.21	-	-	240	5.86	23	12	5.87	1.208
2	- 3.38	17	6.04	-	-	440	5.74	24	16	5.77	1.606
3	- 4.19	17	5.79	6.3	18	200	5.98	27	13	5.16	1.455
4	- 13.1	20	5.98	6.5	21	340	5.83	27	13	6.26	2.739
5	- 7.8	18	5.13	6.01	19	250	6.21	27	26	5.53	9.398
6	0.13	13	3.52	6.77	19	350	6.13	27	26	4.37	2.103
7	6.27	18	6.91	-	-	160	6.08	29	15	6.92	2.458
8	- 9.07	20	6.43	-	-	1.370	5.76	22	14	5.53	4.436
9	- 1.11	18	4.68	6.32	19	580	6.14	29	25	4.30	2.210
10	- 8.37	20	7.18	6.21	21	290	5.96	24	13	5.96	1.806

## 6 结 语

利用聚类分解,可将大规模W-BSF问题分成可以实际求解的小规模问题。通过相关性分析,聚类分解比直接按区域分块划分得到的各子问题的相关性有所降低。经过比较,基于聚类分解的算法效果较好。 $K$ 均值聚类算法的结果与初始值有关系,并且容易陷入极小点。如何采用更有效的聚类算法来提高分解效果,还需要深入研究。本文采用的网络规划模型比较简单,在今后的研究中应考虑更完善的规划模型,以满足W-CDMA无线网络规划的实际需要。

## 参考文献(References)

- [1] Edoardo Amaldi. Planning UMTS Base Station Location: Optimization Models with Power Control and Algorithms[J]. *IEEE Trans on Wireless Communication*, 2003, 2(5): 939-952.
- [2] Wright M. H. Optimization Methods for Base Station Placement in Wireless Applications[A]. *Proc of 1998 Vehicular Technology Conf* [C]. Ottawa, 1998: 11513-11517.
- [3] Molina A., Athanasiadou G. E., Nix A. R. The Automatic Location of Base-stations for Optimized Cellular Coverage: A New Combinatorial Approach [A]. *IEEE*

*Vehicular Technology Conf* [C]. Amsterdam, 1999: 606-610.

- [4] Tutschku K., Leibnitz K., Phuoc T. G. ICEPT- An Integrated Cellular Network Planning Tool[A]. *IEEE Vehicular Technology Conf* [C]. Phoenix, 1997: 765-769.
- [5] Matthias Gabta, Christian GlaSer, Steffen Reith. A Polynomial-time Approximation Scheme for Base Station Positioning in UMTS Networks[A]. *Proc of the 5th Int Workshop on Discrete Algorithms and Methods for Mobile Computing and Communications* [C]. Rome, 2001: 52 - 59.
- [6] 行小帅, 焦李成. 数据挖掘的聚类方法[J]. *电路与系统学报*, 2003, 8(1): 59-67.  
(Xing X. S., Jiao L. C. Clustering Method in the Field of Data Mining [J]. *J of Circuits and Systems*, 2003, 8(1): 59-67.)
- [7] 张潜, 高立群, 胡祥培, 等. 物流配送路径多目标优化的聚类-改进遗传算法[J]. *控制与决策*, 2003, 18(4): 418-422.  
(Zhang Q., Gao L. Q., Hu X. P., et al. Research on Multi-objective Vehicle Routing Problem of Optimization Based on Clustering Analysis and Improved Genetic Algorithm [J]. *Control and Decision*, 2003, 18(4): 418-422.)

(上接第212页)

- [8] Sanghamitra Bandyopadhyay, Ujjwal Maulik. An Evolutionary Technique Based on  $K$ -means Algorithm for Optimal Clustering in  $R^N$  [J]. *Information Sciences*, 2002, 146(1-4): 221-237.
- [9] Moody J., Darken C. Fast Learning in Networks of Locally-tuned Processing Units[J]. *Neural Computation*,

1989, 1(2): 281-294.

- [10] 王旭东, 邵惠鹤. RBF神经网络理论及其在控制中的应用[J]. *信息与控制*, 1997, 26(4): 272-284.  
(Wang X. D., Shao H. E. The Theory of RBF Neural Network and Its Application in Control [J]. *Information and Control*, 1997, 26(4): 272-284.)