

基于鲁棒优化的 QoE 测试网 不确定选址问题研究

宁颖丹

中国科学院大学 数学科学学院

专业方向：运筹学与控制论

指导教师：杨文国副教授

2015 年 12 月



目录

- ① 学分情况
- ② 研究背景
- ③ 研究进展
- ④ 后续进展
- ⑤ 论文情况
- ⑥ 参考文献



学分情况

总共选修课程：36 学分

其中学位课：21 学分

公共必修课程学分：6 学分

公共选修课学分：4 学分

专业学位课学分：12 学分



QoE 测试网不确定选址问题

QoE (Quality of Experience) 的测量 : QoT (Quality of Terminal) 与 QoD (Quality of Development) 的 KPI (关键网络性能指标) 信息及 KQI (关键业务质量指标) 的监视和测量, 监测和收集用户行为习惯及注册信息等。

QoE 测量网选点问题 : 选择尽可能少的点来代表尽可能多的用户来了解网络中用户接受服务的效果。该问题可以抽象为图论中的集覆盖问题。

集覆盖问题 : NP-Hard, 分支定界, 启发式算法



各测量点的外置探针可能由于某些内在的原因，如结构简单、能量有限，或者外在的因素，如自然或者人为影响，影响其正常测量用户接受服务的效果。

Snyder and Daskin: 给定故障概率下，目标为设施失效后运输成本的期望值最小的选址模型

Berman, Krass and Menezes: 设施按照一定概率中断后，在现存设施中寻求服务的选址模型

Cui 和 Ouyang: 各设施失效概率不同，运输成本之和最小为目标，构建可靠性和经济性供应网络



QoE 测试节点选址模型

QoE 测试节点选址模型可用 0-1 整数规划确切地描述如下：

$$\min \sum_{j \in J} x_j \quad (1)$$

$$s.t. \sum_j a_{ij} x_j \geq 1, \quad \forall i \in I = \{1, 2, \dots, n\} \quad (2)$$

$$x_j \in \{0, 1\}, \quad \forall j \in J = \{1, 2, \dots, m\} \quad (3)$$

x_j : 0-1 决策变量, $x_j = 0$ 表示集合 S_j 未被选中, $x_j = 1$ 表示选取集合 S_j

a_{ij} 表示集合 S_j 是否包含 e_i , 当 $e_i \in S_j$ 时为 1, 否则为 0;

公式 (1) 表示所求备选位置点个数最少;

公式 (2) 表示 E 中任意元素 e_i 至少被集合覆盖一次, 即每个网点服务质量情况至少可以被一个位置点测量;

公式 (3) 为完整性约束。



研究进展

QoE 测试节点鲁棒选址模型

$$\sum_j a_{ij}^t x_j \geq 1, \quad \forall i \in I, t \in T$$

失效选址的鲁棒模型

$$P_{\Gamma_i} \left(\sum_{j=1}^m a_{ij} x_j \geq 1 \right) \geq \alpha_i, \quad i = 1, 2, \dots, n$$



$T = \{1, 2, \dots, \max\}$ 为情景集合, 情景 t ($t \in T$) 下备选位置点集可测试的用户范围记为 S^t

S_j^t 为情景 t 下备选位置点集 S_j 的可覆盖用户

a_{ij}^t 表示情景 t 下备选位置点集 S_j 覆盖各用户的情况

选择尽可能少的点, 使得在不同的情景下, 这些点都可以覆盖全部的用户。

QoE 测试节点鲁棒选址模型

QoE 测试节点鲁棒选址模型可用 0-1 整数规划确切地描述如下：

$$\begin{aligned} \min \quad & \sum_{j \in J} x_j \\ \text{s.t.} \quad & \sum_j a_{ij}^t x_j \geq 1, \quad \forall i \in I, t \in T \\ & x_j \in \{0, 1\}, \quad \forall j \in J \end{aligned}$$



考虑几种特殊情形：

- $S_i^t \subset S_i^{t+1}, \forall i \in I, t \in T$
- $S_i^t \supset S_i^{t+1}, \forall i \in I, t \in T$
- 对任意 $i \in I_1, t \in T, S_i^t \subset S_i^{t+1}$, 对任意 $i \in I_2, t \in T, S_i^t \supset S_i^{t+1}$,
 $I_1 \cup I_2 = I$ 。

在上述几种特殊情形下，求得尽可能少的备选位置点使其在不同情景下均可覆盖全部用户，能够测量各用户使用不同网站的服务效果。



1. 枚举

2. 仅考虑最坏的情景

- 情形 (1) : $t^* = 1$, $S^{t^*} = S^1$
- 情形 (2) : $t = \max$, $S^{t^*} = S^{\max}$
- 情形 (3) : 构造 t^* , $S^{t^*} = S_{I_1}^1 \cup S_{I_2}^{\max}$



启发式算法

对备选位置点集中的集合进行处理，对任意两个集合都取并集得到新的集合，对新集合应用贪心算法的思想得到覆盖集并输出最终解。

算法流程

步骤 1. 输入 E , $S = \{S_j | j \in J\}$;

步骤 2. $X \leftarrow E$, $C = \emptyset$, $\Delta = \{S_{ij} | S_{ij} = S_i \cup S_j, \forall i < j \in J\}$, $k = 1$;

步骤 3. 若 $X \neq \emptyset$ ，判断 Δ 是否为空，若为空集， $C = S$ ，输出 C ；否则 $Y \leftarrow X$ 。选取 S_{ij} 使得其覆盖的 E 元素数目最大，记所得 S_i, S_j 为 u_k, v_k ； $X = X - S_{ij}$, $C = C \cup \{S_i, S_j\}$ ，将 Δ 中下标里含 i 与 j 的元素删掉，重复步骤 3；

步骤 4. 若 $X = \emptyset$ 判断是否 $Y \subset S_j, \forall j \in J$ ，若存在这样的 S_j ， $C = C - \{u_{k-1}\} - \{v_{k-1}\} + \{S_j\}$ 否则 $C = C$ ；

步骤 5. 输出所得 C 。

确定性选址问题算例分析

选取了 3 组需要测试的 QoE 测量网用户集，每组均分别生成 2000 个服务网络中的备选点集，根据鲁棒求解算法，用 Ruby 语言编制程序实现测试点求解并与贪心算法进行比较。

- 用户集 1：分别用本文算法及贪心算法进行求解，其中156次本文算法找的解比贪心算法更优，1829次本文与贪心算法所得解相同；
- 用户集 2：分别用本文算法及贪心算法进行求解，其中246次本文算法找的解比贪心算法更优，1709次本文与贪心算法所得解相同；
- 用户集 3：分别用本文算法及贪心算法进行求解，其中194次本文算法找的解比贪心算法更优，1774次本文与贪心算法所得解相同。

通过实算的结果来看，相对贪心算法而言，本文算法所得解的效果更好。



随机失效选址模型

随机失效 QoE 测试节点选址模型可用如下的规划来确切地描述：

$$\begin{aligned} \min \quad & \sum_{j=1}^m x_j \\ \text{s.t.} \quad & P\left(\sum_{j=1}^m a_{ij}x_j \geq 1\right) \geq \alpha_i, \quad i = 1, 2, \dots, n \\ & x_j \in \{0, 1\}, \quad j \in \{1, 2, \dots, m\} \end{aligned}$$

式中， x_j 为 0-1 决策变量，当 $x_j = 0$ 表示备选点 S_j 未被选中，当 $x_j = 1$ 表示选取备选点 S_j ，目标函数表示所求备选位置点个数最少，约束条件表示用户 e_i 能被测试到的概率不小于 α_i 。



测量概率矩阵 $P = (p_{ij})$

$$P(a_{ij} = 1) = p_{ij}$$

失效选址模型

$$\begin{aligned} \min \quad & \sum_{j=1}^m x_j \\ \text{s.t.} \quad & 1 - \prod_{j=1}^m (1 - p_{ij}x_j) \geq \alpha_i, \quad i = 1, 2, \dots, n \\ & x_j \in \{0, 1\}, \quad j \in \{1, 2, \dots, m\} \end{aligned}$$



$$(1 - p_{ij}x_j) = (1 - p_{ij})^{x_j} \implies \sum_{j=1}^m x_j \ln(1 - p_{ij}) \leq \ln(1 - \alpha_i)$$

失效选址模型

$$\begin{aligned} \min & \sum_{j=1}^m x_j \\ \text{s.t.} & \sum_{j=1}^m x_j \ln(1 - p_{ij}) \leq \ln(1 - \alpha_i), \quad i = 1, 2, \dots, n \\ & x_j \in \{0, 1\}, \quad j \in \{1, 2, \dots, m\} \end{aligned}$$



失效概率未知的 QoE 测量点选址模型

失效选址鲁棒模型

$$\begin{aligned} \min \quad & \sum_{j=1}^m x_j \\ \text{s.t.} \quad & P_{\Gamma_i} \left(\sum_{j=1}^m a_{ij} x_j \geq 1 \right) \geq \alpha_i, \quad i = 1, 2, \dots, n \\ & x_j \in \{0, 1\}, \quad j \in \{1, 2, \dots, m\} \end{aligned}$$

约束表示在用户 e_i 不多于 Γ_i 个 q_{ij} 的值取到最坏情况 $\bar{q}_{ij} + \hat{q}_{ij}$, 而其余 q_{ij} 均取名义值 \bar{q}_{ij} 的情况下, 用户 e_i 接受服务效果能被测试到的概率不小于 α_i



利用 Dirk Degel 的方法, 可将上述约束转化为如下的混合整数线性规划

$$\begin{aligned}
 & \min \sum_{j=1}^m x_j \\
 & s.t. \sum_{j=1}^m (w_{ij}x_j + \xi_{ij}) + \Gamma_i \eta_i \leq \ln(1 - \alpha_i), \quad i = 1, 2, \dots, n \\
 & \quad \xi_{ij} + \eta_i \geq (w'_{ij} - w_{ij})x_j, \quad \forall i, j \\
 & \quad \xi_{ij} \geq 0, \eta_i \geq 0, \quad i \in \{1, 2, \dots, n\} \\
 & \quad x_j \in \{0, 1\}, \quad j \in \{1, 2, \dots, m\}
 \end{aligned}$$



其中

$$w'_{ij} = \begin{cases} \ln(\bar{q}_{ij} + \hat{q}_{ij}), & \text{if } \bar{q}_{ij} + \hat{q}_{ij} > 0 \\ \ln(1 - \alpha), & \text{if } \bar{q}_{ij} + \hat{q}_{ij} = 0 \end{cases}$$

$$w'_{ij} = \begin{cases} \ln(\bar{q}_{ij}), & \text{if } \bar{q}_{ij} > 0 \\ \ln(1 - \alpha), & \text{if } \bar{q}_{ij} = 0 \end{cases}$$



随机生成用户数与备选位置点数均为 40 的网络，当不考虑测量点失效时，经典的选址问题选择测量点 $\{S_{12}, S_{14}, S_{37}\}$ ，选取的点个数为 3，在区间 $(0.8, 1)$ 上随机生成测量概率矩阵 P 中元素 p_{ij} 。

α	选取的位置点	选取位置点的个数
0.99	$S_2, S_{12}, S_{15}, S_{31}, S_{36}, S_{38}$	6
0.96	$S_8, S_{10}, S_{23}, S_{25}, S_{31}$	5
0.95	$S_{14}, S_{15}, S_{23}, S_{37}$	4
0.90	$S_{10}, S_{23}, S_{27}, S_{34}$	4
0.80	S_{10}, S_{12}, S_{34}	3



在小规模的 QoE 测量网进行测试，备选位置点数为 4，用户数为 5

$\alpha \backslash \Gamma_i$	0	1	2	3
0.8	S_1, S_4	S_1, S_2, S_3	S_1, S_2, S_3	S_1, S_2, S_3
0.7	S_1, S_4	S_1, S_4	S_1, S_3	S_1, S_3
0.6	S_1, S_3	S_1, S_3	S_1, S_3	S_1, S_4
0.5	S_1	S_1, S_3	S_1, S_3	S_1, S_3



后续进展

考虑选取测量点时需要一定的费用，研究费用不确定的 QoE 测试网选址问题。

$$\begin{aligned} \min \quad & \sum_{j \in J} c_j x_j \\ \text{s.t.} \quad & \sum_j a_{ij} x_j \geq 1, \quad \forall i \in I \\ & x_j \in \{0, 1\}, \quad \forall j \in J \end{aligned}$$

$$\begin{aligned} \min \quad & \max_{F \in D} \mathbb{E}_F \left(\sum_{j \in E} c_j x_j \right) \\ \text{s.t.} \quad & \sum_j a_{ij} x_j \geq 1, \quad \forall i \in I \\ & x_j \in \{0, 1\}, \quad \forall j \in J \end{aligned}$$



整理成一篇系统的硕士学位论文。

完成撰写：2016.03

答辩时间：2016.05



论文情况

- 服务网络 QoE 测试节点鲁棒选址问题研究, 网络新媒体技术, 2015.11
- 考虑节点失效的 QoE 测量点鲁棒选址问题研究, 在投



参考文献 I

- [1] Jain R. Quality of experience [J]. IEEE Multimedia, 2004, 11(1): 96-97
- [2] Jeffrey Spiess, Yves T'Joens, Raluca Dragnea, etc. Using Big Data to Improve Customer Experience and Business Performance [J]. Bell Labs Tech. J., 2014, 184.
- [3] Roth.R. Computer solutions to minimum cover problems[J]. OR, 1969.
- [4] Garey M R, Johnson D S. Computers and intractability: a guide to the theory of NP-completeness[M]. NewYork: WHFreeman,1979.
- [5] Carlo Mannino, Antonio Sassano. Solving hard set covering problems [J]. Operations Research Letters,1995,181.
- [6] J.E Beasley, P.C Chu. A genetic algorithm for the set covering problem [J]. European Journal of Operational Research,1996,942.
- [7] Lessing L, DumitrescuI, Stützle T. A Comparison between ACO algorithms for the set covering problem[J]. Lecture Notes in Computer Science, 2004,3172:1-12
- [8] Kouvelis P, Yu G. Robust discrete optimization and its applications[J]. Dordrecht: Kluwer Academic Publishers, 1997.
- [9] Williamson D P. Lecture notes on approximation algorithms. IBM, 1998,3:12-19
- [10] Chvátal V. A greedy heuristic for the set covering problem[J]. Mathematics of Operations Research, 1979, 4 : 233-235.



参考文献 II

- [11] 红叶. 基于移动互联网业务的 QoE 建模与分析 [D]. 北京邮电大学,2013.
- [12] Drezner, z. Heuristic solution methods for two location problems with unreliable facilities. Journal of Operations Research Society,1987,38(6), 509-514.
- [13] Snyder LV, Daskin MS. Reliability models for facility location: The expected failure cost case. Transportation Science ,2005,39:400-416.
- [14] Berman O, Dmitry K, Menezes M. B. C, Facility Reliability Issues in Network p-median problem: Strategic Centralization and Co-Location Effects, Operations Research, 2007, 55(2): 332-350.
- [15] Cui T, Ouyang Y F, Shen Z J M. Reliable facility location design under the risk of disruptions [J]. OR, 2010, 58(4): 998-1011.
- [16] Peng P, Snyder L V, Lim A, et al. Reliable logistics networks design with facility disruptions [J]. Transportation Research Part B, 2011, 45(8): 1190-1211.
- [17] Bertsimas D, Sim M. The price of robustness [J]. OR,2004,52(1).
- [18] Dirk Degel, Pascal Lutter. A Robust Formulation of the Uncertain Set Covering Problem 2013.



谢谢！

