

一种配电网中低压改造项目的优化决策方法

刘 洋¹, 李梦阳¹, 张学新^{2*}, 李 晶¹, 吴永华¹, 叶 青¹,
徐春山¹, 范 典¹, 周 信¹

(1. 国网孝感供电公司, 湖北 孝感 432000; 2. 湖北工程学院 数学与统计学院, 湖北 孝感 432000)

摘 要: 基于孝感市国网供电公司六大系统的数据, 根据不同类别的指标体系, 应用统计方法建立了配电网中低压改造项目的优化模型。首先构建 10 kV 馈线主干长度、主干截面等 7 个问题的故障诊断指标体系, 在层次分析法中利用决策试验与实验评估法(decision making trial and evaluation laboratory, DEMATEL)和交叉增援矩阵法计算项目的改造迫切度, 其次采用两步聚类算法挖掘台区项目改造优先区域, 最后展示了优化模型的应用。

关键词: 中低压配电网; 项目改造; 两步聚类; 优化模型

中图分类号: TM727.2 **文献标志码:** A **文章编号:** 2095-4824(2021)06-0085-05

随着经济发展及人民生活水平的提高, 本地用电负荷持续快速增长, 配电网供电能力与用电需求水平的矛盾日益突出。同时, 配电网中低压项目改造决策存在基础数据点繁多、人工经验预判准确性不高和效率低下等问题, 在配网改造资金有限的情况下, 亟需开展配电网中低压改造项目迫切度分析工作, 为配电网改造和电网规划提供参考, 不断提升配电网的供电能力。

关于配电网中低压改造项目问题的研究, 已有的文献分为两类: 定性的建议与量化的分析。曾丽等^[1]选择配电网 10 kV 馈线重过等 5 个指标, 应用层次分析法讨论了广州某区域中低压配电网改造方案迫切程度。贺少文^[2]、金胜松和周鹏^[3]从电力负荷的预测、站点的选择及配电网网络结构建设方面, 安全可靠, 经济可扩充性原则讨论中低压配电网规划建设的关键问题。铁宏建和朱明政^[4]则补充了合理选择供电半径、选择导线截面和杆型, 科学开展变压器台区改造等中低压配电网规划建设的关键问题。严松^[5]提出抓住网格化, 重点关注结构好的配电网建设改造项目需

求管理的优化思路。曹为^[6]则局限于一个单位, 讨论医院的供配电系统, 在信息化配网基础上提出中低压配网精细化管理。但是对于较大区域, 信息化配网并不便捷。在量化分析方面, 张伟明^[7]分别采用层次分析法和两步聚类算法, 对馈线主干长度和线损率等 7 个指标及配电台区过载相关的 4 个指标, 建立模型分析配电网项目迫切度, 最后将台区分为 3 个子类台区: 老旧和过载问题同时存在的台区; 老旧、过载、低电压、重载问题均不存在的台区; 只有老旧问题的存在台区。本文以生产管理信息系统 PMS2.0、营销管理系统 SG186、调度自动化系统 iES600、设备(资产)运维精益管理系统(GIS 系统)、电能量计量系统 TMR 和用电信息采集系统六大系统的电网参数及运行数据为基础, 采用大数据采集、清洗、整理归类等方法预处理数据, 再综合应用统计方法建立优化模型, 其中结合层次分析法, 采用 DEMATEL 方法和交叉增援矩阵法计算改造迫切度, 层次分析法模型的主要指标包括 10 kV 馈线主干长度、主干截面等 7 个方面问题的故障诊断信息。同时采

收稿日期: 2021-09-12

基金项目: 国网湖北省电力有限公司科技项目(5215K020006R)

作者简介: 刘 洋(1983-), 男, 湖北云梦人, 国网孝感供电公司工程师。

张学新(1966-), 男, 湖北宜城人, 湖北工程学院数学与统计学院教授, 博士, 本文通信作者。

用两步聚类算法挖掘台区项目改造优先区域。最后颗粒化管理项目,使项目储备与星级问题相对应,确保项目需求与电网发展衔接,并将结果融合至 Superset 实现图表实现可视化展示。

1 分析过程

1.1 统计方法

层次分析法是关于目标、多方案优化决策的系统方法,它首先把复杂问题的组成因素拆分成有序的递阶层次结构,再量化人们的主观判断,逐层建立判断矩阵,求解各判断矩阵的权重,最后获得最优方案的排序。

在解决问题的过程中,层次分析法使用了三个原理:优先级的分解,比较和综合。

分解原理应用于构建层次结构,其中同级水平的元素相互独立。从最高的一级目标层开始,到支撑目标的第二级准则,再到第三级子准则等若干准则层,一直到最底层的备选方案措施层。

比较判断的原理应用于构建一个两两比较倒数矩阵 A , 它的元素 $A(i, j)$ 是低层水平上的元素两两之间关于其上层水平的每个元素相对重要性(优先度), A 的主要特征向量给出了优先级。

第三个原理是对某个层的所有元素,以优先度为权重,并进行加权求和,得到综合指数。

可以看出,层次分析法是不使用三段论确实行演绎思维和归纳思维的一个非线性框架。它通过同时考虑多个因素并允许依赖和反馈,并进行数值折衷以获得综合评价,所以用它作为决策分析具有合理的方面。

两步聚类分析是一种统计方法,可以自动识别数据集内的相似人群或对象。相比传统的聚类方法(如均值聚类算法),两步聚类分析也被认为更可靠和更加准确,对连续和分类数据都有效的聚类技术^[8]。两步聚类允许保留数据完整的信息,不需要对数据转换,因而其结果能提供丰富的解释。两步聚类可兼顾处理数值型变量与分类型变量、自动处理异常值,舍弃或归入最近的类,比较适用于大型数据集的挖掘,所以本项目的优化模型第二部分也应用了两步聚类算法,并进行了改进,即对数据的其他非正态分布检验时,我们使用非参数检验代替 t -检验,从统计上检验聚类之间是否存在显著性差异。

两步聚类分析包含预聚类和聚类两步。第一步,通过构建聚类特征树将原始案例聚类。在第

二步,使用标准层次聚类算法,形成分层的聚类,允许探索具有不同数量聚类的解决方案。对于产生的多个解决方案,它能根据 Schwarz 的贝叶斯信息标准(Bayes information criterion, BIC)归纳出最优的聚类数目。一旦形成集群解决方案,就需要实施三种验证。首先,内聚和分离的轮廓度量必须高于设定的 0.05 的显著性水平,以表明集群内距离和集群间距离是有效的。其次,分别对分类变量和连续变量执行 χ^2 检验和 t -检验^[9],以识别聚类中各个变量的重要性,并指出聚类之间的显著差异。如果聚类里某个变量的统计信息的绝对值大于临界值,则该变量被认为在区分某个聚类不同于其他聚类时是重要的。第三,减半后的最终聚类解决方案必须相似(如群集的大小,数量和特征)。

下面详述两步聚类分析方法的统计指标。

在两步聚类分析中使用对数似然标准来揭示数据中的自然分组。设聚类 i 与聚类 j , (ij) 表示它们合并成的聚类,它们含有的样本点个数依次为 n_i 、 n_j 、 $n_{(ij)}$ 。数据集内有 K 个连续变量,其中第 k 个变量, $k = 1, 2, \dots, K$ 在整个数据集、聚类 i 、聚类 j 及 (ij) 的方差估计分别为 $\hat{\sigma}_k^2$ 、 $\hat{\sigma}_{jk}^2$ 、 $\hat{\sigma}_{(ij)k}^2$, 则聚类 i 与聚类 j 的间的对数似然距离 d_{ij} 定义如下:

$$d_{ij} = n_{(ij)} \left(\sum_{k=1}^K \ln \left(\sqrt{\hat{\sigma}_{(ij)k}^2 + \hat{\sigma}_k^2} \right) \right) - n_i \left(\sum_{k=1}^K \ln \left(\sqrt{\hat{\sigma}_{jk}^2 + \hat{\sigma}_k^2} \right) \right) - n_j \left(\sum_{k=1}^K \ln \left(\sqrt{\hat{\sigma}_{jk}^2 + \hat{\sigma}_k^2} \right) \right), \quad (1)$$

显然,对数似然距离是一种基于概率的距离。当两个聚类合并为一个聚类时,它们之间的距离与对数似然性的减少有关。在预聚类步骤中,扫描数据中的所有情况,并测量它们之间的对数似然距离,以某个阈值距离准则确定它们是否将形成预簇。在第二步中,使用凝聚聚类算法,把来自于预聚簇步骤的聚类集聚为最佳数量的聚类。这时第一步的每个子聚类被视为单个记录,如果它们满足(1)中的最小距离阈值,则再被合并为一个记录,直到所有数据记录被集聚成一个群集中。

如果未预先确定所需的聚类数量,则算法可以使用 Schwarz 的贝叶斯准则(BIC)或赤池信息标准(AIC)自动确定最佳数量。设聚类数量为 J , 数据集总观测值个数(总记录) N , 则有:

$$AIC(J) = 2 \sum_{j=1}^J n_j \left(\sum_{k=1}^K \ln(\sqrt{\sigma_{jk}^2 + \sigma_k^2}) \right) + 4KJ, \quad (2)$$

$$BIC(J) = 2 \sum_{j=1}^J n_j \left(\sum_{k=1}^K \ln(\sqrt{\sigma_{jk}^2 + \sigma_k^2}) \right) + 2KJ \ln(N), \quad (3)$$

BIC 被认为是最重要的有用和客观的选择标准之一,因为它避免了传统聚类技术的任意性。在考虑哪些变量要从分析中删除,最好选择 BIC 最低的。学生 t -检验用来确定两组数据的均值彼此是否不同。然而学生 t -检验假定数据服从正态分布,在这里,我们使用了非参数检验,更稳健的替换方法。轮廓系数用于测度簇密度和分离度。设有一个聚类结果 $C = \{C_1, \dots, C_k, \dots, C_m, \dots, C_c\}$, i, j 为样本点, $d(i, j)$ 是它们之间的某个距离, $|C_k|$ 为 C_k 族中样本点的个数。当 $i \in C_k$ 时,设 $a_i = \frac{1}{|C_k|} \sum_{j \in C_k, j \neq i} d(i, j)$, $b_i = \min_{C_m \in C, C_m \neq C_k} \frac{1}{|C_m|} \sum_{j \in C_m, j \neq i} d(i, j)$, 则点 i 轮廓系数定义为

$$s_i = \frac{b_i - a_i}{\max(a_i, b_i)}, \quad (4)$$

由此,所有样本点的平均轮廓系数为

$$\bar{s} = \frac{1}{N} \sum_{i=1}^N s_i, \quad (5)$$

从轮廓系数可以评价聚类效果, s_i 用来衡量一个点 i 是否适宜归为某个聚类程度。如果 $s_i < 0$, 说明点 i 处于不准确的聚类 C_k 里。 $s_i > 0$ 且接近 1, 说明点 i 准确地被分类在聚类 C_k 里。如果大多数样本点都有很高的轮廓系数系数, 即 $\bar{s}$ 比较大, 说明聚类适当, 否则, 说明分类数过多或者过少。实践中一般认为轮廓测量值小于 0.2 时, 聚类比较差; 在 0.2 到 0.5 之间, 聚类尚可; 大于 0.5 代表良好。

下面就 10 kV 馈线迫切度和 10 kV 台区迫切度各自从三个层面介绍以上统计方法的应用结果。

1.2 10 kV 馈线迫切度计算

1) 数据层。配电网的数据流主要由实时数据流、离线数据流和在线数据查询组成, 对配电网馈线主干长度、主干截面、装接容量、重过载、可转供比例、接线模式和线损率等数据进行封装整理, 得到迫切度分析数据样本, 选取其中的 5 条 10 kV 馈线作为测试数据样本。

2) 方法层。中压配电网线路项目涉及线路和设备繁多, 且需综合考虑改造效果、资金计划、供电可靠性、电能质量和经济效益等因素, 传统的数学优化方法难以对其进行统一、整体的建模, 为此, 选择层次分析法及其改进的方法。

通过改进层次分析法, 建立配电网馈线改造项目综合评价体系, 在资金一定的约束下, 对线路存在问题的多少与严重程度进行细致分析, 提出线路改造项目迫切度的概念, 即线路指标不合格对电网安全经济运行影响的严重程度折算值, 并采用基于 DEMATEL 方法及交叉增援矩阵法的改进层次分析法计算指标权重, 将线路指标不合格值与权重值相乘并累加计算线路改造迫切值。

以总项目投资额固定等为约束条件, 以线路迫切度改善最大为目标函数建立优化模型, 可在有限的资金约束下更科学、有效地安排建设与改造项目的优先次序。最后建立了层次分析法 3 层模型, 其结构如图 1 所示。

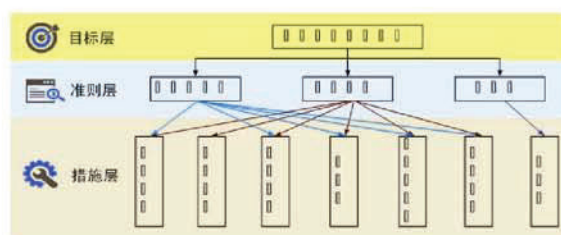


图1 层次分析法模型层次结构

对数据样本计算得到的馈线改造迫切度综合值如表 1 所示。

3) 应用层。把迫切度计算结果以数据可视化的方式在卫星地图上显示。馈线改造迫切度由高

表 1 馈线指标不合格值折算表

馈线名称	主干长度	主干截面	装接容量	重过载	转供比例	接线模式	线损率	综合值
巡 53 八里线	0.00	0.0	0.00	0.00	1.00	1.00	0.12	0.3749
巡 55 桑树线	0.00	0.0	0.00	0.00	1.00	1.00	0.82	0.4274
巡 56 巡店线	0.00	1.0	0.00	0.00	0.00	0.00	0.40	0.1773
巡 58 杨坡线	0.00	0.0	0.00	0.00	1.00	1.00	1.84	0.5039
巡 59 高店线	0.00	0.0	0.00	0.00	0.00	0.00	0.83	0.0622

至低依次为巡 58 杨坡线(红色)、巡 55 桑树线(黄色)、巡 53 八里线(绿色)、巡 56 巡店线(青色)、巡 59 高店线(蓝色),展示效果如图 2 所示。

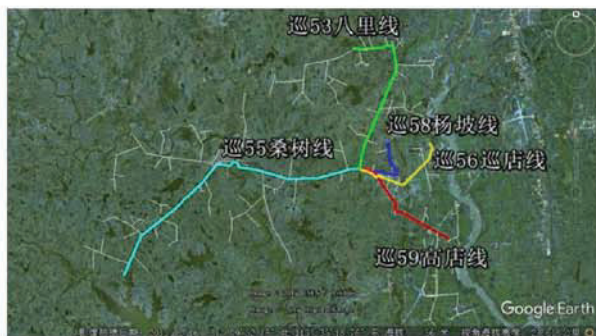


图 2 迫切度分析结果在卫星地图中的展示

1.3 10 kV 台区迫切度计算

1)数据层。从相关应用系统获取台区的经纬度坐标、重过载、运行年限、低电压等信息,建立台区问题聚类数据库,选取云梦公司存在问题的台区作为数据样本。

2)方法层。对台区过载、重载、老旧、低电压等问题采取两步聚类算法,判断是否重复出现低电压、是否成片出现低电压、是否监测数据质量问题、是否属于管理手段可解决的问题、是否确需纳入配电网项目解决以及先前投资安排策略是否合理,确定台区改造迫切度,结果见图 3。

序号	供电所	名称	经纬度坐标	低电压	ST-两步
1	巡桥	永兴村三组	113.7364,30.8560	0	1.000 聚类-1
2	巡桥	公安五号	113.7594,30.8187	0	1.000 聚类-1
3	巡桥	公安五号	113.7487,30.8239	0	0.000 聚类-2
4	巡桥	政府公寓	113.7550,30.8213	0	0.000 聚类-2
5	巡桥	金水花园公	113.7553,30.8182	0	0.000 聚类-2
6	巡桥	公安4号	113.7531,30.8194	0	1.000 聚类-1
7	巡桥	永隆五号	113.7539,30.8806	0	0.000 聚类-2
8	巡桥	永兴村2组	113.7461,30.8607	0	1.000 聚类-1
9	巡桥	明星村2号	113.7496,30.8641	0	1.000 聚类-3
10	巡桥	群利村2号	113.7475,30.8361	0	0.000 聚类-2
11	巡桥	前南村二组	113.7486,30.8448	0	1.000 聚类-1
12	巡桥	前南三组	113.7396,30.8446	0	1.000 聚类-1
13	巡桥	新堤四号	113.7623,30.8217	0	1.000 聚类-1
14	巡桥	永兴四号	113.7415,30.8621	1	0.000 聚类-1
15	巡桥	前北三组	113.7426,30.8472	0	1.000 聚类-1
16	巡桥	烟墩四号	113.7495,30.8804	0	1.000 聚类-1
17	巡桥	永兴五号	113.7383,30.8584	1	0.000 聚类-1
18	巡桥	长兴五号	113.7539,30.8731	1	1.000 聚类-1
19	巡桥	烟墩3号	113.7765,30.8193	1	0.000 聚类-1
20	巡桥	烟墩村	113.7550,30.8817	0	0.000 聚类-2
21	巡桥	长堤村	113.7564,30.8705	1	1.000 聚类-1

图 3 样本聚类结果

3)应用层。云梦公司 375 个台区共聚为 3 类,如图 4 所示。

图 4 中白线为区县行政边界,可以看出问题台区主要分布在县公司的西部和南部,由此可以得到如下结论:

同时存老旧和过载问题的台区为聚类-1,占比 18.4%,此类问题最迫切,共 69 台,应优先解决此类台区,约需改造资金 1700 万元。

样本台区中老旧问题最为突出,聚类-3 为

只存在老旧问题的台区,占比高达 67.7%,对此类问题台区宜可结合负荷发展逐步通过整体统筹方案解决。

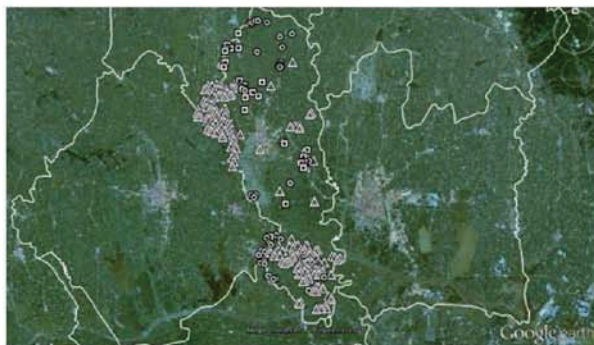


图 4 样本聚类结果空间展示

2 项目应用成效

1)本项目应用于孝感 2017 年第二批配电网项目评审,通过层次分析法计算得到 101 条线路项目迫切度,如图 5 所示。

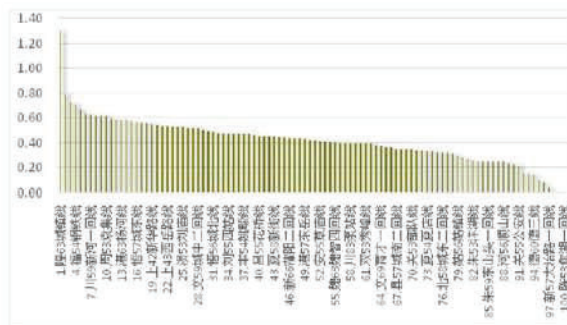


图 5 孝感 2017 年第二批配电网线路项目迫切度分析

通过本项目的实施可以为配电网线路项目评审提供既兼顾了专家经验又较为客观且符合投资导向的迫切度评价。

2)配电网规划。电网规划以问题和需求为导向,通过项目迫切度数据分析,明确问题和需求所处的区域,确定配电网规划的重点区域。

3)中低压项目评审。以馈线迫切度改善最大为目标函数,总投资等为约束条件,建立配电网优化改造模型,以使资金得到最大限度使用,有效安排项目改造顺序。

4)中低压项目后评价。对已投产中低压项目进行项目后评价,可改进投资决策管理,完善相关政策措施,提高科学管理水平,为今后中低压项目建设提供经验。

3 讨论

本项目的统计方法目前已经应用于如下多个

方面。

1)已应用 Ambari 2.6.0.0 在 VMware 虚拟机上安装 Hortonworks 公司 HDP-2.6.3.0 版大数据 4 机集群测试平台。

在台区风险预警方面,与区县公司开展台区所属单相用户相位识别研究合作,针对三相不平衡台区列出所有单相用户相位明细,根据偏相严重程度下达台区风险预警。

2)已实现按台区自动召测该台区下所有用户的电压。

3)未来可在电网现状分析方面,组建规划大数据工作室,搭建大数据平台,通过网络爬虫获取线路与台区的运行数据,根据电网诊断指标,形成电网现状分析报告。还需要与台区风险预警方面进行整合。

[参 考 文 献]

[1] 曾丽,荆朝霞,赖水生,等.基于层次分析法低压配电网改造方案迫切程度研究[J].南方电网技术,

2011,5(1):65-69.

[2] 贺少文.中低压配电网规划建设关键问题分析[J].企业技术开发,2013,32(14):97-98.

[3] 金胜松,周鹏.中低压配电网规划建设关键问题分析[J].山东工业技术,2014(19):172.

[4] 铁宏建,朱明政.中低压配电网工程规划建设关键问题分析[J].通信电源技术,2018,35(12):229-230.

[5] 严松.配电网建设改造项目需求管理思路的分析优化[J].中国新通信,2020,22(16):236.

[6] 曹为.信息化配网下中低压配电网精细化管理分析和探讨[J].科技创新与应用,2019(1):183-184.

[7] 张伟明.基于层次分析法和两步聚类算法的配电网项目迫切度分析[D].宜昌:三峡大学,2018.

[8] NORUSIS M J. SPSS 16.0 Guide to Data Analysis [M]. 2nd ed. Upper Saddle River: Prentice Hall, 2008.

[9] NORUSIS M J. IBM SPSS Statistics 19 Procedures Companion[M]. Texas: Addison Wesley, 2011.

An Optimal Decision Method of the Medium and Low Voltage Renovation Project for Distribution Network

Liu Yang¹, Li Mengyang¹, Zhang Xuexin^{2*}, Li Jing¹, Wu Yonghua¹, Ye Qing¹,
Xu Chunshan¹, Fan Dian¹, Zhou Xin¹

(1. State Grid Xiaogan Power Supply Company, Xiaogan, Hubei 432000, China;

2. School of Mathematics and Statistics, Hubei Engineering University, Xiaogan, Hubei 432000, China)

Abstract: Based on the data of the six major systems of the Xiaogan State Grid Power Supply Company, and for different types of index systems, an optimal model for the medium and low voltage transformation projects of the distribution network was established by applying statistical methods. Firstly, the researchers build a 10 kV feeder trunk length, trunk section and other seven trouble shooting index systems, and apply the DEMATEL(decision making trial and evaluation laboratory) method and the cross reinforcement matrix method to the analytic hierarchy process to calculate the urgency of project transformation. Secondly, the two-step clustering algorithm is used to mine the transformation priority area of the project and finally the application of the model is demonstrated.

Key Words: medium and low voltage distribution network; project transformation; two-step clustering; optimal model

(责任编辑:熊文涛)