

Early Assessments of Electric Power Information System Based on Projection Pursuit Cluster Method Integration

Diangang Wang¹, Jie Liu¹, Long Cheng¹, Junyong Liu², Zhen Ruan², Xiaodong Shen²

¹State Grid Sichuan Information and Communication Company, Chengdu Sichuan

²School of Electrical Engineering and Information, Sichuan University, Chengdu Sichuan

Email: liujy@scu.edu.cn

Received: Apr. 16th, 2017; accepted: Apr. 27th, 2017; published: Apr. 30th, 2017

Abstract

Scientific power information system of early assessment will provide data for the decision-making of leadership. So to find an efficient and accurate assessment of power information system method which is combined the data of equipment operation, safety indicators, network and operation mode, task supervision, room facilities, construction management, etc., has important significance. In this paper, a new method, Projection Pursuit Cluster, is introduced into the early assessment of power information system, and taking a provincial power company as an example, based on the index of fuzzy comprehensive evaluation method, we have made a longitudinal comparison and a lateral comparison of the company's power information system by using Projection Pursuit Cluster Method, overcoming the shortcomings of fuzzy comprehensive evaluation method, such as low flexibility and subjectivity and providing a new idea for the early assessment of power information system.

Keywords

Power Information System, Comprehensive Evaluation, Projection Pursuit Cluster Method

基于投影寻踪聚类法的电力信息系统早会评估

王电钢¹, 刘捷¹, 陈龙¹, 刘俊勇², 阮振², 沈晓东²

¹国网四川省电力公司信息通信公司, 四川 成都

²四川大学电气信息学院, 四川 成都

Email: liujy@scu.edu.cn

收稿日期: 2017年4月16日; 录用日期: 2017年4月27日; 发布日期: 2017年4月30日

摘要

科学的电力信息系统早会评估为领导决策提供了数据支撑，因此找到一种高效准确的评估方法，从设备运行情况、安全指标、网络与运行方式、任务监督、机房设施、建设管理等方面评价电力信息系统综合情况，具有重要意义。本文将一种新的方法：投影寻踪聚类法引入电力信息系统早会评估，以某省电力公司为例，在模糊综合评价法建立的指标得分基础上，运用投影寻踪聚类法对该公司电力信息系统早会综合情况进行纵向横向评估对比，克服了模糊综合评价法灵活性差，主观性强的缺点，为电力信息系统早会评估提供了一种新的思路。

关键词

电力信息系统，综合评估，投影寻踪聚类

Copyright © 2017 by authors and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

国家电网公司正在大力建设并稳步推进的智能电网是以信息平台为支撑，实现“电力流、信息流、业务流”的高度一体化融合的坚强智能电网。地区供电公司早会已成为供电公司日常生产安排、运营事务处理、重要问题会商的管理模式关键枢纽。随着管理精益化的逐步深入，地区供电公司早会已成为供电公司日常生产安排、运营事务处理、重要问题会商的管理模式关键枢纽。如何从各部门信息体系中准确收集、处理、钻取管理层在早会中关注的关键信息，并成体系的自动发布、推送，是供电公司早会走向系统化、信息化、智能化和更加精益化的必然趋势。早会信息支撑技术与一般专业部门的信息化相比，具有更强的综合性和宏观指导性，对数据结构设计、数据流程组织效率以及功能划分均有更高要求[1]。目前专业口径的电力信息化技术已较为普遍，针对专业应用，部分技术已基本实用化。国内外研究人员对电力信息系统从资源整合方法到风险及脆弱性评估方面已经做了大量的研究[2] [3] [4]。但是针对供电公司信息系统综合分析、运营态势评估的高层决策会商辅助系统，从理论到具体技术研究较少。该会商系统中，对各地区供电公司电力信息系统综合评估是对系统综合分析的基础，也是高层决策的重要依据。

目前对电力信息系统安全性风险性等评估常采用模糊评判法，层次分析法等方法，基于专家拟定的评分标准和公式，对各项指标打分后综合评估[5] [6] [7] [8]。但对信息系统综合评估时，评估对象数据维度较高，采用专家评价法，模糊综合评价法时使用的多为经验公式，操作繁琐，灵活性较差，实际操作不易实现，且客观性较差，局限性[9]。而投影寻踪法(projection pursuit, PP)是一种 20 世纪 70 年代初发展起来的多元数据分析的新数学方法[10]，相对于模糊评判法，物元分析法等方法，投影寻踪法具有精确处理高维、非线性、非正态问题的能力，计算简便、结果稳定等优点。该方法目前已经在水质评价、环境监测、灾情评估、水资源评价以及综合国力评估等诸多领域取得了广泛应用[11] [12] [13] [14]，但在电力信息系统建模中应用很少。由于电力信息系统综合评估的复杂性和特殊性，在模糊综合评价法建立的指标得分基础上，使用投影寻踪法建立的状态评估模型可以有效地对电力信息系统涉及到的多个指标进行多因素综合考察，为评价信息系统整体运行效果提供一个新的评价准则。

2. 投影寻踪聚类模型

投影寻踪聚类法(Projection pursuit cluster model, PPCM)的基本思想是把高维数据通过某种组合, 投影到低维(1~3 维)子空间上, 通过极大(小)化某个投影指标, 寻找出能反映高维数据结构或特征的投影, 在低维空间上对数据结构进行分析, 以达到研究和分析高维数据的目的。实质是一种数值方法求最优解的优化方法, 即利用已知特征指标值, 寻找最大程度地暴露高维数据某一类特征结构的最佳投影方向[15] [16]。建模过程如下:

1) 样本评价指标集的归一化处理。假设样本集合为 $\{x^*(i, j) | i=1, 2, \dots, n, j=1, 2, \dots, p\}$, 其中 $x^*(i, j)$ 表示第 i 个样本第 j 个指标值, n, p 分别表示样本的个数和指标的数目。

对于越大越优的指标

$$x(i, j) = \frac{x^*(i, j) - x_{\min}(j)}{x_{\max}(j) - x_{\min}(j)} \quad (1)$$

对于越小越优的指标

$$x(i, j) = \frac{x_{\max}(j) - x^*(i, j)}{x_{\max}(j) - x_{\min}(j)} \quad (2)$$

其中, $x_{\max}(j)$ 和 $x_{\min}(j)$ 分别为第 j 个指标的最大值和最小值。

2) 构造投影指标函数。把 p 维数据 $\{x^*(i, j) | j=1, 2, \dots, p\}$ 综合成以 $a = (a(1), a(2), a(3), \dots, a(p))$ 为投影方向的一维投影值 $z(i)$, 即

$$z(i) = \sum_{j=1}^p a(j)x(i, j), i=1, 2, \dots, n \quad (3)$$

选择不同的投影方向时, 评价指标数据以不同的权重组合方式得出不同的综合评价指标值。综合投影指标值 $z(i)$ 的散布特征的要求为: 投影点在局部上尽可能密集, 尽量凝聚成若干个点团, 而整体上应尽可能使点团散开[17]。因此, 投影指标函数可以表述为

$$Q(a) = S_z D_z \quad (4)$$

其中, S_z 为投影值 $z(i)$ 的标准差; D_z 为投影值 $z(i)$ 的局部密度。 $Q(a)$ 值越大, 聚类结果就越合理。其中:

$$S_z = \sqrt{\frac{\sum_{i=1}^n (z(i) - E(z))^2}{n-1}} \quad (5)$$

$$D_z = \sum_{i=1}^n \sum_{j=1}^n (R - r(i, j)) \mu(R - r(i, j)) \quad (6)$$

式中 $E(z)$ 为投影值 $\{z(i) | i=1, 2, \dots, n\}$ 的平均值。 R 为计算局部密度的窗口半径, 它的选取决定了包含在窗口内投影点的平均个数, 因而既不能太大, 也不能太小[18]。 $R = r_{\max} + p/2$ 是聚类的结果趋向于相对稳定的临界值点, 即为密度窗宽可取的最小值, 在实际运算中 R 的取值范围通常选取为 $r_{\max} + p/2 \leq R \leq 2p$ 。 $r(i, j)$ 样本之间的距离, $r(i, j) = |z(i) - z(j)|$ 。 $\mu(t)$ 为单位阶跃函数。

3) 优化投影指标函数。

当各指标值的样本给定时, 投影指标函数 $Q(a)$ 只随投影方向 a 的变化而变化, 不同的投影方向反映不同的数据结构特征。要确定最佳投影方向, 就是要找到能最大暴露高维数据分类排序结构特征的投影方向。因此可通过求解投影指标函数最大化问题来估计最佳投影方向[19], 即:

$$\max Q(a) = S_z D_z \quad (7)$$

$$s.t. \sum_{j=1}^p a^2(j) = 1 \quad (8)$$

4) 等级评价

把由 3) 中优化得到的最佳投影方向 a^* 带入(3)式, 计算的各样本的投影值 $z(i)$, 将 $z(i)$ 和 $z(j)$ 进行比较, 二者的大小越接近, 表示样本 i 与 j 越倾向于同一类。按照 $z(i)$ 的大小排序, 就可以得出待评估样本从优到劣的排序。

3. 应用分析

3.1. 评估指标参数确定

电力信息系统早会评估指标主要从各地市公司电力信息系统日常运行情况中汇总获得。日常运行情况主要包括以下六个方面一级指标:

- A) 设备运行情况。服务器, 网络设备, 链路的运行情况。
- B) 安全指标。弱口令数、违规外联数、保密安装率等电力信息系统运行的安全指标。
- C) 网络与运行方式。网络告警统计、数据刷新情况等指标。
- D) 任务监督。任务业绩指标及重大活动任务跟踪。
- E) 机房设施。各电力信息系统机房设施指标。
- F) 建设管理。科技、生产项目管理情况。

评估指标参数如表 1 所示。

以上评估指标具有种类多样, 难以量化的特点。运用模糊综合评价法, 得出个一级指标综合得分。对每个方面的二级指标采用 5 级评价机制 $V = \{v_1, v_2, v_3, v_4, v_5\} = \{\text{很好, 好, 中, 差, 很差}\}$, 采用德尔菲法, 根据专家的知识、经验分析、判断, 对各方面二级指标主观赋权值, 建立模糊评估矩阵, 经过计算得出各一级指标模糊综合评价结果。将该评价结果运用上节 1) 中方法进行样本评价指标集的归一化处理, 得出省公司各个年份纵向比较综合指数得分及各个地市公司横向比较的综合指数得分。其中纵向比较指数综合得分为该省电力公司 2012 年至 2016 年各指标评估得分, 横向比较指数得分为该省电力公司不同地市公司 2016 年各指标评估得分。数据如下表 2、表 3 所示。

3.2. 评价步骤及结果

根据表 2、表 3 中各项目的指数得分, 借助 Matlab2012 软件, 按照上节 2) 3) 步骤构造投影指标函数, 优化计算得出纵向比较及横向比较的最佳投影方向向量:

$$a_1 = (0.2263, 0.6281, 0.3098, 0.4566, 0.4306, 0.2837)$$

$$a_2 = (0.3701, 0.3651, 0.3834, 0.4149, 0.1632, 0.6202)$$

再由上节 4) 中等级评价方法分别计算 2012 至 2016 年投影特征值 $z_1(i)$, 及地区 1 至地区 8 投影特征值 $z_2(i)$

$$z_1(i) = \sum_{j=1}^6 a_1(j) x(i, j), i = 1, 2, \dots, 5$$

$$z_2(i) = \sum_{j=1}^6 a_2(j) x(i, j) j = 1, 2, \dots, 8$$

$$z_1(i) = \{1.4150, 1.5150, 1.7176, 1.8483, 1.8794\}$$

$$z_2(i) = \{1.5068, 1.8151, 1.2442, 1.2980, 1.7606, 1.4650, 1.1285, 0.7893\}$$

Table 1. Evaluation parameters system
表 1. 评估指标参数

一级指标	二级指标
运行设备情况	服务器 运行率、告警统计、CPU、内存
	网络设备 告警统计、运行率
	链路 运行率
安全指标	弱口令数、违规外联数、保密安装率、系统注册率、敏感检查率、补丁分发率、防毒安装率
网络与运行方式	告警统计、数据刷新情况
任务监督	业绩指标(MAS 系统)、任务跟踪、重大活动
机房设施	机柜空间、空调系统(漏水)、电源、温湿度、UPS、消防、视频
建设管理	科技项目、生产项目

Table 2. Longitudinal comparative index score
表 2. 纵向评估指数得分

年份	指标	运行设备	安全指标	网络运行方式	任务监督	机房设施	建设管理
2012		0.19	0.98	0.51	0.74	0.48	0.19
2013		0.19	0.96	0.49	1	0.48	0.19
2014		0.31	1	0.29	0.90	1	0.31
2015		0.62	0.89	1	1	0.48	0.62
2016		0.69	0.96	0.43	0.79	1	0.69

Table 3. Lateral comparative index score
表 3. 横向评估指数得分

地区	指标	运行设备	安全指标	网络运行方式	任务监督	机房设施	建设管理
地区 1		1.00	0.94	0.26	0.84	0.90	0.32
地区 2		0.97	0.72	0.27	0.75	0.97	1.00
地区 3		0.15	0.50	0.61	0.90	1.00	0.38
地区 4		0.48	0.91	0.49	0.66	0.21	0.47
地区 5		0.80	1.00	1.00	0.53	0.19	0.75
地区 6		0.98	0.37	0.20	1.00	0.71	0.58
地区 7		0.60	0.32	0.16	0.62	0.34	0.67
地区 8		0.21	0.29	0.49	0.66	0.16	0.19

Table 4. Longitudinal comparative result
表 4. 纵向评估结果

年份	投影特征值 $z_i(i)$	排序结果
2012	1.4150	5
2013	1.5150	4
2014	1.7176	3
2015	1.8483	2
2016	1.8794	1

Table 5. Lateral comparative result
表 5. 横向评估结果

地区	投影特征值 $z_2(i)$	排序结果
地区 1	1.5068	3
地区 2	1.8151	1
地区 3	1.2442	6
地区 4	1.2980	5
地区 5	1.7606	2
地区 6	1.4650	4
地区 7	1.1285	7
地区 8	0.7893	8

根据投影特征值的大小进行排序的结果于下表 4、表 5 所示。

由表 4 数据可知,纵向比较该省电力公司信息系统的早会平台总体评估情况逐年向好。2012、2013 两年投影特征值 $z_1(i)$ 较低,2014 年以后逐年升高,情况转好。由表 5 数据可知,横向比较该省电力公司各地市总体评估情况,除地区 8 情况较差, $z_2(i)$ 投影结果低于 1 以外,其他 7 个地区大体相差不大,其中地区 2 情况最好。

4. 结论

本文在运用模糊综合评价法建立的电力信息系统评估指标参数的基础上,利用投影寻踪聚类模型能将高维评价指标转化为一维投影指标、充分挖掘电力信息系统特征信息的特点,运用投影寻踪聚类法对某省电力公司信息系统的早会做了综合评估比较。该方法既解决了电力信息系统早会评估指标种类多样、难以量化的问题,相比于二阶模糊综合评价法等方法又最大限度的降低了专家主观因素影响,为电力信息系统早会评估提供了一种有效的新思路。

参考文献 (References)

- [1] 李博. 电力系统信息安全研究综述[J]. 中国科技信息, 2014(11): 200-201.
- [2] 郭创新, 陆海波, 俞斌, 等. 电力二次系统安全风险评估研究综述[J]. 电网技术, 2013, 37(1): 112-118.
- [3] 王继业, 张崇见. 电力信息资源整合方法综述[J]. 电网技术, 2006, 30(9): 83-87.
- [4] 石立宝, 简洲. 基于动态攻防博弈的电力信息物理融合系统脆弱性评估[J]. 电力系统自动化, 2016, 40(17): 99-105.
- [5] 孙绍辉. 基于智能电网的黑龙江省电力信息化建设评价研究[D]: [硕士学位论文]. 北京: 华北电力大学, 2012.
- [6] 杨恒, 王勇, 裴秀高, 等. 基于层次分析和模糊评价的电力信息系统风险评估方法[J]. 上海电力学院学报, 2013, 29(6): 518-522.
- [7] 冯小安, 解鸿斌, 刘艳平. 基于模糊综合评判法的电力网络信息系统安全评估[J]. 电网技术, 2008(23): 40-43.
- [8] 梁丁相, 陈曦. 基于模糊综合评判理论的电力信息系统安全风险评估模型及应用[J]. 电力系统保护与控制, 2009, 37(5): 61-64.
- [9] 李文武, 游文霞, 王先培. 电力系统信息安全研究综述[J]. 电力系统保护与控制, 2011(10): 140-147.
- [10] 付强, 赵小勇. 投影寻踪模型原理及其应用[M]. 北京: 科学出版社, 2006: 46-47.
- [11] 梁仲燕, 戴本林. 基于投影寻踪聚类法的某近岸海域水质评价[J]. 广州化工, 2016, 44(19): 148-151.
- [12] 程毛林, 韩云. 基于投影寻踪主成分分析法的综合国力评价模型研究[J]. 淮阴师范学院学报(自然科学版), 2015(1): 1-4.

- [13] 邵磊, 周孝德, 杨方廷, 等. 基于自由搜索的投影寻踪水质综合评价方法[J]. 中国环境科学, 2010, 30(12): 1708-1714.
- [14] 于国荣, 叶辉, 夏自强, 等. 投影寻踪分类模型的改进及在水质评价中的应用[J]. 四川大学学报(工程科学版), 2008, 40(6): 27-32.
- [15] 楼文高, 乔龙. 投影寻踪分类建模理论的新探索与实证研究[J]. 数理统计与管理, 2015, 34(1): 47-58.
- [16] 周静, 熊素琴, 苏斌, 等. 一种电力通信网络运行质量量化评估方法及应用[J]. 电网技术, 2012, 36(9): 168-173.
- [17] 孙兵, 张亚平, 戴银华. 基于改进三角模糊数的网络安全风险评估方法[J]. 计算机应用研究, 2009(6): 2131-2135.
- [18] 刘斐. 网络安全指标体系合理性评估研究[C]//中国计算机学会计算机安全专业委员会. 全国计算机安全学术交流会议论文集: 第二十五卷. 中国计算机学会计算机安全专业委员会, 2010: 6.
- [19] 卓先德. 网络安全评估的仿真与应用研究[J]. 计算机仿真, 2011(6): 177-180.

期刊投稿者将享受如下服务:

1. 投稿前咨询服务 (QQ、微信、邮箱皆可)
2. 为您匹配最合适的期刊
3. 24 小时以内解答您的所有疑问
4. 友好的在线投稿界面
5. 专业的同行评审
6. 知网检索
7. 全网络覆盖式推广您的研究

投稿请点击: <http://www.hanspub.org/Submission.aspx>

期刊邮箱: sg@hanspub.org