

考虑质量标签的多数据源配电网状态估计算法

徐玮韡^{1,2}, 刘东¹, 柳劲松¹, 黄玉辉¹, 鲁跃峰¹, 陆融³, 邱继东³, 潘华明³

(1. 上海交通大学 电气工程系 电力传输与功率变换控制教育部重点实验室, 上海 200240;

2. 浙江省电力试验研究院, 浙江 杭州 310014; 3. 上海市电力公司, 上海 200122)

摘要: 针对配电网多数据源的现状, 在电力系统综合数据平台收集多个异构系统信息的基础上, 利用质量标签和决策树理论对量测数据进行综合评价, 提出一种实用的配电网状态估计方法。利用配电网数据规律对配电网量测数据进行逻辑判断, 结合决策树理论框架建立数据质量评估体系; 并对不良数据进行修补以提高状态估计的输入数据质量; 利用质量标签修改权重, 对加权最小二乘法状态估计算法进行改进, 减少不良数据对状态估计结果的影响。某中心供电区域的 2 个算例表明了该方法的有效性和可行性。

关键词: 配电网; 状态估计; 决策树; 质量标签; 多数据源

中图分类号: TM 711

文献标识码: A

文章编号: 1006-6047(2011)04-0078-04

0 引言

状态估计是当代能量管理系统(EMS)的重要组成部分。它以系统量测数据为输入量, 求解系统的状态量(电压和相角)。这些状态量是电力系统一系列高级应用的数据基础。在状态估计器的估计准则方面, 探索和总结出了很多宝贵经验。

状态估计问题还涉及到不良数据检测与辨识、拓扑参数辨识以及系统可观性等问题。静态状态估计依据同一时刻各量测量之间空间的联系识别不良数据。动态状态估计是从时间和空间 2 个角度建立测量之间联系的二维状态估计方法, 利用各量测量随时间的变化识别不良数据等不正常事件^[1-2], 文献[3]提出了一种交直流混合状态估计方法。

以往配电网缺少数据量测设备和监测手段, 在此情况下, 配电网状态估计往往基于一系列假设条件而被转化为潮流匹配问题。文献[4-6]提出了多种求解配电网潮流的方法。随着电力系统自动化的不断实施, 对状态估计又有了新的要求。文献[7]提出了一种量测评估和配置优化的方法, 选出排序靠前的量测类型和安装位置, 从而得到量测装置的配置优化方案。文献[8]提出了一种适用于状态估计的面向对象的抽象电力网络拓扑模型。文献[9-10]在总结北京输电网 EMS 工程实践基础上提出“多源数据”的概念来表示同一个实时数据点来源多样化, 并根

据可信度、优先级、控制码及优选规则, 从多源数据中选取最优质量的数据作为 EMS 监控和历史统计的基础数据。如今, 随着配电网自动化的不断发展^[11], 配电网数据监测水平得到了显著提高, 由于配电网高级应用分析软件的需要, 配电网状态估计的必要性和可能性大幅提高。

配电网状态估计需要两大类数据, 即量测数据和线路参数数据。量测数据包括遥信数据和遥测数据, 前者主要来自于 DSCADA 的遥信信息, 它是配电网拓扑分析的基础, 后者主要来自于 DSCADA 的遥测数据, 此外还需要基于其他数据采集系统的历史数据以及电度量数据等冗余数据作为数据修补依据, 以达到状态估计对数据冗余度的要求。由于配电网的量测数据量巨大, 而且含有不良数据, 因此需要相应的数据过滤技术进行数据预处理, 即不良数据的检测与辨识, 以及数据修补。

与此同时, 在数据处理方面各种新技术和新理论不断涌现, 如专家系统、神经网络、遗传算法、机器学习等, 这为解决状态估计问题的智能计算和分析提供了可能。

本文立足于配电网量测配置现状, 以状态估计实用化为目标, 针对实际工程面临的问题, 利用决策树理论, 提出了一种新的配电网状态估计方法。

1 配电网量测的配置现状

配电网 DSCADA 系统所采集的配电网信息的完整性、正确性、采集频率直接影响到系统应用功能的有效性。随着上海电力各基础信息系统以及监控系统的相继建设, 配电网的量测水平在不断提高, 这为配电系统运行监控提供了有力支撑。

以上海市某中心城区为例, 其中低压配电网数据采集现状示意图如图 1 所示。目前其 35 kV 变电站都装有综合自动化系统; 10 kV 的大用户都装有集

收稿日期: 2010-05-10; 修回日期: 2011-02-15

基金项目: 国家重点基础研究发展计划 973 计划资助项目(2009CB219703); 上海市科技攻关重点项目(10dz1202800); 教育部新世纪优秀人才支持计划(NCET-07-0547)

Project supported by the Major State Basic Research Development Program of China (973 Program) (2009CB219703), Shanghai Science and Technique Foundation(10dz1202800) and the Program for New Century Excellent Talents in University of China(NCET-07-0547)

抄系统;10 kV 的 P 型站、K 型站、杆变、箱变和柱上开关分别装有 DTU、TTU 和 FTU,因此三相电流、电压、电量、开关遥信变位信号等数据都可以及时采集、上传和保存。各采集系统均实现与电力系统综合数据平台 IDP(Integrated Data Platform)的接口。

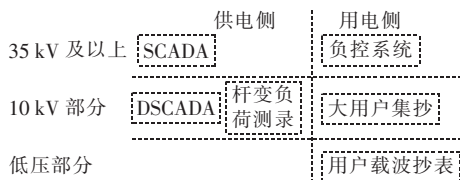


图 1 中低压配电网数据采集现状示意图

Fig.1 Multiple data sources of mid- and low-voltage distribution network

智能型调度在保持电力系统安全、经济、灵活运行方面具有重要的作用与意义^[12],而数据的准确与可用是智能型调度的基础,目前的配电网也面临多数据源的现状,需要一系列措施来对采集数据进行优选和估计,从而从源头上保证配电网监控管理数据的有效性,同时多数据源的存在客观上也为配电网状态估计提供了可能性。

2 基于质量标签的配电网状态估计

2.1 质量标签的定义

随着数据采集设备在配电系统中的广泛应用,获取实时数据的渠道越来越多,使得同一个实时数据点的信息可以从不同渠道、多个来源得到,即称为多数据源数据。由于各种渠道数据采集质量良莠不齐,影响后续的高级应用,所以有必要对多源数据进行可靠性判断和优选。因此对每个实时数据的每项来源数据均设质量标志,代表当前数据的质量状况。

定义量测点的数据质量标签为 $Q(i)$,它的取值为 0~1, $Q(i)$ 越大则表示数据越可信。可以将 $Q(i)$ 分为 $Q_v(i)$ 、 $Q_i(i)$ 和 $Q_p(i)$,分别反映该量测点处的电压、电流和功率的数据质量。

$Q(i)$ 初始值为 1。当某个量测点的量测数据违背下文中提到的配网数据规律,该量测点的数据质量标签每次减少相应的罚值,罚值=(1-标志位)×权值,但是不使数据质量标签值小于 0,即

$$Q(i) = 1 - (1 - f_{\text{tag}}) \times W \quad (1)$$

其中, f_{tag} 为下述规律符合情况的标志位; W 为每条规律相应的权重。

本文选取试点区域配网的部分数据作为样本,基于样本数据的质量标签赋值结果,结合决策树理论,对海量数据的质量标签进行初始化,经过数据修补模块修补的数据及其相应的数据质量标签作为状态估计的输入数据进行状态估计。

2.2 配网不良数据检测的依据

对于开环运行的配电线路,显然具有以下规律。

规律 1:电压在标称值范围之内。

规律 2:从电源至末梢方向,沿线电压逐渐降低^[13]。

规律 3:量测点电流满足

$$I_m > \sum_{n \in M} I_n \quad (2)$$

其中, M 表示量测点 m 的下游相邻量测点的集合^[13]。

规律 4:基于本文的研究对象,电度量与相应有功量测满足

$$P_{\text{Mi}} = [(P_{i-1} + P_i) / 2 \pm \xi] / (3.6 \times 10^6) \quad (3)$$

其中, ξ 为量测误差。

规律 5:经典的量测量突变检测为

$$\frac{|z_{i-1} - z_i|}{z_i} < \zeta \quad (4)$$

其中, ζ 为突变阈值。

规律 6:历史数据校核为

$$|z_{\text{di}} - z_{\text{da}}| < \omega \quad (5)$$

其中, z_{di} 为回溯近期 7 个相同日类型同一个时刻量测数据的均值; ω 为经验值。

经过不良数据检测模块之后生成的质量标签对应到相应的量测点数据下,作为输入数据代入数据修补模块。

2.3 不良数据检测规则的决策树算法

决策树学习算法中最有影响的是 ID3 算法^[1]。其基本原理是一种以递归的方式来划分输入变量所跨越的空间,最终得到的是一个分类。它用概率的函数来描述事件的不确定性;用熵表示信息的不纯度;用信息增益度来描述事件发生所提供的信息量。

设 S 是一组训练样本,该样本含甲、乙两大类数据, p_+ 是样本中甲类数据的所占比例, p_- 是样本中乙类数据的所占比例。则 S 的熵为

$$E(S) = -p_+ \log_2 p_+ - p_- \log_2 p_- \quad (6)$$

设 A 是数据的一个属性, V 是属性 A 的值的集合。用信息增益度 G 表示根据属性 A 的值进行分类求解的熵的减少值,即

$$G(S, A) = E(S) - \sum_{v \in V(A)} \frac{S_v}{S} E(S_v) \quad (7)$$

最优决策在 G 的最大值处取得。在每个子树的量测点上反复进行该迭代计算,直到树完整生成。

考虑估计不良数据分类问题,以第 2.2 节所述的 6 个规律为属性,分别简称为标么值规律、电压降规律、KCL、电量校核规律、突变量检测、历史数据校核。

理想条件下,当某量测点下的数据符合 6 个属性时,它必然是高质量数据,反之亦然。但是由于部分属性是通过和相关数据进行比较得出的,而不良数据的存在使得数据属性不能绝对地反映数据质量。作为训练样本的数据一定程度上反映了特定配电网及其数据获取条件下,数据属性与数据质量之间的关系。ID3 根据实际样本数据的属性和分类生成最小决策树,该决策树能够藉由最短路径对相同配电网及数据获取条件下的海量数据进行正确的分类。表 1 列出了已知质量标签分类的部分数据样本。图 3 的决策树表示了表 1 中的分类,这棵树能对表

中所有对象正确分类。在决策树中,每个内部节点表示对某些属性的测试,如标么值规律或电压降规律,这个属性的每个可能的值对应树的一条边。叶子节点表示类别,如低或中等质量。未知类别的数据通过遍历这棵树来进行分类;在每个内部节点,测试该数据此属性的值,取相应的边。继续这一操作,直到到达叶子节点,就找到了该对象的分类。

表 1 决策树模块中的部分样本数据

Tab.1 Partial sample data of decision tree

编号	标签	标么值规律	电压降规律	KCL	电量校核规律	突变变量检测	历史数据校核
1	0.2	不符	不符	未知	不符	符合	符合
2	0.2	不符	不符	未知	未知	不符	符合
3	0.6	符合	不符	未知	未知	符合	不符
4	0.2	未知	未知	不符	未知	不符	不符
5	1	符合	不符	未知	未知	符合	符合
6	1	符合	符合	未知	未知	符合	符合
7	0.2	未知	未知	不符	不符	符合	符合
8	0.6	符合	符合	未知	不符	符合	符合
9	1	未知	未知	符合	符合	符合	符合
10	1	未知	未知	符合	符合	不符	不符
11	0.2	不符	不符	未知	符合	符合	不符
12	0.6	未知	未知	较符	符合	符合	不符
13	0.6	不符	不符	未知	未知	符合	符合
14	0.2	不符	不符	未知	不符	不符	不符

KCL=0:0 (26.0/1.3)

KCL=1:2 (41.0/8.3)

KCL=2

标么值规律=0:0 (15.0/1.3)

标么值规律=1:2 (7.0/1.3)

标么值规律=2

电压降规律=0:0 (7.0/1.3)

电压降规律=1:0 (0.0)

电压降规律=2:2 (6.0/1.2)

电压降规律=3:2 (3.0/2.1)

标么值规律=3

电压降规律=0:2 (0.0)

电压降规律=1:0 (5.0/2.3)

电压降规律=3:2 (8.0/1.3)

电压降规律=2

电量校核规律=0:0 (0.0)

电量校核规律=1:2 (3.0/2.1)

电量校核规律=2:0 (0.0)

电量校核规律=3:0 (3.0/2.1)

图 2 样本数据分类决策结果

Fig.2 Decision tree grown from sample data

2.4 基于质量标签的不良数据修补

在数据修正过程中对质量标签进行更新,当某个量测点的量测数据不符合 2.2 节中配网数据规律 1~6,该量测点的数据质量标签每次减少相应的罚值,罚值=(1-标志位)×权值,但是不使数据质量标签值小于 0。当某个量测点的量测数据符合上述规律,相应地其质量标签每次加上相应的罚值,但是不使质量标签值大于 1。

设置质量标签分类阈值 χ_{bad} ,当某量测点量测数据的质量标签小于该值时,则该数据为可疑数据,继续数据修正程序;设置质量标签分类阈值 χ_{good} ,当某个量测点量测数据的质量标签大于该值时,该数据是可以信赖的,数据修正程序结束。

下面分别介绍电压、电流数据的修正方程。

假设量测点 s 的数据 k 需要修正,若为电压数据,则

$$U_k^{\text{re}} = U_k^{\text{mea}} Q_U(k) + \frac{U_j^{\text{mea}} + \max_{n \in K} U_n^{\text{mea}}}{2} [1 - Q_U(k)] \quad (8)$$

其中 j 为 k 的父节点; K 为 k 的子节点(下行节点)集合; U_k^{re} 为 U_k^{mea} 的修正值。

假设量测点 s 的数据 k 需要修正,若为电流数据,则

$$I_k^{\text{re}} = I_k^{\text{mea}} Q_I(k) + \frac{\left[\left(I_j^{\text{mea}} - \sum_{m \in J, m \neq k} I_m^{\text{mea}} \right) + \sum_{n \in K} I_n^{\text{mea}} \right]}{2} [1 - Q_I(k)] \quad (9)$$

其中 j 为 k 的父节点; J 为 j 的子节点集合; I_k^{re} 为 I_k^{mea} 的修正值。

3 算法总体流程及算例验证

算法的总体流程如图 3 所示。多源数据经过状态重获接口到统一的数据平台下;不良数据检测模块基于 6 个规律对数据进行判断,结果输入到决策树处理模块;经过样本训练的决策树对海量数据进行快速分类,分类结果赋以相应的质量标签;数据修补模块的结果作为状态估计的输入数据,其中质量标签作为加权最小二乘法的权重系数,对传统的做法做了进一步改进。

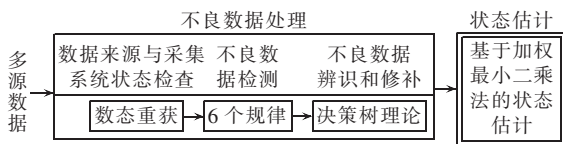


图 3 算法总体流程示意图

Fig.3 Overall process

为了说明本状态估计算法优越性,截取试点线路 2008 年 6 月的时间断面上的量测量进行验证。表 2 给出了 2 种权重体系下状态估计的效率比较。

表 2 迭代效率比较

Tab.2 Comparison of iteration efficiency

项目	传统 WLS	考虑质量标签的 WLS
迭代次数	2	2
最后 2 次迭代状态量差值	2.17811×10^{-6}	3.64992×10^{-7}
辅助判别量	0.00189865	9.80697×10^{-5}
迭代次数	3	3
最后 2 次迭代状态量差值	3.48544×10^{-7}	1.8459×10^{-9}
辅助判别量	3.69672×10^{-5}	2.17811×10^{-6}

以 46 节点的试点线路数据作为输入数据,由表 2 可知,基于质量标签的加权最小二乘估计经过 2 次迭代,其状态量的差值为 3.64992×10^{-7} ,精度高于传统最小二乘估计的 2.17811×10^{-6} 。在第 3 次迭代中,这种优势仍然明显。可见考虑质量标签的 WLS 发挥了其过滤不良数据和修补数据的作用,提高了状

态估计精度。虽然在既定的迭代结束条件下,传统 WLS 第 2 次迭代也可以收敛,但是从估计结果可以明显看出,考虑质量标签的状态估计估计结果数据具有更好的相容性。因此考虑质量标签的 WLS 在估计精度上优于传统 WLS,具有更广阔的应用前景。

4 结语

本文立足于目前配电自动化试点区域的量测配置情况和多数据源现状,对配电网量测数据进行估计前辨识,尽可能阻止错误数据的影响进入到状态估计程序中,克服了传统状态估计把局部的数据错误平摊到全部节点上的弊端。在不良数据检测与辨识模块中把配电网量测数据的自然属性结合到决策树理论中,克服了传统不良数据检测与辨识的残差淹没问题,减少了数据的误判概率。同时综合利用上、下电压等级的冗余量测数据,结合质量标签,对部分不良数据进行修补,提高状态估计输入数据质量。

试点线路的状态估计应用了本文提出的基于质量标签的最小二乘估计方法,结果表明该方法有利于提高高质量数据的稳定性,减少不良数据的影响,改善估计结果的准确性,是一种实用的配电网状态估计方法。

参考文献:

- [1] 周苏荃,柳焯. 新息图法拓扑错误辨识[J]. 电力系统自动化,2000,24(4):23-27.
ZHOU Suquan,LIU Zhuo. An innovation graph approach to topology error identification[J]. Automation of Electric Power Systems,2000,24(4):23-27.
- [2] 周苏荃,柳焯. 负荷突变与拓扑错误及坏数据三者交叠情况下的识别问题[J]. 中国电机工程学报,2002,22(6):6-10.
ZHOU Suquan,LIU Zhuo. Identification of three simultaneous anomalies of sudden load change and topology error and bad data[J]. Proceedings of the CSEE,2002,22(6):6-10.
- [3] 孙国强,李育燕,卫志农,等. 含 VSC-HVDC 的交直流混合系统状态估计[J]. 电力自动化设备,2010,30(9):6-12.
SUN Guoqiang,LI Yuyan,WEI Zhinong,et al. State estimation of power system with VSC-HVDC[J]. Electric Power Automation Equipment,2010,30(9):6-12.
- [4] 刘健,马莉,韦力,等. 复杂配电网潮流的降规模计算[J]. 电网技术,2004,28(8):45-49.
LIU Jian,MA Li,WEI Li,et al. Analysis of large scale distribution networks with reduced calculation[J]. Power System Technology,2004,28(8):45-49.
- [5] 蔡中勤,郭志忠. 基于逆流编号法的辐射型配电网牛顿法潮流[J]. 中国电机工程学报,2000,20(6):13-16.
CAI Zhongqin,GUO Zhizhong. Newton load flow for radial distribution network based on upstream labeling technique[J]. Proceedings of the CSEE,2000,20(6):13-16.
- [6] 孙宏斌,张伯明,相年德. 配电网潮流前推回推法的收敛性研究[J]. 中国电机工程学报,1999,19(7):26-29.
SUN Hongbin,ZHANG Boming,XIANG Niande. Study on convergence of back/forward sweep distribution power flow[J]. Proceedings of the CSEE,1999,19(7):26-29.
- [7] 徐臣,余贻鑫. 提高配电网状态估计精度的量测配置优化方法[J]. 电力自动化设备,2009,29(7):17-21.

- XU Chen,YU Yixin. Evaluation and optimization of meter placement to enhance distribution state estimation[J]. Electric Power Automation Equipment,2009,29(7):17-21.
- [8] 祝滨,仇晋. 基于矢量坐标搜索法的地区电网网络拓扑方法的研究[J]. 电力自动化设备,2003,23(2):63-65.
ZHU Bin,QIU Jin. Research of a network topology method based on vector coordinates[J]. Electric Power Automation Equipment,2003,23(2):63-65.
- [9] 卢建刚,温柏坚,姜彩玉,等. 基于“多源数据”处理的 EMS 实时数据选优技术及其实现[J]. 电力系统自动化,2002,26(23):73-75.
LU Jiangang,WEN Bojian,JIANG Caiyun,et al. Research of multi-source data and its quality control in energy management system[J]. Automation of Electric Power Systems,2002,26(23):73-75.
- [10] 叶锋,何桦,顾全,等. EMS 中负荷预测不良数据的辨识与修正[J]. 电力系统自动化,2006,30(15):85-88.
YE Feng,HE Hua,GU Quan,et al. Bad data identification and correction for load forecasting in energy management system[J]. Automation of Electric Power Systems,2006,30(15):85-88.
- [11] 刘东,丁振华,滕乐天. 配电自动化实用化关键技术及其进展[J]. 电力系统自动化,2004,28(7):16-19.
LIU Dong,DING Zhenhua,TENG Letian. Key technology and its development in the practicability of distribution automation systems[J]. Automation of Electric Power Systems,2004,28(7):16-19.
- [12] 艾琳,华栋. 电力系统智能型调度[J]. 电力自动化设备,2008,28(10):83-88.
AI Lin,HUA Dong. Power system intelligent dispatch[J]. Electric Power Automation Equipment,2008,28(10):83-88.
- [13] 刘健,毕鹏翔,董海鹏. 复杂配电网简化分析与优化[M]. 北京:中国电力出版社,2002:147-157.
- [14] 刘峰. EMS 中状态估计模块生产运行及实用化问题分析[J]. 电力自动化设备,2007,27(4):118-123,126.
LIU Feng. Operation and practicability of state estimation module in EMS[J]. Electric Power Automation Equipment,2007,27(4):118-123,126.
- [15] 刘辉乐,刘天琪. 电力系统动态状态估计的研究现状和展望[J]. 电力自动化设备,2004,24(12):73-77.
LIU Huile,LIU Tianqi. Status and prospect of power system dynamic state estimation[J]. Electric Power Automation Equipment,2004,24(12):73-77.

(编辑:李育燕)

作者简介:

徐玮韡(1983-),女,浙江温州人,硕士研究生,主要研究方向为配电自动化及配电管理系统(E-mail:vvvvei@hotmail.com);

刘东(1968-),男,江苏滨海人,研究员,博士研究生导师,博士,教育部新世纪优秀人才,主要研究方向为智能电网、调度自动化及计算机信息处理、配电自动化及配电管理系统(E-mail:liudong@ieee.org);

柳致松(1971-),男,安徽庐江人,高级工程师,在站博士后,研究方向为配电自动化及配电管理系统;

黄玉辉(1983-),男,山东威海人,博士研究生,主要研究方向为配电自动化及配电管理系统(E-mail:hyh8307@sjtu.edu.cn);

鲁跃峰(1985-),男,安徽阜阳人,硕士研究生,主要研究方向为配电自动化及配电管理系统;

陆融(1969-),男,上海人,高级工程师,研究方向为电力系统规划、配电自动化。

(下转第 86 页 continued on page 86)

State estimation based on quality tag for distribution network with multiple data sources

XU Weiwei^{1,2}, LIU Dong¹, LIU Jinsong¹, HUANG Yuhui¹,

LU Yuefeng¹, LU Rong³, QIU Jidong³, PAN Huaming³

(1. Key Laboratory of Control of Power Transmission and Conversion, Ministry of Education,
Department of Electrical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China;

2. Zhejiang Electric Power Test and Research Institute, Hangzhou 310014, China;

3. Shanghai Power Company, Shanghai 200122, China)

Abstract: A practical SE(State Estimation) method is proposed based on the data collected from multiple sources by the integrated data platform, which are evaluated with the quality tag based on the decision tree theory. The collected data are logically judged according to the rules of distribution network data and the data quality evaluation system is established combined with the decision tree theory. The quality of input data for state estimation is improved by data repair, and the quality tag is used to modify the weights to reduce the influence of bad data on the least-square state estimation. Two calculation cases of a distribution network prove the effectiveness and applicability of the proposed method.

Key words: distribution network; state estimation; decision tree; quality tag; multiple data sources