

ISODATA 方法在配网状态估计不良数据辨识中的应用

卫志农¹, 张云岗¹, 郑玉平²

(1. 河海大学电气工程学院, 江苏 南京 210098; 2. 国家电力公司自动化研究院, 江苏 南京 210003)

摘要:采用 ISODATA 技术(迭代自组织数据分析技术 A), 利用标准残差 R_n 和相邻采样时刻量测值之差 Δz 作为特征值, 对量测数据进行模糊聚类分析, 并根据隶属度的大小来辨识其是否属于不良数据. 数字仿真表明, 该计算方法简单、快速、可靠.

关键词:不良数据; 辨识; 模糊聚类分析

中图分类号: TM715

文献标识码: A

文章编号: 1000-1980(2002)02-0097-04

不良数据的检测与辨识是电力系统状态估计的重要功能之一, 其目的在于排除量测采样数据中偶然出现的少量不良数据, 以提高状态估计的可靠性^[1].

传统的不良数据的检测与辨识研究主要集中在输电系统, 常用的方法有残差搜索法、非二次准则法、零残差法、估计辨识法等. 这些方法主要是将加权残差 R_w 或标准残差 R_n 作为特征值, 假设其服从某一概率分布, 并按照一定的置信度水平确定一个门槛值, 进行假设检验. 找到可疑量测数据后, 将其从量测数据中排除或减小其权值, 得到新的状态估计值. 这类方法的缺点是很可能出现残差污染和残差淹没现象, 从而造成漏检或误检, 影响了辨识的效果. 在配网中, 由于量测数据少, 量测数据也不如输电系统来得准确、可靠, 残差污染或淹没现象更严重, 必须寻求更有效的方法.

许多学者采用了一些新的理论来解决这个问题. 文献[2]构造了一个基于 GMDH(group method of data handling)的神经网络, 用正则信息作为输入变量检测与辨识不良数据; 文献[3]用反向传播神经网络进行估计前的滤波, 用典型工况的正确量测作为训练样本, 以便在实时监控时能正确辨识不良数据, 但神经网络的不足之处在于训练样本的选定及其代表性将直接影响最后辨识的效果.

本文在文献[4]所提出的分解雅可比状态估计算法的基础上, 应用模糊聚类分析方法 ISODATA (iterative self-organizing data analysis technique A), 寻求良数据和不良数据的聚类中心, 并通过 R_n 和 Δz 两个特征值, 迭代计算每一量测值的隶属度, 再以 0.5 作为分界线, 快速、准确地辨识出不良数据.

1 分解雅可比状态估计算法

1.1 分解雅可比算法^[4]

首先, 将状态估计问题的目标函数按功率量测和电压量测分项列写为

$$J(x) = [Z_s - h_s(x)]^T R_s^{-1} [Z_s - h_s(x)] + [Z_v - h_v(x)]^T R_v^{-1} [Z_v - h_v(x)] \quad (1)$$

式中: x ——状态矢量; Z_s ——功率量测矢量; $h_s(x)$ ——功率量测函数矢量; R_s^{-1} ——功率量测权系数对角矩阵; Z_v ——电压幅值量测矢量; $h_v(x)$ ——电压幅值量测函数矢量; R_v^{-1} ——电压幅值量测权系数矩阵. 由此, 加权最小二乘法状态估计问题转化为下列方程的求解问题:

$$H_s^T R_s^{-1} [Z_s - h_s(x)] + \begin{bmatrix} 0 \\ R_v^{-1} [Z_v - x_v] \end{bmatrix} = 0 \quad (2)$$

式中: x_v ——状态矢量的幅值部分; $H_s^T(x) = \partial h(x) / \partial x$ ——功率量测雅可比矩阵.

由此, 可推得迭代式

收稿日期: 2001-01-02

作者简介: 卫志农(1962—), 男, 江苏江阴人, 教授, 主要从事电力系统运行与控制研究.

$$H_s^T(x^{(k)}) R_s^{-1} H_s(x^{(k)}) \Delta x = H_s^T(x^{(k)}) R_s^{-1} v_s^{(k-1)} + \begin{bmatrix} 0 \\ R_v^{-1} [v_v^{(k-1)} - \Delta x_v^{(k-1)}] \end{bmatrix} \quad (3)$$

式中: Δx ——状态矢量的增量; v_s ——功率量测误差矢量; v_v ——电压幅值量测误差矢量. 从而可以迭代求出状态矢量 x . 为了使该方法能够适用于配电网, 还需引入旋转变换, 具体参见文献[4].

1.2 残差灵敏度矩阵 W 的计算

在输电系统中, 残差矢量 r 与量测误差 v 之间的关系为

$$r = Wv \quad (4)$$

式中 $W = I - H(H^T R^{-1} H)^{-1} H^T R^{-1}$, W 为残差灵敏度矩阵.

在分解雅可比算法中, 将目标函数分为两部分. 与之类似, 也可将残差方程改写为

$$\begin{bmatrix} r_s \\ r_v \end{bmatrix} = W \begin{bmatrix} v_s \\ v_v \end{bmatrix} \quad (5)$$

式中: r_s ——功率量测残差矢量; r_v ——电压幅值量测残差矢量;

$$W = \begin{bmatrix} W_{ss} & W_{sv} \\ & I \end{bmatrix}$$

$$W_{ss} = I - H_s(H_s^T R_s^{-1} H_s)^{-1} H_s^T R_s^{-1}$$

$$W_{sv} = -H_s(H_s^T R_s^{-1} H_s)^{-1} R_v^{-1}$$

式中 I 为单位矩阵. 由此, 可以求出残差方程、加权残差以及标准残差等.

2 ISODATA 方法

2.1 模糊 ISODATA 聚类法基本原理^[5]

聚类分析属于数理统计多元分析的一支. 设有限样本集为

$$X = (x_1, x_2, x_3, \dots, x_n) \quad (6)$$

每个元素抽取 m 个特征

$$x_i = (x_{i1}, x_{i2}, \dots, x_{im}) \quad i = 1, 2, \dots, n \quad (7)$$

将它们分为 c 类: $A_1, A_2, \dots, A_c, 2 \leq c < n$. 这样的分类可以用一个 $c \times n$ 阶模糊矩阵 U 来表示:

$$\begin{matrix} & x_1 & x_2 & \dots & x_n \\ \begin{matrix} 1 \\ 2 \\ \dots \\ c \end{matrix} & \begin{bmatrix} u_{11} & u_{12} & \dots & u_{1n} \\ u_{21} & u_{22} & \dots & u_{2n} \\ \dots & \dots & \dots & \dots \\ u_{c1} & u_{c2} & \dots & u_{cn} \end{bmatrix} \end{matrix} \quad (8)$$

其中: $\forall k, i \quad u_{ki} \in [0, 1]; \forall i \quad \sum_{k=1}^c u_{ki} = 1; \forall k \quad 0 < \sum_{i=1}^n u_{ki} < n$.

这样的分类可以判断某一样本以某种程度属于这一类, 以另一种程度属于另一类. 模糊 ISODATA 算法就是找出在某种条件下的最优分类矩阵 U .

令

$$u_{ki} \in (0, 1)$$

$$V_k = \sum_{i=1}^n (u_{ki})^y \cdot x_i / \sum_{i=1}^n (u_{ki})^y \quad (9)$$

为第 k 类的聚类中心, 构造出泛函为

$$J(U, V) = \sum_{i=1}^n \sum_{k=1}^c (u_{ki})^y \|x_i - V_k\|^2 \quad (10)$$

其中 $\|\cdot\|$ 是 R^m 空间中的任一种范数. Dunn^[5]证明了上述泛函极小值的问题是可行的.

由此, 按最大隶属度原则, 即 x_i 对哪一类的隶属度最大, 就将 x_i 归入哪一类.

2.2 不良数据的模糊 ISODATA 辨识法

对于配网不良数据的辨识, 采用功率量测的标准残差 R_n 和相邻采样时刻量测值之差 ΔZ 两个特征值,

得到样本数据集.

采用量测突变检测来形成初始分类矩阵 U^0 . 当本采样时刻的量测值和前一时刻的预测值之差大于某一门槛值时,取对应项的 Z_i 对于不良数据集的隶属度 $U_A(Z_i)$ 为一个 0.5 和 1 之间的数,对于良数据集的隶属度 $U_B(Z_i)$ 为一个 0 和 0.5 之间的数,且满足 $U_A(Z_i) + U_B(Z_i) = 1$;否则相反.

得到初始分类矩阵 U^0 后,即可进行迭代计算,求出不良数据集.

最后,将得到最优解 (U, V) . 然后可以利用清晰化的方法,得到 Z 的普通分类,辨识出量测数据中的不良数据.

3 仿真算例及分析

采用 5 个节点、4 条支路的辐射型配网模型如图 1 所示.

采用有功量测 18 个,无功量测 18 个,功率量测总测点数为 36. 在 P_{1a} 点设不良数据为 $\pm 10\text{ kW}$,取 $\gamma = 2$,当 $R_n \geq 2.81$ 时,取 $U_A(Z_i) = 0.854$, $U_B(Z_i) = 0.146$;反之,取 $U_A(Z_i) = 0.146$, $U_B(Z_i) = 0.854$. 辨识结果以 0.5 为界,分别判定其隶属于不良数据集或良数据集的隶属度,如表 1 所示.

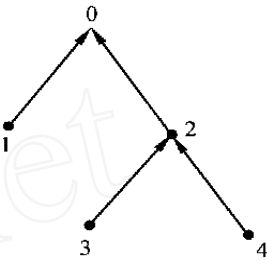


图 1 配网模型接线
Fig. 1 5-bus test system

表 1 各功率量测值对良数据集的隶属度

Table 1 Membership degree of measured power belonging to good data set

测点功率	P_{1a}	P_{1b}	P_{1c}	P_{3a}	P_{3b}	P_{3c}	P_{4a}	P_{4b}	P_{4c}
隶属度	0.1939	0.7894	0.8505	0.9816	0.8109	0.8986	0.9859	0.9878	0.9728
测点功率	P_{02a}	P_{02b}	P_{02c}	P_{23a}	P_{23b}	P_{23c}	P_{24a}	P_{24b}	P_{24c}
隶属度	0.9390	0.8755	0.9920	0.9746	0.8430	0.9920	0.9920	0.9695	0.9524
测点功率	Q_{1a}	Q_{1b}	Q_{1c}	Q_{3a}	Q_{3b}	Q_{3c}	Q_{4a}	Q_{4b}	Q_{4c}
隶属度	0.8886	0.9488	0.8499	0.9515	0.9659	0.9516	0.9905	0.9906	0.9674
测点功率	Q_{02a}	Q_{02b}	Q_{02c}	Q_{23a}	Q_{23b}	Q_{23c}	Q_{24a}	Q_{24b}	Q_{24c}
隶属度	0.8932	0.9626	0.9891	0.9512	0.9706	0.9920	0.9920	0.9785	0.9752

为简便计,当不良数据为 $\pm 20\text{ kW}$, $\pm 50\text{ kW}$, $\pm 100\text{ kW}$ 时,表 2 将只列出 P_{1a} 的隶属度.

算例表明,该方法能正确地辨识出不良数据. 其他算例同样说明了此方法的有效性,考虑到篇幅,这里不再举例说明.

表 2 不同不良数据值下的 P_{1a} 的隶属度

Table 2 Membership degree of P_{1a} under different bad data

不良数据值	$\pm 20\text{ kW}$	$\pm 50\text{ kW}$	$\pm 100\text{ kW}$
隶属度	0.1279	0.1065	0.0862

4 结 论

本文采用模糊聚类分析方法,对配网状态估计不良数据辨识问题进行了研究. 仿真算例表明:(a) 模糊 ISODATA 辨识法能快速、有效地辨识出不良数据,且较其他辨识法简单、可靠. (b) 该方法是在一定条件下的最优解,改变这些条件,如迭代初值、矩阵范数、指数等,可能得到其他的局部最优解. 如何得到全局最优解,还需继续研究. (c) 该方法是一种较理想的不良数据辨识法.

参考文献:

[1] 于尔铿. 电力系统状态估计[M]. 北京:水利电力出版社,1985. 105 ~ 165.
[2] Souza J C S, Leite da Silva A M, Alves da Silva A P. Information debugging in forecasting-aided state estimation using a pattern analysis approach[A]. In: Proc of 12th PSCC[C]. Dresden, Germany:Berlin in Press,1996. 45 ~ 49.
[3] Salehfar H, Zhao R. A neural network pre-estimation filter for bad data detection and identification in power system state estimation[J]. Electric Power System Research,1995,34:127 ~ 134.
[4] 卫志农,顾桦,鞠平,等. 三相辐射配网状态估计方法[J]. 中国电机工程学报,2000,20(3):84 ~ 87.
[5] 张俊福. 应用模糊数学[M]. 北京:地质出版社,1988. 5 ~ 48.

Application of ISODATA to Bad-Data Identification in Distribution Networks

WEI Zhi-nong¹, ZHANG Yun-gang¹, ZHENG Yu-ping²

(1. College of Electrical Engineering, Hohai Univ., Nanjing 210098, China;

2. Nanjing Automation Research Institute, Nanjing 210003, China)

Abstract: In the process of distribution system state estimation, bad-data identification can make estimation more reliable. An ISODATA method for bad-data identification is presented in this paper. In this method, the normalized residual and the difference between data measured at successive sampling times are taken as characteristic values, and by use of the fuzzy clustering analysis method, the bad-data are identified based on the membership degree of measured data belonging to good data set. Digital test shows that the method is simple, fast, and reliable.

Key words: bad-data; identification; fuzzy clustering analysis

《水资源保护》征订启事

(CN32 - 1356/ TV, ISSN1004 - 6933, 季刊, 自办发行)

《水资源保护》是由环境水利研究会与河海大学共同主办的,以技术性为主,兼顾学术性和管理性的技术性期刊,是全国唯一的水资源保护方面的专业性期刊。1985 年创刊,国内外公开发行。本刊以宣传贯彻国家的水资源保护方针政策,反映国内外水资源保护的先进科学技术,促进我国水资源保护和管理以及水环境保护工作为宗旨。其主要任务是探讨我国水资源保护的技术政策、科学技术、宏观管理及生产实践中的重大问题;及时反映国内外在水资源保护领域中的新颖实用技术和科技信息;推广全国各地在水资源保护工作中的成果和经验;报道有关国内外学术会议、技术交流等消息动态。

主要刊登内容:水资源保护、管理、评价、监测、优化配置、节约以及水环境污染控制和水环境监测仪器等方面的文章,主要栏目有科技政策、综合述评、学术研究、工程措施、成果推广、经验交流、专题讲座、国外动态等。

主要读者对象:从事水资源保护工作的水利、环保及相关领域的工作者,有关工程技术、科研人员、管理干部以及大专院校的师生。

《水资源保护》全年定价 30 元(含邮资在内)。每季末月 20 日出版。自办发行。

订阅方式:

(1) 邮局汇款:请将订单和邮签与汇款单仔细填好后寄回本刊编辑部(请注明订阅《水资源保护》杂志),未收到订单者亦可向编辑部函索订单。

(2) 银行汇款:开户行中国工商银行南京宁海路分理处,河海大学 4301011409001024513 (请注明订阅《水资源保护》杂志)。编辑部收款后即开具正式发票。

编辑部地址:210098 南京市西康路 1 号 河海大学。联系电话:(025) 3786642。