

基于多元线性回归的单相电表相别判断方法

张丽强¹, 丛伟¹, 董昱², 缪庆庆³, 孙海彬³, 孙允³

(1. 山东大学 电网智能化调度与控制教育部重点实验室, 山东 济南 250061;

2. 国网山东省电力公司, 山东 济南 250001; 3. 国网山东省电力公司济宁供电公司, 山东 济宁 272100)

摘要:单相电表的相别信息对提高负荷管理和建模水平、改善供电质量和供电可靠性具有重要意义。依据单相电表采集电气量与供电台区关口电表采集电气量间的关联关系,提出一种基于多元线性回归的单相电表相别判断方法,构建了以关口电表电压、电流和单相电表电流为自变量,以单相电表电压为因变量的多元线性回归方程,对每一个单相电表分别计算其与关口电表A、B、C相构成的3个回归方程的决定系数,根据决定系数的大小来判断相别。利用实际抄表数据对所提方法进行了算例分析,分析结果表明所提方法具有较高的准确性和可靠性。

关键词:单相电表;相别判断;多元线性回归;决定系数;高级量测系统

中图分类号:TM 933

文献标志码:A

DOI:10.16081/j.epae.202003023

0 引言

单相电表的相别信息对精准地进行负荷建模与供电网络线损计算、改善供电质量、提高供电可靠性等工作具有重要意义。基于单相电表的相别信息还可以开展负荷不平衡检测与判断,为准确有效的负荷不平衡治理提供依据。但是在当前的用电信息管理系统中并不包含单相电表的相别信息,由于用户电表数量庞大,用电和接线情况复杂,采用人工查找、确认的方式建立单相电表相别信息库难度较大^[1],需要研究一种单相电表相别自动判断方法。

高级量测系统AMI(Advanced Metering Infrastructure)近年来快速发展,提供了较为全面的用电量测数据^[2-4]。智能电表作为AMI的终端设备,承担着用户用电数据计量、采集、上传等任务,是实现用户用电信息集成和计算分析的基础。它根据预先设定的时间间隔(如15、30 min等)上传多种计量信息(如电能、有功功率、无功功率、电压、电流等),这些信息可以反映供电网络中每个节点的用电时间断面信息^[5],也是对供电网络拓扑结构、负荷、线损等要素的综合体现,在数学上具有特定的、可描述的关联关系,因此可以通过电表量测数据对单相电表的相别关系进行判断。

目前单相电表相别的判断方法主要有以下几类。

(1)基于载波通信的相别判断方法^[6-7]。利用I型集中器发出特定的载波信号,通过各单相电表对载波信号的响应进行相别判断,该方法适用于配置I型集中器的场合,要求集中器和电表均具备载波通信功能,对信号处理有较高要求。

(2)利用整数规划的相别判断方法。文献[8-9]采用0-1整数线性规划,使用分支和边界搜索算法进行相别判断,该方法需要台区下所有电表同时参与计算,计算量随问题规模的增加而快速膨胀。文献[10]采用线性规划和二次规划相结合的方法进行相别判断,在用户数据相似时会影响判断结果的准确性。

(3)基于聚类分析的相别判断方法。基于电压时间序列,文献[11]利用带约束的K-means聚类对电表连通性进行识别,但该方法的判别准确率随月份而波动。文献[12]利用频谱聚类的方法来验证和校正现有系统模型中的用户相别。该方法展示了机器学习在利用AMI仪表产生的大数据方面的潜在用途。

(4)基于相关性分析的相别判断方法。文献[13-14]提出利用电表电压测量值间的相关性进行台区识别的方法。文献[15-18]基于电表数据对配电网地理信息系统拓扑正确性进行校验,该方法受线路模型不确定性的影响较大。

本文提出一种基于多元线性回归的单相电表相别判断方法,以关口电表电压、电流和单相电表电流为自变量,以用户单相电表测得的电压为因变量,建立多元线性回归方程,通过计算单相电表与关口电表各相组成的回归方程的决定系数来判断单相电表的相别。利用现场实测数据进行算例分析,对所提方法有效性和准确性进行了验证。

收稿日期:2019-03-27;修回日期:2020-01-19

基金项目:国家自然科学基金资助项目(51877126);山东省自然科学基金资助项目(ZR2019MEE098)

Project supported by the National Natural Science Foundation of China(51877126) and the Natural Science Foundation of Shandong Province(ZR2019MEE098)

1 基于多元线性回归的单相电表相别判断模型

单相电表相别判断的多元线性回归模型为:

$$y = \beta_0 + \beta_1 x_1 + \cdots + \beta_m x_m + \varepsilon \quad (1)$$

其中, $x_j (j=1, 2, \cdots, m)$ 为自变量, m 为自变量个数; y 为因变量; β_0 为常数; β_j 为回归系数; ε 为随机误差项, 表示不由自变量决定的部分, $\varepsilon \sim N(0, \sigma^2)$ 。

将历史样本数据代入式(1), 用最小二乘法求出 $\beta_0, \beta_1, \cdots, \beta_m$ 的估计量 $\hat{\beta}_0, \hat{\beta}_1, \cdots, \hat{\beta}_m$, 得到回归方程为:

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \cdots + \hat{\beta}_m x_m \quad (2)$$

其中, $\hat{y}$ 为方程拟合值。

为确保所建立线性回归方程的准确性与合理性, 需要对式(2)进行如下检验。

(1) 自变量多重共线性检验。常用条件指数和方差比例对线性回归方程进行共线性诊断, 当某一维度的条件指数大于等于 30, 且对应的自变量方差比例大于 0.5 时, 则表明这些自变量间存在明显共线性(维度、条件指数和方差比例具体定义及多重共线性的判断标准见附录 A), 需进行变量的取舍选择。在多要素所构成的系统中, 控制其他自变量的线性影响来分析该自变量与因变量的线性相关性, 计算两者的偏相关系数, 根据其大小舍去对因变量影响较小的自变量, 保留对因变量解释程度更高的自变量。

(2) F 检验。检验所有自变量从整体上对因变量影响的显著性: 若 $F \geq F_{\alpha}(m, n-m-1)$, 其中 α 为显著性水平, n 为样本数量, 则认为回归效果显著; 反之则认为回归效果不显著, 需重新寻找自变量^[19]。

(3) t 检验。检验每个自变量对因变量影响的显著性, 若 $|t| \geq t_{\alpha/2}(n-m-1)$, 则认为该自变量影响显著; 反之则认为不显著, 将该变量从方程中剔除。

(4) 拟合优度检验。计算决定系数 R^2 来评价回归方程对样本观测值的拟合优度, 即:

$$R^2 = \sum_{q=1}^n (\hat{y}_q - \bar{y})^2 / \sum_{q=1}^n (y_q - \bar{y})^2 \quad (3)$$

其中, $\hat{y}_q, y_q, \bar{y}$ 分别为样本 q 的方程拟合值、观察值、样本观察值的平均值; R^2 为决定系数, 描述回归平方和占总离差平方和的比重, 值域为 0~1, 数值越大表示回归方程与样本拟合得越好。

R^2 随自变量的数目增加而增大, 可采用调整后的决定系数 $\bar{R}^2$ 代替 R^2 , 即:

$$\bar{R}^2 = 1 - (1 - R^2) \frac{n-1}{n-m-1} \quad (4)$$

有别于常规利用多元线性回归模型预测因变量, 本文基于线性回归的基本原理, 采用符合线性回

归特征的自变量和因变量构建多元线性回归方程, 通过式(4)计算各回归方程的决定系数 $\bar{R}^2$, 根据决定系数大小判断自变量对因变量解释程度的高低, 据此对单相电表的相别进行判断。

2 供电网络拓扑结构及变量分析

2.1 供电网络拓扑结构

在 400 V 供电网络中, 一般以供电台区为单位对用户电表进行管理。在变压器出线处装有三相关口电表, 能够采集并记录三相电压、电流、有功功率、无功功率、功率因数和电能等较为完整的电气量参数。下游的用户单相电表分别接在 A、B、C 三相, 一般只测量记录电压、电流、有功功率和有功电能等重要参数。完整的电表测量电气量种类信息如附录 B 中表 B1 所示。简化的 400 V 供电网络拓扑结构如图 1 所示。

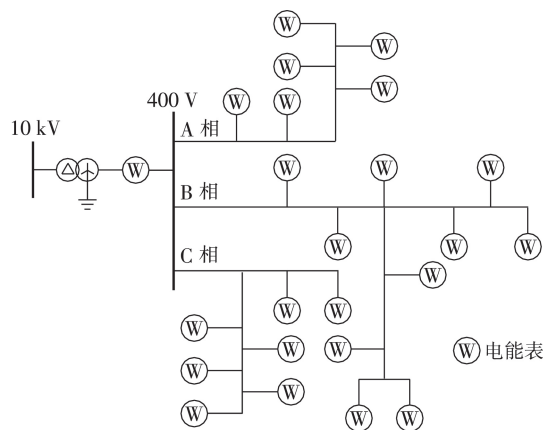


图1 简化的400 V供电网络拓扑结构图

Fig.1 Simplified topological structure of 400 V power supply network

400 V 供电网络呈辐射状结构, 关口电表与各单相电表在电路中是上下游关系, 二者电气量之间存在固定的、可描述的数学关系。但由于网络拓扑、供电线路参数等重要信息无法获取, 且电表只能提供某一时间断面而非全时间周期的量测信息, 故这种客观存在的数学关系难以直接描述。

低压配电变压器的三相电压难免存在一定的差异, 各相电压及负荷的差异会造成各相电流差异, 进一步导致各相线路上损耗的差异, 因此一般情况下不同相单相用户的电压会有区别, 且与关口表电压的变化有趋同性。考虑到单相电表与关口电表对应相量测值之间存在较强的相关性, 具备构建线性回归方程的条件, 可以采用多元线性回归模型描述上述变量间的关系。

2.2 供电网络变量关系分析

基于各类电表采集记录的电气量信息类型和内

容,本文建立的多元线性回归方程以用户单相电表的电压作为因变量,关口电表各相电压、电流、功率和用户单相电表电流、功率等参数中的部分或全部作为自变量。

图2为供电台区的等效电路图。图中, $j=A,B,C$; $i=1,2,\dots,r$; P_j, I_j, U_j, R_j 分别为关口电表 j 相有功功率、电流、端口电压幅值和线路电阻; $P_{ji}, I_{ji}, U_{ji}, R_{ji}$ 分别为 j 相用户 i 电表(即用户电表 ji) 的有功功率、电流、端口电压幅值和线路电阻; 低压配电线路一般较短,台区关口电表和各用户间的线路电抗值相对较小,因此可以将线路视为纯阻性;此外在供电网络中功率因数一般不低于0.9,因此可以忽略线路上的无功功率。

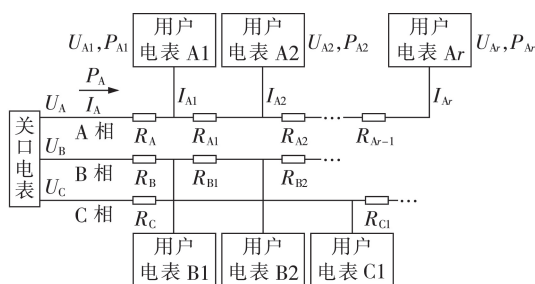


图2 供电台区电表连接等效电路图

Fig.2 Equivalent circuit of meter connection in power supply region

以A相为例分析关口电表和用户电表A2之间的电压关系。

$$U_{A2} = U_A - I_A R_A - \sum_{i=2}^r I_{Ai} R_{Ai} \quad (5)$$

$$U_{A2} = U_A - \frac{P_A R_A}{U_A} - \sum_{i=2}^r \frac{P_{Ai}}{U_{Ai}} R_{Ai} \quad (6)$$

由式(5)可见,用户电表电压 U_{ji} 与关口电表电压 U_j 和用户电表电流 I_{ji} 之间存在确定的数学关系;由式(6)可见,用户电表电压 U_{ji} 与关口电表有功功率 P_j 和用户电表有功功率 P_{ji} 之间也存在确定的数学关系。将上述电气量均作为回归方程的自变量,得到的线性回归方程为:

$$U_{zi} = \beta_0 + \beta_1 U_j + \beta_2 I_j + \beta_3 I_{zi} + \beta_4 P_j + \beta_5 P_{zi} \quad (7)$$

其中, $z=A,B,C$ 。

由式(7)可见,对任一未知相别的单相电表,都可以与关口电表建立3个如式(7)所示的线性回归方程。

3 共线性诊断及多元线性回归方程构建

由电路理论可知,在式(7)中自变量有功功率 P 和电流 I 之间存在共线性关系,因此,式(7)不能直接用来进行电表相别判断,对比式(5)和式(6),电流 I 和用户电表电压 U 的表达式更符合线性回归方程

的一般形式,但这是在忽略线路上的无功功率 Q 的情况下得到的结论,还需要根据实际情况,采用数据分析法选择恰当的自变量。

首先根据实际数据进行自变量共线性诊断以保留对因变量解释程度更高的自变量,基于式(7)得到适合本文场景和需求的回归方程表达式。以某地市供电公司的抄表数据为例说明自变量共线性关系的分析诊断过程。选择的供电台区1共有77个用户单相电表,A、B、C三相下分别接有25、28、24个用户。样本数据包含2018年10月10日00:00到10月20日00:00共240个时间点的数据,数据刷新周期为1h。关口电表和用户电表参数、电表部分实测数据分别如附录B中表B2和表B3所示。

自变量的共线性分析需要对每个用户的数据进行计算,计算量较大但计算过程较为相似。限于篇幅,本文以用户电表C3为例介绍自变量共线性的诊断过程。选取用户电表C3和关口电表的采集数据按式(7)所示方程建立线性回归方程,其中用户电表C3的测量电压为因变量,关口电表C相电压、电流、有功功率和用户电表C3的电流、有功功率为自变量,共线性诊断结果如附录B中表B4所示。

可以看出,维度5和6条件指数均大于30,其中维度5上 I_{C3} 和 P_{C3} 的方差比例均为0.7左右,维度6上 I_C 和 P_C 的方差比例为0.8左右,表明电流和有功功率存在明显的共线性关系。在存在共线性关系的自变量中通过偏相关系数检验进行变量选择,分别计算出 $U_C, I_C, P_C, I_{C3}, P_{C3}$ 与 U_{C3} 的偏相关系数,结果如表1所示。

表1 各自变量与因变量的偏相关系数

Table 1 Partial correlation coefficients between independent variables and dependent variable

因变量	偏相关性系数				
	U_C	I_C	P_C	I_{C3}	P_{C3}
U_{C3}	0.922	-0.835	0.760	-0.373	0.226

自变量 I_C 和因变量 U_{C3} 的偏相关系数绝对值为0.835,高于 P_C 和 U_{C3} 的偏相关系数0.760; I_{C3} 和 U_{C3} 的偏相关系数绝对值为0.373,高于 P_{C3} 和 U_{C3} 的偏相关系数绝对值0.226。可见相较于有功功率,电流和用户电压的相关程度更高。采用该供电台区下其他单相电表的数据进行分析结论相同,表明在自变量中保留电流变量、舍弃功率变量更加合理。

获取了其他供电台区的抄表数据(详见附录B中表B5),偏相关系数分析结果如表2所示,其结论与表1相同。

为进一步验证上述自变量共线性分析结论的合理性,应用SPSS软件将电流和有功功率分别引入回归模型方程计算 F 检验结果值、显著性计算结果值

表2 各自变量与因变量的偏相关系数

Table 2 Partial correlation coefficients between independent variables and dependent variable

因变量	偏相关系数				
	U_B	I_B	P_B	I_{B2}	P_{B2}
U_{B2}	0.913	-0.885	0.793	-0.122	0.109

和决定系数 $\bar{R}^2$ 计算结果值,如表3所示。

表3 模型运算结果

Table 3 Calculation results of model

模型	自变量	因变量	F值	显著性	$\bar{R}^2$
1	U_C	U_{C3}	360.9	0	0.601
2	U_C, I_C, I_{C3}	U_{C3}	1271.6	0	0.941
3	U_C, P_C, P_{C3}	U_{C3}	694.8	0	0.897

可见,经过 F 检验,3个模型的显著性值均为0,表示结果在0.01的水平上是显著的。模型 t 检验结果如附录B中表B6所示,在0.1的显著性水平下,模型2和3的自变量均通过了 t 检验,表明所选变量是有解释程度的。同时由表3可见,模型2保留电流作为自变量,得到的 $\bar{R}^2$ 为0.941;模型3保留有功功率作为自变量,得到的 $\bar{R}^2$ 为0.897,决定系数越高说明自变量对因变量的解释程度越好,所构成的线性回归方程也越合理。故通过计算决定系数值也进一步验证了在线性回归方程中保留变量电流而舍弃变量有功功率更加合理。

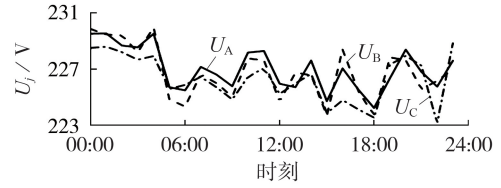
分别对该供电台区下77个单相用户电表决定系数 $\bar{R}^2$ 进行计算分析,结果表明:选择功率作为自变量时回归方程的 $\bar{R}^2$ 平均值为0.883,选择电流作为自变量时回归方程的 $\bar{R}^2$ 平均值为0.932。则在式(7)基础上本文建立的用于单相电表相别判断的线性回归方程为:

$$U_{xi} = \beta_0 + \beta_1 U_j + \beta_2 I_j + \beta_3 I_{xi} \quad (8)$$

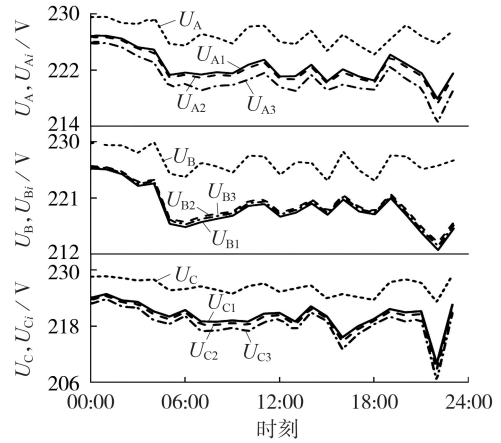
4 算例分析

采用第3节所述的供电台区1的电表采集数据进行算例分析。图3给出了某日该台区24 h的电压曲线。由图3(a)可见关口电表三相电压幅值及变化趋势在观测时间段内存在一定的差别;由图3(b)可见各单相电表电压幅值的变化趋势与关口电表对应相电压幅值的变化趋势较为相似,表明二者之间存在明显的相关性,这为建立以用户电压幅值为因变量,以关口电表电压幅值为自变量的线性回归方程对单相电表进行相别判断提供了可能。

限于篇幅,本文以用户电表A3和用户电表B1的相别判断过程为例进行分析。首先根据电压幅值时间序列的相关性进行相别判断,将用户电表A3、B1的电压幅值分别与关口电表三相电压幅值做相关性分析,结果如图4所示。



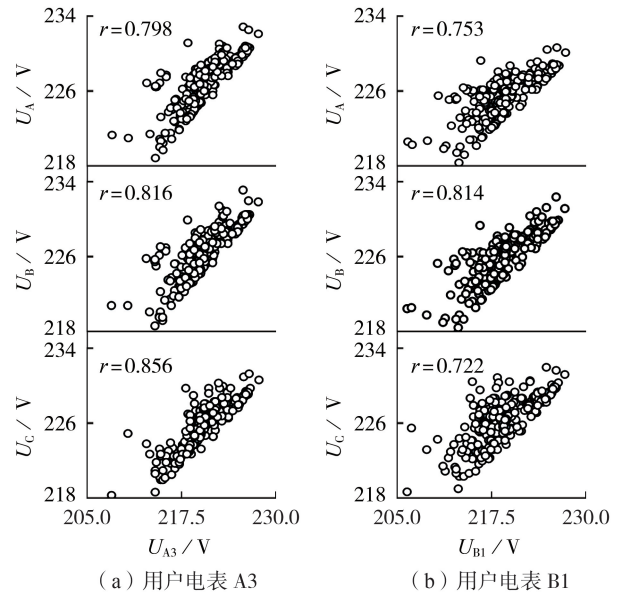
(a) 关口电表三相电压幅值



(b) 关口电表及用户电表电压幅值

图3 电压曲线

Fig.3 Voltage curves



(a) 用户电表 A3

(b) 用户电表 B1

图4 用户电表与关口电表三相电压的相关性比较

Fig.4 Comparison between user electric meter voltage and three-phase voltages in gateway meter

图4给出了单相用户电表电压幅值分别与关口电表三相电压幅值的散点图,并给出了二者电压幅值的相关系数值。可见,由于关口电表各相电压幅值的变化趋势较为相似,每个用户电表电压与关口电表三相电压的散点图中点的分布规律存在较大的相似性。由计算的相关系数可知,用户电表B1与关口电表B相的相关系数最高,为0.814,据此判断用户电表B1位于B相,判断结果正确;而用户电表A3

与关口电表C相的相关系数最大,为0.856,高于与关口电表A相的相关系数0.798,会发生误判,因此仅根据电压幅值变化趋势的相似性无法保证给出可靠的判断结果。

采用K-means聚类的判断方法,计算公式为:

$$D_{j-zi} = \sqrt{\sum_{t=1}^n (U_{jt} - U_{zit})^2} \quad (9)$$

其中, t 为采样点序号; U_{jt} 为关口电表 j 相的第 t 个采样点电压; U_{zit} 为用户电表 zi 的第 t 个采样点电压; D_{j-zi} 表示 U_j 和 U_{zi} 的 n 个采样点的等效空间距离。计算结果如表4所示。

表4 K-means聚类结果

Table 4 K-means clustering results

电表电压	D_{j-zi}		
	U_A	U_B	U_C
U_{A3}	103.04	96.28	85.27
U_{B1}	126.69	115.09	116.86

本文的聚类分析是通过计算用户电表电压幅值与关口电表三相电压幅值的等效空间距离值实现的,距离值越小表示越有可能位于同一相别上。由表4中的数据可见,用户电表A3与关口电表C相的距离值最小,判断结果不正确;用户电表B1与关口电表B相的距离值最小,但与关口电表C相的距离相比,区分度不明显,存在一定的误判可能。

采用本文所提多元线性回归方法进行计算,通过计算单相电表与关口电表三相分别建立回归方程的决定系数 $\bar{R}^2$ 值来判断相别,计算结果如表5所示。

表5 多元线性回归方程决定系数计算结果

Table 5 Calculation results of determination coefficients for multiple linear regression equations

相别	$\bar{R}^2$ (自变量为 U, I)		$\bar{R}^2$ (自变量为 U, P)	
	用户电表A3	用户电表B1	用户电表A3	用户电表B1
A	0.924	0.812	0.897	0.836
B	0.859	0.939	0.866	0.917
C	0.803	0.563	0.814	0.575

可见,对用户电表A3而言,与关口电表A相建立的线性回归方程的决定系数 $\bar{R}^2$ 最大, $\bar{R}^2=0.924$;对用户电表B1而言,与关口电表B相建立的线性回归方程的决定系数 $\bar{R}^2$ 最大, $\bar{R}^2=0.939$,判断结果准确、可靠。此外,还将采用电流 I 作为自变量得到的结果与采用功率 P 作为自变量得到的结果进行对比,可见,二者的计算结果都可以正确判断出用户电表所在相别,但采用电流 I 作为自变量计算结果更加可靠,区分度更加明显。

对算例中包含的77个单相用户电表分别进行多元线性回归计算分析,通过与用户的实际相别信

息核对,计算该方法的判断准确率。同时与采用电压幅值相关系数、K-means聚类等方法计算得到的判别结果进行对比,结果如表6所示。

表6 各种方法的判断准确率对比

Table 6 Comparison of judgment accuracy of

various methods		
编号	判别方法	判别准确率/%
1	相关系数	81.8
2	K-means聚类	79.2
3	多元线性回归(自变量为 U, P)	85.7
	多元线性回归(自变量为 U, I)	89.6

由判别结果可以看出基于多元线性回归的判断方法具有较高的准确率。分析3种方法的时间复杂度,方法1和方法2算法时间复杂度均为 $O(n)$,方法3的时间复杂度为 $O(n \times (m+1)^2)$ 。可见在自变量数量不大的情况下,上述3种方法的时间复杂度相当(具体计算分析过程见附录C),但本文所提方法具有更高的判断准确率。

通过分析多元线性回归方法自身特点和实际算例结果发现,以下因素会影响判断结果的准确率。

(1)关口电表和各用户电表数据采集的同步性。目前AMI无法保证各电表的数据采集是严格同步的,往往存在分钟级时间误差。非同步数据会影响各变量间的关联关系,进而影响判断结果的正确性。

(2)用户负荷及线路损耗的变化。用户负荷的变化会影响线路损耗,进而会抵消用户与总表各相之间的差异,增加了拟合优度的不确定性。

(3)数据样本的数量。如果数据样本过少会导致判别结果有较大的偶然性,但数据样本过多会增加计算量,因此需要确定适当的样本数据量。

(4)关口电表三相电压的不平衡度。如果台区三相电压基本平衡,用户电压与关口电表三相电压之间相关性差异会不够明显,进而影响判别结果的准确率。利用实际数据对方法的判别准确率进行分析,发现判别准确率随着不平衡度的增加而提高,当不平衡度在1%左右时,判别准确率能达到90%,具体分析过程见附录D。

5 结论

智能电表是位于电网末端的测量终端。AMI的覆盖为电网运行和管理人员提供了海量的用户端量测数据,充分利用这些数据,挖掘数据间的关联信息,能够实现多种附加功能和增值服务功能。

本文提出了一种基于多元线性回归的单相电表相别判断方法,该方法以400 V供电网络中AMI量测数据为基础,利用用户电表测量电压与供电台区

关口电表电压之间的关联性,采用多元线性回归方程描述二者间的关系。通过计算、比较用户电表与关口电表三相分别建立的线性回归方程决定系数大小,对单相电表所在相别进行判断,算例分析结果表明,与电压幅值相关系数法、K-means聚类法相比,本文所提方法具有较高的可靠性和准确性。本文的研究工作对进一步开展用户侧负荷精确建模、用户用电行为分析、用户供电异常判断与快速供电恢复、三相不平衡治理等应用具有重要意义。

附录见本刊网络版(<http://www.epae.cn>)。

参考文献:

- [1] 刘利平,蔡广林,陈旭,等.考虑用电可靠性的客户用电信息管理系统[J].电力自动化设备,2018,38(4):204-209.
LIU Liping, CAI Guanglin, CHEN Xu, et al. Electricity information management system considering power utilization reliability[J]. Electric Power Automation Equipment, 2018, 38(4): 204-209.
- [2] 张国荣,陈夏冉.能源互联网未来发展综述[J].电力自动化设备,2017,37(1):1-7.
ZHANG Guorong, CHEN Xiaran. Future development of energy internet[J]. Electric Power Automation Equipment, 2017, 37(1): 1-7.
- [3] 佟为明,梁建权,李中伟,等.基于数据可恢复的AMI无线传感器网络安全数据聚合技术[J].电力自动化设备,2017,37(5):211-216,223.
TONG Weiming, LIANG Jianquan, LI Zhongwei, et al. Security data aggregation based on data recoverability for wireless sensor network of AMI[J]. Electric Power Automation Equipment, 2017, 37(5): 211-216, 223.
- [4] 郝然,艾芊,肖斐.基于多元大数据平台的用电行为分析构架研究[J].电力自动化设备,2017,37(8):20-27.
HAO Ran, AI Qian, XIAO Fei. Architecture based on multi-variate big data platform for analyzing electricity consumption behavior[J]. Electric Power Automation Equipment, 2017, 37(8): 20-27.
- [5] 王思彤,周晖,袁瑞铭,等.智能电表的概念及应用[J].电网技术,2010,34(4):17-23.
WANG Sitong, ZHOU Hui, YUAN Ruiming, et al. Concept and application of smart meter[J]. Power System Technology, 2010, 34(4): 17-23.
- [6] 夏水斌,余鹤,董重重,等.基于集抄系统深化应用的低压台区电网拓扑重构方案[J].电测与仪表,2017,54(20):110-115.
XIA Shuibin, YU He, DONG Chongchong, et al. Low-voltage district topology reconstruction scheme based on extended application of remote centralized meter reading system[J]. Electrical Measurement & Instrumentation, 2017, 54(20): 110-115.
- [7] 张剑,徐向阳,陈文福.基于台区电表相位自动识别的动态三相相线损计算方法[J].华北电力技术,2012(6):9-11.
ZHANG Jian, XU Xiangyang, CHEN Wenfu. Dynamic three-phase split-phase line loss calculation method based on energy meter phase automatic recognition[J]. North China Electric Power, 2012(6): 9-11.
- [8] ARYA V, SEETHARAM D, KALYANARAMAN S, et al. Phase identification in smart grids[C]//2011 IEEE International Conference on Smart Grid Communications. Brussels, Belgium: IEEE, 2011: 25-30.
- [9] DILEK M. Integrated design of electrical distribution systems: phase balancing and phase prediction case studies[D]. Blacksburg, USA: Virginia Polytechnic Institute and State University, 2001.
- [10] KUMAR P, ARYA V, BOWDEN D A, et al. Leveraging DERs to improve the inference of distribution network topology[C]//2017 IEEE International Conference on Smart Grid Communications. Dresden, Germany: IEEE, 2017: 52-57.
- [11] WANG W, YU N P, FOGGO B, et al. Phase identification in electric power distribution systems by clustering of smart meter data[C]//2016 15th IEEE International Conference on Machine Learning and Applications (ICMLA). Anaheim, USA: IEEE, 2016: 259-265.
- [12] BLAKELY L, J-RENO M, FENG W. Spectral clustering for customer phase identification using AMI voltage timeseries[C]//2019 IEEE Power and Energy Conference at Illinois (PECI). Champaign, USA: IEEE, 2019: 1-7.
- [13] 徐大青,栾文鹏,王鹏,等.智能电表数据分析方法及应用[J].供用电,2015,32(8):25-30.
XU Daqing, LUAN Wenpeng, WANG Peng, et al. Analysis and applications of smart meter data[J]. Distribution & Utilization, 2015, 32(8): 25-30.
- [14] 栾文鹏,赵磊,王兵,等.智能电表数据分析元及示例[J].南方电网技术,2016,10(1):1-5.
LUAN Wenpeng, ZHAO Lei, WANG Bing, et al. Smart meter data analytics and application examples[J]. Southern Power System Technology, 2016, 10(1): 1-5.
- [15] LUAN W P, PENG J, MARAS M, et al. Distribution network topology error correction using smart meter data analytics[C]//Power and Energy Society General Meeting. Vancouver, Canada: IEEE, 2013: 1-5.
- [16] LUAN W P, PENG J, MARAS M, et al. Smart meter data analytics for distribution network connectivity verification[J]. IEEE Transactions on Smart Grid, 2015, 6(4): 1964-1971.
- [17] BOLOGNANI S, BOF N, MICHELOTTI D, et al. Identification of power distribution network topology via voltage correlation analysis[C]//52nd IEEE Conference on Decision and Control. Florence, Italy: IEEE, 2013: 1659-1664.
- [18] SHORT T A. Advanced metering for phase identification, transformer identification, and secondary modeling[J]. IEEE Transactions on Smart Grid, 2013, 4(2): 651-658.
- [19] 李昉,罗汉武.基于多元线性回归理论的河南省用电量预测[J].电网技术,2008(增刊1):124-126.
LI Fang, LUO Hanwu. Electric consumption forecasting of Henan province based on multiple linear regression[J]. Power System Technology, 2008(Supplement 1): 124-126.

作者简介:



张丽强

张丽强(1995—),男,山东威海人,硕士研究生,主要研究方向为智能电表数据挖掘与应用(E-mail: zlq-sdu@foxmail.com);

丛伟(1978—),男,山东威海人,副教授,博士,通信作者,主要研究方向为电力系统继电保护(E-mail: weicong@sdu.edu.cn);

董昱(1967—),男,山东济南人,高级经济师,主要研究方向为电网规划与运行(E-mail: hlliang1995@163.com)。

(编辑 王欣竹)

(下转第187页 continued on page 187)

Transient characteristic analysis of converter bridge arm short circuit fault in MMC-HVDC system

ZHANG Fang¹, DU Xuejing¹, CHEN Kun²

(1. Key Laboratory of Smart Grid of Ministry of Education, Tianjin University, Tianjin 300072;
2. Electric Power Research Institute, State Grid Hubei Electric Power Company, Wuhan 430077, China)

Abstract: Bridge arm short circuit is a severe fault in MMC-HVDC (Modular Multilevel Converter based High Voltage Direct Current) system. Transient characteristics of bridge arm short circuit fault are analyzed when the converter is unblocked or blocked. Aiming at the case that the converter is unblocked, the transient characteristics of bridge arm electrical quantity of the converter at both ends are given, and the composition of the bridge arm short circuit current is analyzed. For the other case that the converter is blocked, the circuit model of bridge arm short circuit current is established. Meanwhile, the analytical equation of bridge arm short circuit current is deduced, and the dynamic variations of voltage and current on AC and DC side are analyzed. The protection configuration scheme of bridge arm short circuit fault is also given. A two-terminal MMC-HVDC simulation model is built based on PSCAD / EMTDC. The simulation results validate the correctness of transient characteristic analysis and the effectiveness of the protection configuration. The transient characteristic analysis of the bridge arm short circuit fault can provide references for converter protection strategies of MMC-HVDC system.

Key words: MMC-HVDC; fault analysis; bridge arm short circuit; converter blocking; transient characteristic

(上接第149页 continued from page 149)

Method for single-phase electric meter phase identification based on multiple linear regression

ZHANG Liqiang¹, CONG Wei¹, DONG Gang², MIAO Qingqing³, SUN Haibin³, SUN Yun³

(1. Key Laboratory of Power System Intelligent Dispatch and Control of Ministry of Education, Shandong University, Jinan 250061, China; 2. State Grid Shandong Electric Power Company, Jinan 250001, China; 3. Jining Power Supply Company of State Grid Shandong Electric Power Company, Jining 272100, China)

Abstract: The phase information of single-phase electric meter is crucial to improve the level of load management, the quality of modeling, as well as the quality and the reliability of power supply. According to the relationship between the electric quantities collected by the single-phase electric meter and the gateway meter in the power supply region, a method for single-phase electric meter phase identification based on multiple linear regression is proposed. The multiple linear regression equation takes the single-phase electric meter voltage as the dependent variable. Meanwhile, it takes the gateway meter voltage, the gateway meter current and the single-phase electric meter current as the independent variables. Each single-phase electric meter forms three regression equations with the gateway meter phase A, B and C respectively. The determination coefficients of the three regression equations are calculated. According to the value of the coefficients, the phase of single-phase electric meter is determined. An example is given to analyze the proposed method with actual energy meter reading data. The analysis results show that the proposed method has higher accuracy and reliability.

Key words: single-phase electric meter; phase identification; multiple linear regression; determination coefficient; advanced metering infrastructure

附录 A

维度和条件指数是共线性诊断结果里的指标。维度即主成分，原则上如果有 p 个变量，则最多可以提取 p 个最大主成分，即 p 个维度，按照携带的信息量大小从 1 维到 p 维排列。

维度的计算方法为：从自变量 X 的协方差阵 Σ_X 出发，先求出 Σ_X 的特征值，并依从大到小的顺序排列成 $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq 0$ ，再求出相应于特征值 λ_1 、 λ_2 、 $\dots$ 、 λ_p 的单位正交特征向量 $\mathbf{a}_1$ 、 $\mathbf{a}_2$ 、 $\dots$ 、 $\mathbf{a}_p$ ，于是就可获得 X 的 p 个主成分 $\mathbf{y}_k = \mathbf{a}_k^T \mathbf{x}$ ，且 $\text{Var}(\mathbf{y}_k) = \lambda_k$ ，其中 $\mathbf{a}_k$ ($k=1,2,\dots,p$) 称为 X 的第 k 主轴即第 k 维度， $p=m+1$ ，为自变量的数目（计及常数项）， m 为自变量的数目（不计常数项）。

条件指数定义为最大的主成分与当前主成分比值的算术平方根，所以第一个主成分对应的条件指数永远是 1，最后一个主成分由于提取不来信息，条件指数自然会大。

条件指数的计算公式为：

$$h_d = \sqrt{\lambda_{\max} / \lambda_d} \quad d = 1, 2, \dots, p \quad (\text{A1})$$

方差比例定义为回归模型中（包括常数项）的变异能被各主成分所能解释的比例，即各主成分对模型中各项的贡献。

方差比例计算过程如下。

第 d 个主成分为：

$$F_d = u_{1d}x_1 + u_{2d}x_2 + \dots + u_{pd}x_p \quad (\text{A2})$$

第 g 个自变量为：

$$x_g = u_{g1}F_1 + u_{g2}F_2 + \dots + u_{gp}F_p \quad (\text{A3})$$

则：

$$\text{Var}(x_g) = \text{Var}(u_{g1}F_1 + u_{g2}F_2 + \dots + u_{gp}F_p) \quad (\text{A4})$$

$$u_{g1}^2\lambda_1 + u_{g2}^2\lambda_2 + \dots + u_{gd}^2\lambda_d + \dots + u_{gp}^2\lambda_p = \sigma_g^2 \quad (\text{A5})$$

其中， λ_d 为自变量 X 的协方差矩阵 Σ_X 的第 d 个特征根； σ_g^2 为自变量 x_g 的方差； $u_{gd}^2\lambda_d$ 是为第 g 个原始变量的方差中 F_d 能解释的部分。则 $u_{gd}^2\lambda_d / \sigma_g^2$ 为 F_d 提取的第 g 个原始变量信息的比重，即为第 d 个维度下 x_g 的方差比例。

多重共线性的判断标准为如果多个变量在同一主成分中贡献的信息量都很多，则说明这几个变量在这个主成分（方向）上存在较多的共同信息，也就是这几个变量存在较强的共线性关系。一般而言，在第 d 个维度下，当条件指数 $h_d \geq 30$ 且对应的自变量方差比例大于 0.5 时，可认为多重共线性严重存在。

附录 B

表 B1 400 V 供电网络中电表测量的电气量种类

Table B1 Types of electric quantities measured by electric meters in a 400 V power supply network

电表类型	电压	电流	零序电流	有功功率	无功功率	正向有功总电能	正向无功总电能	测量点正向有功总电能示值	总功率因数	各相功率因数
关口表（三相）	✓	✓	✓	✓	✓	✓	✓		✓	✓
用户表（单相/三相）	✓	✓		✓		✓		✓		

注：“✓”表示测量此电气量信息；空白表示不测量此电气量信息。

表 B2 某供电台区下关口表和用户电表参数

Table B2 Parameters of gateway meter and user meter in power supply region

电表类型	电压等级/kV	计量点容量/(kV·A)	综合倍率	计量方式	计量点所属侧	接线方式
关口电表	交流 10	160	30	高供低计	用户侧	三相四线
用户电表	交流 0.22	3/5/10	1	低供低计	用户侧	单相/三相四线

表 B3 2018 年 10 月 10 日供电台区一电表测量数据
Table B3 Measurement data of meters in power supply region 1 on October 10, 2018

时刻	U_A/V	I_A/A	P_A/kW	U_{A1}/V	I_{A1}/A	P_{A1}/kW	U_{A2}/V	I_{A2}/A	P_{A2}/kW	U_{A3}/V	I_{A3}/A	P_{A3}/kW
00:00	229.8	0.933	0.196	226.9	1.691	0.360	226.6	0.744	0.154	225.7	1.491	0.307
01:00	229.5	0.849	0.180	226.9	1.424	0.303	226.6	0.867	0.185	225.8	1.332	0.282
02:00	229.6	1.016	0.215	226.4	1.805	0.375	226.2	0.901	0.190	225.2	1.588	0.336
03:00	228.5	1.046	0.220	225.2	1.655	0.340	225.0	1.341	0.284	224.0	1.487	0.301
04:00	229.7	1.575	0.336	224.9	2.795	0.587	224.5	1.622	0.340	223.1	2.216	0.467
05:00	225.7	1.349	0.276	221.3	2.142	0.436	221.0	1.599	0.317	219.6	1.812	0.365
06:00	225.5	1.194	0.246	221.7	1.672	0.347	221.4	1.711	0.349	220.1	1.494	0.299
07:00	227.1	1.784	0.368	221.4	3.03	0.616	220.9	1.505	0.311	219.2	2.315	0.456
08:00	226.6	1.531	0.318	221.7	2.638	0.543	221.3	1.293	0.266	219.8	2.202	0.452
09:00	225.8	1.299	0.266	221.6	2.218	0.453	221.2	1.377	0.281	219.9	1.784	0.354
10:00	228.1	1.675	0.347	222.8	3.072	0.633	222.3	1.457	0.293	220.7	2.375	0.486
11:00	228.3	1.472	0.303	223.5	2.441	0.490	223.1	1.640	0.342	221.6	2.038	0.404
12:00	224.9	1.192	0.244	221.1	1.571	0.311	220.8	1.869	0.378	219.6	1.455	0.293
13:00	226.7	1.680	0.341	221.2	2.693	0.537	220.7	1.527	0.301	219.1	2.252	0.442
14:00	226.6	1.241	0.262	222.8	1.691	0.355	222.5	1.948	0.410	221.3	1.548	0.324
15:00	223.7	1.040	0.214	220.4	1.362	0.271	220.2	1.767	0.359	219.2	1.253	0.258
16:00	228.0	1.887	0.398	222.2	3.834	0.795	221.7	1.742	0.356	220.0	3.000	0.622
17:00	225.5	1.383	0.286	221.2	2.392	0.487	220.8	1.482	0.295	219.5	2.121	0.437
18:00	223.9	1.042	0.213	220.5	1.538	0.320	220.3	1.655	0.343	219.3	1.341	0.265
19:00	228.3	1.307	0.273	224.2	2.179	0.452	223.8	0.856	0.180	222.5	1.769	0.360
20:00	228.4	1.671	0.344	222.9	2.805	0.583	222.5	0.755	0.151	220.8	2.128	0.426
21:00	226.8	1.626	0.336	221.6	2.355	0.478	221.2	0.580	0.116	219.5	1.901	0.384
22:00	226.7	2.796	0.582	218.0	5.951	1.206	217.3	0.505	0.103	214.6	3.942	0.793
23:00	227.6	1.845	0.378	221.6	2.384	0.484	221.1	0.474	0.094	219.2	1.910	0.380
时刻	U_B/V	I_B/A	P_B/kW	U_{B1}/V	I_{B1}/A	P_{B1}/kW	U_{B2}/V	I_{B2}/A	P_{B2}/kW	U_{B3}/V	I_{B3}/A	P_{B3}/kW
00:00	229.8	0.824	0.174	225.7	0.752	0.157	226.1	1.527	0.322	225.9	0.452	0.094
01:00	229.4	0.775	0.164	225.5	0.860	0.182	225.9	1.290	0.265	225.7	0.428	0.089
02:00	229.4	0.910	0.191	224.8	0.880	0.181	225.2	1.537	0.314	225.0	0.683	0.143
03:00	228.2	1.041	0.219	222.9	1.316	0.274	223.4	1.493	0.304	223.2	0.983	0.207
04:00	229.9	1.327	0.281	223.3	1.630	0.343	223.9	2.206	0.463	223.7	0.495	0.103
05:00	224.9	1.610	0.334	216.8	1.632	0.331	217.6	1.839	0.376	217.3	2.344	0.473
06:00	224.3	1.529	0.308	216.3	1.639	0.320	217.1	1.489	0.297	216.7	2.480	0.487
07:00	226.6	1.895	0.394	217.1	1.436	0.283	218.0	2.331	0.463	217.6	2.189	0.451
08:00	226.0	1.706	0.358	217.6	1.270	0.257	218.4	2.223	0.460	218.1	2.487	0.508
09:00	225.0	1.355	0.278	218.1	1.349	0.270	218.8	1.847	0.379	218.5	1.542	0.307
10:00	227.7	1.581	0.329	219.7	1.479	0.302	220.5	2.326	0.466	220.1	1.518	0.316
11:00	227.6	1.465	0.300	220.0	1.601	0.326	220.7	2.030	0.400	220.4	1.312	0.262
12:00	224.7	1.333	0.276	218.0	1.884	0.384	218.6	1.484	0.304	218.4	1.214	0.242
13:00	226.7	1.619	0.337	218.6	1.566	0.317	219.4	2.357	0.484	219.1	0.741	0.147
14:00	226.5	1.289	0.269	220.1	1.925	0.401	220.7	1.520	0.312	220.4	0.814	0.165

15:00	223.7	1.084	0.224	218.3	1.789	0.369	218.8	1.254	0.258	218.6	0.640	0.125
16:00	228.4	1.534	0.323	220.7	1.753	0.360	221.4	2.983	0.615	221.1	0.738	0.153
17:00	225.4	1.293	0.266	218.8	1.529	0.314	219.4	2.075	0.418	219.2	0.988	0.200
18:00	223.7	1.045	0.213	218.3	1.584	0.314	218.9	1.374	0.278	218.7	0.657	0.131
19:00	227.8	1.338	0.276	220.9	0.830	0.169	221.6	1.731	0.345	221.3	1.575	0.318
20:00	227.6	1.869	0.388	218.2	0.779	0.161	219.0	2.192	0.452	218.6	2.482	0.487
21:00	225.6	1.990	0.403	215.3	0.578	0.115	216.2	1.920	0.391	215.7	3.025	0.587
22:00	226.1	2.703	0.558	212.7	0.491	0.098	213.8	3.885	0.771	213.3	3.658	0.737
23:00	227.0	2.128	0.436	216.1	0.469	0.092	217.1	1.899	0.376	216.6	2.018	0.400
时刻	U_C/V	I_C/A	P_C/kW	U_{C1}/V	I_{C1}/A	P_{C1}/kW	U_{C2}/V	I_{C2}/A	P_{C2}/kW	U_{C3}/V	I_{C3}/A	P_{C3}/kW
00:00	228.5	1.189	0.245	224.0	1.599	0.323	223.5	1.487	0.305	222.7	1.408	0.290
01:00	228.6	1.025	0.215	224.8	1.368	0.279	224.4	1.292	0.266	223.7	1.238	0.259
02:00	228.2	1.294	0.272	223.4	1.794	0.371	223.0	1.564	0.326	222.1	1.533	0.313
03:00	227.6	1.239	0.260	223.1	1.646	0.336	222.6	1.514	0.312	221.8	1.538	0.322
04:00	227.9	1.844	0.384	221.1	2.681	0.540	220.4	2.136	0.434	219.2	2.293	0.474
05:00	225.6	1.524	0.315	220.0	2.164	0.446	219.4	1.807	0.363	218.4	1.798	0.366
06:00	225.9	1.207	0.249	221.4	1.648	0.337	220.9	1.486	0.296	220.1	1.426	0.287
07:00	226.4	2.003	0.412	218.9	3.085	0.638	218.2	2.306	0.453	216.9	2.313	0.455
08:00	225.7	1.811	0.370	218.9	2.566	0.514	218.2	2.162	0.435	217.0	2.251	0.452
09:00	224.8	1.584	0.331	219.2	2.252	0.467	218.6	1.839	0.376	217.6	1.901	0.387
10:00	226.4	2.047	0.425	218.9	2.985	0.598	218.2	2.382	0.489	216.8	2.485	0.495
11:00	227.1	1.769	0.370	220.6	2.52	0.522	220.0	2.072	0.417	218.8	2.048	0.420
12:00	225.2	1.203	0.247	220.7	1.612	0.328	220.3	1.437	0.286	219.5	1.495	0.306
13:00	226.0	1.881	0.386	218.9	2.727	0.551	218.2	2.254	0.443	217.0	2.212	0.451
14:00	226.6	1.201	0.247	222.1	1.628	0.329	221.7	1.541	0.321	220.9	1.408	0.282
15:00	223.9	1.024	0.210	220.1	1.372	0.275	219.7	1.226	0.247	219.0	1.306	0.270
16:00	224.8	2.409	0.484	215.5	3.624	0.710	214.7	2.805	0.544	213.0	2.716	0.532
17:00	224.1	1.634	0.330	217.9	2.369	0.478	217.3	2.024	0.398	216.2	1.924	0.378
18:00	223.6	1.105	0.225	219.4	1.511	0.309	219.0	1.347	0.267	218.3	1.345	0.268
19:00	227.3	1.526	0.315	221.6	2.141	0.436	221.0	1.728	0.344	220.0	1.911	0.384
20:00	227.9	1.877	0.388	220.9	2.687	0.535	220.2	2.169	0.443	218.9	2.344	0.468
21:00	227.2	1.621	0.333	221.1	2.306	0.458	220.5	1.876	0.373	219.4	2.024	0.419
22:00	223.2	3.486	0.696	209.9	5.640	1.083	208.6	3.738	0.714	206.3	3.831	0.735
23:00	228.8	1.688	0.351	222.5	2.408	0.493	221.9	1.997	0.416	220.8	2.040	0.413

表 B4 共线性诊断

Table B4 Collinearity diagnosis									
模型	因变量	维度	条件指数	常量	方差比例				
					U_C	I_C	P_C	I_{C3}	P_{C3}
1	U_{C3}	1	1.000	0	0	0	0	0	0
		2	6.820	0	0	0	0	0	0
		3	49.510	0	0	0	0	0.01	0.03
		4	157.432	0.01	0.01	0.01	0.08	0.18	0.11
		5	374.511	0.22	0.23	0.21	0.13	0.68	0.70
		6	571.691	0.76	0.76	0.77	0.79	0.13	0.16

表 B5 2018 年 12 月 13 日供电台区二电表测量数据
Table B5 Measurement data of meters in power supply region 2 on December 13, 2018

时刻	U_A/V	I_A/A	P_A/kW	U_{A1}/V	I_{A1}/A	P_{A1}/kW	U_{A2}/V	I_{A2}/A	P_{A2}/kW	U_{A3}/V	I_{A3}/A	P_{A3}/kW
00:00	232.0	0.942	0.200	229.0	1.650	0.341	228.7	0.776	0.168	227.8	1.560	0.334
01:00	230.5	0.843	0.177	227.8	1.386	0.285	227.5	0.858	0.181	226.7	1.292	0.264
02:00	230.7	1.020	0.216	227.4	1.831	0.383	227.2	0.883	0.183	226.2	1.545	0.316
03:00	229.6	1.056	0.223	226.2	1.693	0.350	226.0	1.352	0.289	224.9	1.522	0.311
04:00	232.7	1.577	0.336	227.7	2.902	0.616	227.3	1.583	0.324	225.8	2.205	0.453
05:00	228.8	1.392	0.293	224.4	2.275	0.482	224.1	1.621	0.326	222.7	1.878	0.385
06:00	226.4	1.215	0.254	222.6	1.726	0.364	222.3	1.726	0.355	221.0	1.502	0.298
07:00	228.2	1.826	0.381	222.4	3.124	0.635	221.9	1.515	0.315	220.1	2.473	0.508
08:00	228.9	1.556	0.324	223.9	2.715	0.561	223.5	1.285	0.263	221.9	2.294	0.480
09:00	226.7	1.339	0.279	222.5	2.334	0.491	222.1	1.346	0.268	220.8	1.899	0.390
10:00	229.1	1.706	0.356	223.7	3.111	0.640	223.3	1.477	0.302	221.6	2.481	0.518
11:00	229.4	1.521	0.321	224.6	2.608	0.545	224.2	1.666	0.353	222.7	2.145	0.438
12:00	229.0	1.232	0.259	225.1	1.716	0.366	224.8	1.857	0.373	223.6	1.526	0.317
13:00	230.7	1.761	0.372	225.2	2.854	0.587	224.7	1.623	0.340	223.0	2.405	0.492
14:00	229.6	1.243	0.261	225.6	1.698	0.352	225.3	1.950	0.411	224.1	1.540	0.316
15:00	224.8	1.048	0.216	221.4	1.433	0.294	221.2	1.782	0.366	220.2	1.242	0.249
16:00	230.9	1.906	0.401	224.8	3.833	0.780	224.3	1.790	0.375	222.6	3.003	0.614
17:00	228.6	1.412	0.295	224.1	2.505	0.528	223.8	1.562	0.327	222.4	2.162	0.445
18:00	227.8	1.076	0.224	224.4	1.553	0.316	224.1	1.663	0.346	223.1	1.449	0.301
19:00	232.4	1.333	0.284	228.1	2.260	0.476	227.8	0.872	0.187	226.5	1.908	0.408
20:00	230.4	1.709	0.357	224.9	2.889	0.602	224.5	0.770	0.157	222.7	2.236	0.461
21:00	229.7	1.640	0.345	224.6	2.406	0.493	224.1	0.595	0.122	222.5	2.001	0.417
22:00	231.0	2.777	0.576	222.0	6.002	1.191	221.3	0.510	0.105	218.5	3.968	0.788
23:00	229.0	1.723	0.357	223.4	2.375	0.475	223.0	0.478	0.096	221.2	1.948	0.391
时刻	U_B/V	I_B/A	P_B/kW	U_{B1}/V	I_{B1}/A	P_{B1}/kW	U_{B2}/V	I_{B2}/A	P_{B2}/kW	U_{B3}/V	I_{B3}/A	P_{B3}/kW
00:00	231.7	0.810	0.169	227.5	0.736	0.151	227.9	1.495	0.306	227.7	0.457	0.096
01:00	230.4	0.779	0.167	226.6	0.829	0.169	226.9	1.356	0.291	226.8	0.435	0.092
02:00	230.3	0.926	0.197	225.7	0.903	0.191	226.1	1.596	0.337	225.9	0.673	0.138
03:00	229.1	1.036	0.216	223.8	1.308	0.271	224.3	1.498	0.301	224.1	0.964	0.199
04:00	232.9	1.333	0.286	226.3	1.658	0.355	226.9	2.21	0.455	226.6	0.489	0.101
05:00	227.9	1.594	0.329	219.7	1.603	0.319	220.5	1.909	0.398	220.1	2.313	0.461
06:00	225.4	1.572	0.325	217.5	1.692	0.341	218.2	1.562	0.323	217.9	2.511	0.499
07:00	227.6	1.899	0.395	218.1	1.454	0.29	219.0	2.365	0.465	218.6	2.187	0.450
08:00	227.7	1.662	0.339	219.0	1.231	0.241	219.8	2.241	0.458	219.4	2.394	0.471
09:00	226.2	1.378	0.288	219.3	1.374	0.28	220.0	1.926	0.401	219.7	1.583	0.324
10:00	228.8	1.586	0.329	220.7	1.455	0.292	221.4	2.426	0.495	221.1	1.502	0.309
11:00	228.6	1.465	0.300	220.9	1.606	0.328	221.7	2.088	0.415	221.4	1.327	0.268
12:00	228.7	1.352	0.284	221.9	1.898	0.389	222.6	1.466	0.293	222.3	1.273	0.266
13:00	230.9	1.665	0.357	222.7	1.613	0.336	223.5	2.482	0.524	223.2	0.777	0.162
14:00	229.4	1.287	0.269	222.8	1.936	0.405	223.5	1.507	0.303	223.2	0.817	0.167
15:00	224.7	1.065	0.217	219.1	1.739	0.348	219.7	1.256	0.254	219.5	0.661	0.133
16:00	231.3	1.569	0.336	223.6	1.807	0.383	224.3	3.091	0.650	224.0	0.743	0.155

17:00	228.3	1.317	0.275	221.5	1.564	0.328	222.2	2.138	0.435	221.9	0.967	0.192
18:00	227.7	1.063	0.219	222.2	1.645	0.339	222.8	1.401	0.281	222.5	0.671	0.137
19:00	231.9	1.352	0.285	225.0	0.854	0.179	225.6	1.863	0.389	225.3	1.598	0.327
20:00	229.6	1.855	0.383	220.0	0.774	0.159	220.9	2.152	0.427	220.5	2.510	0.498
21:00	228.7	2.014	0.418	218.5	0.592	0.121	219.4	1.939	0.392	219.0	3.119	0.624
22:00	229.8	2.624	0.536	216.1	0.495	0.099	217.3	3.917	0.768	216.8	3.624	0.723
23:00	228.4	1.964	0.410	218.6	0.493	0.102	219.5	1.931	0.384	219.1	2.087	0.428
时刻	U_C/V	I_C/A	P_C/kW	U_{C1}/V	I_{C1}/A	P_{C1}/kW	U_{C2}/V	I_{C2}/A	P_{C2}/kW	U_{C3}/V	I_{C3}/A	P_{C3}/kW
0:00	230.4	1.226	0.258	225.9	1.660	0.345	225.4	1.526	0.319	224.6	1.439	0.301
1:00	229.6	1.028	0.215	225.7	1.431	0.304	225.4	1.296	0.266	224.7	1.208	0.245
2:00	229.3	1.332	0.285	224.5	1.851	0.391	224.1	1.604	0.340	223.2	1.591	0.334
3:00	228.6	1.262	0.265	223.9	1.680	0.345	223.5	1.555	0.324	222.6	1.537	0.317
4:00	230.8	1.897	0.397	223.7	2.803	0.574	223.0	2.191	0.447	221.7	2.278	0.457
5:00	228.3	1.551	0.321	222.4	2.183	0.444	221.9	1.916	0.401	220.8	1.789	0.356
6:00	226.9	1.245	0.261	222.4	1.719	0.361	221.9	1.560	0.322	221.1	1.477	0.304
7:00	227.2	2.063	0.426	219.5	3.186	0.661	218.8	2.426	0.489	217.4	2.365	0.465
8:00	227.5	1.896	0.396	220.5	2.655	0.536	219.9	2.247	0.460	218.6	2.363	0.486
9:00	225.6	1.597	0.330	219.7	2.260	0.460	219.2	1.823	0.359	218.1	1.929	0.391
10:00	227.2	2.076	0.427	219.4	3.071	0.623	218.6	2.367	0.471	217.2	2.570	0.516
11:00	227.9	1.827	0.386	221.3	2.607	0.545	220.7	2.185	0.454	219.5	2.067	0.419
12:00	229.1	1.234	0.256	224.4	1.675	0.349	224.0	1.470	0.294	223.2	1.532	0.317
13:00	229.8	1.964	0.412	222.5	2.907	0.609	221.8	2.336	0.464	220.5	2.311	0.481
14:00	229.6	1.236	0.258	225.0	1.645	0.331	224.6	1.527	0.311	223.7	1.478	0.306
15:00	224.7	1.039	0.212	220.8	1.404	0.282	220.4	1.251	0.252	219.7	1.265	0.249
16:00	227.7	2.476	0.504	218.3	3.799	0.766	217.3	2.86	0.557	215.7	2.744	0.535
17:00	227.2	1.722	0.361	220.9	2.473	0.514	220.3	2.182	0.453	219.1	2.012	0.405
18:00	227.4	1.143	0.234	223.1	1.544	0.313	222.6	1.400	0.281	221.9	1.392	0.279
19:00	231.1	1.585	0.331	225.2	2.165	0.436	224.6	1.816	0.370	223.5	2.022	0.418
20:00	229.6	1.919	0.396	222.3	2.778	0.557	221.6	2.193	0.443	220.3	2.431	0.491
21:00	230.1	1.653	0.341	223.8	2.367	0.477	223.1	1.949	0.396	2220	1.980	0.393
22:00	226.8	3.650	0.747	213.1	6.057	1.212	211.8	3.876	0.752	209.4	3.879	0.738
23:00	229.6	1.721	0.362	223.2	2.477	0.516	222.6	2.031	0.425	221.4	2.084	0.426

表 B6 t 检验结果
Table B6 t -test results

模型	自变量	因变量	未标准化系数	t 值	显著性
1	常量	U_{C3}	-64.447	-4.146	0
	U_C	U_{C3}	1.263	17.516	0.000
2	常量	U_{C3}	-0.832	-0.959	0.339
	U_C	U_{C3}	1.004	261.624	0
	I_C	U_{C3}	-0.159	-41.984	0
	I_{C3}	U_{C3}	-1.454	-1.868	0.063
3	常量	U_{C3}	-9.289	-6.775	0
	U_C	U_{C3}	1.042	171.599	0
	P_C	U_{C3}	-0.842	-36.309	0
	P_{C3}	U_{C3}	1.435	1.795	0.074

附录 C

为度量方法的运算量和效率,文章从事前分析估算(渐进复杂度)和事后统计 2 个角度进行分析。

角度 1: 事前分析估算。

计算算法的渐进复杂度来对比分析 3 种方法。假设矩阵 X 为 $n \times p$ 的矩阵, Y 为 $n \times 1$ 的矩阵, 其中 $p=m+1$, $n > p$, m 是自变量个数, n 是样本数量。

文中所列举方法的计算过程如下。

方法 1: 相关系数法。

相关系数的计算公式为:

$$\rho_{xy} = \frac{\text{Cov}(x, y)}{\sqrt{D(x)}\sqrt{D(y)}} = \frac{E[(x - E[x])(y - E[y])]}{\sqrt{E[(x - E[x])^2]}\sqrt{E[(y - E[y])^2]}} \quad (\text{C1})$$

算法的时间复杂度为 $O(n)$ 。

方法 2: K-Means 聚类法。

计算公式见文中式(9)。算法的时间复杂度为 $O(n)$ 。

方法 3: 多元线性回归法。

(1) 首先计算回归系数:

$$\hat{\beta} = (X^T X)^{-1} X^T Y = \left(\sum_{i=1}^n x_i x_i^T \right)^{-1} \left(\sum_{i=1}^n x_i y_i \right) \quad (\text{C2})$$

则算法的时间复杂度为:

$$\begin{array}{c} O(n \times p^2) + O(p^3) + O(p \times n) + O(p^2) = O(n \times p^2) \\ \downarrow \quad \downarrow \quad \downarrow \quad \downarrow \\ X^T X \quad \text{求逆} \quad X^T Y \quad (X^T X)^{-1} \text{和 } X^T Y \text{ 相乘} \end{array} \quad (\text{C3})$$

(2) 计算决定系数, 计算公式见原文公式(3)。算法的时间复杂度为 $O(n)$;

(3) 计算调整决定系数, 计算公式见原文公式(4)。算法的时间复杂度为 $O(1)$ 。

综上, 多元线性回归方法的算法复杂度为 $O(n \times p^2)$ 。

方法 1 和方法 2 算法时间复杂度均为 $O(n)$, 方法 3 的时间复杂度为 $O(n \times p^2)$, 比方法 1 和方法 2 的时间复杂度稍高一些, 因为方法 1 和方法 2 只需要计算电压, 而方法 3 还需要对电流进行计算, 增加了计算量。但对本文而言, 最终的判别表达式(8)中自变量的个数 $m=3$, 取 10 天共 240 组的数据进行计算。因此 $p=4$, $n=240$, $n \gg p$, 可见方法 3 的计算量并未显著提高, 但是从原文表 6 可以看出判别准确率能提高 10% 左右。

角度 2: 事后统计。

对上述 3 种方法编制程序, 在配置为 CPU 为 i5-7500, 频率为 3.4 GHz, 内存 8 GB 的台式机上用 MATLAB R2016a 运行, 运行时间如表 C1 所示。

表 C1 不同判别方法的运行时间

Table C1 Run time of different identification methods

判别方法	运行时间/s	判别方法	运行时间/s	判别方法	运行时间/s
相关系数	3.762×10^{-3}	K-Means 聚类	1.215×10^{-3}	多元线性回归	6.777×10^{-3}

其中运行时间表示确定一个单相用户相位所需的时间。

2 个角度的分析结果一致, 因此认为本文所提方法仍具有一定优势。

附录 D

IEC 精确定义的利用负序电压复数与正序电压复数的比值计算三相电压负序不平衡度, 而在实际测量中, 电压相位随时间变化, 获取难度较大, 为满足工程实际需要, 本文采用国家标准 GB/T 15543—2008《电能质量 三相电压允许不平衡》推荐的算法, 该方法只要知道相电压 U_A 、 U_B 、 U_C 的大小即可, 负序不平衡度 ε_L 计算公式如下:

$$\varepsilon_L = \sqrt{\frac{1 - \sqrt{3 - 6L}}{1 + \sqrt{3 - 6L}}} \times 100\% \quad (D1)$$

其中， $L = (U_A^4 + U_B^4 + U_C^4) / (U_A^2 + U_B^2 + U_C^2)^2$ 。

电压不平衡度仅根据一组数据点即可计算，而本文所提多元线性回归的判别方法需要一段时间的多组数据，为了将不平衡度与判别准确率相对应，本文参考前文的国家标准，取实测值的 95%概率大值作为这段时间的电压不平衡度。例如本文在计算时，样本数据为 10 d 共 240 组的数据，计算每组数据的不平衡度，按由大到小次序排列，舍弃前面 5%的大值取剩余实测值中的最大值。以台区一的数据为例计算其不平衡度，结果如图 D1 所示。

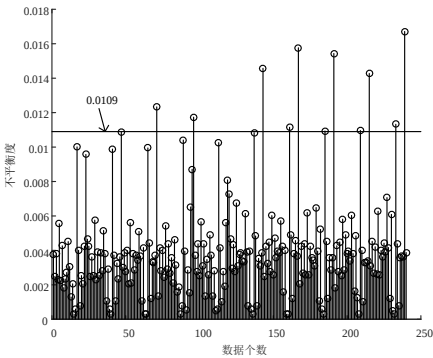


图 D1 240 组数据的不平衡度针状图

Fig.D1 Needle graph of unbalance of 240 sets of data

从图中可以看出，舍弃 $240 \times 5\% = 12$ 个大值，剩余实测值中的最大值大概为 1.09%，即可认为该台区的不平度为 1.09%。国家标准 GB/T 15543—2008《电能质量 三相电压允许不平衡》中明确规定，在正常情况下电网各级电压的三相不平衡度不大于 2%，每个用户在公共连接点引起的三相电压不平衡度不得超过 1.3%。本文从现场选取不同不平衡度台区的数据，计算本文所提方法的准确率，结果表 D1 所示。分析可知所提方法的判别准确率随着不平衡度的增加而提高，当不平衡度在 1%左右时，判别准确率能达到 90%。

表 D1 不同平衡度台区下相别判断准确率

Table D1 Phase identification accuracy rate under power supply regions of different unbalance							
台区不平衡度/%	台区用户数	正确判别数	判别准确率/%	台区不平衡度/%	台区用户数	正确判别数	判别准确率/%
0.22	95	62	65.3	1.43	86	78	90.7
0.46	69	56	81.2	1.92	82	75	91.5
1.09	77	69	89.6				

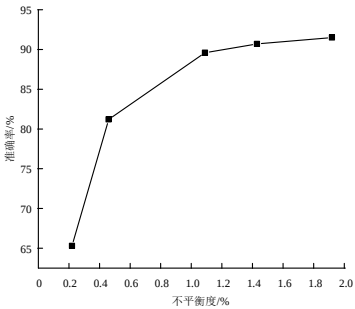


图 D2 相别判断准确率和不平衡度的关系

Fig.D2 Relationship between phase identification accuracy and unbalance