

考虑 PMU 数据质量问题的电力系统扰动检测方法

李子康¹, 刘灏^{2*}, 毕天姝², 杨奇逊²

(1. 国网上海市区供电公司, 上海市 虹口区 200080;

2. 新能源电力系统国家重点实验室(华北电力大学), 北京市 昌平区 102206)

Power System Disturbance Detection Method Considering PMU Data Quality Problems

LI Zikang¹, LIU Hao^{2*}, BI Tianshu², YANG Qixun²

(1. State Grid Shanghai Urban Electric Power Supply Company, Hongkou District, Shanghai 200080, China;

2. State Key Laboratory of Alternate Electrical Power System with Renewable Energy Sources

(North China Electric Power University), Changping District, Beijing 102206, China)

ABSTRACT: Fast and accurate power system disturbance detection can provide effective guidance information for subsequent disturbance analysis, and the wide area measurement system (WAMS) is widely used to provide a powerful data base for disturbance detection. Based on PMU measurement data, this paper proposes a disturbance event detection method considering PMU bad data. First, the behavioral characteristics of PMU abnormal data are analyzed to reveal the differential characteristics of disturbance events and PMU bad data. Furthermore, a PMU abnormal data initial screening method based on the combination of differential Teager-Kaiser energy operator and 3Sigma criterion is proposed to avoid the problems of low intensity disturbance miss detection and repeated detection of disturbances. Then, the dynamic time warping and the maximal information coefficient are used to calculate the spatio-temporal similarity among different PMUs and the correlation among different measurements within the same PMU, respectively. And it is used as features to characterize the differences of disturbance events and PMU bad data. Finally, the obtained comprehensive metrics are analyzed by a local outlier probability algorithm to achieve accurate detection of disturbance events in scenarios containing PMU bad data. Based on the IEEE 39 system, the actual grid model and the filed PMU data, the proposed method is verified to have good accuracy, real-time and generalization capability.

KEY WORDS: synchronous phasor measurement; disturbance detection; data quality problems; bad data

摘要: 快速准确的电力系统扰动检测能够为后续扰动分析提

供有效的指导信息,而广域测量系统(wide area measurement system, WAMS)的广泛应用为扰动检测提供了有力的数据基础。基于 PMU 量测数据,该文提出一种考虑 PMU 不良数据的扰动事件检测方法。首先分析 PMU 异常数据行为特性,揭示扰动事件与不良数据的差异性特征。进一步,提出一种基于差分 Teager-Kaiser 能量算子与 3Sigma 准则相结合的 PMU 异常数据初筛方法,避免了低强度扰动漏检和扰动的重复检测问题。接着,利用动态时间规整和最大互信息系数分别计算不同 PMU 间的时空相似性,以及同一台 PMU 内不同量测间的相关性,并以此作为表征扰动事件和不良数据差异的特征。最后,通过局部离群概率算法对得到的综合度量指标进行分析,可实现在含有不良数据场景下的扰动事件准确检测。基于 IEEE 39 系统,实际电网模型以及 PMU 实测数据,验证所提方法具有较好准确性、实时性以及泛化能力。

关键词: 同步相量测量; 扰动检测; 数据质量问题; 不良数据

0 引言

随着系统复杂程度的增加与系统规模的增长,电力系统安全稳定问题愈发突出^[1-5]。实时准确的电力系统扰动检测能够为系统状态监测提供有效的指导信息。目前,基于 PMU 量测数据的电力系统扰动检测问题已被广泛研究^[6-14]。

数据驱动方法主要分为指标法、传统机器学习以及深度学习法。对于指标法,文献[6]利用经验模态分解的固有模函数以及其谱峰度表来检测电力系统中的事件。但是需要进行复杂而耗时的数学变换,算法实时性较差。文献[7]利用递归量化分析法(recurrence quantification analysis, RQA)对 PMU 量测数据进行非线性分析,通过复现率(recurrence

基金项目: 国家自然科学基金项目(51725702)。

Project Supported by National Natural Science Foundation of China (51725702).

rate, RR)和确定性(determinism, DET)来表征系统中扰动事件发生数量,从实现扰动检测。文献[8-9]提出一种基于最小封闭椭球(minimum volume enclosing ellipsoid, MVEE)的电力系统事件检测方法,通过使用最小椭球的体积变化率、中心、偏心率等指标来检测系统扰动事件。然而,寻找MVEE首先需要解决一个复杂的数学优化问题,这限制了MVEE算法在大容量电力系统中的实时应用。文献[10]提出一种基于Teager-Kaiser能量算子(Teager-Kaiser energy operator, TKEO)的多事件检测方法。它对信号的瞬态变化有很强的响应,但对扰动后续短时间内的机电振荡也较为敏感,从而容易导致事件的重复检测。

对于传统机器学习方法,文献[11]通过主成分分析对电力系统进行线性化近似,并定义一种用于电力系统事件检测的指标,从而实现电力系统事件的实时检测;文献[12]通过引入移动时间窗,提出一种基于主成分分析法(principal components analysis, PCA)的多重电网事件检测方法。然而,PCA会将非线性时变变量的相互关系线性化,这将导致事件检测结果的精确性较低。为了考虑电力系统的非线性;文献[13-14]分别提出核主成分分析法(kernel PCA, KPCA)和基于Nystrom理论的主成分分析法(Nystrom PCA, NPCA),它们通过构造非线性函数将PMU量测数据映射到高维空间,从而更好地表征电力系统的动态特性以实现事件检测。KPCA和NPCA虽然可以考虑非线性特性,但很难选择一个合适的非线性映射函数;文献[15]利用局部异常因子(local outlier factor, LOF)方法对电力系统中的异常事件进行检测;文献[16]提出一种集成学习方法来实现PMU数据中数据异常检测,同时能够区分不良数据与扰动事件。

为充分挖掘事件数据中的信息,一些基于深度学习的检测方法被提出。文献[17]通过将频率变化率和相对角度偏移转换为图像,并利用卷积神经网络来实现事件检测;文献[18]将最大重叠离散小波变换和叠加稀疏去噪自动编码器相结合,提高了扰动事件检测的抗噪性。但上述基于深度学习的扰动检测方法需要较长的数据窗以及计算时间。

值得注意的是,基于PMU数据的扰动检测方法严重依赖于其数据质量。然而,PMU量测数据均存在一定的数据质量问题^[19-20]。不良数据的存在,会使得现有扰动事件检测方法会发生误判,从

而严重影响其在后续扰动辨识、定位以及控制措施中的指导意义。

本文首先通过PMU实测现场数据分析扰动事件与不良数据在PMU数据异常时刻的差异性特征。并根据扰动时刻量测数据会发生瞬态变化的特点,提出基于差分Teager-Kaiser能量算子与3Sigma准则相结合的PMU异常数据初筛方法。进一步,为防止初筛结果为PMU不良数据,从而引起扰动事件误判。提出基于综合度量指标的扰动事件检测方法,利用不良数据与真实扰动数据的差异性特征构造综合度量指标,并通过局部离群概率算法实现两者地有效区分。算例结果表明,该方法能够降低PMU不良数据对电力系统扰动事件检测的影响,为后续进一步扰动分析提供有效的指导信息。

1 PMU异常数据行为特性分析

为探究扰动事件与不良数据下PMU量测数据行为特性,现对现场实测扰动数据和不良数据进行分析,总结扰动事件与PMU不良数据的差异性,为后续扰动事件的准确检测奠定基础。

示例一:西北某地区在2016年6月1日16时51分18秒发生扰动,扰动发生位置附近的多台PMU量测频率和电压幅值数据如图1所示。

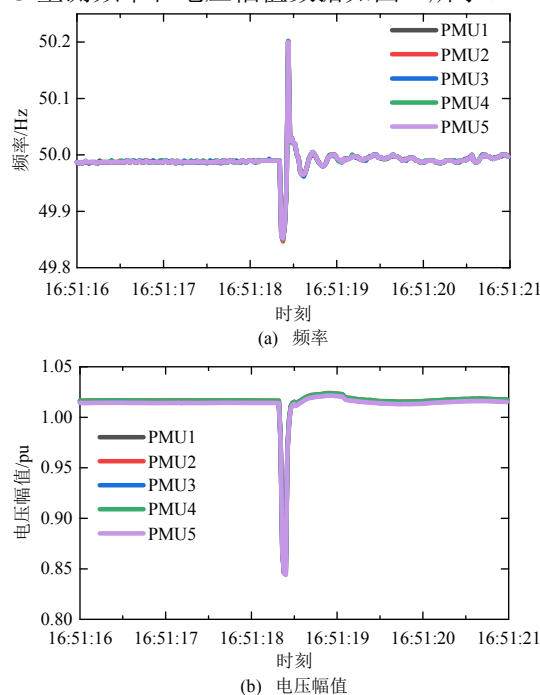


图1 PMU实测扰动数据

Fig. 1 PMU measured disturbance data

由图1可知,当发生扰动事件时,扰动发生位置附近的多台PMU量测数据都会在扰动时刻产生信号的突变,并且在扰动发生后的短时间内伴随

着机电振荡现象。

示例二: 西北某地区在 2015 年 9 月 18 日 23 时 7 分至 10 分 PMU 量测发生不良数据现象, 该地区内多台 PMU 量测频率和电压幅值数据如图 2 所示。

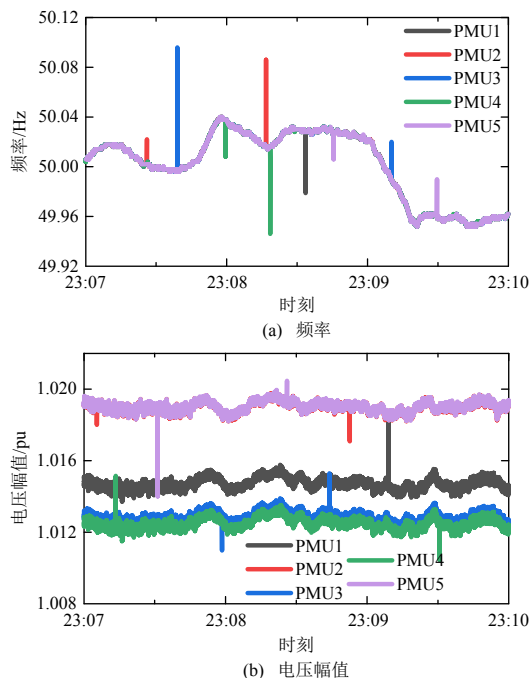


图 2 PMU 实测不良数据
Fig. 2 PMU measured bad data

由图 2 可知, 数据质量问题会导致 PMU 量测数据发生突变, 且在每一时刻, 不良数据一般只发生于单台 PMU 内的某个量测量, 并且不会对系统运行状态产生影响, 即在扰动后的短时间内不会伴随机电振荡现象。

示例三: 西北某地区在 2015 年 9 月 3 日 20 时 6 分 27 秒至 32 秒出现频率和电压幅值长时间丢失现象, 该地区内多台 PMU 量测频率和电压幅值数据如图 3 所示。

由图 3 可知, 由于通信、设备故障等原因, 导致某一时间段内变电站的相量数据集中器(phasor data concentrator, PDC)PDC 发生错误, 从而引起该 PDC 中所有 PMU 量测数据在短时间内同时发生数据丢失现象。

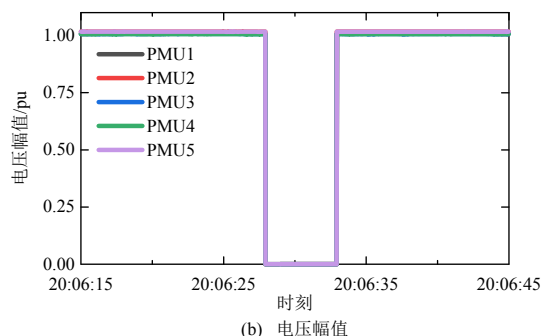
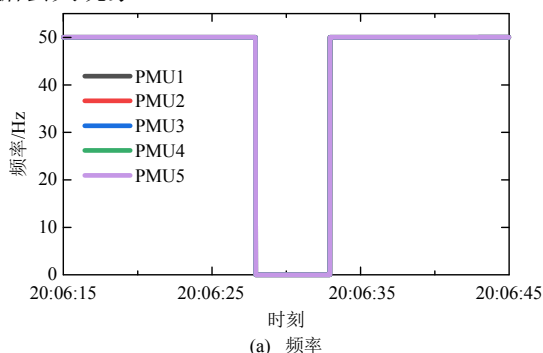


图 3 PMU 实测 PDC 错误现象
Fig. 3 PMU measured PDC error

综上所述, 本文将扰动事件, 不良数据以及 PDC 错误^[21]归结为 PMU 数据异常现象。由进一步分析可知, 扰动事件通常会引起系统功率不平衡, 并在扰动传播中使得扰动附近的多台 PMU 中的不同量测发生数据异常, 并表现出相似的波动趋势。因此, 在发生扰动时, 扰动附近的多台 PMU 间以及同一台 PMU 中的不同量测间都具有较强的时空相似性和相关性。对于不良数据而言, 由于数据质量问题是由通信或者硬件设备故障等外界因素引起的, 因此不良数据不会对系统运行状态产生影响。同时, 当某台 PMU 的某个量测量出现不良数据时, 当前时刻的多台 PMU 以及同一台 PMU 中的不同量测都表现出较弱的时空相似性和相关性。对于 PDC 错误而言, 由于多台 PMU 会在短时间内同时发生数据丢失, 因此发生 PDC 错误时多台 PMU 以及同一台 PMU 内不同量测具有极高的时空相似性与相关性。

2 基于 dTKEO-3Sigma 的异常数据初筛方法

基于上述 PMU 数据的不同异常现象差异性分析, 首先需要对 PMU 数据异常现象进行检测, 其目的是将疑似扰动事件的 PMU 异常数据初筛出来, 为后续扰动事件的进一步检测奠定基础。

2.1 基于 dTKEO 的异常特征增强方法

在检测 PMU 数据异常现象时, 希望存在一个指标使得在异常发生时刻具有极强响应, 而在准稳态或扰动后暂态过程具有微弱响应。考虑到 Teager-Kaiser 对信号瞬态变化的快速响应和机电振荡过程中的微小变化, 文献[10]提出一种基于 Teager-Kaiser 能量算子的事件检测方法。但原始的 TKEO 对扰动初期的机电振荡过程较为敏感, 量测信号的较大波动使得同一扰动事件会被重复检测, 影响扰动事件的正确判断。

针对上述问题,提出一种基于差分 TKEO 的异常特征增强方法。原始的 TKEO 算子在离散域中表示为

$$\psi(x(t)) = x^2(t) - x(t-1)x(t+1) \quad (1)$$

式中: $x(t)$ 为一个离散信号; $\psi[x(t)]$ 为其在第 t 时刻的瞬时能量响应。

当 $x(t)$ 为频率量测时, TKEO 算子能够反映每一时刻系统频率的能量响应。为了增大扰动时刻与非扰动时刻间的能量响应差异,本节引入差分算子 $x'(t) = x(t) - x(t-1)$, 代入到原始 TKEO 算子中, 可得到差分 TKEO:

$$\psi_{\text{diff}}[x(t)] = [x(t) - x(t-1)]^2 - [x(t-1) - x(t-2)][x(t+1) - x(t)] \quad (2)$$

则差分 TKEO 在扰动发生时刻 t_0 产生的瞬时脉冲响应可表示为

$$\psi_{\text{diff}}^{\text{Mom}}[f(t_0)] = [f(t_0) - f(t_0-1)]^2 - [f(t_0-1) - f(t_0-2)][f(t_0) - f(t_0-1)] \quad (3)$$

式中: $\psi_{\text{diff}}^{\text{Mom}}[f(t_0)]$ 为扰动发生时刻的差分 TKEO 响应; $f(t_0)$ 为扰动发生时刻的 PMU 量测频率; $f(t_0-2)$ 、 $f(t_0-1)$ 和 $f(t_0+1)$ 分别为扰动发生前后时刻的 PMU 量测频率。

由于稳定运行的系统在发生扰动前的频率变化 $f(t_0-2) - f(t_0-1) \approx 0$, 因此扰动发生时刻的频率量测差分 TKEO 响应 $\psi_{\text{diff}}^{\text{Mom}}[f(t_0)]$ 约为 $[f(t_0) - f(t_0-1)]^2$ 。当以 $\Delta f = \pm 0.01 \text{ Hz}$ 作为扰动检测的阈值时, $\psi_{\text{diff}}^{\text{Mom}}[f(t_0)]$ 的数量级在 10^{-4} 左右。

在扰动发生过的机电振荡阶段, 由于扰动引起的 PMU 频率信号变化近似为阻尼振荡, 即 $f(t) = A_t \cos(\Omega_t t + \phi_t)$ [22]。当在信号的采样频率远高于信号自身振荡频率的前提下, 其在一定时间内的波形可以近似地用一系列固定参数近似描述。即:

$$f'(t) = f(t) - f(t-1) = A_t [\cos(\Omega_t t + \phi_t) - \cos(\Omega_t(t-1) + \phi_t)] \quad (4)$$

此时, 在机电振荡过程的频率量测差分 TKEO 如下:

$$\psi_{\text{diff}}^{\text{Osc}}[f(t)] = 4A_t^2 \sin^2 \frac{\Omega}{2} \sin^2 \Omega \approx A_t^2 \Omega^4 \quad (5)$$

式中 A_t 和 Ω_t 分别为频率信号在机电振荡过程中的振幅和振荡频率。

电力系统中, 由于扰动引起的系统频率在 2~5 个振荡周期内, 其振幅大多降低到 $\pm 0.01 \text{ Hz}$ 以下, 而振荡频率大概在 0~2 Hz 的范围内 [22]。意味着在

机电振荡过程中 $\psi_{\text{diff}}^{\text{Osc}}[f(t)]$ 的数量级在 10^{-8} 左右。

而在机电振荡过程中, 频率量测在原始 TKEO 的能量响应约为 $A_t^2 \Omega^2$, 数量级在 10^{-6} 左右。综上所述, 有:

$$\psi_{\text{diff}}^{\text{Mom}}[f(t_0)] \approx \psi_{\text{diff}}^{\text{Mom}}[f(t_0)] \gg \psi_{\text{diff}}^{\text{Osc}}[f(t)] > \psi_{\text{diff}}^{\text{Osc}}[f(t)] \quad (6)$$

因此, 所提差分 TKEO 算子能够有效增强扰动时刻 PMU 频率量测的异常特征, 解决了低强度扰动漏检问题。此外, 所提方法能够在扰动后的机电振荡过程中获得更小的响应, 从而更有助于清晰地区分扰动发生时刻与扰动后的暂态过程, 解决了扰动检测判据的重复启动问题。

2.2 基于 3Sigma 的 PMU 异常数据检测方法

经差分 TKEO 算子映射后的 PMU 量测频率在 PMU 异常发生时刻具有极强响应, 而在准稳态或扰动后暂态过程具有微弱响应的特点。然而, 差分 TKEO 算子响应信号自身并不能实现 PMU 异常数据的检测, 因此需要额外的异常检测方法实现 PMU 异常数据检测。现有的 PMU 数据异常检测方法大都是基于人为设置的阈值实现的, 考虑到不同扰动强度的差异性, 基于阈值的异常检测方法可靠性可能会具有较大的局限性。因此, 本节利用一种能够根据自身数据分布特性的非阈值异常检测方法 3Sigma 准则, 来检测 PMU 异常数据。

3Sigma 准则是基于正态分布特性来实现异常数据的检测的。在正态分布中, σ 代表数据的标准差, μ 代表数据的均值。则有:

数值分布在 $(\mu - \sigma, \mu + \sigma)$ 中的概率为 0.682 7;

数值分布在 $(\mu - 2\sigma, \mu + 2\sigma)$ 中的概率为 0.954 5;

数值分布在 $(\mu - 3\sigma, \mu + 3\sigma)$ 中的概率为 0.997 3。

从上可以看出, 对于一个满足正态分布的数据集合而言, 正负 3 倍标准差之间的概率为 99.74%, 大于和小于 3 倍标准差总概率只有 0.26%, 因此在这范围内的任何数值都可以看为是整体数据集合里的异常值。

针对 PMU 数据异常现象检测问题, 设经差分 TKEO 算子映射后的 PMU 量测频率响应为 $\psi_{\text{diff}} = \{\psi_{\text{diff}}^1, \psi_{\text{diff}}^2, \dots, \psi_{\text{diff}}^n\}$, 其中 n 为用于实现 PMU 异常数据初筛的频率量测数据窗长。

首先对 ψ_{diff} 进行归一化处理, 使其满足正态分布:

$$\bar{\psi}_{\text{diff}} = (\psi_{\text{diff}} - \mu) / \sigma \quad (7)$$

式中: μ 和 σ 为 ψ_{diff} 的均值和标准差; $\bar{\psi}_{\text{diff}}$ 为经归一化处理后满足标准正态分布 $N(0,1)$ 的差分 TKEO 算子响应向量。基于拉依达准则, 如果某一时刻 t 满足 $|\bar{\psi}_{\text{diff}}^t| > 3$ 时, 则认为在 t 时刻的 PMU 数据存在异常现象。

3 基于综合度量指标的扰动事件检测方法

当 PMU 发生数据质量问题时, PMU 量测信号同样会产生异常现象, 导致误判。因此, 需要依据 1 节中对扰动事件及不良数据下的 PMU 量测特性, 对初筛到的 PMU 数据异常现象进一步分析, 来判断是否为扰动事件。

3.1 基于 DTW 的时空相似性度量方法

由于扰动一般会引起多台 PMU 的量测数据异常, 而不良数据往往只会造成单台 PMU 数据出现异常。因此, 发生扰动时, 不同 PMU 呈现出较强的时空相似性, 而不良数据的 PMU 时空相似性较弱。基于上述分析, 可以将 PMU 数据异常时刻与附近 PMU 量测数据间的时空相似性作为区分扰动事件与不良数据的重要依据。

为有效度量 2 台 PMU 间的时空相似性, 选取合适的度量衡指标至关重要。现有时间序列的相似性度量指标主要包括欧氏距离、余弦相似性以及动态时间规整(dynamic time warping, DTW)等。然而, 欧氏距离与余弦相似性只是简单地计算 2 个 PMU 量测序列对应时间点之间的距离, 具有较大的局限性。而 DTW 采用动态规划思想, 通过调整 PMU 序列中不同时刻对应量测之间的关系, 找到一条最优弯曲路径, 使路径的距离最小, 从而能够更好地度量 2 台 PMU 量测数据时间序列间的时空相似性^[23]。DTW 路径示意图如图 4 所示。

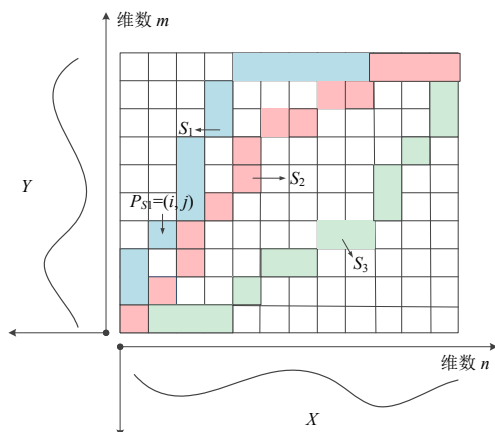


图 4 DTW 路径示意图

Fig. 4 Schematic of DTW path

给定两台 PMU 量测数据 $X=(x_1, x_2, \dots, x_n)$ 和

$Y=(y_1, y_2, \dots, y_m)$, X 、 Y 可以同为电压幅值或频率。本文以幅值为例, 构造了一个距离矩阵 $D_{n \times m}$, 其元素为

$$D(i, j) = \sqrt{(x_i - y_j)^2}, \quad i = 1, \dots, n; j = 1, \dots, m \quad (8)$$

式中: x_i, y_j 表示每个测量点的量测值; n, m 是每台 PMU 数据的测量点数。

式(8)表示 2 个 PMU 数据点 x_i 和 y_j 的欧几里德距离。矩阵 D 中每组相邻元素(图 4 中的 S_1, S_2)的集合称为弯曲路径, 需要同时满足边界性、连续性以及单调性约束。满足上述约束条件的弯曲路径有多条, 可以表示为 $S=\{S_1, S_2, \dots, S_s, \dots, S_k\}$, 其中 k 是路径中元素的总数。元素 P_{s1} 是路径 S_1 上某个点的坐标, 即 $P_{s1}=(i, j)$

DTW 将最短路径的距离来衡量两个 PMU 量测时间序列间的相似性。规整路径距离越小, 相似度越高。DTW 的最短路径距离计算方法如下:

$$D_{\text{TW}}(X, Y) = \min \sum_{s=1}^k D(S_s) \quad (9)$$

因此, 利用式(9)可以度量任意一台 PMU 数据与其他 PMU 数据在一段时间内的时空相似性, 并以此作为区分扰动事件和不良数据的重要依据。

3.2 基于 MIC 的量测间相关性度量方法

在同一台 PMU 中, 如果频率量测发生数据质量问题的, 并不意味着电压幅值, 有功功率等其他量测也会产生不良数据。因此, 在扰动发生时, 同一台 PMU 中的不同量测会具有较强的相关性。而发生数据质量问题时, 不同量测的相关性较低。基于此, 提出一种基于最大互信息系数(maximal information coefficient, MIC)的量测间相关性度量方法来表征不同量测间的相关性。

3.2.1 互信息

互信息是衡量变量之间相关程度的指标。给定变量 $A=\{a_i; i=1, 2, \dots, n\}$ 和 $B=\{b_i; i=1, 2, \dots, n\}$, n 为数据点数量, 则其互信息为

$$M_I(A, B) = \sum_{a \in A} \sum_{b \in B} p(a, b) \log \frac{p(a, b)}{p(a)p(b)} \quad (10)$$

式中: $p(a, b)$ 为 A 和 B 的联合概率密度; $p(a)$ 和 $p(b)$ 分别为 A 和 B 的边缘概率密度。

3.2.2 最大互信息系数

最大互信息系数是由 Reshef 等人于 2011 年首次在 Science 上提出, 可有效挖掘两变量间的非线性甚至非函数的依赖关系^[24]。因此, 本节利用最大

互信息系数来度量同一台 PMU 内不同量测量之间的相关性。

最大信息系数主要是通过互信息和网格划分法对其进行计算的。给定某台 PMU 的不同量测数据 $Q=(q_1, q_2, \dots, q_n)$ 和 $E=(e_1, e_2, \dots, e_n)$, Q 、 E 可以分别为频率、电压幅值。本文以频率 f 和电压幅值 U 为例, 假设 $D_{\text{pair}}=\{(f_i, U_i), i=1, 2, \dots, n\}$ 为一个有序量测数据对集合。定义 G 为将 PMU 量测 f 的值域等分成 a 段, U 的值域等分成 b 段后的网格。由于 a 和 b 不唯一, 网格 G 的划分方式不同。通过 3.1 节的互信息计算公式求出当前划分下频率 f 和电压幅值 U 间的互信息 $I(D_{\text{pair}}|G)$, 并将不同划分方式中的最大互信息值作为划分 G 的互信息值:

$$M_1^*(D_{\text{pair}}, a, b) = \max M_1(D_{\text{pair}} | G) \quad (11)$$

式中 $(D_{\text{pair}}|G)$ 表示量测数据对 D_{pair} 在网格 G 下对应的划分。则频率 f 和电压幅值 U 间的最大信息系数定义为

$$M_{\text{IC}}(f, U) = \max_{a, b < K} \frac{M_1^*(D_{\text{pair}}, a, b)}{\log \min\{a, b\}} \quad (12)$$

式中 K 为网格划分的上限值一般设置数据量 n 的 0.6 次方。

由式(12)得到的 $\text{MIC}(f, U) \in [0, 1]$, MIC 值越大则说明 PMU 数据异常时刻频率 f 和电压幅值 U 之间的相关性越强, 该异常点为扰动事件的概率越大。

3.2.3 基于综合度量的局部离群概率扰动检测方法

基于不同 PMU 间量测的时空相似性以及单台 PMU 内不同量测间的相关性构造的综合度量指标, 提出一种基于局部离群概率 (local outlier probabilities, LoOP) 的扰动事件检测方法。

局部离群因子能够通过局部密度比有效实现异常点的检测。然而对于不同的 PMU 数据异常场景, 其计算得到异常得分没有统一的取值范围, 因此在扰动检测时难以设置合理的阈值。相较于局部离群因子, 局部离群概率能够将每个样本离群程度与所有训练样本的属性相结合, 并将概率得分以异常概率的形式映射到 $[0, 1]$ 的范围内, 使其更直观的反应 PMU 量测数据在时空维度以及量测量间的离群程度, 避免了由于扰动类型与扰动强度的差异性导致阈值设置可靠性不足的问题, 从而能够有效实现扰动事件的检测^[25]。

图 5 中, 每个点表示系统中不同位置 PMU 的量测数据分布。 C_1 、 C_2 代表扰动发生时的 PMU 簇,

o_1 、 o_2 代表 PMU 不良数据。可以看到, 如果一台 PMU 的局部密度比较小, 且它周围 PMU 量测数据的局部密度比较高时, 那么该台 PMU 量测数据的离群概率越高, 即为 PMU 不良数据的概率越高。

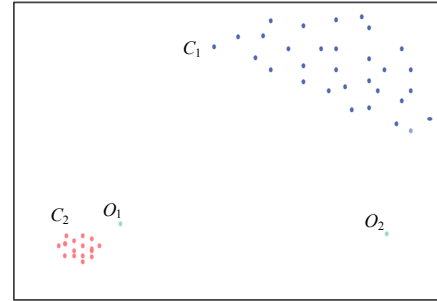


图5 局部离群概率示意图

Fig. 5 Schematic diagram of local outlier probabilities

设 $C=[c_1, c_2, c_3, \dots, c_N]$ 为 N 台 PMU 的综合度量指标向量。其中 $c_i=(d_i, m_i)$, d_i 为第 i 台 PMU 与其余 PMU 的量测频率 DTW 距离, m_i 为第 i 台 PMU 量测频率和电压幅值间的 MIC 相关性。接下来, 需要利用综合度量指标向量 C 来表征每一台 PMU 的局部密度以实现扰动数据检测。

首先通过观测第 i 台 PMU 到其近邻参考 PMU 集合 $R \in C$ 的概率距离 $p_{\text{dist}}(\lambda, i, R)$ 来反应第 i 台 PMU 的量测数据与其附近 PMU 间的相关程度, 其中概率距离 $p_{\text{dist}}(\lambda, i, R)$ 需满足如下条件:

$$\forall r \in R: P[d(i, r) \leq p_{\text{dist}}(\lambda, i, R)] \geq \varphi \quad (13)$$

式中: R 为被观测的第 i 台 PMU 的近邻参考 PMU 集合; λ 是权值系数。直观来说, P 为 R 中 PMU 量测数据被包含在以第 i 台 PMU 为球心, $p_{\text{dist}}(\lambda, i, R)$ 为半径的球中的概率。因此, $p_{\text{dist}}(\lambda, i, R)$ 能够反应第 i 台 PMU 与其参考集合 R 的局部密度。

概率距离由下式得出:

$$p_{\text{dist}}(\lambda, i, R) = \lambda \cdot \sqrt{\left[\sum_{r \in R} d(i, r)^2 \right] / |R|} \quad (14)$$

式中 $d(i, r)$ 是第 i 台和第 r 台 PMU 间的欧几里得距离。

进一步, 当以 k 台近邻 PMU 作为参考集时。以概率距离表征的概率局部离群因子 (probabilistic local outlier factor, PLOF) 定义为

$$P_{\text{LOF}, \lambda, k}(i) = \frac{p_{\text{dist}}[\lambda, i, N_{Nk}(i)]}{E_{r \in N_{Nk}(i)} [p_{\text{dist}}(\lambda, r, N_{Nk}(i))]} - 1 \quad (15)$$

式中: E 表示均值; $N_{Nk}(i)$ 表示第 i 台 PMU 的 k 近邻参考集。

最后采用高斯误差函数 erf 将概率局部离群因

子映射到范围为[0,1]的离群概率,以避免由于扰动类型与扰动强度的差异性导致异常得分阈值设置可靠性不足的问题:

$$L_{\text{OP},\lambda,k}(t) = \max \{0, e_{\text{tr}}(\frac{P_{\text{LOF},\lambda,k}(t)}{\sqrt{2 \cdot nP_{\text{LOF}}}})\} \quad (16)$$

式中 $nP_{\text{LOF}} = \lambda \cdot \sqrt{E[(P_{\text{LOF}})^2]}$ 。

通过 LoOP 映射可得到维度为 PMU 台数 n 的概率集合 P , P 中每一个元素在[0,1]范围内,代表每台 PMU 量测可能为离群点的概率。当在数据异常时刻,具有最大 dTKEO 响应的 PMU 量测数据获得的异常得分较大时,将其认为是 PMU 不良数据。

对于 PDC 错误而言,由于多台 PMU 同时发生数据丢失(量测数据均为零),多台 PMU 以及 PMU 的不同量测间具有极高的时空相似性与相关性,因此其获得的异常得分将会非常低。针对该问题,可以通过进一步计算检测数据窗内频率量测的均值大小来实现扰动事件的检测。

4 方法流程

所述方法的流程图如图 6 所示,具体步骤为:

- 1) 利用 3Sigma 准则对 N 台 PMU 的 dTKEO 响应进行分析,实现的 PMU 数据异常现象初筛;
- 2) 当发生异常时,选取最大 5 台(top-5)dTKEO 响应所对应的 PMU 量测频率和电压幅值来计算时空相似性向量以及量测相关性向量。并将上述组合得到综合度量指标向量作为 LoOP 的输入以获取具有最大 dTKEO 响应的 PMU 量测异常得分;
- 3) 当数据异常时刻具有最大 dTKEO 响应的 PMU 量测被识别为异常点时,其为不良数据。否则,通过数据窗内频率量测数据的均值来判断发生

的是扰动事件还是 PDC 错误。

5 算例分析

5.1 数据集构造

基于 IEEE 39 节点系统,利用 Digsilent 仿真软件对文献[16]中考虑的 8 类扰动进行仿真,仿真步长 20 ms,噪声等级 60 dB。在不同的可再生能源渗透率下共得到 10 576 组扰动数据,并以 0.8 和 0.2 划分为训练集和测试集。

5.2 评价指标

算法的性能通过以下指标进行评估:

- 1) 异常误检率(abnormal false discovery rate, AFDR): 实际未发生数据异常但被检测的数目/数据异常事件总数;
- 2) 异常漏检率(abnormal misdetection rate, AMDR): 实际发生数据异常但漏检的数目/数据异常事件总数;
- 3) 触发次数(trigger number, TN): 同一个异常事件的平均检测次数;
- 4) 扰动事件误检率(disturbance false discovery rate, DFDR): 实际未发生扰动事件但被检测的数目/扰动事件总数;
- 5) 扰动事件漏检率(disturbance misdetection rate, DMDR): 实际发生扰动事件但漏检的数目/扰动事件总数;
- 6) 扰动检测计算时间(check time, CT): 扰动检测算法计算时间。

5.3 数据异常检测结果

为了证明所提 PMU 异常数据初筛方法的优越性,基于可再生能源渗透率为 20%的 IEEE-39 节点系统上仿真一个多扰动事件算例,其 PMU 频率量测信号如图 7(a)所示。图 7(b)、(c)分别为该算例下

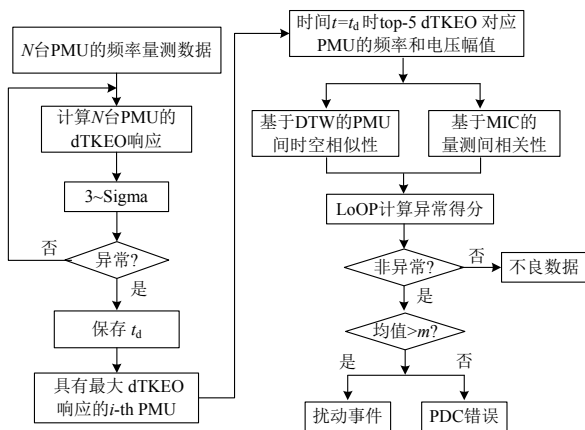
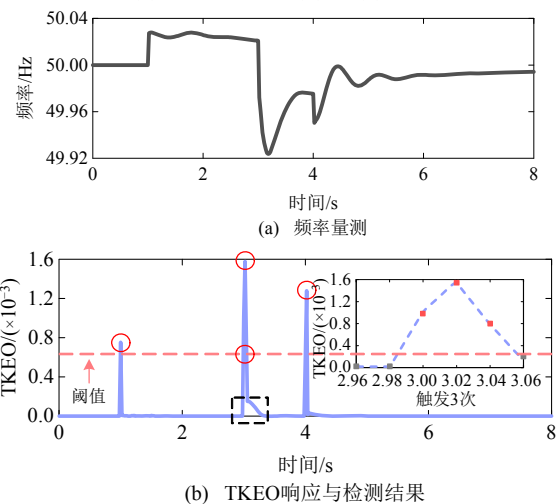


图 6 所提扰动事件检测方法流程图

Fig. 6 Flowchart of the proposed disturbance event detection method



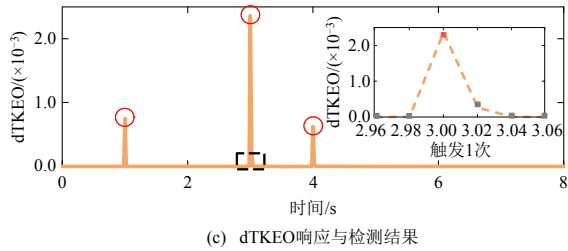


图 7 多扰动事件频率量测与不同方法的检测结果

Fig. 7 Plot of frequency and detection results of different method for the multiple disturbances

基于原始 TKEO-阈值与所提 dTKEO-3Sigma 的 PMU 数据异常检测方法的测试结果。

由图 7 可知，上述两种方法都能在 PMU 数据异常时刻获得较强响应。然而，与 TKEO 相比，dTKEO 响应能够有效屏蔽扰动初期量测信号较大波动的影响，避免了重复启动检测判据。此外，基于 3Sigma 准则的异常检测方法能够根据异常时刻的数据自身分布特性来检测异常，避免了文献[10]中的认为设置阈值造成的检测算法可靠性较低的问题，降低了漏检率。

为进一步突出所提 dTKEO-3Sigma 异常数据检测方法优势，表 1 显示了不同方法在不同可再生能源渗透率下的 PMU 异常数据检测结果。由表 1 可知，相较于基于 TKEO-阈值的检测方法，所提方法的检测结果具有更低的漏检率与触发次数，保证扰动准确检测同时，避免了对同一扰动事件重复检测。

表 1 不同 PMU 数据异常检测方法的对比结果

Table 1 Comparison results of different PMU data anomaly detection methods %						
渗透率/%	TKEO-阈值			dTKEO-3Sigma		
	AFDR	AMDR	TN	AFDR	AMDR	TN
0	0.11	0.91	1.25	0.16	0.75	1.001 1
20	0.21	0.7	1.86	0.16	0.54	1.001 6
40	0.38	0.54	2.53	0.32	0.43	1.003 2
60	0.7	0.27	3.67	0.54	0.16	1.007 5

5.4 扰动事件检测结果

进一步，为表明所提方法能够获得高度可分的映射空间，在考虑数据质量问题下将所提基于综合度量的局部离群概率扰动事件检测方法对扰动事件、异常数据以及 PDC 错误 3 种 PMU 数据异常现象的异常得分分布进行分析，如图 8 所示。

由图 8 可知，扰动事件的异常得分较小，在 0.4 以下。这是因为扰动会影响多个 PMU 量测数据，此时多台 PMU 以及 PMU 的不同量测间具有较大时空相似性与相关性，因此其异常得分较低。相较于

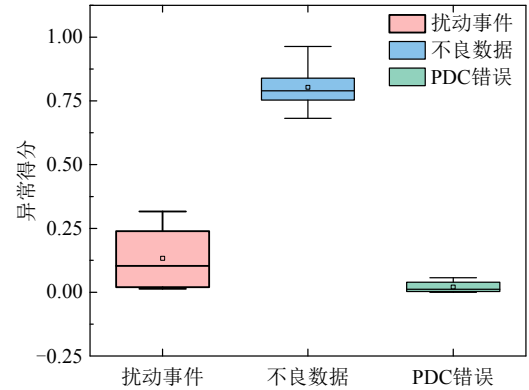


图 8 扰动事件、不良数据及 PDC 错误的异常得分分布

Fig. 8 Abnormal score distribution of disturbance events, bad data and PDC errors

扰动事件，不良数据的异常得分通常大于 0.6。这是因为不良数据是由通信或者硬件设备故障等外界因素引起的。当发生数据质量问题时，多台 PMU 间以及同一台 PMU 中的不同量测间分别表现出较弱的时空相似性和相关性，因此其获得的异常得分较高。对于 PDC 错误，由于多台 PMU 同时发生数据丢失，因此多台 PMU 以及 PMU 的不同量测间具有极高的时空相似性与相关性，其获得的异常得分非常低，在 0.1 以内。虽然，扰动事件和 PDC 错误的异常得分分布具有一定重叠，但仍能通过计算频率量测均值来实现扰动检测。为了降低扰动事件误检率，将异常得分大于 0.4 的异常现象视作 PMU 不良数据。

为进一步突出所提扰动事件检测方法对不良数据具有较高鲁棒性，将所提方法与现有扰动检测方法在不同不良数据比例下进行比较。其中 SS-LOF 为基于相似度搜索与局部离群因子相结合的检测算法^[10]，ENS 为集成学习方法的扰动检测算法^[16]，DW-DAE 为最大重叠离散小波变换与去噪自编码器相结合的检测算法^[12]，所有方法时间窗长都设置为 0.5 s。表 2、3 分别为各方法对应的扰动事件误检率以及漏检率，其中误检率偏差和漏检率偏

表 2 不同方法在增长的不良数据比例下的扰动事件误检率

Table 2 Disturbance false detection rate of different methods for increased levels of bad data %						
方法	不良数据比例					误检率偏差
	0%	5%	10%	15%	20%	
LOF	3.78	5.40	7.92	11.90	16.588	4.63
MVEE	4.51	5.83	8.51	11.61	16.08	4.17
DW+DAE	1.28	2.49	4.41	7.33	11.246	3.59
本文方法	2.11	2.29	2.62	2.96	3.47	0.49

表 3 不同方法在增长的不良数据比例下的
扰动事件漏检率

Table 3 Disturbance missing detection rate of different methods for increased levels of bad data %						
方法	不良数据比例					漏检率 偏差
	0%	5%	10%	15%	20%	
LOF	1.73	4.53	5.77	8.02	13.60	3.99
MVEE	3.04	4.85	7.28	7.84	12.80	3.30
DW+DAE	1.28	1.78	2.53	3.11	5.28	1.39
本文方法	1.67	2.08	2.42	2.77	3.25	0.54

差反映该方法对于不良数据的敏感性。

由仿真结果可知，当存在不良数据时，所提方法具有更低的误检率；而当不良数据比例大于 5% 时，所提方法具有更低的漏检率。因此，本文所提扰动检测方法在含有 PMU 不良数据的场景下时，能够有效避免不良数据引起的扰动误检问题；在具有较高 PMU 不良数据比例时，能够从扰动数据以及不良数据中精确识别出扰动发生时刻，避免扰动漏检；同时，所提方法的误检率偏差和漏检率偏差更低。

综上，所提方法对于不良数据具有更高的鲁棒性，能够在含有较高比例不良数据的 PMU 量测中准确检测扰动事件。

5.5 计算时间结果

在配置为 NVIDIA RTX3080 GPU、i9-9900k 3.6GHz CPU 和 16GB 内存的服务器上对不同方法的计算时间进行对比。由表 4 所知，所提方法具有较低的计算时间，只需要 0.115 s。因此，可以满足扰动检测的实时性需求。

表 4 不同方法的计算时间

Table 4 Computation time of various methods				
方法	SS-LOF	ENS	DW+DAE	本文方法
CT	0.326 s	0.351 s	0.037 s	0.115 s

6 实际系统算例测试

为验证所提方法的泛化性，本节以中国西部某实际区域电网模型仿真的 3 个算例对所提方法进行测试，其中新能源渗透率为 23.7%。实际区域电网单线图如图 9 所示，220 kV 及以上的变电站均安装有 PMU。

算例一：线路 750C-750D 在 5 s 发生三相短路，0.1 s 后清除故障。并按一定比例下对该算例进行损坏后的各节点量测的频率和电压信号如图 10 所示。

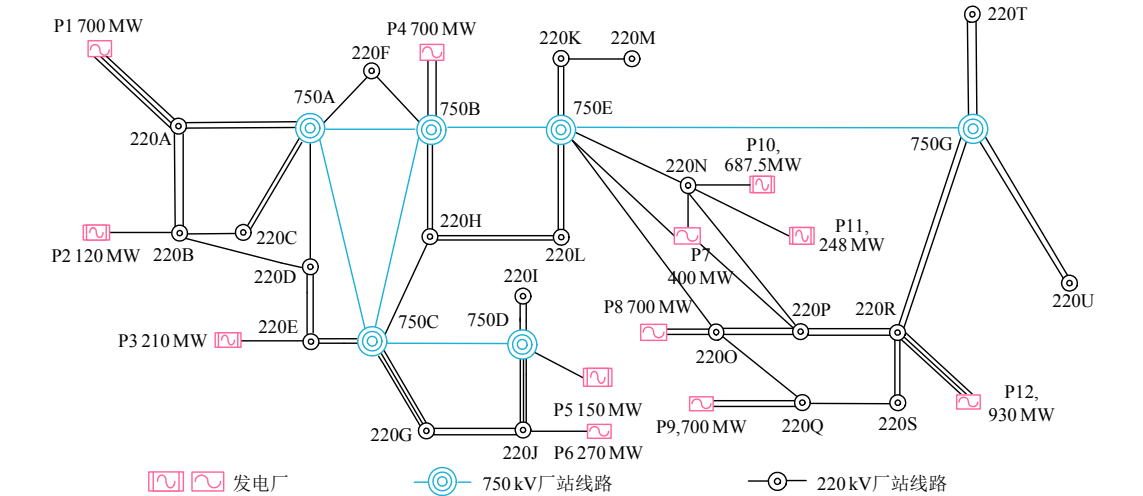


图 9 实际电网地理接线图

Fig. 9 Single line diagram of an actual power grid

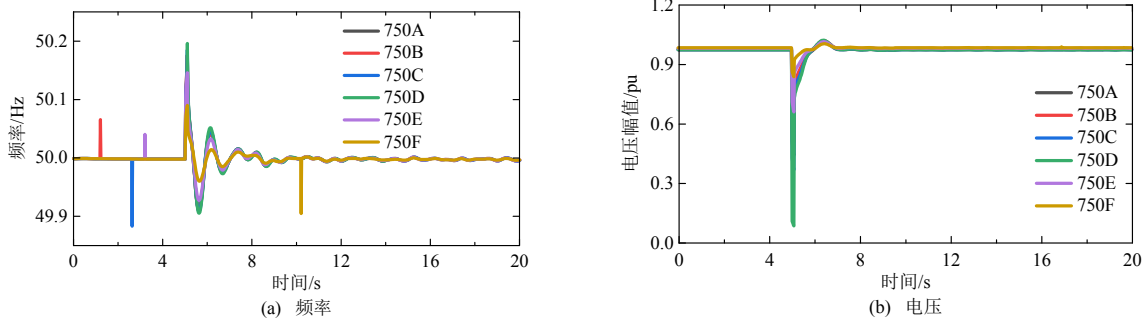


图 10 三相短路的频率和电压量测信号

Fig. 10 Frequency and voltage measurement signals of three-phase short circuit

所提 dTKEO-3Sigma 的异常检测方法验证结果如图 11 所示。由图可知, 所提检测方法能够实现频率量测信号内的所有 PMU 数据异常现象检测。

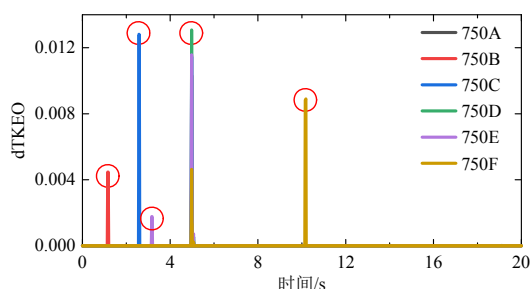


图 11 三相短路事件的 dTKEO 响应与检测结果
Fig. 11 dTKEO response and abnormal detection results of three-phase short circuit

此外, 在被检测的 5 个数据异常时刻的 LoOP 异常得分分别为 0.696 1、0.748 5、0.632 7、0.187 1 和 0.420 1, 如图 12 所示。即, 只有第 4 个异常现象被检测为扰动事件, 其余异常数据皆被所提扰动检测方法正确识别为 PMU 不良数据。

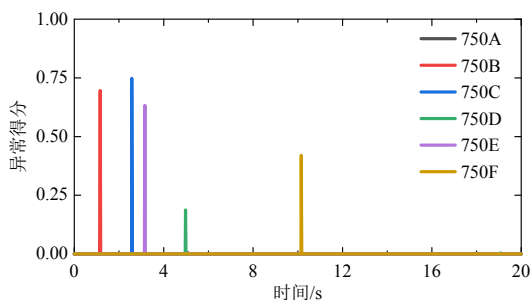


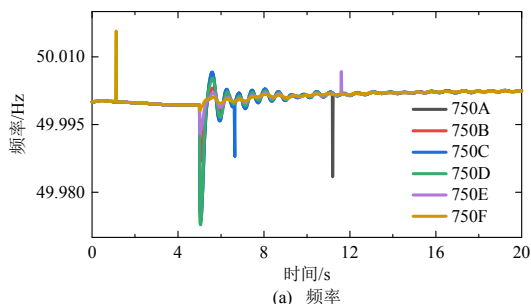
图 12 三相短路事件的异常得分结果

Fig. 12 Abnormal score result of three-phase short circuit

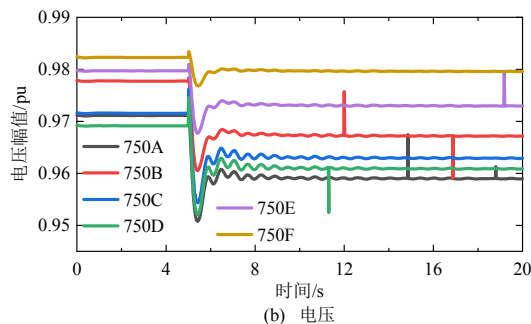
算例二: 线路 750C-750D 在 5 s 发生线路跳闸。按一定比例下对该算例进行损坏后的各节点量测的频率和电压信号如图 13 所示。

利用所提 dTKEO-3Sigma 的异常检测方法验证结果如图 14 所示。由图可知, 所提检测方法能够实现频率量测信号内的所有 PMU 数据异常现象检测。

此外, 在被检测的 5 个数据异常时刻的 LoOP 异常得分分别为 0.730 1、0.205 4、0.685 4、0.816 3 和 0.505 3, 如图 15 所示。即只有第 2 个异常现象



(a) 频率



(b) 电压

图 13 线路跳闸的频率和电压量测信号

Fig. 13 Frequency and voltage measurement signals of line tripping

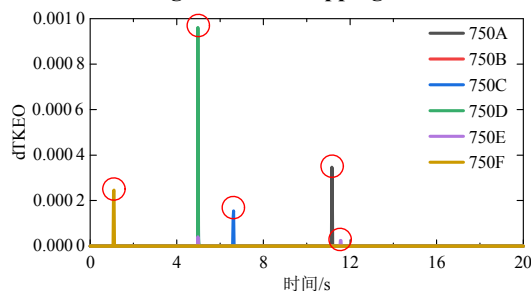


图 14 线路跳闸事件的 dTKEO 响应与检测结果

Fig. 14 Dtkeo response and abnormal detection results of line tripping

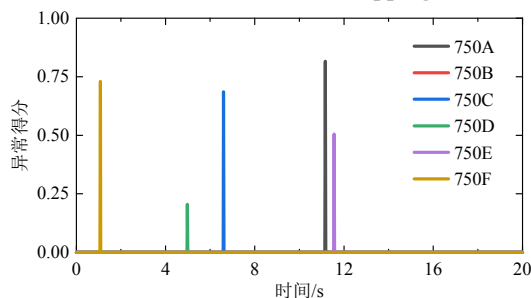


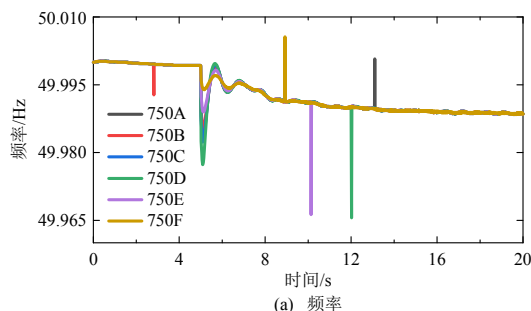
图 15 线路跳闸事件的异常得分结果

Fig. 15 Abnormal score result of line tripping

被检测为扰动事件, 其余异常数据皆被正确识别为 PMU 不良数据。

算例三: 节点 220J 在 5 s 降低发电机出力 80MW。按一定比例下对该算例进行损坏后的各节点量测的频率和电压信号如图 16 所示。

利用所提 dTKEO-3Sigma 的异常检测方法验证结果如图 17 所示。由图可知, 所提检测方法能够实现频率量测信号内的所有 PMU 数据异常现象检测。



(a) 频率

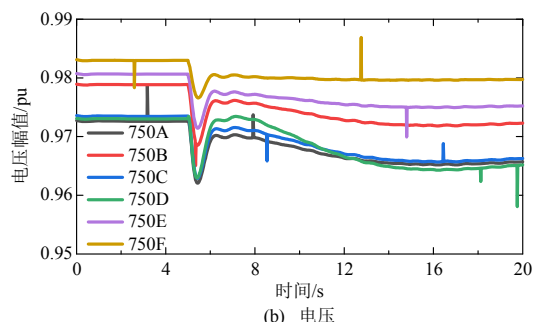


图 16 发电机出力降低的频率和电压量测信号

Fig. 16 Frequency and voltage measurement signals of generation loss

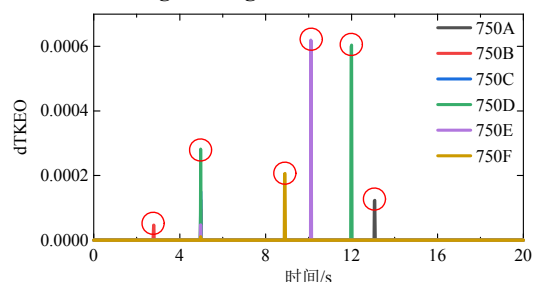


图 17 发电机出力降低事件的 dTKEO 响应与检测结果

Fig. 17 Dtkeo response and abnormal detection results of generation loss

同时, 在被检测的 6 个数据异常时刻的 LoOP 异常得分分别为 0.654 6、0.181 0、0.697 7、0.759 6、0.772 1 和 0.671 9, 如图 18 所示。即, 只有第 2 个异常现象被检测为扰动事件, 其余异常数据皆被所提扰动检测方法正确识别为 PMU 不良数据。

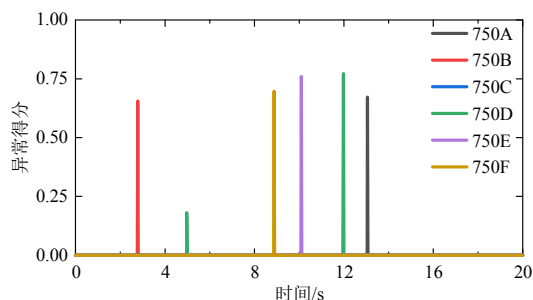


图 18 发电机出力降低事件的异常得分结果

Fig. 18 Abnormal score result of generation loss

由以上算例的仿真结果说明, 所提 dTKEO 的异常特征增强方法能够清晰的区分 PMU 异常数据与其他正常状态, 从而为后续的异常检测打下坚实的基础。此外, 由不同异常时刻下的异常得分可以看出, 由所提基于综合度量指标的扰动事件检测方法能够实现扰动的精确检测, 对不良数据具有较高的鲁棒性的同时, 也具有较好的泛化性。

为进一步验证所提方法实时性, 利用该区域电网对所提方法进行测试。该实际区域电网占中国国土面积 11%, 并且 220 kV 及以上的变电站均安装

有 PMU。在 i9-9900k 3.6GHz CPU、NVIDIA RTX3080 GPU 和 16GB 内存的计算机实验平台下, 所提扰动检测方法的计算时间为 0.126 s, 因此认为所提方法在大规模电网中仍具有较高的实时性及有效性。

7 PMU 实测数据验证

为进一步验证所提方法适用性, 利用中国西部地区的实测数据来对所提方法进行测试, 如图 19 所示。

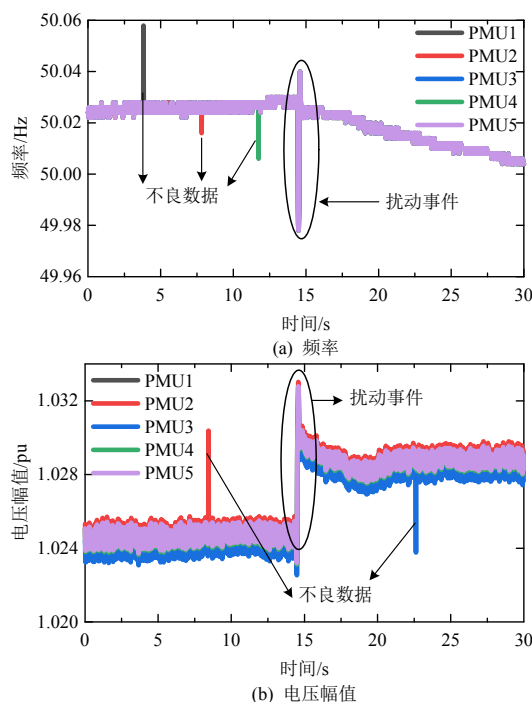


图 19 PMU 现场数据

Fig. 19 PMU field data

利用所提 dTKEO-3Sigma 的异常检测方法验证结果如图 20 所示。由对 dTKEO 信号的 3Sigma 响应可知, 所提检测方法能够实现实测 PMU 数据异常现象的检测。

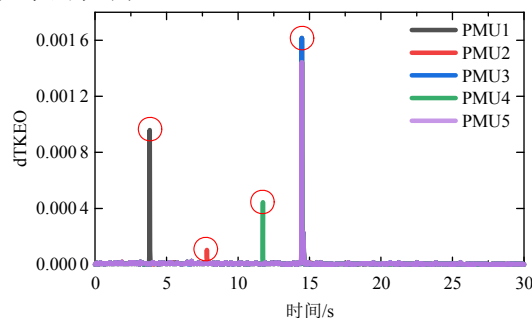


图 20 PMU 实测数据的 dTKEO 响应与异常检测结果

Fig. 20 Dtkeo response and abnormal detection results of PMU real data

在被检测的 4 个数据异常时刻的 LoOP 异常得分分别为 0.832 2、0.753 2、0.607 和 0.195 3, 如

图 21 所示。即, 只有第 4 个异常现象被检测为扰动事件。因此所提方法能够实现现场实测数据中扰动事件的准确检测, 并具有较强的泛化性。

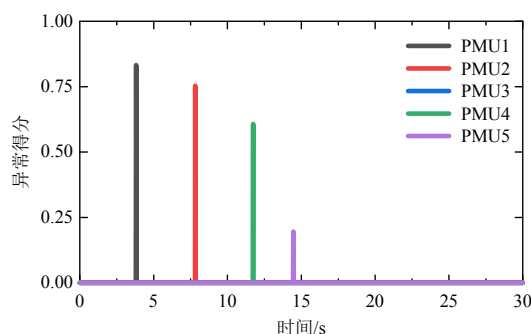


图 21 PMU 实测数据的异常得分结果

Fig. 21 Abnormal score result of PMU real data

8 结论

本文提出一种考虑 PMU 不良数据的电力系统扰动检测方法。该方法利用差分 Teager-Kaiser 能量算子来实时检测 PMU 数据异常事件。然后通过局部离群概率算法对由基于不同 PMU 间的时空相似性以及同一台 PMU 内的量测间相关性差异构成的综合度量指标进行分析, 实现扰动事件的准确检测。由 IEEE 39 节点系统、实际电网模型和 PMU 现场数据的测试结果表明, 所提方法具有较高的扰动检测精度和计算效率, 同时能够在含有较高不良数据比例下实现扰动的准确检测。

参考文献

- [1] 李琳, 冀鲁豫, 张一驰, 等. 巴基斯坦“1·9”大停电事故初步分析及启示[J]. 电网技术, 2022, 46(2): 655-661.
LI Lin, JI Luyu, ZHANG Yichi, et al. Preliminary analysis and lessons of blackout in Pakistan power grid on January 9, 2021[J]. Power System Technology, 2022, 46(2): 655-661(in Chinese).
- [2] 常忠蛟, 刘云. 巴西“3·21”大停电后电网恢复情况分析[J]. 电网技术, 2021, 45(3): 1078-1088.
CHANG Zhongjiao, LIU Yun. Analysis on Brazilian power grid restoration after “March 21” blackout[J]. Power System Technology, 2021, 45(3): 1078-1088(in Chinese).
- [3] 刘云. 巴西“9·13”远西北电网解列及停电事故分析及启示[J]. 中国电机工程学报, 2018, 38(11): 3204-3213.
LIU Yun. Analysis on and inspiration of the “9·13” islanding and outage of Brazilian remote northwest power grid[J]. Proceedings of the CSEE, 2018, 38(11): 3204-3213(in Chinese).
- [4] 李子康, 刘灏, 毕天姝, 等. 基于数据驱动的鲁棒性电

力系统扰动识别[J]. 中国电机工程学报, 2021, 41(21): 7261-7274.

LI Zikang, LIU Hao, BI Tianshu, et al. Data-driven robust power system disturbance identification[J]. Proceedings of the CSEE, 2021, 41(21): 7261-7274(in Chinese).

- [5] 秦晓辉, 毕天姝, 杨奇逊. 基于广域同步量测的电力系统扰动识别与定位方法[J]. 电网技术, 2009, 33(12): 41-48.
QIN Xiaohui, BI Tianshu, YANG Qixun. Study on WAMS based power system disturbance identification and location approach[J]. Power System Technology, 2009, 33(12): 41-48(in Chinese).
- [6] NEGI S S, KISHOR N, UHLEN K, et al. Event detection and its signal characterization in PMU data stream [J]. IEEE Transactions on Industrial Informatics, 2017, 13(6): 3108-3118.
- [7] BHUI P, SENROY N. Application of recurrence quantification analysis to power system dynamic studies [J]. IEEE Transactions on Power Systems, 2016, 31(1): 581-591.
- [8] DAHAL O P, BRAHMA S M. Preliminary work to classify the disturbance events recorded by phasor measurement units[C]//2012 IEEE Power and Energy Society General Meeting. San Diego, USA: IEEE, 2012: 1-8.
- [9] DAHAL O P, BRAHMA S M, CAO Huiping. Comprehensive clustering of disturbance events recorded by phasor measurement units[J]. IEEE Transactions on Power Delivery, 2014, 29(3): 1390-1397.
- [10] YADAV R, PRADHAN A K, KAMWA I. Real-time multiple event detection and classification in power system using signal energy transformations [J]. IEEE Transactions on Industrial Informatics, 2019, 15(3): 1521-1531.
- [11] LIU X, LAVERTY D M, BEST R J, et al. Principal component analysis of wide-area phasor measurements for islanding detection — a geometric view[J]. IEEE Transactions on Power Delivery, 2015, 30(2): 976-985.
- [12] RAFFERTY M, LIU Xueqin, LAVERTY D M, et al. Realtime multiple event detection and classification using moving window PCA[J]. IEEE Transactions on Smart Grid, 2016, 7(5): 2537-2548.
- [13] LI Zhe, KRUGER U, XIE Lei, et al. Adaptive KPCA modeling of nonlinear systems[J]. IEEE Transactions on Signal Processing, 2015, 63(9): 2364-2376.
- [14] LIU Guohong, CHEN Hong, SUN Xiaoying, et al. Low-complexity nonlinear analysis of synchrophasor

- measurements for events detection and localization [J]. IEEE Access, 2018, 6: 4982-4993.
- [15] LIU Shengyuan, ZHAO Yuanxuan, LIN Zhenzhi, et al. Data-driven event detection of power systems based on unequal-interval reduction of PMU data and local outlier factor[J]. IEEE Transactions on Smart Grid, 2020, 11(2): 1630-1643.
- [16] ZHOU Mengze, WANG Yuhui, SRIVASTAVA A K, et al. Ensemble-based algorithm for synchrophasor data anomaly detection[J]. IEEE Transactions on Smart Grid, 2019, 10(3): 2979-2988.
- [17] ZHU Yongli, LIU Chengxi, SUN Kai. Image embedding of PMU data for deep learning towards transient disturbance classification[C]//2018 IEEE International Conference on Energy Internet(ICEI). Beijing, China: IEEE, 2018: 169-174.
- [18] XIAO Fei, LU Tianguang, WU Mingli, et al. Maximal overlap discrete wavelet transform and deep learning for robust denoising and detection of power quality disturbance[J]. IET Generation, Transmission & Distribution, 2020, 14(1): 140-147.
- [19] California ISO. Five year synchrophasor plan[R]. Folsom: California ISO, 2011.
- [20] QI Wenbin. Comparison of differences between SCADA and WAMS real-time data in dispatch center[C]// Proceedings of the 12th International Workshops Electric Power Control Centers. Bedford: Springs, 2013.
- [21] KHALEDIAN E, PANDEY S, KUNDU P, et al. Real-time synchrophasor data anomaly detection and classification using isolation Forest, KMeans, and LoOP[J]. IEEE Transactions on Smart Grid, 2021, 12(3): 2378-2388.
- [22] KAMWA I, PRADHAN A K, JOÓ S G. Robust detection and analysis of power system oscillations using the Teager-Kaiser energy operator[J]. IEEE Transactions on Power Systems, 2011, 26(1): 323-333.
- [23] ITAKURA K. Minimum prediction residual principle applied to speech recognition[J]. IEEE Transactions on Acoustics, Speech, and Signal Processing, 1975, 23(1): 67-72.
- [24] RESHEF D N, RESHEF Y A, FINUCANE H K, et al. Detecting novel associations in large data sets [J]. Science, 2011, 334(6062): 1518-1524.
- [25] KRIEGEL H P, KRÖGER P, SCHUBERT E, et al. LoOP: local outlier probabilities[C]//Proceedings of the 18th ACM Conference on Information and Knowledge Management. Hong Kong, China: ACM, 2009: 1649-1652.



李子康

在线出版日期: 2023-05-18。

收稿日期: 2022-08-03。

作者简介:

李子康(1993), 男, 博士研究生, 主要研究方向为基于数据驱动的电力系统扰动辨识, zkli@ncepu.edu.cn;

*通信作者, 刘灏(1985), 男, 副教授, 主要研究方向为同步相量测量技术, hliu@ncepu.edu.cn;

毕天姝(1973), 女, 教授, 博士生导师, 主要研究方向为电力系统保护与控制、广域同步相量测量技术及应用, tsbi@ncepu.edu.cn。

(责任编辑 邱丽萍)

Power System Disturbance Detection Method Considering PMU Data Quality Problems

LI Zikang¹, LIU Hao^{2*}, BI Tianshu², YANG Qixun²

(1. State Key Laboratory of Alternate Electrical Power System with Renewable Energy Sources (North China Electric Power University); 2. State Grid Shanghai Urban Electric Power Supply Company)

KEY WORDS: synchronous phasor measurement; disturbance detection; data quality problems; bad data

Fast and accurate power system disturbance detection can provide effective guidance information for subsequent disturbance analysis, and the wide area measurement system (WAMS) is widely used to provide a powerful data base for disturbance detection. However, the existence of PMU bad data will make the existing disturbance event detection methods misjudge, which will seriously affect their guiding significance in subsequent disturbance identification, location and control measures.

To solve this problem, a disturbance event detection method considering PMU bad data is proposed. Firstly, the behavioral characteristics of PMU abnormal data are analyzed to reveal the differential characteristics of disturbance events and PMU bad data. Furthermore, a PMU abnormal data initial screening method based on the combination of differential Teager-Kaiser energy operator and 3Sigma criterion is proposed to avoid the problems of low intensity disturbance miss detection and repeated detection of disturbances. Then, the dynamic time warping and the maximal information coefficient are used to calculate the spatio-temporal similarity among different PMUs and the correlation among different measurements within the same PMU, respectively. And use it as features to characterize the differences of disturbance events and PMU bad data. Finally, the obtained comprehensive metrics are analyzed by a local outlier probability algorithm to achieve accurate detection of disturbance events in scenarios containing PMU bad data. The detailed steps of the proposed method are as follows.

Step 1: Analysis of the dTKEO responses of N PMUs using the 3Sigma criterion, and the initial screening of PMU data anomalies is realized.

Step 2: When anomalies occur, the PMU measurement frequency and voltage amplitude corresponding to the maximum 5 (top-5) dTKEO responses are selected to calculate the spatio-temporal similarity vector and the measurement correlation vector. The above combination is used as the input to LoOP to

obtain the PMU measurement anomaly score with the maximum dTKEO response.

Step 3: If the PMU measurement with the maximum dTKEO response at the moment of data anomaly is identified as anomaly, it is bad data. Otherwise, the mean value of the frequency measurement data in the data window is used to determine whether a disturbance event or a PDC error occurs.

Based on IEEE 39 bus system, eight kinds of disturbances related to active power are analyzed. It includes three-phase short circuit (3- ϕ Flt), single-phase ground fault (ϕ -g Flt), generation loss (GL), load switch off (L-off), load switch on (L-on), shut capacitor switch off (SC-off), shut capacitor switch on (SC-on) and line trip (LT).

Table 1 and 2 show the false detection rate and missed detection rate of the different methods under different levels of bad data. From the table, compared with other methods, the proposed method has a lower false detection rate (FDR) when bad data are present, and a lower miss detection rate (MDR) when the percentage of bad data is greater than 5%.

Table 1 Disturbance false detection rate (%) of different methods for increased levels of bad data

Methods	Bad data levels					FDR
	0%	5%	10%	15%	20%	Deviation
LOF	3.78	5.40	7.92	11.90	16.588	4.63
MVEE	4.51	5.83	8.51	11.61	16.08	4.17
DW+DAE	1.28	2.49	4.41	7.33	11.246	3.59
Proposed	98.34	98.16	97.83	97.48	96.97	0.49

Table 2 Disturbance missing detection rate (%) of different methods for increased levels of bad data

Methods	Bad data levels					MDR
	0%	5%	10%	15%	20%	Deviation
LOF	1.73	4.53	5.77	8.02	13.60	3.99
MVEE	3.04	4.85	7.28	7.84	12.80	3.30
DW+DAE	1.28	1.78	2.53	3.11	5.28	1.39
Proposed	1.67	2.08	2.42	2.77	3.25	0.54

Therefore, the proposed method has higher robustness to bad data and can accurately detect disturbance events in PMU measurements that contain a high percentage of bad data.