

DOI: 10.19783/j.cnki.pspc.241523

基于 DBSCAN 的配电网同步测量坏数据检测方法

刘灏¹, 陈容¹, 毕天姝¹, 赵丹^{1,2}, 张一鸣^{1,3}

(1. 新能源电力系统全国重点实验室(华北电力大学), 北京 102206; 2. 国网保定供电分公司, 河北 保定 071051;
3. 中国华能集团清洁能源技术研究院有限公司, 北京 102209)

摘要: 配电网环境复杂, 配电网同步相量测量装置(distribution network synchronous phasor measurement unit, D-PMU)容易受到干扰而产生坏数据, 进一步影响基于测量数据的应用效果。为了提高 D-PMU 数据质量, 提出一种不依赖系统拓扑的基于密度的噪声应用空间聚类(density-based spatial clustering of applications with noise, DBSCAN)的配电网同步测量坏数据检测方法。首先利用基于密度的聚类算法 DBSCAN 进行异常数据检测。通过轮廓系数和邓恩指数对 DBSCAN 的聚类结果进行综合评价。利用麻雀搜索算法实现自适应参数调整, 解决检测时需要预先处理训练、标记数据的问题。在此基础上, 将时间序列聚类的 K-Medoids 算法和动态时间规整算法相结合, 通过衡量不同时间序列之间的相似性, 解决了 D-PMU 在电气联系较弱时对抗扰动数据与坏数据的区分问题, 增强了数据处理的准确性与噪声环境下的稳健性。仿真和实际数据的测试结果表明, 所提方法能有效区分真实扰动数据并准确识别 D-PMU 坏数据。

关键词: 坏数据检测; DBSCAN 算法; 动态时间规整; K-Medoids 算法

A DBSCAN-based bad data detection method for distribution network synchronous measurement

LIU Hao¹, CHEN Rong¹, BI Tianshu¹, ZHAO Dan^{1,2}, ZHANG Yiming^{1,3}

(1. State Key Laboratory of Alternate Electrical Power System with Renewable Energy Sources (North China Electric Power University), Beijing 102206, China; 2. State Grid Baoding Electric Power Supply Company, Baoding 071051, China; 3. Huaneng Clean Energy Research Institute, Beijing 102209, China)

Abstract: The distribution network environment is complex, and distribution network synchronous phasor measurement units (D-PMU) is susceptible to interference, resulting in bad data that further impacts applications relying on measurement data. In order to improve the data quality of D-PMU, this paper proposes a density-based spatial clustering of applications with noise (DBSCAN) based bad data detection method for distribution network synchronous measurements that does not depend on system topology. First, the DBSCAN is used for anomaly data detection. The clustering results of DBSCAN are comprehensively evaluated using the silhouette coefficient and Dunn index. To address the need for preprocessing training and labeling data during detection, the sparrow search algorithm is used for adaptive parameter adjustment. On this basis, the K-Medoids algorithm for time series clustering is combined with the dynamic time warping algorithm to measure the similarity between different time series, thus solving the difficulty of distinguishing between disturbance data and bad data in D-PMU when the electrical connection is weak. This enhances both the accuracy of data processing and the robustness under noisy environments. Simulation and real data tests show that the proposed method can effectively distinguish real disturbance data and accurately identify bad D-PMU data.

This work is supported by the Science and Technology Project of State Grid Corporation of China (No. 5100-202199530A-0-5-ZN and No. 5211DS21N013).

Key words: bad data detection; DBSCAN algorithm; dynamic time warping; K-Medoids algorithm

0 引言

长期以来, 配电网存在线路数量多、量测装置

不足及信息传输阻塞等情况^[1-2]。随着光伏、风电等分布式电源以及电动汽车等柔性负荷的接入, 配电网电力电子化趋势日益凸显, 可测、可观、可控等问题更加严重^[3-6]。

同步相量测量单元(synchrophasor measurement units, PMUs)具有实时性、快速性、高精度的特点, 是实时监测电力系统状态的有效工具。随着电力电

基金项目: 国家电网有限公司科技项目(5100-202199530A-0-5-ZN, 5211DS21N013)“新能源电力系统协同控制保护系统及应用”

子设备渗透率的增加,配电网机理特性、运行控制模式和分析方法正在经历深刻变革^[7-10]。美国田纳西大学 Yilu Liu 团队研发的频率扰动分析仪(frequency disturbance recorder, FDR)可测量配电网的电压幅值、相角和频率等电气量^[11],实现配电网层面的广域同步测量。华北电力大学研究团队根据不同场景的测量需求,搭建了源网荷全景同步测量系统(full-view synchronized measurement system, SYMS)^[12],实时收集全国各地 70 余台负荷侧同步测量装置(synchronized measurement device for load, SMD-L)上传的电压幅值、相角、频率等数据,初步实现了配电网层面的可测与可观。

然而,由于通信阻塞、卫星信号授时不稳定、装置硬件故障等原因,D-PMU 存在不同程度的数据质量问题^[13]。报告显示,北美地区约 10%~17%的 PMU 数据存在数据质量问题^[14],我国约 20%~30%的 PMU 数据存在数据质量问题^[15]。数据质量问题会导致基于 PMU 数据的状态估计、扰动辨识等应用准确性下降,甚至得到错误的计算结果,严重影响配电网的状态监测。因此,D-PMU 坏数据的实时准确检测至关重要。

目前,电力系统测量装置的坏数据检测方法主要包括模型驱动方法和数据驱动方法。传统的模型驱动方法主要基于状态估计中的不良数据检测与辨识,包括残差分析法^[16-17]、非二次准则检测法^[18-19]等。文献[17]利用线性加权最小二乘的状态估计算法,通过计算实际测量值与状态估计值之间的最大归一化残差,达到检测异常点的目的。文献[18]中非二次准则检测法利用不同方式构造目标函数,提高结果的鲁棒性和准确性。除此之外,文献[20]基于状态估计器,利用增强状态相量方法纠正系统数据偏差。文献[21]利用相量测量单元的测量数据建立支路两端之间的联系,对错误的参数进行辨识和估计。文献[22-23]以卡尔曼滤波为基础实现坏数据检测。以上模型驱动方法均需要系统拓扑和模型参数信息,而这些信息在实际配电网中是未知的,此时模型驱动方法不适用于 D-PMU 的坏数据检测。

随着人工智能技术的迅速发展,数据驱动方法受到研究人员的广泛关注。其主要包括有监督学习、无监督学习等。有监督学习需要大量经过标记的正常数据和坏数据作为训练数据。文献[24]利用递归神经网络在处理时间序列数据方面的记忆特性进行 PMU 坏数据检测。文献[25]使用长短期记忆网络对 PMU 异常数据实现分解重构和有效检测。上述方法利用神经网络强大的特征提取能力,对 PMU 坏数

据实现有效检测。但是,PMU 实测数据中出现坏数据属于小概率事件,难以构造大量带标签样本。而无监督方法不需要标记训练数据。文献[26]使用多台电气距离较近的 PMU 数据构造 Hankel 矩阵,利用 Hankel 结构的低秩性检测和修正坏数据。文献[15,27]利用 PMU 实时数据的时空相关性,采用局部离群因子(local outlier factor, LOF)方法对异常数据进行检测。文献[28]使用主成分分析法区分坏数据与事件数据,实现不良数据注入检测。上述方法计算速度较快,能够满足实时应用需求,但是需要较多电气距离相近的 PMU 测点,在配电网量测冗余不足的场景中难以应用。文献[29]利用高斯混合模型(Gaussian mixed model, GMM)对局部离群因子检测算法进行改进,实现异常数据检测阈值的准确自动选择。文献[30]对传统 K-means 聚类选择初始聚类中心的策略进行调整以提高异常检测准确率。文献[31]利用基于密度的空间聚类方法,对 PMU 不良数据进行在线检测、分类和恢复。文献[32]提出一种基于谱聚类的不良数据检测方法,该方法摆脱了耗时的离线学习,具有较强的实时性。文献[33]采用爬山-蚁群-FCM 模糊聚类算法克服传统聚类算法在选取聚类数目和聚类中心时的偶然性。上述基于聚类的坏数据检测方法实现了 PMU 坏数据的实时检测,但是在处理不同坏数据类型时存在参数适应性不强的问题。此外,配电网含有的大量噪声分量也将影响聚类检测结果。

综上,由于配电网的拓扑结构和参数难以完全获得、量测冗余不足(各 PMU 布点电气距离相对较远)、随机波动强且噪声含量大等特点,导致上述数据驱动方法在配电网场景下难以实现 D-PMU 的坏数据检测。

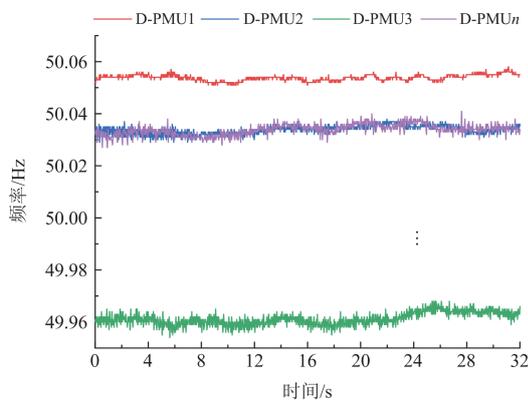
因此,本文提出了一种考虑噪声影响的适用于 D-PMU 数据的无监督自适应坏数据检测方法。利用基于密度的噪声应用空间聚类(density-based spatial clustering of applications with noise, DBSCAN)算法对异常值不敏感的特点,对 D-PMU 数据进行聚类处理,将轮廓系数和邓恩(Dunn)指数作为聚类结果的综合评价指标,采用麻雀搜索算法对最优参数进行自适应寻找。对于坏数据与扰动数据难以区分的问题,本文考虑 D-PMU 数据的时空相关性,使用动态时间规整(dynamic time warping, DTW)算法衡量两个时间序列之间的相似性,划分聚类算法的分类器选择 K-Medoids,二者结合实现时间序列分类。仿真和实测数据均表明,本文所提方法能够实时、准确检测 D-PMU 坏数据,并区分扰动数据。

1 D-PMU 数据特征

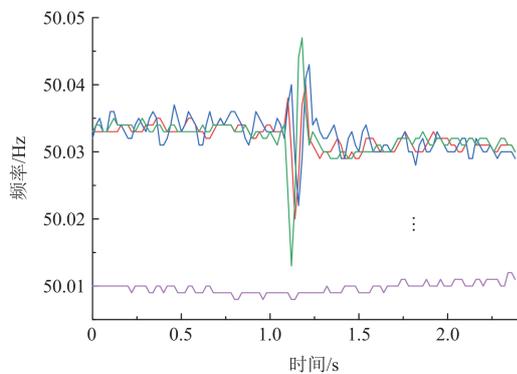
选取 SYMS 系统装置中一定时间内的 D-PMU 实测频率数据, 分别对正常数据、扰动数据和坏数据特征进行分析。

图 1(a)为 D-PMU1, D-PMU2, ..., D-PMU_n 的正常频率数据; 图 1(b)为位于同一特高压交流线路的某一 D-PMU 发生线路跳闸事件的频率数据; 图 1(c)为同一时间窗内 D-PMU2 出现坏数据。图 1(b)中出现扰动时, 扰动频率偏差在 0.02 Hz 内, 持续时间在 0.1 s 内; 图 1(c)中所示坏数据跳变幅度在 0.12 Hz 内, 持续时间在 0.08 s 内, 与前述扰动事件频率数据变化和时间变化均十分接近, 在单通道检测时很容易混淆扰动数据和坏数据, 给坏数据的检测带来困难。

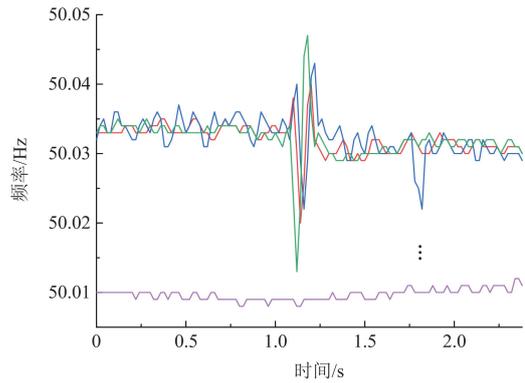
进一步分析, 图 1(b)位于同一特高压交流线路的 D-PMU1、D-PMU2 和 D-PMU3 的数据表现出相同的波动趋势, 时序相关性强; 而 D-PMU_n 与上述几台 D-PMU 位于不同特高压交流线路, 扰动发生时波形变化不一致。当位于同一特高压交流线路中某一台 D-PMU 出现坏数据时, 如图 1(c)中 D-PMU2, 其余两台 D-PMU 为正常数据, 即个别 D-PMU 出现时序异常, 与其他装置数据表现出不同



(a) 正常数据



(b) 扰动数据



(c) 坏数据

图 1 实测 D-PMU 频率正常数据、扰动数据和坏数据对比
Fig. 1 Comparison of measured D-PMU frequency normal data, disturbance data and bad data

的波动趋势, 虽然坏数据偏差较小, 但此时其他通道并未出现该时序特征, 时序相关性较差。基于上述分析可知, 结合这种时序特征, 可以实现扰动数据和坏数据的区分。

2 异常数据检测方法

2.1 基于 DBSCAN 的异常数据检测方法

传统的 PMU 坏数据检测方法大多需要提前标记样本, 对数据预先进行离线训练。DBSCAN 算法是一种基于密度聚类的数据异常检测方法。与其他聚类算法相比, 该算法不需要对样本进行训练得到先验知识、预先设置簇数, 只需要根据邻域半径 ϵ 和邻域样本数阈值 N_{MinPts} 两个参数就可以进行聚类分析, 自动识别出不属于任何簇的异常点。

利用 DBSCAN 聚类算法, 取 N_{MinPts} 为 3 时, 将 D-PMU 数据点分为 3 类: 核心点、边界点和噪声点。利用图 2 所示的某一 D-PMU 实测频率数据说明该方法。图中每个圆的半径为 ϵ , 红色点在半径 ϵ 内包含不少于 N_{MinPts} 数目的点, 红色点即为正常频率数据所在的核心点; 从红色点出发能到达蓝色点, 且

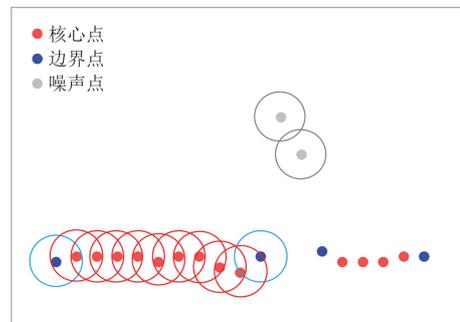


图 2 DBSCAN 原理图

Fig. 2 Schematic diagram of DBSCAN

在以蓝色点为中心的邻域内包含的点数小于 N_{MinPts} , 蓝色点即为边界点; 灰色点无法由其他点到达, 此种处于稀疏区域的异常频率数据点即为噪声点。除灰色点外, 红色点与蓝色点组成了聚类簇。

2.2 基于 DBSCAN 的异常数据检测方法

邻域半径和邻域样本数阈值两个参数的合理选择对 D-PMU 数据聚类效果尤为重要, 是 DBSCAN 聚类的难点。

配电网正常运行时, 由于负荷变化、发电机输出波动和其他因素会引起 PMU 所测数据存在正常安全范围内的波动。如图 3 所示, 在选择距离阈值参数 ε 时, 若 ε 过大, 异常点可能会落入邻域半径内; 若 ε 过小, 簇中正常波动的数据可能会丢失。参数 N_{MinPts} 决定了簇的密度, 若 N_{MinPts} 选择过大, 一些异常数据点可能会误判为正常点; 若 N_{MinPts} 选择过小, 会导致核心点变多, 影响异常数据识别。因此, 选择参数时要同时考虑这两个参数, 共同配合调参, 寻找最优聚类结果。

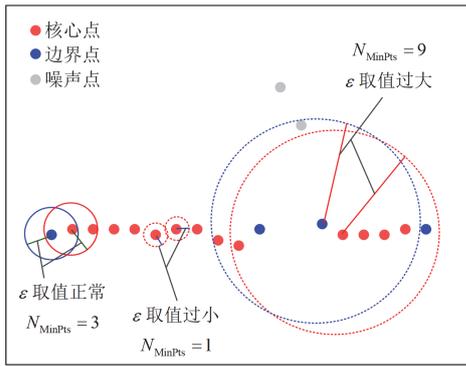


图 3 参数选择情况

Fig. 3 Parameters' selection situation

该聚类算法没有标签, 对聚类效果的评价需从簇内与簇间的关系出发。轮廓系数和 Dunn 指数是评价聚类可靠性的重要指标。轮廓系数依据簇内的凝聚程度和簇间的离散程度来评价聚类的效果, 簇内距离越小, 簇间距离越大, 聚类效果越好。轮廓系数的评价结果适用性强且易于解释, 但它假设簇是具有一定形状的, 在聚类簇形状较为复杂时可能表现不佳, 且对噪声敏感。而 Dunn 指数度量的是簇间最小距离与簇内最大距离的比值。通过比较簇间和簇内的距离, 能够较好地抵抗噪声的影响。与轮廓系数相比, Dunn 指数对簇的形状没有太多假设, 能够处理形状较复杂的簇。因此, 同时将轮廓系数和 Dunn 指数作为评价指标以全面评价聚类效果。

对于轮廓系数, 假设待检测的 D-PMU 数据集经过聚类形成簇 $C_i (i = 1, 2, \dots, k)$, 选定数据点 $i (i \in C_i)$ 。

计算簇内不相似度 a_i 和簇间不相似度 b_i 。

$$a_i = \frac{1}{|C_i - 1|} \sum_{j \in C_i, i \neq j} d_{ij} \quad (1)$$

$$b_i = \min_{k \neq i} \frac{1}{C_k} \sum_{j \in C_k} d_{ij} \quad (2)$$

式中: d_{ij} 是簇 C_i 中数据点 i 和 j 的距离。

由簇内不相似度 a_i 和簇间不相似度 b_i 可得单一数据点轮廓系数 S_i 为

$$S_i = \frac{b_i - a_i}{\max\{a_i, b_i\}} \quad (3)$$

所有数据点的轮廓系数 S 即为待检测的 D-PMU 数据集的轮廓系数。

$$S = \frac{1}{n} \sum_{i=1}^n S_i \quad (4)$$

式中: 轮廓系数 S 的取值范围为 $[-1, 1]$, 该值越接近 1 时, b_i 远大于 a_i , 该样本点与所在簇关系越密集, 异常数据识别效果越好; 当该值越接近 -1 时, a_i 远大于 b_i , 分类结果不理想。

对于 Dunn 指数, 对每个簇 C_i 计算簇内所有点之间的距离, 找到该簇内的最大距离为

$$D_1(C_i) = \max_{m, n \in C_i} d_{mn} \quad (5)$$

式中: d_{mn} 为簇 C_i 内两个点 m 和 n 的距离。

找到所有簇内最大距离中的最大值 $D_{1\max}$, 该值反映了簇内的最大紧密度。

$$D_{1\max} = \max_{0 \leq i \leq k} D_1(C_i) \quad (6)$$

对任意两个簇 C_i 和 C_j , 计算两个簇之间的最小距离。

$$D_2(C_i, C_j) = \min_{m \in C_i, n \in C_j} d'_{mn} \quad (7)$$

式中: d'_{mn} 是簇 C_i 中的点 m 与簇 C_j 中的点 n 之间的距离。

找到所有簇间最小距离中的最小值 $D_{2\min}$, 该值反映了最具分离度的簇对。

$$D_{2\min} = \min_{0 \leq i < j \leq k} D_2(C_i, C_j) \quad (8)$$

Dunn 指数 D_i 即为簇间最小距离与簇内最大距离的比值。其目标是最大化簇之间的分离度, 同时最小化簇内的紧密度。簇间最小距离越大, 簇内最大距离越小, Dunn 指数越大, 表示聚类的分离度越好, 簇内的紧凑度越高。

$$D_i = \frac{D_{2\min}}{D_{1\max}} \quad (9)$$

由于轮廓系数取值范围为 $[-1, 1]$, Dunn 指数的取值范围为大于 0, 为使轮廓系数和 Dunn 指数在同

一评价尺度, 将 Dunn 指数的取值范围归一化至轮廓系数的同一尺度。因此, 综合评价得分 M 为

$$M = S + \bar{D}_l \quad (10)$$

式中: \bar{D}_l 为归一化后的 Dunn 指数。即 M 值越大, 聚类效果越好。

将综合评价得分作为 D-PMU 的聚类指标后, 需要根据该指标进行快速的参数寻优。传统的粒子群算法、灰狼算法在面对高维复杂问题时, 容易陷入局部最优, 全局搜索能力可能较弱。而麻雀搜索算法(sparrow search algorithm, SSA)是一种具有快速收敛和较强的寻优能力的群智能优化算法^[34]。它通过模拟麻雀的觅食行为, 在全局搜索和局部搜索之间保持平衡, 并通过逃逸行为防止陷入局部最优。这种平衡机制使得麻雀搜索算法在复杂的优化问题中具有较强的适应性, 可以处理较为复杂且高维的优化问题。

对于给定的一组 ε 和 N_{MinPts} , DBSCAN 聚类完成后计算综合评价得分 M , 并以其值作为优化的目标, 目标函数形式为

$$f(\varepsilon, N_{\text{MinPts}}) = M(\varepsilon, N_{\text{MinPts}}) \quad (11)$$

通过最大化 $M(\varepsilon, N_{\text{MinPts}})$, 即可找到最优的 ε 和 N_{MinPts} 。

初始麻雀种群, 每个个体代表一个 DBSCAN 参数组合 $(\varepsilon, N_{\text{MinPts}})$, 个体随机分布在参数空间中。

$$X_u^t = (\varepsilon_u^t, N_{\text{MinPts},u}^t) \quad (12)$$

麻雀种群通过侦察者行为探索邻域范围内的解, 该过程负责局部搜索, 更新策略为

$$X_{u,v}^{t+1} = \begin{cases} X_{u,v}^t \cdot e^{-u/(T \cdot I_m)} & r_2 < S_{\text{Threshold}} \\ X_{u,v}^t + Q \cdot L & r_2 \geq S_{\text{Threshold}} \end{cases} \quad (13)$$

式中: t 为当前迭代次数; $X_{u,v}^{t+1}$ 为第 $t+1$ 次迭代中第 u 只麻雀在 v 维的位置; $T \in (0,1]$ 为控制搜索步长; I_m 为最大迭代次数; Q 为标准正态分布随机变量, 引入随机性; L 是 v 维的单位矩阵; $S_{\text{Threshold}} \in [0.5,1]$ 为安全阈值; $r_2 \in (0,1)$ 为控制行为的随机性。当 $r_2 < S_{\text{Threshold}}$ 时, 种群安全, 侦察者可以在该范围内搜索更优的适应度; $r_2 \geq S_{\text{Threshold}}$ 时, 侦察者发现危险, 调整搜索策略。

采用觅食者行为探索更广泛的参数空间, 该过程专注于全局搜索, 更新策略为

$$X_{u,v}^{t+1} = \begin{cases} Q \cdot e^{(X_{\text{worst}}^t - X_{u,v}^t)/u^2} & u > \frac{n}{2} \\ X_p^{t+1} + |X_{u,v}^t - X_p^{t+1}| A^+ L & u \leq \frac{n}{2} \end{cases} \quad (14)$$

式中: X_{worst}^t 为当前种群中全局最差位置; X_p^{t+1} 是

当前种群中的最优位置; A 矩阵的元素为随机值 1 或 -1, $A^+ = A^T(AA^T)^{-1}$ 。当 $u > n/2$ 时, 适应度较差的觅食者需要前往其他适应度更高的区域觅食; 当 $u \leq n/2$ 时, 觅食者可以在当前最优位置进行全局探索。侦察者行为的局部搜索和觅食者的全局搜索交替进行, 保证了在参数空间的广泛搜索。

在迭代计算过程中, 如果陷入局部最优, 则通过警戒者的警戒行为引导个体跳出当前局部最优解, 增强全局搜索能力。

$$X_{u,v}^{t+1} = \begin{cases} X_{\text{best}}^t + \alpha |X_{u,v}^t - X_{\text{best}}^t| & f_u > f_g \\ X_{u,v}^t + K \left(\frac{|X_{u,v}^t - X_{\text{worst}}^t|}{(f_u - f_w) + \gamma} \right) & f_u = f_g \end{cases} \quad (15)$$

式中: X_{best}^t 为当前种群中全局最佳位置; α 为均值为 0、方差为 1 的随机数, 控制步长; f_u 为麻雀个体适应度; f_g 和 f_w 分别为当前最佳和最差适应度; K 为 $[-1,1]$ 内的随机数; γ 为避免分母为 0 的最小常数。

每个个体的适应度由综合评价得分决定, 通过 DBSCAN 算法计算当前参数的综合评价得分, 作为适应度评估依据。

$$\text{Fitness}(X_u^t) = M(\varepsilon_u^t, N_{\text{MinPts},u}^t) \quad (16)$$

侦察者行为和觅食者行为通过更新策略迭代更新位置, 每次迭代后, 算法都会评估每个个体的适应度, 并更新全局最优的 ε 和 N_{MinPts} 组合。当迭代次数达到最大或适应度不再显著提高时, 停止迭代; 输出一组最优的 ε 和 N_{MinPts} , 使得综合评价得分 M 最大化, 即获得最优参数。

对原始 D-PMU 数据进行标准化预处理之后, 根据经验设定参数: 麻雀种群数量为 30; 最大迭代次数为 50; 邻域半径 ε 搜索区间为 $[0.1,4]$; 邻域样本数 N_{MinPts} 搜索范围为 $[2,30]$; 局部搜索步长调节因子 T 为 0.01; 局部搜索安全阈值 $S_{\text{Threshold}}$ 为 0.5; 随机因子 Q 为 0.1; 控制警戒者更新的步长 α 为 1.5。按上述参数优化方法, 当综合评价得分 M 最高时, 聚类效果最优。

2.3 D-PMU 数据的异常数据识别流程

本文利用 DBSCAN 对异常数据进行检测, 以轮廓系数和 Dunn 指数作为综合评价指标, 采用麻雀搜索算法得到不同参数组合的最佳得分情况, 从而实现自适应调参, 参数寻优的流程如图 4 所示。

首先, 输入待检测的 D-PMU 测量数据, 对数据进行标准化处理; 给定麻雀搜索算法需要搜索的两个参数 ε 和 N_{MinPts} 的取值范围, 初始化麻雀搜索算法; 利用 DBSCAN 对该 D-PMU 数据进行聚类,

得到综合评价得分 M , 当聚类结果最优时, 输出最优参数和检测结果。

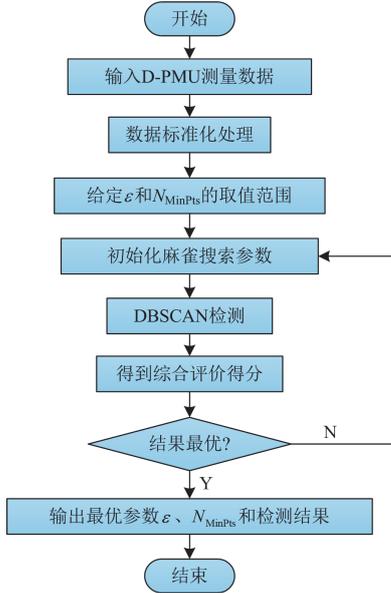


图4 采用自适应参数调整的DBSCAN检测流程图

Fig. 4 DBSCAN detection flowchart with adaptive parameter adjustment

3 基于时间序列聚类的扰动区分方法

3.1 无监督 K-Medoids + DTW 方法

前述 DBSCAN 算法能够区分出正常数据与异常数据, 而异常数据中包含扰动数据与坏数据, 为了进一步区分 D-PMU 中的扰动数据与坏数据, 提出基于时间序列的无监督 K-Medoids + DTW 扰动区分方法。

由于 D-PMU 电气联系较弱, 且配电网拓扑和扰动发生位置未知, 扰动发生时并不是所有 D-PMU 都能明显捕捉到信息, 使得 D-PMU 数据往往包含噪声和不完整信息; D-PMU 数据是按时间顺序排列的一组数值序列, 挖掘 D-PMU 数据的时序特征, 能更好地揭示其中潜在的真实信息, 为后续分析和应用提供支持。基于上述特点, 聚类作为一种常用的无监督数据挖掘方法, 能够通过分析时序特征, 将相似的数据点分组; 同时针对噪声和不完整信息, 增强对数据分布的理解, 提供更稳健的异常数据检测能力, 因此, 在处理 D-PMU 数据时尤其有效。然而, 与常规的聚类方法不同, D-PMU 数据具有强烈的时间顺序属性, 需要采用能处理时间序列的聚类技术。DTW 算法常用于评估两个时间序列的相似性, 通常结合分类器来进行时间序列分类; K-Medoids 因其在处理数据噪声和异常值方面表现更佳, 适合作为划分聚类算法的分类器。因此, 本文利用基于时

间序列聚类的 K-Medoids + DTW 算法实现 D-PMU 坏数据与扰动数据的区分。

DTW 的核心思想是通过构建一个累积距离矩阵, 计算两台 D-PMU 时间序列数据之间的最小匹配成本, 寻找一条最优对齐路径, 使总距离最小, 距离越小则相似度越高。路径规划过程中, DTW 距离矩阵的填充示意图如图 5 所示。

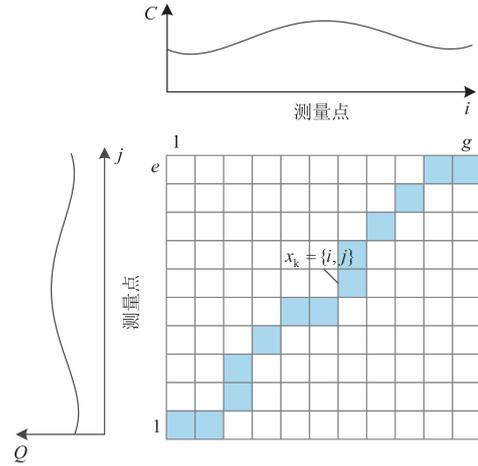


图5 DTW 距离矩阵填充示意图

Fig. 5 DTW distance matrix filling schematic

选取两台 D-PMU 的电压幅值或频率数据 $Q = \{q_1, q_2, \dots, q_e\}$ 和 $C = \{c_1, c_2, \dots, c_g\}$, 构造距离矩阵 D , 采用欧式距离进行度量, 此时两个数据点之间的距离 D_{ij} 为

$$D_{ij} = \sqrt{(q_i - c_j)^2} \quad (17)$$

式中: $i=1, 2, \dots, e, j=1, 2, \dots, g$, q_i 和 c_j 为两台 D-PMU 测量数据的第 i 个和第 j 个数据点, e 和 g 为两台装置的测量点数。

图 5 中蓝色区域的集合 x_k 表示归整路径, 此路径需满足边界性、连续性和单调性 3 个条件。满足这些约束条件的路径有多条, 记为 X , $X = \{x_1, x_2, \dots, x_k\}$, 需要根据式(18)找到一条最短路径。

$$H_{DTW, QC} = \min_X \sum_1^N D(x_k) \quad (18)$$

式中: $H_{DTW, QC}$ 是 Q 和 C 的最短路径, 即 DTW 距离; x_k 为路径上该蓝色方格对应的两个时间序列的坐标, $x_k = (i, j)$; N 为图中归整路径对应的蓝色方格总数。

通常与 DTW 配合使用的分类器是 K-means 算法。K-Medoids 的步骤与 K-means 相似, 实质上是对其进行了改进和优化, 解决了 K-means 对异常值敏感的问题, 从而能够快速完成聚类。通过使得数据点到簇中心的距离平方和最小, 找到最优的聚类

分配方案，其目标函数为

$$\min_C \sum_{l=1}^k \sum_{C_i=l} \|x_i - z_l\|^2 \quad (19)$$

式中： C 为数据点的聚类分配方案； l 为簇的索引， $l=1,2,\dots,k$ ， k 为指定的簇数量； $C_i=l$ 表示数据点 i 被分到第 l 个簇； z_l 为第 l 个簇的中心点。

K-Medoids 算法的步骤如下：1) 从数据集中随机选择 k 个初始中心点；2) 计算其余数据点到各中心点的距离，并将它们归入最近的中心点所在的簇；3) 重新计算每个簇中心点，选择距离中心最近的数据点作为新的中心点；4) 重复步骤2)和3)，直到聚类结果达到预期的效果。

给定包含 11 台装置的 D-PMU 数据集，在检测过程中发现 D-PMU1 数据存在异常。据此，将这 11 台装置在同一时间段内的对应时间序列数据一并送入 K-Medoids + DTW 算法中进行深入分析，如图 6 所示。在图 6 中呈现了算法的计算过程，红色点表示图 7(a)中的 3 条异常波动数据，这 3 台装置的数据明显两两之间 DTW 距离接近，而与其他装置的 DTW 距离相差较远。

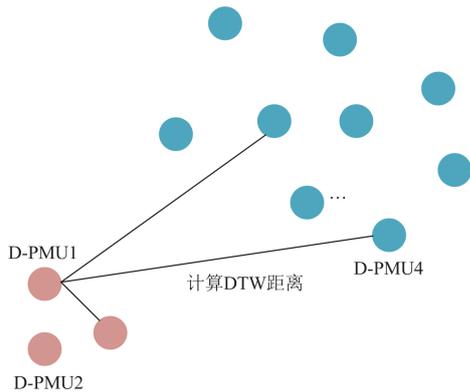
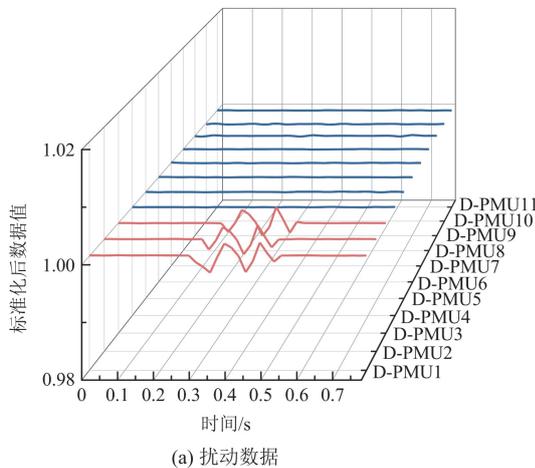


图 6 K-Medoids + DTW 算法
Fig. 6 K-Medoids + DTW algorithm



(a) 扰动数据

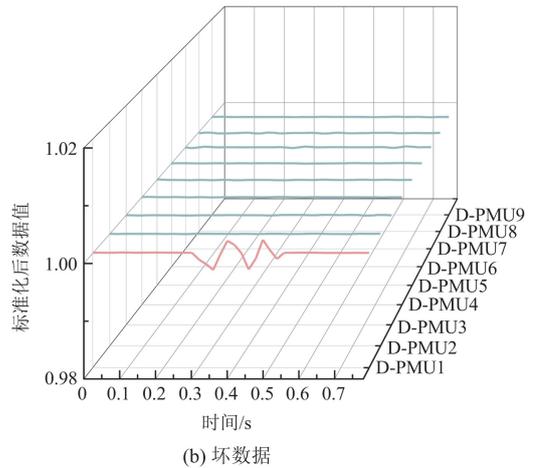


图 7 K-Medoids + DTW 算法聚类效果

Fig. 7 Clustering effect of K-Medoids+DTW algorithm

在应用该算法后，可以观察到图 7(a)中 D-PMU2 和 D-PMU3 两台设备的时间序列数据与 D-PMU1 呈现出高度一致的波动特征，可判定这三台 D-PMU 的数据波动并非偶然坏数据，而是共同受到了某种扰动因素的影响，故将 D-PMU1、D-PMU2 和 D-PMU3 的数据合理判断为扰动数据。

给定 9 台装置的 D-PMU 数据集，当检测到 D-PMU1 设备出现了异常数据现象时，将这 9 台 D-PMU 在同一时间段内的对应时间序列数据送入 K-Medoids + DTW 算法进行综合分析处理。然而，在对整个数据集应用该算法后，在图 7(b)中未发现其余 8 台 D-PMU 设备的时间序列与 D-PMU1 表现出任何相似波动的情况，可以判断 D-PMU1 的异常数据并非源自系统扰动因素，而是属于独立的、随机的坏数据。

3.2 坏数据检测方法整体流程

利用上述过程可以实现无监督 D-PMU 坏数据识别，并区分扰动数据，该算法的流程图如图 8 所示。首先输入待检测 D-PMU 测量数据，对数据进行标准化处理，随后送入 DBSCAN 算法检测异常数据，得到异常数据后选取有关多台 D-PMU 对应数据，进行扰动数据区分，采用 K-Medoids + DTW 算法进行时间序列聚类，若存在同类簇则判定为扰动数据，否则为坏数据。

4 算例分析

利用仿真和实测数据对本文方法的有效性进行验证，并与改进 K-means 方法^[30]、谱聚类法^[32]、LOF 方法^[15]和 GMM 改进 LOF 算法(LOF-GMM)方法^[29]进行了比较。

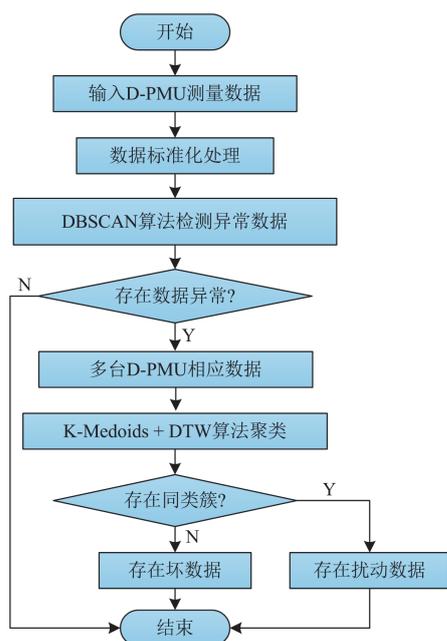


图 8 坏数据检测算法流程图

Fig. 8 Flowchart of bad data detection algorithm

4.1 仿真测试

在 IEEE 10 机 39 节点系统中进行仿真, 人为设置了多个不同偏差范围的单个或连续坏数据, 利用 5 种方法对坏数据进行检测。

1) 稳态情况下的坏数据

在稳态情况下, 设置偏差范围在 0.3%~1.5% 的单点和连续坏数据, 图 9 为检测结果, 其中红色点为检测出的坏数据。结果表明, 利用 LOF-GMM 方法可以检测出 0.3% 以上的单点坏数据和连续坏数据, 利用改进 K-means 方法和谱聚类方法可以检测出偏差在 0.5% 以上的单点或者连续多点坏数据, 但它们对于偏差较小的坏数据检测效果不理想。利用

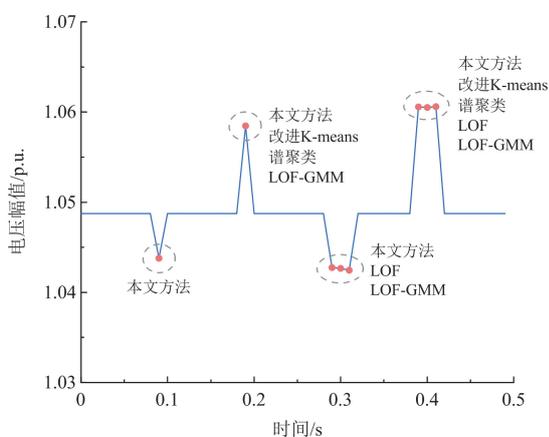


图 9 静态情况下仿真坏数据检测结果

Fig. 9 Simulated bad data detection results under static conditions

LOF 方法可以较好地检测出偏差在 0.5% 和 1.2% 的连续多点坏数据, 但对于单点坏数据检测效果不理想。利用本文所提方法可以检测出偏差在 0.3%~1.5% 之间的单点或者连续多点坏数据。

2) 动态情况下的坏数据

在动态情况下, 设置偏差范围在 0.3%~1.5% 的单点和连续坏数据, 图 10 为检测结果, 红色点为动态情况下检测出的坏数据。结果表明, 利用 LOF-GMM 方法可以较好地检测出偏差在 0.3% 以上的单点或者连续多点坏数据。利用改进 K-means 方法和谱聚类方法可以检测出偏差在 0.5% 以上的单点或者连续多点坏数据, 但对于偏差较小或者处于振荡的波峰和波谷间的坏数据检测效果不理想。利用 LOF 方法可以较好地检测出动态情况下的连续多点坏数据, 但对于单点坏数据的检测效果不理想。利用本文所提方法可以检测出动态情况下, 偏差在 0.3%~1.5% 之间的单点或者连续多点坏数据。

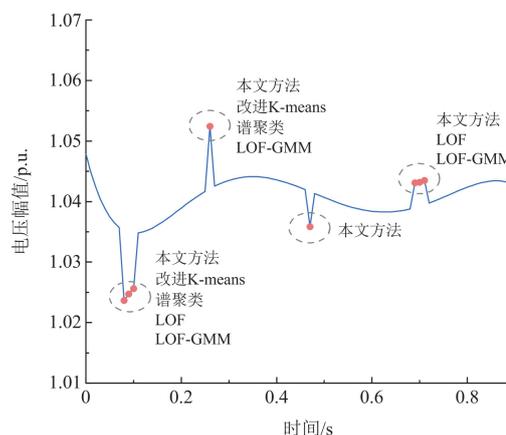


图 10 动态情况下仿真坏数据检测结果

Fig. 10 Simulated bad data detection results in dynamic condition

3) 5 种方法检测能力对比

将 70 组数据分别设置不同偏差的坏数据, 利用 5 种方法分别对其进行坏数据检测, 其结果对比如表 1 所示。在表 1 中, 所提方法对偏差为 0.5% 及更小的坏数据具有较高的准确率, 而其他方法则在偏差为 0.5% 以内时检测效果明显变差, 原因是坏数据偏差太小容易和正常数据混淆, 其他方法在偏差太小时检测效果存在局限性, 因此所提方法在检测能力上相较于其他方法具有一定的优势。

当坏数据的比例和位置发生改变时, 利用 5 种方法进行检测, 检测能力对比如表 2 所示。从表 2 可以看出, 当坏数据比例大于 5% 时, 改进 K-means 方法检测效果不佳。当坏数据的比例大于 15% 时, LOF-GMM 方法和谱聚类方法不能很好地检测到坏数据。当坏数据比例小于 5% 时, LOF 方法检测能力

有限。而本文方法可以检测坏数据比例为1%~20%。

表 1 各类方法不同坏数据偏差下的检测能力对比
Table 1 Comparison of detection capabilities of various methods under different bad data biases

检测方法	不同坏数据偏差下的准确率/%				
	5%	1%	0.5%	0.4%	0.2%
本文方法	99.25	98.91	98.19	97.89	97.43
改进 K-means	97.28	96.83	96.15	95.42	93.57
谱聚类	98.92	98.75	97.01	96.44	94.30
LOF	98.72	98.51	97.23	96.95	95.71
LOF-GMM	99.03	98.79	97.84	97.28	96.02

表 2 各类方法不同坏数据比例下的检测能力对比
Table 2 Comparison of detection capabilities of various methods under different bad data ratios

检测方法	不同坏数据比例下的检测能力				
	1%	2.5%	5%	15%	20%
本文方法	√	√	√	√	√
改进 K-means	√	√	√	×	×
谱聚类	√	√	√	√	×
LOF	×	×	√	√	√
LOF-GMM	√	√	√	√	×

4) 5 种方法检测时间对比

实验的计算机是 Windows11(64 bit)系统、Intel i7 处理器、16 G 内存，仿真使用 MATLAB 2021b 进行。表 3 中比较了 5 种检测方法在不同时间窗下的计算时间，其中每 0.5 s 数据包含 25 个数据点。对每种方法下不同时间窗长的数据各进行 50 次计算，求取平均时间，该时间不包括坏数据和扰动事件区分过程的时间。改进 K-means 方法和 LOF 方法计算量较小，计算时间较短，但相较于其他方法的检测效果不佳。LOF-GMM 方法与本文方法检测效果接近，但随着数据量的增加，LOF-GMM 算法的拟合逐渐复杂，计算量增加，计算时间较长。本文方法相较于谱聚类方法和 LOF-GMM 方法的计算时间更短。当数据量为 200 时，检测时间在 0.018 s 内，满足实时性要求。

表 3 各类方法不同时间窗下的计算时间对比

Table 3 Comparison of computation time of various methods under different time windows

检测方法	各类方法不同时间窗下的计算时间/s				
	0.5 s	1.0 s	2.0 s	3.0 s	4.0 s
本文方法	0.0023	0.0045	0.0087	0.0126	0.0175
改进 K-means	0.0018	0.0033	0.0064	0.0096	0.0125
谱聚类	0.0026	0.0050	0.0097	0.0146	0.0195
LOF	0.0021	0.0040	0.0079	0.0116	0.0155
LOF-GMM	0.0041	0.0079	0.0772	0.1408	0.1941

4.2 D-PMU 实测数据验证

通过 SYMS 装置^[12]获得部分地区 D-PMU 的频率等实测数据。在实测数据中，扰动发生时，多为过程持续时间短、幅度偏差小、监测到的 D-PMU 数量有限的情况，给坏数据的识别带来了困难，但本文所提方法仍可准确检测。

1) D-PMU 坏数据检测

选取 D-PMU 实测频率数据对本文方法的有效性进行验证，检测结果如图 11、图 12 所示。图 11 为利用 K-Medoids + DTW 算法进行聚类的结果，未发现河北石家庄频率数据中红色异常数据段与其他装置数据表现一致，则该地区标记的异常数据为坏数据。图 12 按时间顺序依次显示了对河北石家庄所标记的 6 处异常数据进行实测频率坏数据的检测效果。在该段数据中，160 s 前为静态数据，160 s 后为动态数据，静态和动态条件下分别包含了单点及连续坏数据，且数据偏差范围均在 0.5% 以内，本文方法可以准确检测出所包含的坏数据。

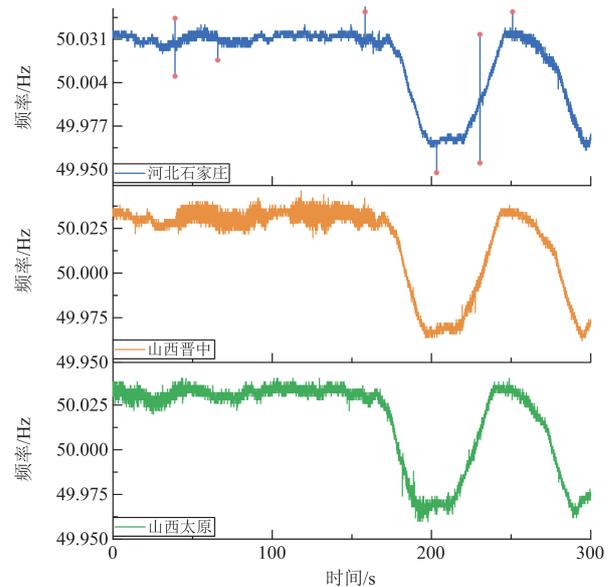
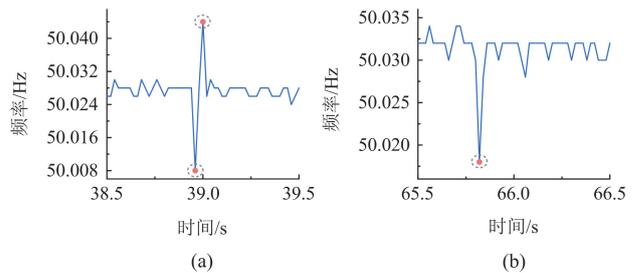


图 11 聚类结果

Fig. 11 Clustering results



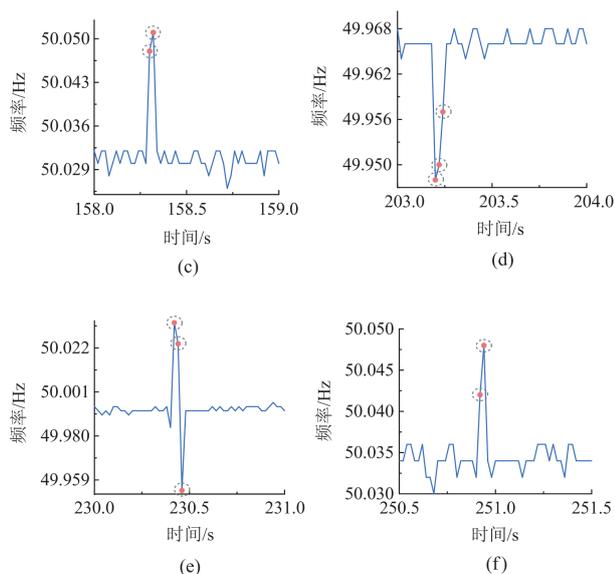


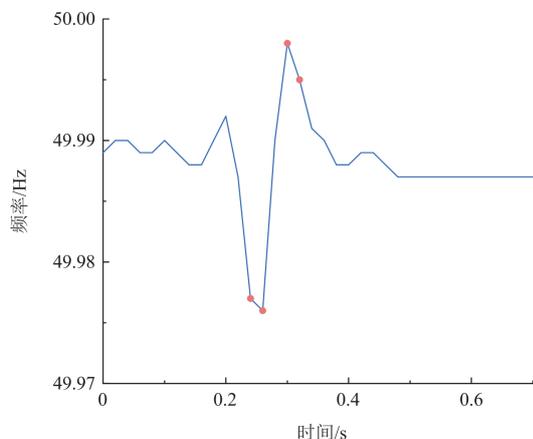
图 12 DBSCAN 算法检测结果

Fig. 12 Detection results of DBSCAN algorithm

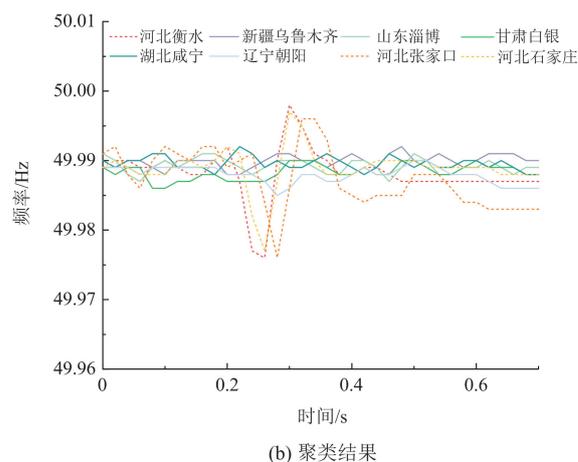
2) D-PMU 坏数据与扰动数据区分

在 D-PMU 捕捉到的扰动事件中,有些扰动发生时持续时间短,偏差小,容易和坏数据混淆。本文方法能够对抗动数据和坏数据进行有效区分,实现坏数据检测。首先,当 DBSCAN 算法检测到某台 D-PMU 存在异常数据时,为进一步区分坏数据和扰动数据,此时与其他 D-PMU 的相应数据段一起送入 K-Medoids + DTW 算法中进行时间序列聚类。若有同类装置数据与异常数据段表现一致,则该段数据为扰动数据,反之为坏数据。

在图 13 中,经过 DBSCAN 算法检测,发现河北衡水的 D-PMU 数据的某一时段存在异常数据即图 13(a)中红色点,这时选取多台有关 D-PMU 的对应数据段组成数据组,利用时间序列聚类方法进行快速划分簇类,聚类结果显示河北衡水、张家口、



(a) DBSCAN 算法检测结果



(b) 聚类结果

图 13 实测坏数据检测效果

Fig. 13 Detection effect of measured bad data

石家庄趋势一致,即图 13(b)中虚线,故此时发生了扰动,而不是坏数据。由于此情况数据偏差较小,其他几种方法具有局限性,并不能很好地区分该类扰动数据和坏数据。

5 结语

本文针对 D-PMU 场景中的实测数据,提出了一种无监督的 D-PMU 坏数据检测方法。该方法没有离线训练的负担,不用人工提前标记数据,且可以避免对于扰动数据的误判。本文结论如下。

1) 提出了一种基于 D-PMU 实测数据的坏数据检测方法。利用不同 D-PMU 数据之间的相似性,对坏数据与扰动数据进行了有效区分。

2) 利用 DBSCAN 算法进行异常数据检测,检测准确率高,相较于其他聚类方法更适用于 D-PMU 数据应用场景,且不需要对样本进行训练得到先验知识。

3) 提出的基于综合评价得分和麻雀搜索算法的方法解决了 DBSCAN 需要手动设置参数的问题,与其他方法相比,该方法实现了自适应参数优化,提高了算法的准确性。

4) 提出了基于 K-Medoids + DTW 的聚类算法,解决了目前 D-PMU 测量装置提供的信息量少,其他检测方法会因为数据量不足造成结果不准确的问题,可以有效区分扰动数据和坏数据。

本文未考虑当单台 D-PMU 发生扰动情况时如何区分坏数据的情况;同时,在之后的研究中,将进一步探讨配电网中坏数据的修正方法。

参考文献

- [1] 中华人民共和国国家发展改革委员会. “十四五”可再生能源发展规划(发改能源〔2021〕1445号)[Z]. 2022.

- National Development and Reform Commission of the People's Republic of China. Renewable energy development plan for the 14th five-year plan period (NDRC Energy [2021] No. 1445)[Z]. 2022.
- [2] 张忠静, 刘磊, 蔡永翔, 等. 新型电力系统下大规模新能源接入配电网的线损变化趋势分析[J]. 电力大数据, 2024, 27(2): 11-22.
ZHANG Zhongjing, LIU Lei, CAI Yongxiang, et al. Analysis of line loss change trend of large-scale new energy access distribution network under new power system[J]. Power Systems and Big Data, 2024, 27(2): 11-22.
- [3] 朱蜀, 刘开培, 秦亮, 等. 电力电子化电力系统暂态稳定性分析综述[J]. 中国电机工程学报, 2017, 37(14): 3948-3962, 4273.
ZHU Shu, LIU Kaipei, QIN Liang, et al. A review of transient stability analysis of power electronics-based power systems[J]. Proceedings of the CSEE, 2017, 37(14): 3948-3962, 4273.
- [4] 李华瑞, 李文博, 李铮, 等. 基于生成对抗网络与度量学习的数据驱动频率安全评估[J]. 电力系统保护与控制, 2024, 52(18): 101-111.
LI Huarui, LI Wenbo, LI Zheng, et al. Data-driven frequency security assessment based on generative adversarial networks and metric learning[J]. Power System Protection and Control, 2024, 52(18): 101-111.
- [5] 胡加伟, 王彤, 王增平. 直流闭锁后系统暂态稳定紧急协同控制策略研究[J]. 电力系统保护与控制, 2023, 51(4): 43-52.
HU Jiawei, WANG Tong, WANG Zengping. Collaborative emergency control strategy of system transient stability after DC blocking[J]. Power System Protection and Control, 2023, 51(4): 43-52.
- [6] 李金鹏, 冯华, 陈晓刚, 等. 计及多重不确定性的规模化电动汽车接入配电网调度方法[J]. 电力系统自动化, 2024, 48(7): 138-146.
LI Jinpeng, FENG Hua, CHEN Xiaogang, et al. Dispatching method for large-scale electric vehicles connecting to distribution network considering multiple uncertainties[J]. Automation of Electric Power Systems, 2024, 48(7): 138-146.
- [7] 刘正谊, 帅勇, 袁文, 等. 基于随机森林算法的线路接地在线诊断模型[J]. 电力大数据, 2021, 24(3): 68-75.
LIU Zhengyi, SHUAI Yong, YUAN Wen, et al. On-line diagnosis model of line grounding based on random forest algorithm[J]. Power Systems and Big Data, 2021, 24(3): 68-75.
- [8] 王增平, 林一峰, 王彤, 等. 电力系统继电保护与安全控制面临的挑战与应对措施[J]. 电力系统保护与控制, 2023, 51(6): 10-20.
WANG Zengping, LIN Yifeng, WANG Tong, et al. Challenges and countermeasures for power system relay protection and security control[J]. Power System Protection and Control, 2023, 51(6): 10-20.
- [9] 沈赋, 李施伟, 王健, 等. 融合储能的光伏发电系统并网逆变器建模与稳定性分析[J]. 电力系统保护与控制, 2024, 52(19): 131-143.
SHEN Fu, LI Shiwei, WANG Jian, et al. Modeling and stability analysis of a photovoltaic grid-connected inverter integrated with an energy storage system[J]. Power System Protection and Control, 2024, 52(19): 131-143.
- [10] VERMA N, KUMAR N, GUPTA S, et al. Review of sub-synchronous interaction in wind integrated power systems: classification, challenges, and mitigation techniques[J]. Protection and Control of Modern Power Systems, 2023, 8(2): 277-302.
- [11] ZHANG Y, MARKHAM P, XIA T, et al. Wide-area frequency monitoring network (FNET) architecture and applications[J]. IEEE Transactions on Smart Grid, 2010, 1(2): 159-167.
- [12] LIU H, BI T, XU S, et al. A full-view synchronized measurement system for the renewables, controls, loads, and waveforms of power-electronics-enabled power distribution grids[J]. IEEE Transactions on Smart Grid, 2022, 13(5): 3879-3890.
- [13] SUNDARARAJAN A, KHAN T, MOGHADASI A, et al. Survey on synchrophasor data quality and cybersecurity challenges, and evaluation of their interdependencies[J]. Journal of Modern Power Systems and Clean Energy, 2019, 7(3): 449-467.
- [14] California ISO. Five year synchrophasor plan[R]. 2011.
- [15] 刘灏, 朱世佳, 毕天姝. 基于局部离群因子的 PMU 连续坏数据检测方法[J]. 电力系统自动化, 2022, 46(1): 25-32.
LIU Hao, ZHU Shijia, BI Tianshu. Continuous bad data detection method for PMU based on local outlier factor[J]. Automation of Electric Power Systems, 2022, 46(1): 25-32.
- [16] 马志强. 系统的可观测性与不良数据的再估计识别[J]. 电网技术, 1979(2): 66-79.
MA Zhiqiang. Observability of the system and re-estimation identification of bad data[J]. Power System Technology, 1979(2): 66-79.
- [17] KORRES G N, MANOUSAKIS N M. State estimation and bad data processing for systems including PMU and SCADA measurements[J]. Electric Power Systems Research, 2011, 81(7): 1514-1524.
- [18] HANDSCHIN E, SCHWEPPE F C, KOHLAS J, et al. Bad data analysis for power system state estimation[J].

- IEEE Transactions on Power Apparatus and Systems, 1975, 94(2): 329-337.
- [19] 于尔铿. 电力系统状态估计中不良数据的零残差辨识法[J]. 电网技术, 1981(2): 66-73.
YU Erkeng. Zero residual identification method for bad data in power system state estimation[J]. Power System Technology, 1981(2): 66-73.
- [20] GHIODEL S G, CHOW J H, STEFOPOULOS G, et al. Phasor-measurement-based state estimation for synchrophasor data quality improvement and power transfer interface monitoring[J]. IEEE Transactions on Power Systems, 2013, 29(2): 881-888.
- [21] 陈晓刚, 易永辉, 江全元, 等. 基于 WAMS/SCADA 混合量测的电网参数辨识与估计[J]. 电力系统自动化, 2008, 32(5): 1-5.
CHEN Xiaogang, YI Yonghui, JIANG Quanyuan, et al. Grid parameter identification and estimation based on WAMS/SCADA hybrid measurements[J]. Automation of Electric Power Systems, 2008, 32(5): 1-5.
- [22] YANG J, ZHANG W, GUO F. Distributed Kalman-like filtering and bad data detection in the large-scale power system[J]. IEEE Transactions on Industrial Informatics, 2022, 18(8): 5096-5104.
- [23] JONES K D, PAL A, THORP J S. Methodology for performing synchrophasor data conditioning and validation[J]. IEEE Transactions on Power Systems, 2015, 30(3): 1121-1130.
- [24] AYAD A, FARAG H E Z, YOUSSEF A, et al. Detection of false data injection attacks in smart grids using recurrent neural networks[C] // 2018 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT), February 19-22, 2018, Washington, DC, USA: 1-5.
- [25] 杨智伟, 刘灏, 毕天姝, 等. 基于长短期记忆网络的 PMU 不良数据检测方法[J]. 电力系统保护与控制, 2020, 48(7): 1-9.
YANG Zhiwei, LIU Hao, BI Tianshu, et al. PMU bad data detection method based on long short-term memory network[J]. Power System Protection and Control, 2020, 48(7): 1-9.
- [26] HAO Y, WANG M, CHOW J H, et al. Modelless data quality improvement of streaming synchrophasor measurements by exploiting the low-rank Hankel structure[J]. IEEE Transactions on Power Systems, 2018, 33(6): 6966-6977.
- [27] WU M, XIE L. Online detection of low-quality synchrophasor measurements: a data-driven approach[J]. IEEE Transactions on Power Systems, 2016, 32(4): 2817-2827.
- [28] MAO Z, XU T, OVERBYE T J. Real-time detection of malicious PMU data[C] // 2017 19th International Conference on Intelligent System Application to Power Systems (ISAP), September 17-20, 2017, San Antonio, TX, USA: 1-6.
- [29] 梅玉杰, 李勇, 周王峰, 等. 基于机器学习的配电网异常缺失数据动态清洗方法[J]. 电力系统保护与控制, 2023, 51(7): 158-169.
MEI Yujie, LI Yong, ZHOU Wangfeng, et al. Dynamic cleaning method for abnormal missing data in distribution networks based on machine learning[J]. Power System Protection and Control, 2023, 51(7): 158-169.
- [30] 吴蕊, 张安勤, 田秀霞, 等. 基于改进 K-means 的电力数据异常检测算法[J]. 华东师范大学学报(自然科学版), 2020(4): 79-87.
WU Rui, ZHANG Anqin, TIAN Xiuxia, et al. Power data anomaly detection algorithm based on improved K-means[J]. Journal of East China Normal University (Natural Science), 2020(4): 79-87.
- [31] WANG X, SHI D, WANG J, et al. Online identification and data recovery for PMU data manipulation attack[J]. IEEE Transactions on Smart Grid, 2019, 10(6): 5889-5898.
- [32] YANG Z, LIU H, BI T, et al. Bad data detection algorithm for PMU based on spectral clustering[J]. Journal of Modern Power Systems and Clean Energy, 2020, 8(3): 473-483.
- [33] 甘迪, 金岩磊, 葛立青, 等. 基于爬山-蚁群-FCM 模糊聚类的不良负荷数据辨识[J]. 电力大数据, 2022, 25(1): 1-8.
GAN Di, JIN Yanlei, GE Liqing, et al. Identification of bad load data based on hill climbing-ant colony-FCM fuzzy clustering[J]. Power Systems and Big Data, 2022, 25(1): 1-8.
- [34] XUE J, SHEN B. A novel swarm intelligence optimization approach: sparrow search algorithm[J]. Systems Science & Control Engineering, 2020, 8(1): 22-34.

收稿日期: 2024-11-14; 修回日期: 2025-03-26

作者简介:

刘 灏(1985—), 男, 博士, 副教授, 研究方向为广域同步相量测量技术; E-mail: hliu@ncepu.edu.cn

陈 容(1999—), 女, 硕士研究生, 研究方向为数据驱动的电力系统扰动识别; E-mail: 930033105@qq.com

毕天姝(1973—), 女, 通信作者, 博士, 教授, 博士生导师, 研究方向为电力系统保护与控制、WAMS 应用研究等。E-mail: tsbi@ncepu.edu.cn

(编辑 石晋美)