[大数据]

基于大数据分析的配电网实时故障定位技术

文/肖月
国网四川省电力公司大英县供电分公司

传统的故障定位方法往往依赖于经验判断和简单的硬件传感数据，在处理复杂、动态变化的故障情况时常显得力不从心。随着信息技术的飞速发展，大数据技术已引入配电网管理中，提供一种全新的解决方案。根据分析大规模数据集，包括实时数据流、历史故障记录和环境数据，大数据技术能够揭示故障发生的模式和规律，从而实现更快速、准确的故障定位。

在现代电力系统运营中，实时故障定位技术的高效实施对于保障供电可靠性和系统稳定性具有至关重要的作用；随着配电网结构的复杂化及负荷需求的不断增加，系统故障的频发不可避免地对电力供应与经济运行造成严重影响。配电网故障定位的传统方法，如基于过电流保护的定位技术，虽然在历史上发挥了重要作用，但在应对多源馈电、分布式发电和电网自动化等新兴结构时显得力不从心，常常由于定位准确度不高和响应速度缓慢导致故障修复时间延长，进而增加了电网的经济损失和系统风险。在这种背景下，大数据技术的引入为配电网故障定位技术的创新提供了新的动力和可能性。

大数据技术概述

大数据技术在配电网实时故障定位领域的应用正日益成为提升电网运维效率与准确性的关键因素，尤其是在处理复杂网络结构和不断变化的电网环境中显示出其独特优势。通过集成来自智能电表、传感器网络、地理信息系统（GIS）以及操作历史数据库的海量数据，大数据技术能够为配电网的故障检测、诊断和定位提供深入的见解和强大的支持，这些数据不仅涵盖了电流、电压和功率因数等电气参数，还包括天气条件、设备老化状态和用户用电行为等非电气信息，它们共同构成了对配电网运行状态的全面描述。借助高级数据分析工具和算法，如流数据处理、模式识别和机器学习，专家们能够实时分析这些数据，准确识别出故障发生的初步信号，并快速定位故障源。这一过程显著减少了系统的响应时间和服务中断时间，从而大幅提高了电网的可靠性和经济效益。实际上，将大数据技术应用于配电网实时故障定位不仅仅是技术革新的需求，也是应对日益增加的电网负荷和复杂性的必然选择，它使得电网系统运维从传统的反应式管理转变为基于数据驱动的主动管理模式，极大地提升了配电网故障处理的效率和准确性，这对于建设可持续和高效的现代电力系统具有重要意义。

配电网实时故障定位的现状

配电网实时故障定位技术目前处于快速发展阶段，旨在准确迅速地识别和定位电网中的故障点，以最小化电力供应中断的影响并优化配电网的运维管理。当前的实时故障定位方法主要依赖于集成的传感器系统、故障指示器和智能电表等设备来收集关键的运行数据，如电压、电流和相位角等信息，

这些数据实时监控系统被传输至中心处理单元，其中应用预设的故障诊断算法进行分析以识别故障特征和故障类型。此外，现有的故障定位系统在处理大规模数据和实时分析方面存在局限，如数据处理延迟和算法的适应性问题，这些问题可能导致故障定位的准确率和效率受到影响。针对这些挑战，业界和学术界正在探索包括机器学习和大数据分析技术在内的新方法，这些方法能够从海量数据中学习故障模式，提高故障诊断的智能化水平并优化故障响应策略。例如，通过实施高级数据分析技术，可以在故障发生前对潜在的故障点进行预测，从而实现更加主动的故障管理。

基于大数据分析的配电网实时故障定位技术

1. 大数据采集与处理

大数据采集与处理是实现配电网实时故障定位的基础环节，关键在于从多源数据流中提取有价值的信息，以支持高效且准确的故障分析和定位。在配电网中，数据采集主要依赖于广泛部署的智能传感器、智能电表、继电保护设备以及其他监测设备，这些设备能够实时捕获电网的运行状态，如电流、电压、频率和温度等参数，同时还记录了天气、设备负荷和网络拓扑变化等数据，所有这些数据共同构成了配电网的“大数据”，为故障分析提供了全面的视角。数据处理环节涉及数据的预处理、清洗、集成和存储，通过预处理去除数据中的噪声和异常值，接着使用数据清洗技术标准化数据格式和质量，保证数据的一致性和准确性。随后，数据集成技术用来融合来自不同源的数据，建立统一的数据模型，使得数据分析可以在一个整合的数据集上进行，增加数据的可用性和有效性。

2. 数据驱动的故障检测算法

数据驱动的故障检测算法在配电网实时故障定位技术中扮演着至关重要的角色，通过精确分析大量收集的电网数据来识别潜在故障。本研究采用基于支持向量机（SVM）的算法，该算法通过构建一个分类模型来区分正常运行状态与各种故障状态。SVM 是一种监督学习的方法，特别适用于高维数据集，能够有效处理非线性问题并具有良好的泛化能力。

SVM的目标是找到一个最优的超平面，使得不同类别的数据在该平面两侧，并且距离平面的间隔最大化。可以通过以下优化问题表达

$$\min_{w,b}\frac{1}{2}\|w\|^2+C\sum_{i=1}^{n}\xi_i \tag{1}$$

式中，w是超平面的法向量；b是偏置项；ξ_i是松弛变量；用于处理非线性可分的情况；C是正则化参数，用于控制间隔与分类错误之间的权衡。

目标函数中的第一项$1\frac{1}{2}\|w\|^2$旨在最大化间隔，第二项$C\sum_{i=1}^{n}\xi_i$则尽可能减少分类错误。

约束条件为

$$y_i(w\cdot x_i+b)\geq 1-\xi_i,\xi_i\geq 0,i=1,\ldots,n \tag{2}$$

式中，x_i和y_i分别是训练样本及其标签，标签y_i的值为+1或−1，分别代表不同的类别。

这些约束确保所有数据点都正确分类或处于超平面正确的一侧，除非允许一定的误差ξ_i。

3. 故障定位的实时数据分析模型

在配电网实时故障定位领域，模型的准确性至关重要，通常通过实时数据分析模型来实现，尤其是采用基于行波理论的故障距离估计算法，该算法通过捕捉和分析传导过程中的电气波形来实现故障点的精准定位。模型接收来自各监测节点的同步时间戳数据，通过行波速度v和到达时差Δt计算行波到达故障点的距离，表述为

$$d=\frac{1}{2}v\Delta t \tag{3}$$

式中，d为从监测点到故障点的距离；v为行波在电缆中的传播速度，通常取决于电缆材质和构造；Δt为行波从发生点到达两个监测节点的时间差。

继而，系统通过实时数据来校准行波速度，利用电缆参数和温度影响修正传播速度，确保距离估计的准确性，修正后的行波速度为

$$v'=v(1+\alpha\Delta T) \tag{4}$$

式中，α为温度对行波速度的影响系数；ΔT为电缆温度变化值。

基于式（3）和式（4），逐步缩小故障定位的范围，进而利用机器学习算法，如支持向量机（SVM）来对故障类型进行分类，算法的基本形式为

$$f(x)=\mathrm{sgn}(w\phi(x)+b) \tag{5}$$

式中，x为从行波数据中提取的特征向量；$\phi(x)$是将数据映射到高维空间的核函数；w和b是SVM模型的参数。

此分类不仅能够区分故障的类型，还可以估计

故障严重性，为配电网运维提供决策支持。在实时故障定位系统中，综合运用以上模型和算法，可实现对故障点精确定位及其性质的快速判断，大幅提高配电网的可靠性和响应速度。此外，通过实时数据分析，还可以动态调整行波速度的估计值，以适应环境变化和网络配置的实时更新，保证故障定位系统在各种操作条件下的高效率和高准确度。

4. 数据可视化与结果反馈

数据可视化与结果反馈在基于大数据分析的配电网实时故障定位技术中起着至关重要的作用，为操作员提供直观的数据解读和决策支持。配电网的运行数据量庞大且复杂，包括来自各类传感器的电压、电流、温度和振动等参数，此外还涉及时间、位置和设备状态等信息。通过有效的数据可视化技术，这些数据不仅被简化和结构化，还能以易于理解的形式呈现，使得故障分析变得更加迅速和准确。可视化工具如散点图、线图、热图和地理信息系统（GIS）映射广泛用于显示实时数据和历史趋势，这些工具可以揭示不同设备和网络组件之间的相互作用及其随时间的变化趋势。为进一步提高故障诊断的效率，结果反馈机制在故障定位系统中设计为多层次的。这种分层的反馈机制确保了从初步识别到深入诊断的连续性，极大地缩短故障响应时间，并提高处理的精确度。此外，反馈系统还包括对操作员的动态指导，如通过实时数据更新推荐的维修措施或调整操作策略，以及通过模拟结果展示可能的故障解决方案的预期效果。表1是一个示例数据表格，展示了配电网实时故障定位系统中的故障诊断结果及其可视化呈现。

表1　配电网实时故障定位系统应用数据

设备ID	故障发生时间	故障类型	置信度（%）	地理位置	可视化图表链接
001	2022-10-03 15:22:05	超温故障	95	经度 28.538, 纬度 -81.378	温度趋势图
002	2022-10-03 15:24:10	电压下降	90	经度 28.538, 纬度 -81.378	电压监测图
003	2022-10-03 15:30:45	断线故障	85	经度 28.540, 纬度 -81.370	GIS位置映射
004	2022-10-03 15:35:21	接地故障	88	经度 28.541, 纬度 -81.366	电流异常图

表1中列出了各故障设备的具体ID、故障发生时间、故障类型、置信度、地理位置和对应的可视化图表链接，通过点击链接，操作员可以直接查看详细的图表分析，这些图表基于实时更新的数据生成，能够提供关于故障动态的直观信息，辅助技术人员快速准确地进行故障处理和决策。

技术应用测试

在本次测试中采用的数据集是通过模拟得到的，旨在确保测试的全面性和适应性。模拟数据集的生成涉及故障类型、发生时间和环境因素等多样化参数的配置，以充分测试系统在不同条件下的表现。所有生成的数据均用于系统的功能测试，主要是因为系统的故障检测算法在开发阶段已经完成了预训练，测试阶段主要集中于评估模型的实际应用性能。数据集总共包括了1 000 000条记录，每条记录模拟一个独立的故障事件，以测试系统对不同故障情况的响应速度和准确性。每条数据记录中包含的内容极为详细，涵盖了配电网运行中关键的电气测量参数，如电压、电流和频率，以及故障具体发生的时间。测试共执行了五次，以确保结果的一致性和可重复性。 测试主要围绕以下几个关键指标进行：故障检测准确率、系统响应时间、CPU利用率和内存使用。这些指标共同定义了系统性能的多个方面：故障检测准确率反映了系统识别和定位故障的能力；系统响应时间衡量从故障发生到系统响应所需的时间；CPU利用率和内存使用则显示了系统运行的资源效率。测试数据见表2。

表2　测试数据

测试编号	故障检测准确率（%）	系统响应时间 /s	CPU利用率（%）	内存使用 / GB
1	98.43	0.45	72.34	5.64
2	99.12	0.42	68.55	5.72
3	98.97	0.47	70.12	5.67
4	98.79	0.43	69.97	5.58
5	98.89	0.44	70.82	5.51

结果显示，系统的故障检测准确率始终保持在98.43%～99.12%，说明算法能高效准确地识别和定位故障。系统响应时间介于0.42～0.47 s，反映出系统能够迅速处理并响应故障事件。CPU利用率和内存使用在测试中也显示出系统资源管理的高效性，保持在合理范围内。这些测试结果证明了基于大数据分析的配电网实时故障定位系统在实际电力应用中的有效性和可靠性，能够满足工业级的需求。EA

（收稿日期：2024-12-15）