

中国电机工程学报
Proceedings of the CSEE
ISSN 0258-8013, CN 11-2107/TM

《中国电机工程学报》网络首发论文

题目: 基于概率分布联合训练时空自编码器的同调发电机在线辨识及其可解释性
作者: 刘丰瑞, 李柏潼
网络首发日期: 2021-08-10
引用格式: 刘丰瑞, 李柏潼. 基于概率分布联合训练时空自编码器的同调发电机在线辨识及其可解释性. 中国电机工程学报.
<https://kns.cnki.net/kcms/detail/11.2107.TM.20210809.1643.005.html>



网络首发: 在编辑部工作流程中, 稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定, 且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式 (包括网络呈现版式) 排版后的稿件, 可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定; 学术研究成果具有创新性、科学性和先进性, 符合编辑部对刊文的录用要求, 不存在学术不端行为及其他侵权行为; 稿件内容应基本符合国家有关书刊编辑、出版的技术标准, 正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性, 录用定稿一经发布, 不得修改论文题目、作者、机构名称和学术内容, 只可基于编辑规范进行少量文字的修改。

出版确认: 纸质期刊编辑部通过与《中国学术期刊 (光盘版)》电子杂志社有限公司签约, 在《中国学术期刊 (网络版)》出版传播平台上创办与纸质期刊内容一致的网络版, 以单篇或整期出版形式, 在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊 (网络版)》是国家新闻出版广电总局批准的网络连续型出版物 (ISSN 2096-4188, CN 11-6037/Z), 所以签约期刊的网络版上网络首发论文视为正式出版。

基于概率分布联合训练时空自编码器的同调发电机在线辨识及其可解释性

刘丰瑞¹, 李柏潼^{2*}

(1. 东北电力大学 电气工程学院, 吉林省 吉林市 132012; 2. 康奈尔大学, 纽约州 纽约市 10041)

Online Coherency Identification Based on Joint Training of Spatial Temporal Auto-Encoder With Probability Distribution and Its Interpretability Analysis

LIU Fengrui¹, LI Baitong^{2*}

(1. School of Electrical Engineering, Northeast Electric Power University, Jilin 132012, Jilin Province, China; 2. Cornell Tech University, New York City 10041, New York State, America)

ABSTRACT: This study proposed an improved Auto-Encoder (IAE) scheme for online coherency identification by employing synchro phasor measurements obtained by phasor measurement units (PMUs) in the power system. First, deep neural network layers were constructed for feature extraction and jointly trained with the clustering layer. Second, a distance matrix that presents the similarity index between each pair of generators was calculated according to complexity-invariant distance (CID), considering the characteristic of similar rotor movement trends of coherent generator groups. Third, Gaussian Mixture Model (GMM), improved by KL divergence, was applied to implement probabilistic data-driven clustering and optimize parameters in both feature extraction and clustering processes. The proposed approach was evaluated and validated on the simulation cases of 4-machine 11-bus two-area system and 16-machine 68-bus 5-area system, as well as actual data gathered through China Southern Power Grid (CSG) and Yunnan Power Grid. Also, the high-dimensional features of power angle were visualized and explained via adopting the class activation map and the idea of transfer learning. The results demonstrate that this approach is not only robust under noise and avoid deficiency of pivotal features while mining data, but also realizes prompt and accurate online coherency identification, all the while and the explaining of its results' interpretability without a manually defined number of generator groups or pre-trained off-line labeled data.

KEY WORDS : wide-area measurement; coherency identification; spatial temporal neural network; joint training; transfer learning; interpretability analysis

摘要: 提出一种改进的自编码器(Improved Auto-Encoder, IAE)框架, 借助电力系统广域量测信息, 在线辨识电力系统同调机群。设计时空神经网络搭建特征提取层, 与聚类层联

合训练; 考虑同调机组转子运动趋势相似的特点, 计算数据特征的复杂度不变性距离(Complexity-Invariant Distance, CID)表征机组相似度; 基于 Kullback-Leibler 散度(KL 散度)改进高斯混合模型(Gaussian Mixture Model, GMM)进行概率性数据驱动的聚类, 优化聚类 and 特征提取过程中的参数。将所提方法应用于典型 4 机 11 节点 2 区域电力系统、典型 16 机 68 节点 5 区域电力系统、中国南方电网进行分析、验证, 基于类激活映射(Class Activation Map, CAM)和迁移学习思想进行功角高维特征的可视化解释, 结果表明: 该方法在噪声干扰下具有较强鲁棒性, 避免了数据关键特征缺失、不依赖于人工定义的机群数量、不需要训练离线标注数据, 实现了同调机群快速准确在线辨识和结果的有效解释。

关键词: 广域量测信息; 同调辨识; 时空神经网络; 联合训练; 迁移学习; 可解释性

0 引言

随着我国碳达峰、碳中和目标的提出, 区域互联电网规模不断扩大, 以新能源为主体的新型电力系统运行方式更加复杂多变, 意外干扰、元件故障和人为失误都可能导致全系统停电, 给电力系统的安全稳定运行带来严峻的挑战^[1]。

电力系统的发电机同调性反映了系统中发电机在受扰后的动态过程中具有相同运动趋势, 每组同调发电机保持几乎恒定的转子角度差^[2]。快速搜索和辨识同调机群是电网实时紧急控制、对关键支路主动解列以防止长时间区域间功率振荡诱发连锁故障进而造成大停电事故的首要任务^[3-4]。

随着广域同步相量测量单元(phasor measurement unit, PMU)的广泛应用, 研究人员借助数据挖掘技术, 从大量广域量测信息中提取可表

征发电机间同调性的关键信息来划分同调机群^[5-8], 数据驱动的同调辨识主要有无监督和有监督两种方法。有监督方法先通过仿真生成大量带事件标注的数据, 再利用支持向量机(Support Vector Machine, SVM)^[9]、人工神经网络(Artificial Neural Network, ANN)^[10-12]和决策树^[13-14]等机器学习技术识别发电机的特征, 其预测速度较快、具有较高的精确度且训练样本歧义性低, 但依赖大量的离线标注数据, 且训练样本和实际数据的分布以及特征往往不统一, 使得模型泛化能力差, 离线测试效果良好但是实际运用效果不佳。针对有监督方法的诸多问题, 无监督方法在同调发电机识别领域受到关注: 传统的特征工程方法利用统计学人工构建派生特征, 如文献[6]先计算发电机间的相关系数, 再根据阈值划分机群; 文献[7]在文献[6]的基础上, 基于相关系数构建拉普拉斯矩阵, 再使用谱图聚类; 文献[8]先将高维时序数据投影到低维空间, 再利用遗传算法搜索出最优投影方向。这些人工构造的特征缺乏科学指导, 依赖于研究者的先验知识, 难以评估其构造的效果。另一类方法通过筛选特征进行数据降维, 再对发电机聚类。常用的特征筛选方法包括主成分分析(Principal Component Analysis, PCA)^[15]、核主成分分析(Kernel PCA, KPCA)^[16-17]、独立成分分析(Independent Component Analysis, ICA)^[18], 可以将高维时序数据嵌入低维欧氏空间, 再使用支持向量聚类^[16]、亲和聚类^[17]、k-means^[18]、层次聚类^[19]等方法辨识同调机组。但是, 这类方法不但对离群值敏感, 且在处理相对长的观察窗时所需的高处理能力不足, 在多工况及噪声干扰下鲁棒性低。以上无监督方法在提取和构造特征时往往使用线性映射或低非线性映射, 会丢失有价值的信息, 缺乏对数据整体模式的表征能力, 也难以确定合适的特征组合输入, 且计算方法复杂耗时, 具有较大的局限性。

以上应用于同调发电机识别问题的方法均采用人工特征提取结合浅层机器学习的方式, 无法全面挖掘广域量测数据的未知特征, 无法在线处理复杂、高阶、结构化、多特征的无标签数据, 难以拟合高度非线性的复杂映射, 且难以设计较好的去噪声环节。现阶段暂无文献应用深度学习(Deep Learning, DL)来提取发电机功角信息进行同调性辨识, 这一领域尚属空白。另外, 全部机器学习方

法都类似于“黑箱”, 欠缺可解释性是阻碍其实际应用的主要原因之一。

针对上述问题, 本文提出一种基于概率分布联合训练时空自编码器的同调发电机在线辨识方法并进行可视化解释, 在挖掘大量数据特征的同时实现快速聚类。创新点如下:

1) 本文在无监督的同调发电机在线聚类问题中, 结合统计学推导出以 KL 散度为损失函数、沿着梯度下降方向更新聚类参数的方法(通常 KL 散度只能用于有监督学习, 而且传统聚类算法无法建立输入特征到输出组别的映射, 因此无法计算梯度)。相较于传统方法, 能够更准确地确定机群数量和特征中心。

2) 本文设计的训练方式是联合训练特征损失(均方误差, 即 Mean Square Error, MSE)和聚类损失(KL 散度), 预训练特征提取层后微调参数, 更有利于得到表达时序未标记数据分割成多个类别的潜在维度; 基于自编码思想, 采用 Adam 自适应优化算法训练, 自适应调节学习率, 适用于非常嘈杂和稀疏梯度的非稳态问题; 考虑到网络的复杂性, 引入 Dropout 和 early stop 技术, 有效防止过拟合并降低模型对部分超参数的敏感度。

3) 针对发电机转子摇摆曲线的特点, 设计了结合卷积神经网络(Convolutional Neural Networks, CNN)和双向长短期记忆网络(Bidirectional Long Short Term Memory, BiLSTM)的自编码器(Auto-Encoder, AE)结构, 论述其能够捕捉同调机群时空特征的能力, 逐层训练深度神经网络以捕捉发电机转子角的长时依赖、短距离波动及空间相关性, 提取更全面、更具代表性的特征。高非线性映射将时序数据自动嵌入低维空间, 数据降维的同时最大程度保留有效信息, 具有较高的拟合精度和较强的泛化能力。

4) 本文定义的发电机相似函数考虑了同调机群响应趋势, 由于同调机组响应趋势相近, 但其欧式距离可能较大, 且量测数据含噪声, 可能造成错误聚类, 本文构建的响应趋势复杂度的衡量标准提高了机群划分的准确度。

5) 本文基于 CAM 对发电机功角高维特征进行了热度图可视化解释, 消除了深度学习模型的“黑箱”性质, 提高了可信度。

1 同调发电机的聚类指标

1.1 同调发电机的基本判据

发电机功角表征了发电机运动的位置和速度等基本信息, 由 PMU 测量的功角数据反映了影响发电机同调性的主要因素, 包括网络拓扑和扰动的影响。在关注的时间尺度 $t \in [0, \tau]$ 之内, 如果任意时刻发电机 i 和 j 的功角增量的偏差都满足下式, 那么发电机 i 和 j 在此时段同调^[20]。

$$\max_{t \in [0, \tau]} |\Delta \delta_i(t) - \Delta \delta_j(t)| \leq \varepsilon \quad (1)$$

其中, $\Delta \delta_i(t) = \delta_i(t) - \delta_i(0)$, $\Delta \delta_j(t) = \delta_j(t) - \delta_j(0)$; $\delta_i(t)$ 和 $\delta_j(t)$ 是发电机 i, j 在时刻 t 的转子角; ε 是大于 0 的给定值, 通常是 5° 至 10° 。

上式给出了电力系统同调发电机的基本判据, 但在实际应用中产生许多障碍^[1]。现阶段判别机组同调性的指标是利用各种数学变换作用于发电机动态信号得到的相关性度量值, 如发电机组功角曲线的功角差有效值、结构差异度聚类^[21]和相关系数^[6]等, 但基于机理的方法提取的特征限于已有的知识, 无法挖掘到更加全面的特征, 特征之间存在冗余, 所提出的指标计算式也较为复杂, 难以有效滤除实际系统的噪声信号。因此, 基于深度时空神经网络挖掘发电机同调性特征并降低指标计算的复杂度是具有重要意义的^[22-23]。

1.2 基于广域量测信息的功角差有效值指标

系统发生故障后, 同调发电机群具有相似的时域响应趋势, 发电机功角曲线中含有大量的同调特征信息, 可以用有限次的三角函数多项式表达:

$$\tilde{\delta}(t) = b_0 + \sum_{k=1}^m [a_k \sin(2k\pi t) + b_k \cos(2k\pi t)] \quad (2)$$

式中, $\tilde{\delta}(t)$ 是功角拟合值; b_0 是恒定值; m 是拟合阶数; a_k 、 b_k 是第 k 阶三角函数的幅值。

大量实验证明: 拟合周期取 0.2 秒、拟合阶数取 2 时足以较好地拟合功角曲线。因此, 发电机 i 与 j 的功角差曲线可以表达为:

$$\begin{aligned} \Delta \tilde{\delta}_{ij}(t) &= \Delta b_{ij0} + \\ &\sum_{k=1}^2 [\Delta a_{ijk} \sin(2k\pi t) + \Delta b_{ijk} \cos(2k\pi t)] \end{aligned} \quad (3)$$

其中:

$$\left. \begin{aligned} \Delta b_{ij0} &= b_{i0} - b_{j0}, \\ \Delta a_{ijk} &= a_{ik} - a_{jk}, \\ \Delta b_{ijk} &= b_{ik} - b_{jk}. \end{aligned} \right\} \quad (4)$$

上述功角差拟合曲线是典型的非正弦周期信号, 其重要特征为有效值, 即功角差有效值:

$$EV(i, j) = \sqrt{\Delta b_{ij0}^2 + \sum_{k=1}^2 \left[\left(\frac{\Delta a_{ijk}}{\sqrt{2}} \right)^2 + \left(\frac{\Delta b_{ijk}}{\sqrt{2}} \right)^2 \right]} \quad (5)$$

若 $EV(i, j)$ 小于某一阈值, 发电机 i 与 j 同调。

1.3 基于广域量测信息的 CID 指标

采用上述功角差有效值可以快速得到同调分群结果, 但忽略了时序轨迹的细节特征。在已有实验中发现同调的机组相对于参考机组的功角曲线变化趋势相似, 即曲线的复杂度相近, 但同调机组之间的绝对欧式距离可能较大, 功角差有效值也较大。例如, 采用功角差有效值或欧氏距离计算下图所示的三条曲线的相似度, 得出 L2 与 L3 相似度较高、L1 与其他曲线相似度较低的结论, 但显然将 L2 与 L3 归为一群是不合理的, 原因是功角差有效值与欧氏距离都忽略了曲线的局部特征和运动趋势。

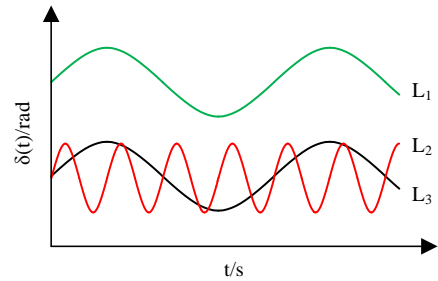


图 1 发电机转子角摇摆曲线示意图

Fig. 1 Swing curves of generator rotor angle

由于上述原因, 本文定义一种考虑发电机功角时序形态的相似度衡量标准, 即 CID。利用关于功角时序形态差异的信息作为相似度的校正因子, 在分类精度上产生显著的提高, 其简单、无参数的特性使这种改进并不损害频繁调用相似度计算的算法效率。定义发电机 i, j 的 CID 为:

$$CID(i, j) = ED(i, j) \times CF(i, j) \quad (6)$$

其中, $ED(i, j)$ 表示发电机 i, j 之间的距离函数, 反映距离特征; $CF(i, j)$ 表示发电机 i, j 之间的复杂度相关系数, 反映形态特征。

由于欧氏距离等传统距离函数在高维空间中的分辨能力和抗噪能力弱, $ED(i, j)$ 表达为一种适用于高维空间的相似度度量($Hsim$ 距离)的倒数:

$$ED(i, j) = \frac{n}{\sum_{i=1}^n \frac{1}{1 + |q_{ii} - q_{ji}|}} \quad (7)$$

$CF(i, j)$ 相当于是复杂度比值, 复杂度越不同, $CF(i, j)$ 越大, CID 就越大。本文将其表达为:

$$CF(i, j) = \frac{\max[CE(i), CE(j)]}{\min[CE(i), CE(j)]} \quad (8)$$

其中, 发电机 i, j 的特征复杂度表征为 $CE(i), CE(j)$ 。

在信息论中, 衡量一个时间序列的复杂度有许多方法, 本文提出一种直观性强、可解释性强且无参数的方法: 一个时间序列越复杂, 当我们把它“拉直”以后, 它就应该越长, 即:

$$CE(i) = \sqrt{\sum_{i=1}^{n-1} (q_i - q_{i+1})^2} \quad (9)$$

其中, q_i 表示功角时序特征的一个特征点。

依据 CID 分析图 1 中的功角曲线, L1 与 L3 归为一群、L2 单独一群, 符合同调性定义。然而, 由于时序的高维性, 良好的相似性度量可能不足以获得最佳的聚类结果, 需要构建与相似性度量兼容的深层特征空间。

2 自编码器与时空神经网络

本文算法的发电机功角特征提取层基于 AE 思想^[25-26], 设计一种由 BiLSTM 和 CNN 构成的新型时空神经网络(由于电力系统中发电机的同调性既包含时序特征、也具有空间关系, 本文据此设计的新型时空神经网络首先通过 CNN 提取发电机转子摇摆曲线的高维空间特征, 然后通过 BiLSTM 捕捉其长时依赖, 由此得到完整表征发电机同调性的信息)。因此, 本节分别介绍 AE、BiLSTM 和 CNN, 以便于下一节搭建完整框架。

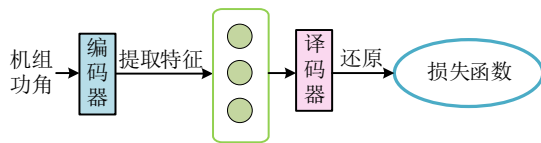


图 2 Auto-Encoder 结构

Fig. 2 Structure of Auto-Encoder

如图 2 所示, AE 思想将输入与输出看作同一层, 编码器提取特征, 解码器试图还原, 若能较好地还原, 说明提取到的是关键、有效的特征。本文采用 MSE 来衡量 AE 的输入和输出之间的差异。

如图 3 的 LSTM 能够学习点和点之间的时序相关性, 由此构建如图 4 的 BiLSTM 将时序特征双向输入并拼接输出, 从而提取时序上下文相关性特征。

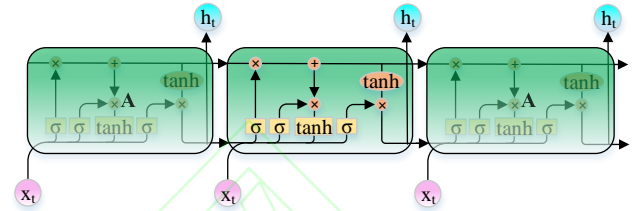


图 3 LSTM 结构

Fig. 3 Structure of LSTM

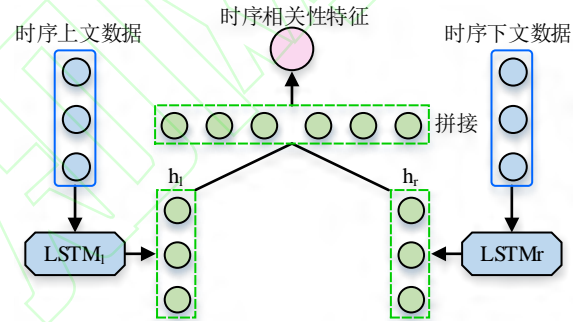


图 4 BiLSTM 结构图

Fig. 4 Structure of Bidirectional LSTM

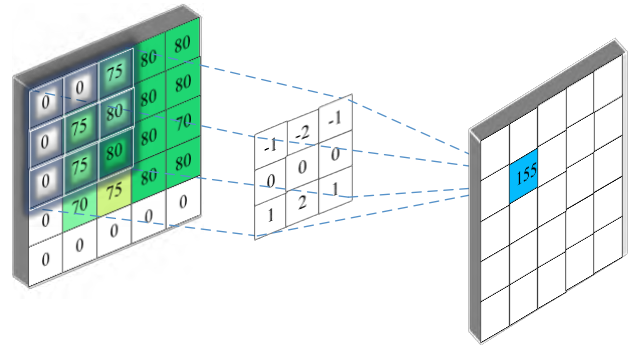


图 5 卷积运算原理图

Fig. 5 Schematic Diagram of Convolution Operation

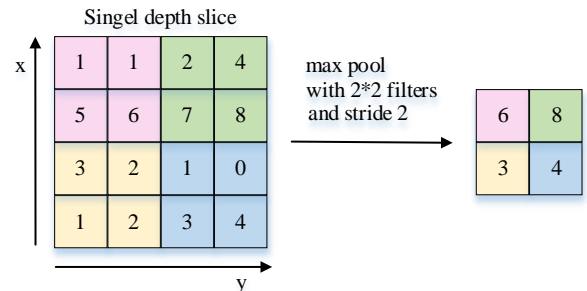


图 6 最大池化原理图

Fig. 6 Schematic Diagram of Max-pooling

本文对于时序短距离波动特征的提取中运用卷积和池化运算，简化了神经网络结构。如图 5 的卷积运算将发电机转子摇摆曲线视为平面进行特征提取，能够捕捉到不同角度的短距离波动特征，将数据拆解成若干边缘特征；如图 6 的池化运算本质上是对数据取最大值，提取区域最显著特征。

3 基于 IAE 的同调发电机识别模型

传统的 AE 运用朴素的全连接层或 LSTM，可以满足数据特征挖掘的准确性，但是无法挖掘高维度特征和复杂映射关系，也无法实现无监督的聚类。

文献[27]在 2018 年提出一种新颖的 AE 结构框架用于处理时序聚类问题，两个创新点是：

- 1) 构建编码器时将 LSTM 加入卷积与池化运算，其特征挖掘效果优越；
- 2) 聚类层与特征提取层联合训练，其特征空间倾向于适合聚类的表达方式(传统的独立训练方法会生成不利于时序分离的特征空间)。

本文设计的 IAE 算法针对同调发电机识别问题，基于文献[27]的结构框架进一步进行三处改进：

- 1) 针对同调发电机在线识别的无监督特性，聚类层推导出以 KL 散度为损失函数的 GMM 训练方法，与特征提取层联合训练，实现概率性数据驱动的无监督聚类，无需离线标注数据和预先确定组数；
- 2) 由于发电机同调性强调动态过程中的运动趋势相似性，聚类层输入特征前基于数据运动趋势采用简单高效的 CID 衡量发电机相似度，概率性驱动本身对噪声的敏感性不高，结合 CID 的抗噪特性，在实际系统的 PMU 噪声干扰下具备了强鲁棒性；
- 3) 实测功角中含有复杂信息和噪声，搭建 BiLSTM 加入卷积与池化运算，挖掘发电机时序相关性与序列内波动特征，与聚类层联合训练得到表达时序未标记数据分割成多个类别的潜在维度。
- 4) 基于类激活映射进行功角特征可视化，通过激活值归一化后绘制的热度图，使本文搭建的模型提取的高维隐向量具备了可解释性。

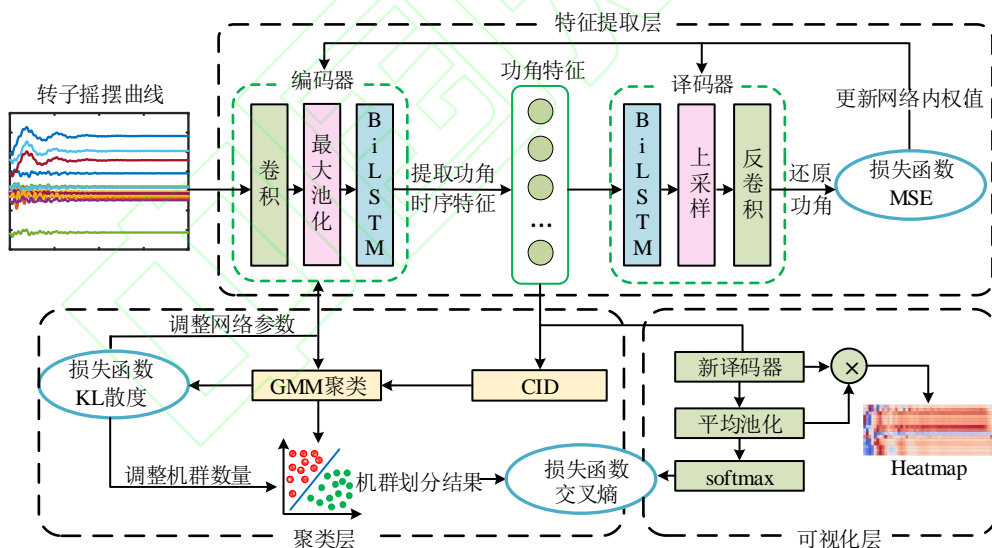


图 7 IAE 结构图

Fig. 7 Structure of IAE

3.1 数据处理与参数设置

发电机慢同调性是在区域间振荡模式(0.1 至 0.8Hz)下定义的，实际电力系统中 PMU 测量中出现的任何更高频率的成分(包括局部振荡、测量噪声)，都可能导致错误的发电机慢同调识别。本文模型输入数据之前加入随机噪声，将输出值与加噪声之前的时序数据进行比较，在 IAE 优越的联合训练机制

下，由此训练的模型具有抗噪声的能力。

IAE 对电力系统广域量测信息进行训练的过程本质上是利用损失函数求解优化问题。深度学习常用的优化算法包括 Momentum 算法、随机梯度下降算法、Nesterov Momentum 算法以及 Adam 算法等。其中，Adam 算法可以自适应地调节学习率，通过估计梯度的一阶矩和二阶矩，计算一阶矩和二阶矩

相除的值更新参数,具体机制可参见文献[28]。本文的 IAE 使用 Adam 优化算法来训练模型,根据先验知识,设置 Adam 基础步长(step size)为 0.001,卷积步长(stride)为 1,最大池化参数为 10,预训练轮次阈值为 300。联合训练的损失函数权重设置为 1:1,若连续 5 次两次迭代之间损失值减小量小于阈值($1e-10$)则停止迭代。对模型有较大影响的参数有:Bi-LSTM 的隐藏层节点个数,卷积池化运算的过滤器个数和批大小。本文使用高效、简便、非结构的网格搜索法^[29]来确定最优超参数组合,即训练停止时损失函数值最小的组合,此时隐藏层提取到最关键的发电机时序特征信息,进而能够得到最准确的聚类结果。选用以上方法的主要原因有:1)此方法计算简单,原理直观,可解释性强,适用于超参数空间离散且较小的情况。2)本文所涉及的数据特征明显、采样均衡,且设计的网络结构能够很好地提取特征并还原数据,在实际实验中仅取经验值即可得到较好的拟合效果,因此,本文网络对超参数不敏感,只需确定超参数的大致量级,而无需精细地确定其数值,这样只需从连续的超参数空间中择取一个较小的离散子空间进行搜索,适用网格搜索法。根据工程经验,分别设置 Bi-LSTM 中的两层隐藏层的搜索空间集合为: {256, 128, 64, 32}, {32, 16, 8, 4}, 卷积池化运算的过滤器个数的搜索空间集合为 {50, 100, 200, 400}, 批大小的搜索空间集合为 {1024, 512, 256, 128, 64}。

3.2 发电机时序特征提取

编码器将机组数 \times 时刻数的数据压缩成一个机组数 \times 特征数的数据,涵盖卷积运算提取的发电机时序内短距离波动特征、池化运算提取的短时间区域关键特征、BiLSTM 提取的时序相关性特征。

上文中编码器得到机组数 \times 特征数的数据,可以直接把每个机组映射到欧式空间中计算 CID,遗留的问题是映射参数的确定。针对以上训练问题,本文构造的译码器部分包括 BiLSTM、上采样层和反卷积,其中, BiLSTM 与编码器中结构一致,仅改变隐藏节点个数;上采样将特征维度还原,采用一种简单插值法,即在时间轴上将每个时间重复若干次;反卷积是一种特殊的卷积运算,从小样本扩大到大样本。步骤是:1)原序列先补 0 来扩大输入数据的尺寸;2)过滤器旋转,再进行常规卷积

运算。

经过以上运算,数据从原来的机组数 \times 特征数还原到机组数 \times 时刻数,比较前后的差异,即 MSE 损失函数,通过译码器的训练,编码器能够学习到映射的参数,与聚类层联合训练调参。

通过上文编码器映射得到的特征向量计算本文定义的 CID,表征发电机时序特征的相似度。

3.3 基于改进 GMM 的同调发电机聚类

上文通过提取的发电机转子摇摆曲线特征计算出机组之间的相似度指标(CID),以此进行机组的聚类。传统聚类常用算法是 k-means,需要预先确定聚类的组数,并且每个样本只能属于一个类,因此算法对初始值十分敏感,且难以评估聚类效果。在同调发电机的识别问题中,聚类技术普遍面临的问题都是如何确定最优的机群数。因此,本文进行概率性数据驱动的聚类,计算每个机组属于某一机群的概率与真实分布对比,从而计算 KL 散度(衡量两个分布之间差异性的标准),反向更改聚类层初始质心和编码器参数,通过网格搜索选取 KL 散度损失值最小的聚类方案,进而自动确定组数。

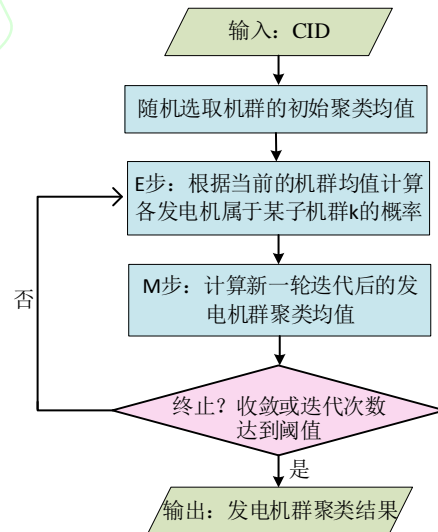


图 8 高斯混合模型算法流程图

Fig. 8 Flow chart of GMM algorithm

基于 GMM 和最大期望算法 (Expectation-maximization algorithm, EM),通过迭代最小化 KL 散度,搜索最优聚类结果,流程图如上所示,图中 E 步预测的发电机概率分布如下:

$$q_{ik} = \frac{\alpha_k \Phi(z_i | \theta_k)}{\sum_{k'=1}^K \alpha_{k'} \Phi(z_i | \theta_{k'})} \quad (10)$$

其中, q_{ik} 表示第 i 个发电机属于机群质心 k 的概率; $\Phi(z_i | \theta_k)$ 表示样本 i 在第 k 个高斯分布下的分布函数值; α_k 是高斯分布系数; z_i 是发电机 i 的隐变量; $\theta_k = (\mu_k, \sigma_k^2)$ 表示第 k 个机群的高斯分布参数。

M 步计算的新一轮迭代参数如下:

新一轮迭代中机群质心 k 的聚类均值为

$$\mu_k^{new} = \mu_k - \eta \frac{\partial Loss}{\partial \mu_k} \quad (11)$$

其中, η 是学习率。本文将损失函数 $Loss$ 定义为预测概率与真实估计概率的 KL 散度, 即

$$Loss = KL = \sum_{i=1}^n \sum_{k=1}^K p_{ik} \log \frac{p_{ik}}{q_{ik}} \quad (12)$$

由于 KL 散度需要将预测的概率分布与真实概率分布进行对比, 通常应用于有监督学习, 难以处理复杂、高阶的结构化、多特征、无标签数据, 本文通过定义辅助分布来估计真实概率分布, 进而作为模型学习的目标来计算 KL 散度, 实现将无监督学习转化为有监督学习, 辅助目标分布为:

$$p_{ik} = \frac{q_{ik}^2 / \sum_{i'=1}^N q_{i'k}}{\sum_{k'=1}^K \left(q_{ik'}^2 / \sum_{i'=1}^N q_{i'k'} \right)} \quad (13)$$

上述定义的优点如下: 1) 提高机群的纯度; 2) 对于置信度较高的机群划分给予较大的权重; 3) 避免尺寸较大的簇扭曲特征空间, 具体证明见附录。

协方差矩阵为

$$\Sigma_k = \frac{\sum_{i=1}^N q_{ik} (z_i - \mu_k^{new})(z_i - \mu_k^{new})^T}{\sum_{i=1}^N q_{ik}} \quad (14)$$

高斯分布系数为

$$\alpha_k = \frac{\sum_{i=1}^N q_{ik}}{N} \quad (15)$$

为了使 KL 散度最小, 关键是要求解损失函数对每个发电机群均值的偏导数, 即沿着负梯度的方

向改变均值。本文对此进行了复杂的重要推导, 最终结果如下, 便于应用于算法核心的训练过程中。

$$\begin{aligned} \frac{\partial KL}{\partial \mu_k} &= \frac{\partial}{\partial \mu_k} \sum_i \sum_{k'} p_{ik'} \log \frac{p_{ik'}}{q_{ik'}} \\ &= \sum_{k'} \sum_i \frac{\partial}{\partial \mu_k} \left(p_{ik'} \cdot \log \frac{p_{ik'}}{q_{ik'}} \right) \\ &= \sum_i (p_{ik} - q_{ik}) \cdot \Sigma^{-1}(z_i - \mu_k) \end{aligned} \quad (16)$$

通过逆向微调编码器的参数和发电机群质心, 与前文编码-译码的训练过程进行联合调节, 达到 KL 散度和 MSE 两个损失函数总和最小(连续 5 次两次迭代之间损失值减小量小于阈值)。

3.4 求解流程

本文算法求解的详细流程图如下所示。

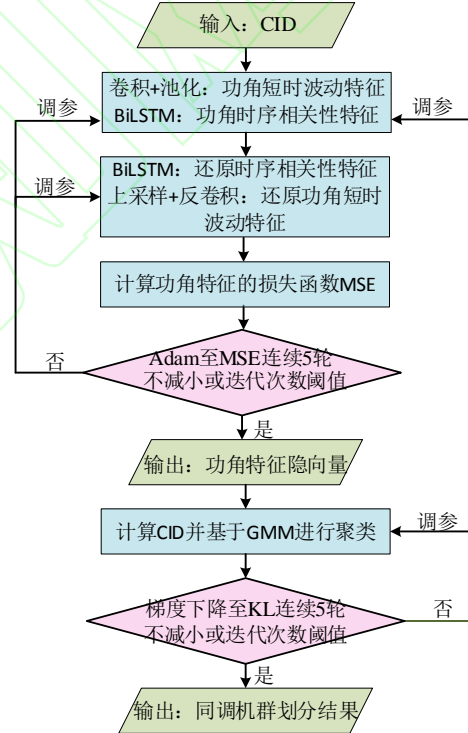


图 9 IAE 同调发电机识别算法流程图

Fig. 9 Flow chart of IAE coherency identification algorithm

本文在系统扰动后的准稳态和机电暂态期间对同调发电机进行分组的算法可以分为两部分:

1) 广域量测信息特征提取: 数据预处理后, 将加入随机噪声的数据输入到编码器结构中提取高维特征隐向量; 高维特征隐向量通过译码器结构还原为发电机功角数据, 与预处理前的发电机功角数据对比并计算损失函数 MSE; 根据 MSE 反向传播从而预训练编码器和译码器参数, 重复前述步骤,

引入 early stop 机制,连续 5 次两次迭代之间损失值减小量小于阈值(1e-10)后停止迭代,得到特征隐向量并根据式(6)至式(9)计算 CID 表征相似度。

2) 同调机群聚类:输入 CID 并根据式(10)计算各发电机属于初始机群质心的概率,根据式(13)至式(15)估计发电机的真实分布;根据式(12)计算 KL 散度并反向传播,沿着式(16)计算的梯度下降方向优化编码器和 GMM 的参数,根据式(11)计算迭代后的机群均值,实现聚类层与特征提取层联合训练,联合训练的损失函数权重设置为 1:1,若连续 5 次两次迭代之间损失值减小量小于阈值(1e-10)则停止迭代,输出同调机群构成。

3.5 基于 CAM 的功角特征可视化解释

本文将计算机视觉领域内前沿的 CAM^[31]迁移到发电机转子摇摆曲线特征提取中,基于迁移学习的思想构建热度图,实现了发电机功角特征的时序区间重要性可视化解释。具体推导过程如下:

首先,构建一个结构和译码器相同的神经网络,即堆叠双层 BiLSTM、上采样层和反卷积层,称为新译码器。将编码器已构建结构和已训练的权重参数直接迁移过来,形成旧编码器-新译码器的结构,其中,最后一个神经网络层即新译码器的反卷积层的卷积核个数和 IAE 确定的组数相同,其输出是一个维度为时刻数×组数的矩阵。再在新译码器的后面加上一个全局平均池化层(Global Average Pooling, GAP)和一个 softmax 层,整体构成一个分类器。将 IAE 得到的同调机群划分结果当作数据的标签输入此分类器在进行有监督学习,因为编码器中的权重是直接迁移得到的,分类器决策的依据是 GAP 的输出,即反卷积层学习出的特征图的均值,所以特征图可以表征对分类的贡献,又因为新译码器不改变原样本时序内部的顺序信息,所以特征图可按时序逐个映射回输入数据,因此可表示某样本在某时刻点数据对分类的贡献,也是其在 IAE 中对聚类的贡献,可以被提取出来并可视化,具体推导过程如下:对于拥有 τ 个时刻点的样本 i ,其数据被旧编码器映射到特征空间中,得到的特征向量为 $h_t(i)$,接着被新译码器还原为长度为 τ 的特征图 $\mathbf{v}=[v_1, v_2, \dots, v_\tau]=f(h_t(i))$,其中特征向量 $v_t(t=1,2,\dots,\tau)$ 的维度为 m , m 既是组数也是卷积核个数。再通过 GAP 后输出 $F_t(i)=\sum \mathbf{v}/m$,是一个维数为 $1 \times m$ 的向量。输

入 softmax 后得到输出 $P_c(i)=\exp(F_t(i))/\sum \exp(F_t(i))$,即样本 i 属于机群 c 的概率,计算出与样本的真实标签间的交叉熵作为分类器的损失函数。本文在此处忽略偏差项:特征偏差项设置为 0,因为它几乎对机群划分没有影响。由前述可知, $F_t(i)$ 即表示长度为 τ 的样本 i 对聚类贡献的权重,因此使用 $F_t(i)$ 和特征图 \mathbf{v}^T 进行点乘得到发电机 i 的 CAM,用 M_i 定义:

$$M_i = F_t(i) \mathbf{v}^T \quad (17)$$

其中, M_i 是一个长度为 τ 的向量,则 $M_{i,t}(t=1,2,\dots,\tau)$ 直接表明了把样本 i 的 t 时刻数据激活对发电机 i 划分为机群 c 的重要性。本文将其归一化处理后表示为热度图,每个发电机对应图中的一行。

4 仿真算例

通过编写 python 代码实现本文算法,在软件配置方面,本算例使用 Tensorflow 框架、Visual Studio 开发环境(包括代码编辑器、编译器、调试器等)。模型训练使用的 Linux 服务器硬件配置为 2.3GHz 八核 Intel Core i9 处理器、32GB 2667MHz DDR4 内存、ubuntu18.04 操作系统。

4.1 4 机 11 节点 2 区域 IEEE 测试系统

搭建 4 机 11 节点 2 区域电力系统如下图所示。

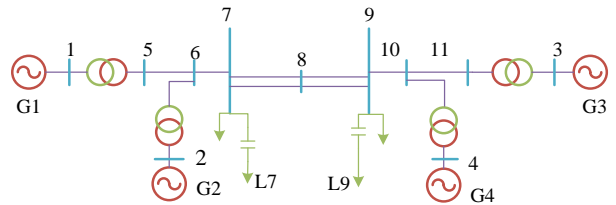


图 10 4 机 11 节点 2 区域电力系统示意图

Fig. 10 4-machine 11 bus 2-area power system

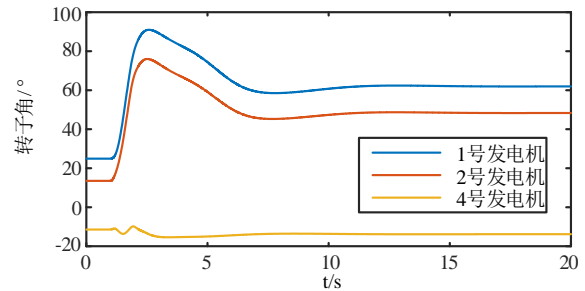


图 11 支路 8-9 单线故障时发电机转子角的摇摆曲线

Fig. 11 Swing curves of generator rotor angle under 8-9 line contingency

以发电机 G3 为参考机,在节点 8、9 之间的某

一交流支路靠近节点 8 侧设置三相永久性故障，故障发生时刻为 1s，故障持续时间 0.1s，0.1s 后故障线路切除，整个仿真时间持续 20s。

由摇摆曲线可知：该三相故障使系统中发电机聚合成 2 组同调机群，其中发电机 G1 和 G2 聚合为一组同调机群；发电机 G4 单独为一组同调机群。

采用本文算法，经过以下两步工作：

1) 功角特征提取：首先进行发电机 1、2、4 的功角预处理，将加入随机噪声的数据输入到编码器结构中提取高维特征隐向量；高维特征隐向量通过译码器结构还原为发电机功角数据，与预处理前的发电机功角数据对比并计算损失函数 MSE；根据 MSE 反向传播从而预训练编码器和译码器参数，重复前述步骤，连续 5 次两次迭代之间损失值减小量小于阈值($1e-10$)后停止迭代，此时 MSE 为 0.0095，说明此时发电机功角特征损失极小；根据式 (6) 至式 (9) 计算发电机相似度指标 CID。

2) 同调机群聚类：输入 CID 并根据式 (10) 计算各发电机属于初始机群质心的概率，根据式 (13) 至式 (15) 估计发电机的真实分布；根据式 (12) 计算 KL 散度并反向传播，沿着式 (16) 计算的梯度下降方向优化编码器和 GMM 的参数，根据式 (11) 计算迭代后的机群均值，实现聚类层与特征提取层联合训练，联合训练的损失函数权重设置为 1:1，连续 5 次两次迭代之间损失值减小量小于阈值($1e-10$)时停止迭代，此时 MSE 为 0.0132、KL 散度为 0.1051，输出同调机群为{1, 2}{4}。

为了解释模型决策结果，本文构建了发电机功角提取的高维特征向量可视化热度图，其颜色表示模型对特征的关注度，关注度经过归一化处理，数值越大则模型对区域特征越关注，自左而右沿时序提取特征，自上而下表示 G1、G2、G4。如下图所示，距离故障点较近的 G4 在故障发生后表现出短促的特征变化而受到模型关注，而 G1、G2 则是在系统拓扑变化后一段时间内都表现出剧烈的特征变化和高关注度，符合图 11 所示的摇摆曲线，当曲线趋于平稳时模型便逐渐不关注其特征。

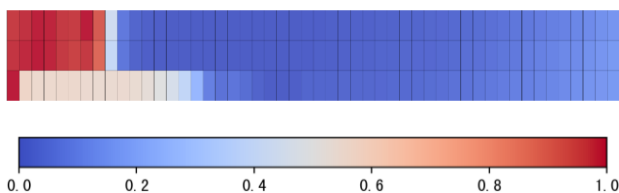


图 12 4 机算例转子角特征热度图

Fig. 12 Heatmap of 4-machine rotor angle features

由功角高维特征计算的 CID 如下表所示，CID 越小，发电机转子运动趋势越相似。例如，G1 与 G2 的 CID 远小于二者与 G4 的 CID，即 G1 与 G2 同调程度较高，与发电机转子角的摇摆曲线分析结论一致，说明模型关注的特征是准确的。

表 1 4 机算例 CID

Tab. 1 4-machine CID

	G1	G2	G4
G1	0	2366	92614
G2	2366	0	71149
G4	92614	71149	0

基于本文总结的深度学习方法进行对照，分别在本算例中使用 CNN(代表经典 FNN)搭建的 AE、LSTM(代表经典 RNN)搭建的 AE、堆栈自编码器(Stacked Auto-encoder, SAE)、变分自动编码器(Variational Auto-Encoder, VAE)和本文设计的 IAE。其中，上述每个模型的特征提取部分均迭代 300 次，CNN-AE 和 LSTM-AE 均采用对称结构，其超参数设置均与本文模型相应超参数一致。因为 IAE 中也采用了 CNN 层和 LSTM 层，使用相同的层结构能控制变量，实验结果更具可信度；SAE 也采用对称结构，编码器的层数为 3，与 IAE 相同，其隐藏层节点个数为{256, 128, 32}；VAE 的隐藏层节点个数为 256。上述超参数均由网格搜索法确定，其搜索空间与 IAE 相同，其训练时间、数据特征损失 MSE 和聚类分布损失 KL 散度如下表所示，可见本文算法的效率和精度明显高于其他算法。

表 2 4 机算例不同算法对比

Tab. 2 Identified 4-machine results for different algorithms

算法	训练时间/s	MSE	KL 散度
CNN-AE	15.6478	0.0782	0.4796
LSTM-AE	19.8652	0.0461	0.4601
SAE	20.6534	0.0580	0.4272
VAE	20.4122	0.0467	0.4192
IAE	6.6127	0.0132	0.1051

数据驱动方法的有效性经常受测量数据的噪声影响。根据文献[32]的实证实验，PMU 的噪声在不同电压下均可看作均值为 0 的高斯噪声，其信噪比可近似为 45dB。为了探究算法的抗噪能力，在四

机算例中加入高斯白噪声。以往研究中常采用 25dB 的信噪比^[18], 因此本文将信噪比设置为 45、25 和 20dB。以量测时间窗口为[1.2s, 6.2s]的发电机转子角作为输入, 辨识结果如下表所示。当信噪比为 45 或 25 时, CNN-AE 会将发电机 1、2、4 划分为一组同调机群, 不符合图 11 的发电机转子运动趋势, 而 LSTM-AE、SAE、VAE 和本文算法 IAE 的分群结果不受噪声影响; 当信噪比为 20 时, CNN-AE、LSTM-AE、SAE 和 VAE 都产生了划分错误, 而 IAE 的分群结果不受噪声影响, 以 1dB 为单位继续降低信噪比直至 6dB 时 IAE 才出现聚类困难, 验证了 IAE 在噪声干扰下的鲁棒性较好。

表 3 4 机算例加入高斯白噪声的辨识结果

Tab. 3 Identified 4-machine results for different algorithms

with white Gaussian noise			
算法	信噪比 45 辨识结果	信噪比 25 辨识结果	信噪比 20 辨识结果
CNN-AE	{1, 2, 4}	{1, 2, 4}	{1, 2, 4}
LSTM-AE	{1, 2}{4}	{1, 2}{4}	{1, 2, 4}
SAE	{1, 2}{4}	{1, 2}{4}	{1, 2, 4}
VAE	{1, 2}{4}	{1, 2}{4}	{1, 2, 4}
IAE	{1, 2}{4}	{1, 2}{4}	{1, 2}{4}

4.2 16 机 68 节点 5 区域 IEEE 测试系统

搭建 16 机 68 节点 5 区域电力系统如下图。

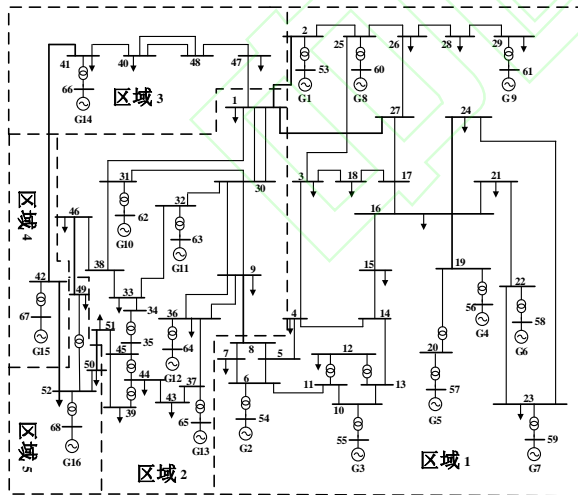


图 13 16 机 68 节点 5 区域电力系统示意图

Fig. 13 16-machine 68-bus 5-area power system

以发电机 G1 为参考机, 0.1s 在支路 41-42 近节点 41 侧设置三相短路故障, 0.15s 时节点 41 侧断路器跳开, 0.2s 时节点 42 侧断路器跳开, 整个仿真时间持续 20s, 采样频率为 100Hz。

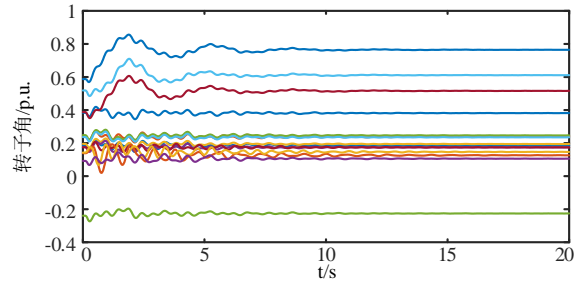


图 14 支路 41-42 单线故障时发电机转子角的摇摆曲线

Fig. 14 Swing curves of generator rotor angle under 41-42 line

contingency

采用本文算法, 经过以下两步工作:

1) 功角特征提取: 首先进行发电机 2 至 16 的功角预处理, 将加入随机噪声的数据输入到编码器结构中提取高维特征隐向量; 高维特征隐向量通过译码器结构还原为发电机功角数据, 与预处理前的发电机功角数据对比并计算损失函数 MSE; 根据 MSE 反向传播从而预训练编码器和译码器参数, 重复前述步骤, 连续 5 次两次迭代之间损失值减小量小于阈值($1e-10$)后停止迭代, 此时 MSE 为 0.0023, 说明此时发电机功角特征损失极小, 根据式 (6) 至 (9) 计算发电机相似度指标 CID, 如表 5 所示。

2) 同调机群聚类: 输入 CID 并根据式 (10) 计算各发电机属于初始机群质心的概率, 根据式 (13) 至式 (15) 估计发电机的真实分布; 根据式 (12) 计算 KL 散度并反向传播, 沿着式 (16) 计算的梯度下降方向优化编码器和 GMM 的参数, 根据式 (11) 计算迭代后的机群均值, 实现聚类层与特征提取层联合训练, 联合训练的损失函数权重设置为 1:1, 连续 5 次两次迭代之间损失值减小量小于阈值($1e-10$)时停止迭代, 此时 MSE 为 0.0028、KL 散度为 0.1335, 输出机群构成为 {2, 3, 4, 5, 6, 7, 8, 9}{10, 11, 12, 13}{14}{15}{16}, 采用式 (1) 进行时域计算验证本文划分的同调机群无误。

经过关注度归一化处理后的特征可视化热度图如下所示, 其热度定义与四机算例相同。

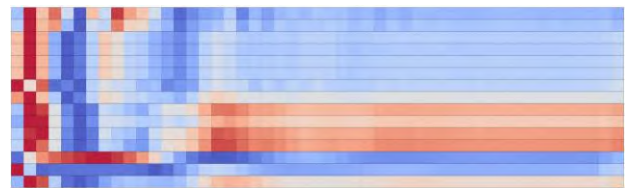


图 15 16 机算例转子角特征热度图

Fig. 15 Heatmap of 16-machine rotor angle features

表 4 16 机量测时间窗口对辨识结果的影响

Tab. 4 Coherent groups of 16-machine generators estimated with different sliding time windows

时间窗口	算法	辨识结果	MSE	KL 散度	训练时间/s
[0.2s, 5.2s]	CNN-AE	{2, 3, 4, 5, 6, 7, 8, 9}{10, 11, 12, 13}{14, 15, 16}	0.0073	0.6487	1.8874
	LSTM-AE	{2, 3, 4, 5, 6, 7, 8, 9}{10}{11}{12, 13}{14}{15}{16}	0.0042	0.6174	2.1452
	SAE	{2, 3, 4, 5, 6, 7, 8, 9}{10, 11, 12, 13}{14, 15, 16}	0.0068	0.7212	3.7523
	VAE	{2, 3, 4, 5, 6, 7, 9}{8}{10}{11, 12, 13}{14, 15, 16}	0.0065	0.6843	5.1288
	IAE	{2, 3, 4, 5, 6, 7, 8, 9}{10, 11, 12, 13}{14}{15}{16}	0.0035	0.1765	1.8573
[15s, 20s]	CNN-AE	{2, 3, 4, 5, 6, 7, 9}{8, 10}{11, 12, 13}{14, 15, 16}	0.0035	0.6876	3.2766
	LSTM-AE	{2, 3, 4, 5, 6, 7, 9}{8}{10}{11, 12, 13}{14, 15, 16}	0.0035	0.6845	2.6754
	SAE	{2, 3, 4, 5, 6, 7, 9}{8, 10}{11, 12, 13}{14, 15, 16}	0.0064	0.6734	4.1365
	VAE	{2, 3, 4, 5, 6, 7, 8, 9}{10, 11, 12, 13}{14, 15, 16}	0.0048	0.6145	4.2378
	IAE	{2, 3, 4, 5, 6, 7, 8, 9}{10, 11, 12, 13}{14}{15}{16}	0.0030	0.1330	1.1787
[0.2s, 20s]	CNN-AE	{2, 3, 4, 5, 6, 7, 9}{8}{10}{11, 12, 13}{14, 15, 16}	0.0034	0.6842	3.1287
	LSTM-AE	{2, 3, 4, 5, 6, 7, 8, 9}{10, 11, 12, 13}{14, 15, 16}	0.0031	0.6733	2.7361
	SAE	{2, 3, 4, 5, 6, 7, 9}{8}{10}{11, 12, 13}{14, 15, 16}	0.0067	0.7017	4.3476
	VAE	{2, 3, 4, 5, 6, 7, 8, 9}{10, 11, 12, 13}{14, 15, 16}	0.0059	0.6982	4.0281
	IAE	{2, 3, 4, 5, 6, 7, 8, 9}{10, 11, 12, 13}{14}{15}{16}	0.0028	0.1335	1.1506

表 5 16 机算例加入高斯白噪声的辨识结果

Tab. 5 Identified 16-machine results for different algorithms with white Gaussian noise

算法	信噪比为 25 的辨识结果	信噪比为 45 的辨识结果
CNN-AE	{2, 3, 5, 6, 9}{4, 7, 8}{10}{11, 12, 13}{14, 15, 16}	{2, 3, 5, 6, 7, 9}{4, 8}{10}{11, 12, 13}{14, 15, 16}
LSTM-AE	{2, 6, 7, 8, 9}{3, 4, 5}{10, 11, 12, 13}{14}{15}{16}	{2, 3, 4, 5, 6, 7, 8, 9}{10, 11, 12, 13}{14}{15}{16}
SAE	{2, 3}{4, 5, 6, 7, 8, 9}{10}{11, 12, 13}{14}{15, 16}	{2, 3}{4, 5, 6, 7, 8, 9}{10}{11, 12, 13}{14, 15, 16}
VAE	{2, 3, 5, 6, 9}{4, 7, 8}{10}{11, 12, 13}{14, 15, 16}	{2, 3, 4, 5, 6, 7, 8, 9}{10}{11, 12, 13}{14, 15, 16}
IAE	{2, 3, 4, 5, 6, 7, 8, 9}{10, 11, 12, 13}{14}{15}{16}	{2, 3, 4, 5, 6, 7, 8, 9}{10, 11, 12, 13}{14}{15}{16}

为了分析量测时间窗口和窗口长度对辨识结果的影响,表 4 给出以量测时间窗口为[0.2s, 5.2s]和[15s, 20s]、[0.2s, 20s]的发电机相对转子角作为输入,不同算法得到的辨识结果、训练时间、数据特征损失 MSE 和聚类分布损失 KL 散度。本文算法的辨识结果不发生改变,其他算法受量测时间窗口的位置与长度影响较大;本文算法的数据特征损失和聚类分布损失始终低于其他算法、训练时间较短。

在 16 机算例中,为了探究算法的抗噪能力,加入信噪比为 25 和 45 的高斯白噪声,以量测时间窗口为[0.2s, 5.2s]的发电机相对转子角作为输入,辨识结果如表 5 所示。CNN-AE、LSTM-AE、SAE 和 VAE 的同调机群辨识结果在加入噪声后都发生了改变,而本文算法不受噪声影响,验证

了复杂系统下本文算法仍对噪声不敏感,具有较高的鲁棒性。

为了评估深度学习算法提取特征的能力,将本文算法和传统降维方法 PCA、KPCA 和 ICA 进行对比。由于这些方法要求样本数大于降维后的特征维数,因此本文只针对 16 机算例进行对比。根据信息论的相关理论,本文采用共享信息(Mutual Information, MI)指标衡量降维后的数据和原数据的相关程度^[33]。将量测窗口为[0.2s, 20s]的 16 机样本采用上述方法降低到 8 维,计算其对应 MI,MI 越高则表明特征提取效果越好,其在相关问题上的有效性在文献[34]中有详细的说明。评估结果如下表所示,自编码器提取特征的能力显著高于传统方法。一方面,AE 以重构样本为目标,能最大化保留原数据信息;另一方面,AE

可拟合任意高阶高非线性函数，学习到传统方法无法挖掘的特征，从而在相同维度下更好地表征原数据。而且，其特征维度理论上可为任意实数，不受样本数量的限制，在实际应用中更能适应各种复杂工况。

表 6 不同方法的特征提取效果

Tab. 6 Performances on feature extraction of different approaches

算法	MI	算法	MI
PCA	1.3170	ICA	1.0884
KPCA	1.5084	IAE	4.1307

改变故障位置，在节点 1 至 2 之间、节点 8 至 9 之间、节点 46 至 49 之间设置三相短路故障，采用本文算法，同调机群的辨识结果如下表所示。

表 7 改变故障位置后同调机群

Tab. 7 Coherent groups of generators after changing the locations of contingency

故障线路	发电机号
1-2	{2, 3, 4, 5, 6, 7, 8, 9}{10}{11}{12, 13}{14}{15}{16}
8-9	{2, 3}{4, 5, 6, 7, 9}{8}{10}{11, 12, 13}{14}{15}{16}
46-49	{2, 3}{4, 5, 6, 7, 8, 9}{10, 11}{12, 13}{14}{15}{16}

由上表可以看出，由于不同故障激发出的主导振荡模式不同，可能激发出高频率的局部振荡模式，同调机群的辨识结果会产生差异。

5 实际算例

本节进一步验证 IAE 在实际电力系统中应用的可行性，基于南方电网某次故障下各变电站的

PMU 实测数据，研究对象包括位于云南的 LP、YS，位于贵州的 XR-I、XR-II，位于广东的 LD、SD、SZ、ZQ、HD、ZZ。已知云南电网与主网(贵州、广东)仅通过直流互联、二者实现异步运行，该故障导致多回直流换相失败，如下图所示，云南电网与主网之间的送电通道被短时切断后两区域内的功率平衡被打破，频率出现较为剧烈的变化。

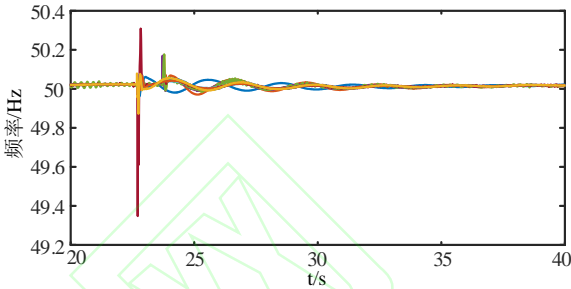


图 16 南方电网各变电站实测频率

Fig. 16 Actually measured frequency of transformer substitutions in CSG

采用本文辨识同调机群的方法，辨识得到位于云南的 LP、YS 和位于广东的 SD、HD 同调，位于贵州的 XR-I、XR-II 和位于广东的 LD 同调，位于广东的 SZ、ZQ、ZZ 同调。历史运行经验中表明南方电网主要存在以云广和云贵为主的区域间振荡模式，即处于云南电网内部的机组相对处于广东电网内部的机组振荡、处于云南电网内部的机组相对处于贵州电网内部的机组振荡，广东电网机组随故障位置不同存在与云南或贵州电网机组同调的可能性，初步符合实验结论。

表 8 南方电网算例 CID

Tab. 8 CID in CSG

	LP	YS	XR-I	XR-II	LD	SD	SZ	ZQ	HD	ZZ
LP	0	1.093	2.843	4.5867	3.26	0.994	26.95	1.242	0.957	3.686
YS	1.093	0	1.278	1.836	1.636	0.656	23.45	0.647	0.588	1.978
XR-I	2.843	1.278	0	0.5756	0.174	1.187	7.816	1.167	1.228	0.650
XR-II	4.586	1.836	0.575	0	0.639	1.742	4.560	1.735	1.791	0.690
LD	3.26	1.636	0.174	0.6393	0	1.486	7.248	1.452	1.546	0.650
SD	0.994	0.656	1.187	1.7423	1.486	0	18.06	0.190	0.082	1.183
SZ	26.95	23.45	7.816	4.5605	7.248	18.06	0	17.42	19.53	5.338
ZQ	1.242	0.647	1.167	1.7354	1.452	0.190	17.42	0	0.186	0.964
HD	0.957	0.588	1.228	1.7919	1.546	0.082	19.53	0.186	0	1.219
ZZ	3.686	1.978	0.650	0.6902	0.650	1.183	5.338	0.964	1.219	0

上表给出了聚类层计算的 CID 指标数据，可

见 LP 和 LD 在云广振荡中参与度极高，LP 和

XR-II 在云贵振荡中参与度极高, 因此: 虽然贵州和广东分别位于西电东送交流通道的送端和受端, 但 LD 在此时仍然被划入贵州同调机群。通过调整机群的聚类均值, 模型将机组分为五群: LP 单独一群, HD 单独一群, YS、SD 同调, XR-I、XR-II、LD 同调, 位于广东的 SZ、ZQ、ZZ 同调, 验证了上述分析。如果要改善系统在云广和云贵振荡模式下的阻尼特性, 应该优先考虑关于 LP 的控制措施。

6 结论

本文提出一种 IAE 算法新框架, 借助电力系统广域同步相量量测信息, 实现电力系统同调机群的快速准确辨识。首先, 搭建特征提取层(AE)将转子摇摆曲线蕴含的信息嵌入低维特征向量, 再用本文定义的 CID 表征机组相似度; 然后, 聚类层与特征提取层联合训练, 迭代优化 GMM 模型进行概率性数据驱动的同调机组辨识; 最后, 基于 CAM 和迁移学习知识进行特征重要性区间的可视化。将所提方法应用于典型 4 机 11 节点 2 区域电力系统、典型 16 机 68 节点 5 区域电力系统、中国南方电网进行分析, 结果表明:

1) 该方法实现简便, 完全由数据驱动, 故可迁移性强、适用范围广。与已有研究相比, 所提方法使用无监督学习即可准确确定初始聚类中心, 不依赖先验知识, 不需要离线标注数据, 无需人工确定机群数量, 能够自动寻找最优的聚类均值。

2) 相较于传统降维方法, 该方法的特征提取部分针对转子摇摆曲线的特性进行设计, 能够提取有效特征, 避免了数据挖掘的关键特征缺失, 克服了维数灾难等问题, 且无需任何人工计算。

3) 和其他 AE 架构相比, 本文方法使用了 Dropout 等机制避免过拟合, 其拟合准确度高、泛化能力强、抗噪能力强、鲁棒性强、对量测窗口敏感度低, 且基于 CAM 和迁移学习知识生成了表达特征重要性区间的热度图, 可解释性强。

本文将深度学习时空特征工程方法、迁移学习思想和概率驱动的机器学习聚类算法相结合, 为同调机群辨识问题提供可靠的工具。未来将继续深入研究在超大规模电网中本文算法的适用性, 以及同调辨识结果的更通用性评估标准。所研究结果对广域量测背景下的电力系统安全稳定在线监测与广域控制具有重要的借鉴意义。

参考文献

- [1] Mazhari S M , Safari N , Chung C Y , et al. A Quantile Regression-Based Approach for Online Probabilistic Prediction of Unstable Groups of Coherent Generators in Power Systems[J]. IEEE Transactions on Power Systems, 2018, PP(3):1-1.
- [2] Naglic M , Popov M , Mart A. M. M. van der Meijden, et al. Synchronized Measurement Technology Supported Online Generator Slow Coherency Identification and Adaptive Tracking[J]. IEEE Transactions on Smart Grid, 2020, 11(4):3405-3417.
- [3] Ariff M A M , Pal B C . Coherency Identification in Interconnected Power System—An Independent Component Analysis Approach[J]. IEEE Transactions on Power Systems, 2013, 28(2):1747-1755.
- [4] Khalil A M , Iravani R . A Dynamic Coherency Identification Method Based on Frequency Deviation Signals[J]. IEEE Transactions on Power Systems, 2016, 31(3):1779-1787.
- [5] 姜涛, 张明宇, 李雪, 陈厚合, 李国庆. 基于正交子空间投影的电力系统同调机群辨识[J]. 电工技术学报, 2018, 33(09): 2077-2088.
Jiang Tao, Zhang Mingyu, Li Xue, Chen Houhe, Li Guoqing, Estimating Coherent Generators from Measurement Responses in Power Systems Using Orthogonal Subspace Projection[J]. Transactions of China Electrotechnical Society, 2018, 33(09): 2077-2088.
- [6] 姜涛, 贾宏杰, 李国庆, 陈厚合, 靳小龙. 基于广域量测信息相关性的电力系统同调辨识[J]. 电工技术学报, 2017, 32(01): 1-11.
Jiang Tao, Jia Hongjie, Li Guoqing, Chen Houhe, Jin Xiaolong. Cross-Correlation Coefficient-Based Coherency Identification in Bulk Power System Using Wide-Area[J]. Transactions of China Electrotechnical Society, 2017, 32(01): 1-11.
- [7] 李雪, 姜涛, 陈厚合, 李国庆. 基于图分割的电力系统同调机群辨识新方法[J]. 中国电机工程学报, 2019, 39(23): 6815-6825+7095.
Li Xue, Jiang Tao, Chen Houhe, Li Guoqing. A Graph Cut Approach for Separating Coherent Groups of Generators in Bulk Power Grid Using Synchrophasors[J]. Proceedings of the CSEE, 2019, 39(23): 6815-6825+7095.
- [8] 姜涛, 黄河, 贾宏杰, 李晓东, 李鹏, 徐光虎. 基于投影寻踪最佳方向的同调机群识别方法[J]. 中国电机工程学报, 2015, 35(02): 359-367.
Jiang Taom Huang He, Jia Hongjie, Li Xiaodong, Li Peng, Xu Guanghu, Coherency Identification Through Projection Pursuit[J]. Proceedings of the CSEE, 2015,

- 35(02): 359-367.
- [9] 文俊, 刘天琪, 李兴源, 任景. 在线识别同调机群的优化支持向量机算法[J]. 中国电机工程学报, 2008(25): 80-85.
Wen Jun, Liu Tianqi, Li Xingyuan, Ren Jing. On-line Identification of Coherent Generator Using Optimized LS-SVM[J]. Proceedings of the CSEE, 2008(25): 80-85.
- [10] Wang M H, Chang H C. Novel clustering method for coherency identification using an artificial neural network[J]. IEEE Transactions on Power Systems, 1994, 9(4): 2056-2062.
- [11] Verma K, Niazi K R. Generator coherency determination in a smart grid using artificial neural network[C]// Power and Energy Society General Meeting, 2012 IEEE. IEEE, 2012.
- [12] Siddiqui S A, Verma K, Niazi K R, et al. Real-Time Identification of Generator Coherent Groups Through Synchrophasor Measurements and ANN[C]// India Conference. IEEE, 2016.
- [13] Paul A, Kamwa I, Joos G. PMU Signals Responses-Based RAS for Instability Mitigation Through On-The Fly Identification and Shedding of the Run-Away Generators[J]. IEEE Transactions on Power Systems, 2020, 35(3):1707-1717.
- [14] Guo T, Jovica V, Milanović. Online Identification of Power System Dynamic Signature Using PMU Measurements and Data Mining[J]. IEEE Transactions on Power Systems, 2016, 31(3):1-9.
- [15] Anaparthi K K, Chaudhuri B, Thornhill N F, et al. Coherency identification in power systems through principal component analysis[J]. IEEE Transactions on Power Systems, 2005, 20(3):1658-1660.
- [16] Babaei M, Muyeen S M, Islam S. Identification of Coherent Generators by Support Vector Clustering With an Embedding Strategy[J]. IEEE Access, 2019, PP(99): 1-1.
- [17] Lin Z, Wen F, Ding Y, et al. WAMS-based Coherency Detection for Situational Awareness in Power Systems with Renewables[J]. IEEE Transactions on Power Systems, 2018:1-1.
- [18] Ariff M, Pal B C. Coherency identification in interconnected power system - an independent component analysis approach[J]. IEEE Transactions on Power Systems, 2013, 28(2):1-1.
- [19] Rovnyak S M, Mei K. Dynamic event detection and location using wide area phasor measurements[J]. European Transactions on Electrical Power, 2011, 21(4):1589-1599.
- [20] Khalil A M, Iravani R. Power System Coherency Identification Under High Depth of Penetration of Wind Power[J]. IEEE Transactions on Power Systems, 2018.
- [21] 张亚洲, 张艳霞, 蒙高鹏, 赵冉, 高瑾. 基于广域信息的同调机群聚类识别方法[J]. 电网技术, 2015, 39(10): 2889-2893.
Zhang Yazhou, Zhang Yanxia, Meng Gaopeng, Zhao Ran, Gao Jin. A Wide Area Information Based Clustering Recognition Method of Coherent Generators[J]. Power System Technology, 2015, 39(10): 2889-2893.
- [22] 王甜婧, 汤涌, 郭强, 黄彦浩, 陈兴雷, 黄河凯. 基于知识经验和深度强化学习的大电网潮流计算收敛自动调整方法[J]. 中国电机工程学报, 2020, 40(08): 2396-2406.
Wang Tianjing Tang Yong, Guo Qiang, Huang Yanhao, Chen Xinglie, Huang Hekai. Automatic Adjustment Method of Power Flow Calculation Convergence for Large-scale Power Grid Based on Knowledge Experience and Deep Reinforcement Learning[J]. Proceedings of the CSEE, 2020, 40(08): 2396-2406.
- [23] 张怡, 张恒旭, 李常刚, 蒲天骄. 深度学习在电力系统频率分析与控制中的应用综述[J/OL]. 中国电机工程学报: 1-15[2020-12-09]. <https://doi.org/10.13334/j.0258-8013.pcsee.201377>.
Zhang Yi Zhang Hengxu Li Changgang Pu Tianjiao. Review on Deep Learning Applications in Power System Frequency Analysis and Control[J/OJ]. Proceedings of the CSEE: 1-15[2020-12-09]. <https://doi.org/10.13334/j.0258-8013.pcsee.201377>.
- [24] Batista, Gustavo & Wang, Xiaoyue & Keogh, Eamonn. A Complexity-Invariant Distance Measure for Time Series[C]. Proceedings of the 11th SIAM International Conference on Data Mining, SDM 2011. 699-710. 10.1137/1.9781611972818.60.
- [25] Hinton, G.E. & Salakhutdinov, R.R. Reducing the Dimensionality of Data with Neural Networks[J]. Science, 2006, PP(313): 504-507.
- [26] LeCun, Y., Bengio, Y. & Hinton, G. Deep learning[J]. Nature, 2015, PP(521): 436-444.
- [27] Madiraju N S, Sadat S M, Fisher D, et al. Deep Temporal Clustering: Fully unsupervised learning of time-domain features[J]. arXiv: Learning, 2018.
- [28] Kingma D, Ba J. Adam: A Method for Stochastic Optimization[J]. Computer Science, 2014.
- [29] Lecun Y, Bottou L, Orr G B. Neural Networks: Tricks of the Trade[J]. Canadian Journal of Anaesthesia, 2012, 41(7):658.
- [30] Xie J, Girshick R, Farhadi A. Unsupervised Deep Embedding for Clustering Analysis[J]. Computer Science, 2015.

- [31] Zhou B, Khosla A, Lapedriza A, et al. Learning Deep Features for Discriminative Localization[C]// CVPR. IEEE Computer Society, 2016.
- [32] Brown M, Biswal M, Brahma S, et al. Characterizing and quantifying noise in PMU data[C]// 2016 IEEE Power and Energy Society General Meeting (PESGM). IEEE, 2016.
- [33] Cover T, Thomas J, Wiley J. Elements of information theory[M]. Tsinghua University Press, 2003.
- [34] 唐振浩, 赵赓楠, 曹生现,等. 基于 SWLSTM 算法的超短期风向预测[J]. 中国电机工程学报, 2019, 39(15).
- Tang Zhenhao, Zhao Gengnan, Cao Shengxian, Zhao Bo. Very Short-term Wind Direction Prediction Via Self-tuning Wavelet Long-short Term Memory Neural Network[J]. Proceedings of the CSEE, 2019, 39(15).

附录 A

辅助的目标分布定义如正文中式(13)所示。

目标分布的第一个优点是可以提升同调机群的纯度,纯度的定义如式(1)所示,是一种常见的聚类效果外部评价指标,评估了每个簇(同调机群)中含有多少种样本(发电机组),样本的种类由其真实的聚类标签决定。

$$\frac{1}{N} \sum_{l \in L} \max_{k \in K} |l \cap k| \quad (1)$$

其中, L 表示广域量测信息的真实标签集合, K 表示同调机组聚类的簇集合, N 表示发电机样本个数。

将同调机群纯度简述为: 如果发电机 i 距离机群 k 不足够近, 则需将 i 排除在 k 之外。本文中定义的目标分布可以被转化为:

$$P_{ik} = \frac{q_{ik}^2 / \sum_{i=1}^N q_{ik}}{q_{ik}^2 / \sum_{i=1}^N q_{ik} + \sum_{k'=1, k' \neq k}^K (q_{ik'}^2 / \sum_{i=1}^N q_{ik'})} \quad (2)$$

$$= \frac{1}{1 + \frac{\sum_{k'=1, k' \neq k}^K (q_{ik'}^2 / \sum_{i=1}^N q_{ik'})}{q_{ik}^2 / \sum_{i=1}^N q_{ik}}}$$

为简化问题, 此处仅考虑机群数为 2, 即 $K=2$ 的情况, 两个机群分别用 k 和 h 表示。上式(2)可以转化为:

$$P_{ik} = \frac{1}{1 + \frac{q_{ih}^2 / \sum_{i=1}^N q_{ih}}{q_{ik}^2 / \sum_{i=1}^N q_{ik}}} \quad (3)$$

考虑发电机样本 i 距机群 k 和 h 的距离较为接近的情况:

$$P_{ik} \approx \frac{1}{1 + \frac{\sum_{i=1}^N q_{ik}}{\sum_{i=1}^N q_{ih}}} \quad (4)$$

由于此情况下发电机样本 i 相对机群 k 不足够近, 即其他发电机样本整体和簇 k 更加接近, 即 $p_{ik} < 0.5$, 发电机样本 i 会被倾向于分配到机群 h 中, 这样可以将距离某个机群 k

不足够近的发电机样本 i 划分到其他机群中, 维持了机群 k 的纯度。

第二个优点是如果发电机样本 i 距离机群 k 足够近, 则给予此次概率预测以较高的权重, 保证发电机 i 能被划分到机群 k 内, 式(2)可以被转化为式(5):

$$P_{ik} = q_{ik} \frac{q_{ik} / \sum_{i=1}^N q_{ik}}{q_{ik}^2 / \sum_{i=1}^N q_{ik} + \sum_{k'=1, k' \neq k}^K (q_{ik'}^2 / \sum_{i=1}^N q_{ik'})} \quad (5)$$

其中, q_{ik} 的系数可以表示为:

$$\frac{q_{ik} / \sum_{i=1}^N q_{ik}}{q_{ik}^2 / \sum_{i=1}^N q_{ik} + q_{ih}^2 / \sum_{i=1}^N q_{ih}} = \frac{1}{q_{ik} + \frac{q_{ih}^2}{q_{ik}} \cdot \frac{\sum_{i=1}^N q_{ik}}{\sum_{i=1}^N q_{ih}}} = \frac{1}{q_{ik} + q_{ih} \frac{q_{ih}}{q_{ik}} \cdot \frac{\sum_{i=1}^N q_{ik}}{\sum_{i=1}^N q_{ih}}} \quad (6)$$

由于 $q_{ik} + q_{ih} = 1$, 仅考虑式(6)分母的第二项中 q_{ih} 的系数, 可转换成式(7):

$$\frac{q_{ih}}{q_{ik}} \cdot \frac{\sum_{i=1}^N q_{ik}}{\sum_{i=1}^N q_{ih}} = \frac{1 + \sum_{i=1, i \neq i}^N \frac{q_{ik}}{q_{ik}}}{1 + \sum_{i=1, i \neq i}^N \frac{q_{ih}}{q_{ih}}} \quad (7)$$

由于发电机样本 i 距离机群 k 足够近, 则距离机群 h 足够远, 因此发电机样本 i 相对其他样本属于机群 k 的概率更大, 属于机群 h 的概率更小, 因此有 $q_{ih} < q_{i'h}$ 和 $q_{ik} > q_{i'k}$, 其中, $i' = 1, 2, \dots, N, i' \neq i$ 。式(7)的值小于 1, 即式(6)的分母小于 1, 即公式(5)中为 q_{ik} 的权重大于 1, 相当于提高了发电机样本 i 属于机群 k 的概率, 保证了预测概率置信度较高的发电机样本 i 更可能被划分到机群 k 中。

第三个优点是避免出现过大的机群, 如果机群 h 中已有较多的发电机样本, 则倾向于将相对于其他发电机样本更远离 h 的发电机样本 i 分配到较小的机群 k 中。在此种情况下, 一定有 $q_{ih} > q_{ik}$, 且 $q_{ih} < q_{i'h}$, $q_{ik} > q_{i'k}$, 其中 $i' = 1, 2, \dots, N, i' \neq i$ 。

因此可得式(8):

$$\sum_{i=1, i \neq i}^N \frac{q_{ih}}{q_{ih}} > \sum_{i=1, i \neq i}^N \frac{q_{ik}}{q_{ik}} \Rightarrow \sum_{i=1, i \neq i}^N \frac{q_{ih}}{q_{ih}} + 1 > \sum_{i=1, i \neq i}^N \frac{q_{ik}}{q_{ik}} + 1 \Rightarrow \sum_{i=1, i \neq i}^N \frac{q_{ih}}{q_{ih}} > \sum_{i=1, i \neq i}^N \frac{q_{ik}}{q_{ik}} \Rightarrow \frac{\sum_{i=1, i \neq i}^N q_{ih}}{q_{ih}} > \frac{\sum_{i=1, i \neq i}^N q_{ik}}{q_{ik}} \Rightarrow \frac{q_{ih}}{\sum_{i=1}^N q_{ih}} < \frac{q_{ik}}{\sum_{i=1}^N q_{ik}} \quad (8)$$

式(2)可被改写为式(9):

$$P_{ik} = \frac{q_{ik}^2 / \sum_{i=1}^N q_{ik}}{q_{ik}^2 / \sum_{i=1}^N q_{ik} + q_{ih}^2 / \sum_{i=1}^N q_{ih}} \quad (9)$$

此时, 虽然 $q_{ih} > q_{ik}$, 但是发电机样本 i 的两机群概率之差是全样本中最小的, 因此有 $q_{ih} - q_{ik} < \xi$, 其中 ξ 是一个较小量, 故可将式(9)放大为式(10), 则:

$$\begin{aligned}
 p_{ik} &\approx \frac{q_{ik}^2 / \sum_{i=1}^N q_{ik}}{q_{ik}^2 / \sum_{i=1}^N q_{ik} + q_{ih} q_{ik} / \sum_{i=1}^N q_{ih}} \\
 &= \frac{q_{ik} / \sum_{i=1}^N q_{ik}}{q_{ik} / \sum_{i=1}^N q_{ik} + q_{ih} / \sum_{i=1}^N q_{ih}}
 \end{aligned} \quad (10)$$

由(8)可知, 式(10)的值大于 0.5, 即 $p_{ik} > p_{ih}$, 即发电机样本 i 应被分到更小的机群 k 以维护机群的尺寸相对平衡。

至此, 本文定义的辅助目标分布所具备的三个优点均证明完毕。

作者简介:



刘丰瑞(1998), 男, 硕士研究生, 研究方向为电力系统安全性与稳定性, lfr0007@aliyun.com;

*通信作者: 李柏橦(1998), 女, 硕士, 研究方向为人工智能在电力系统中的应用, bl745@cornell.edu。