

基于深度信念网络的计量装置故障溯源研究

李 宁¹, 费守江², 刘国亮¹, 杨 琳¹

(1. 国网新疆电力科学研究院, 新疆 乌鲁木齐 830011;
2. 国网新疆巴州供电公司, 新疆 库尔勒 841000)

摘 要: 针对电网海量大数据中存在异常的电表数据, 提出了一种基于深度信念网络的计量装置故障溯源模型。首先, 分析了深度信念网络(DBN)模型基本原理, 提出了一种智能电表故障分类 DBN 结构模型, 并给出了计量装置故障溯源建模流程; 然后, 通过建立离线台账样本库、实时用电曲线故障特征样本库, 实现了计量装置故障样本库构建; 最后, 基于实际电表测试和数据异常识别, 完成了计量装置的台账样本库溯源测试, 并与已有的贝叶斯、决策树、随机森林、决策树提升等分类算法进行比较, 测试结果验证了所提模型和方法的正确性和有效性。

关键词: 深度信念网络; 受限玻尔兹曼机; 计量装置; 故障溯源

中图分类号: TM715

文献标志码: A

文章编号: 2096-4145(2020)07-0118-07

Fault Traceability of Metering Device Based on Deep Belief Network

LI Ning¹, FEI Shoujiang², LIU Guoliang¹, YANG Lin¹

(1. State Grid Xinjiang Electric Power Research Institute, Urumqi 830011, China; 2. State Grid Xinjiang Bazhou Power Supply Company, Kuerle 841000, China)

Abstract: Aiming at the abnormal electricity meter data in massive big data of power grid, a fault traceability model of metering device based on deep belief network is proposed. Firstly, the basic principle of deep belief network (DBN) model is analyzed. A DBN network structure model for intelligent meter fault classification is proposed. And the process of fault traceability modeling of metering device is given. Then, by establishing off-line account sample bank and real-time power curve fault feature sample bank, the fault sample bank of metering device is constructed. Based on actual meter testing and data anomaly recognition, the traceability test of the accounting sample database of the metering device have been completed. Compared with the existing classification algorithms such as Bayes, Decision Tree, Random Forest and Adaboost, the test results verify the correctness and effectiveness of the model and method proposed in this paper.

Key words: deep belief network; restricted Boltzmann machine; metering device; fault traceability

0 引言

电力系统在实际运行过程中产生海量数据, 其高效采集及使用是智能电网大数据亟待解决的一个重要问题^[1-3]。截至 2018 年 1 月, 国家电网公司智能电表及用电信息采集已经覆盖经营范围内 99% 的电力用户, 智能电表所采集到的时间序列数据中包含了电力用户丰富的用电行为信息和特征, 从而为电力系统各项工作的建模与预测提供了科

学的数据库与指导依据, 有助于实现电网企业为电力用户提供个性化可靠供电服务^[4-5]。

供电企业通过对用户电量的测量、采集和汇总进行抄核收工作。若不能及时处理用户电量异常数据, 会产生极其严重的后果。监察人员在用电监察中对用户的实际用电量没有认真核对, 收取电费时则会出现误收, 供电企业或用电用户遭受损失, 甚至引发两者之间的矛盾, 影响企业的社会形象。同时, 用采异常数据也会对配电网运营监测产生很大影响, 例如台区线损率、负载率等指标的计算, 严重影响指标的相关考核, 给运营监测单位的工作带来阻力^[6-7]。因此, 需要对用采数据进行数据清洗, 通过数据清洗筛出异常数据, 从而提高数据的质量和可

基金项目: 国家重点研发计划资助项目(2016YFB0901100)

Project Supported by National Key R & D Project of China (2016YFB0901100)

信度,利于供电企业进行抄核收工作,减少用户和供电企业经济损失。

常用的异常点检测方法主要有基于统计的检测算法^[8]、基于距离或密度的检测算法^[9]以及后来发展的神经网络检测方法^[10]、支持向量的方法^[11]以及聚类分析方法^[12]等。对异常数据产生的原因进行针对性地分析,对存在异常点的电表进行故障溯源,有利于对发生隐蔽故障的电表进行查漏,有利于运营监测中心日常工作的开展,具有工程应用价值。在故障溯源方面,文献[13]研究了海量网管数据快速溯源方法。文献[14]采用故障树方法,对4种故障模式进行了压力故障原因研究。文献[15]通过构建集成化的故障溯源模型,分析复杂机电系统故障溯源。文献[16]分析了电网动态仿真误差溯源方法。文献[17]构建了电力系统设备供应链溯源管理评价体系。在当前的异常点检测和故障溯源算法方面,人工神经网络要优化的目标函数复杂,处理电网大数据存在效率问题。支持向量等传统机器学习方法数据训练时间较长,处理海量数据样本同样面临严峻挑战,尤其对高维特征数据会出现维度灾难。而深度学习作为机器学习神经网络算法的扩展与延伸,能有效处理海量数据分类与回归处理时的高维场景问题^[18-19]。针对传统方法在处理海量数据时处理速度慢、占用内存大、难以处理高维特征等问题,本文在异常数据甄别的基础上,使用深度信念网络(Deep Belief Network, DBN)模型对台账数据样本库进行学习;对训练完成的模型输入样本库所包含的属性数据,实现对计量装置运行状态正常与异常的判断;将检测出异常的电表通过用电特征样本库进行进一步故障溯源,最终得到精确的计量装置故障溯源结果。

1 深度信念网络模型基本原理

DBN是由多个受限玻尔兹曼机(Restricted Boltzmann Machine, RBM)堆叠组成的神经网络模型,其核心是RBM单元。图1为由3层RBM组成的简单DBN结构。其中V为连接观测数据的可见层;H是用于提取输入数据有效特征的隐含层; W 为可见层与隐含层的连接权重^[20]。

RBM是一种基于能量的模型,对于状态 (v, h) 确定的RBM系统所具有的能量可表示为:

$$E(v, h|\theta) = -\sum_{i=1}^n a_i v_i - \sum_{j=1}^m b_j h_j - \sum_{i=1}^n \sum_{j=1}^m v_i w_{ij} h_j \quad (1)$$

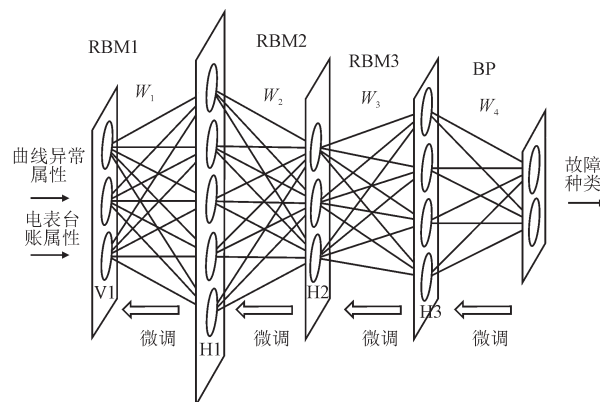


图1 DBN网络结构模型

Fig.1 DBN network structure model

式中: $\theta=(w_{ij}, a_i, b_j)$ 为RBM参数; n, m 分别为可见层与隐含层神经元数量; v 为可见层输入,其表示异常电表的台账信息和电流功率特征等输入; v_i 为可见层神经元 i 的状态,设定其偏置值为 a_i ; h 为隐含层输出,输出为0/1,表示电表未发生故障或电表发生故障; h_j 为隐含层神经元 j 的状态,设定偏置值为 b_j ;神经元 i 和 j 连接权重定义为 w_{ij} 。

由能量函数,可得到 (v, h) 的联合概率分布为:

$$p(v, h|\theta) = \frac{1}{Z(\theta)} \exp(-E(v, h|\theta)) \quad (2)$$

式中: $Z(\theta)$ 为归一化因子,其表达式为:

$$Z(\theta) = \sum_v \sum_h \exp(-E(v, h|\theta)) \quad (3)$$

当学习样本的数量为 N 时,参数 θ 通过学习样本的最大对数似然函数 L 得到,公式如下^[20]:

$$\theta^* = \arg \max_{\theta} L(\theta) = \arg \max_{\theta} \sum_{n=1}^N \log p(v^n|\theta) \quad (4)$$

式中: $p(v|\theta)$ 为观测数据 v 的似然函数,其表达式为:

$$p(v|\theta) = \frac{1}{Z(\theta)} \sum_h \exp(-E(v, h|\theta)) \quad (5)$$

根据可见层神经元状态计算隐含层第 j 个神经元,激活概率为:

$$p(h_j = 1|v, \theta) = \frac{1}{1 + \exp(-b_j - \sum_i v_i w_{ij})} \quad (6)$$

由隐含层重建可见层第 i 个神经元,设定 p 为激活的概率,公式为:

$$p(v_i = 1|h, \theta) = \frac{1}{1 + \exp(-a_i - \sum_j w_{ij} h_j)} \quad (7)$$

由随机梯度上升法求解最大对数似然函数值,各参数变化量计算准则为^[20]:

$$\begin{cases} \Delta w_{ij} = \langle v_i h_j \rangle_{\text{data}} - \langle v_i h_j \rangle_{\text{recon}} \\ \Delta a_i = \langle v_i \rangle_{\text{data}} - \langle v_i \rangle_{\text{recon}} \\ \Delta b_j = \langle h_j \rangle_{\text{data}} - \langle h_j \rangle_{\text{recon}} \end{cases} \quad (8)$$

式中: $\langle \cdot \rangle_{\text{data}}$ 为原始观测数据模型定义分布; $\langle \cdot \rangle_{\text{recon}}$ 为重构后模型定义分布。

考虑学习率 ε 的参数更新准则为:

$$\begin{cases} w_{ij}^{k+1} = w_{ij}^k + \varepsilon \Delta w_{ij} \\ a_i^{k+1} = a_i^k + \varepsilon \Delta a_i \\ b_j^{k+1} = b_j^k + \varepsilon \Delta b_j \end{cases} \quad (9)$$

2 计量装置故障样本库构建

2.1 离线台账样本库

首先需要对计量装置正常工作影响因素进行分析,确定数据挖掘的维度。主要考虑用电环境与气象因素、智能电表的不同类型、运行时长、工程质量、历史用电曲线数据5个方面。前4个方面属于台账数据,历史用电曲线数据属于用电信息采集系统实时数据,为了使模型有充足的样本进行学习,曲线长度取1个月,每日96点。

用电环境与气象因素会对计量装置的可靠性产生显著影响,甚至会引起电表的故障,例如运行环境中含氯化物盐的大气的腐蚀所导致的电能表异常、运行中外力损毁、雷击等^[21]。由于计量装置供应商较多,不同计量装置的内部设计和零件也有所区别,因此计量装置的故障往往呈现出家族性缺陷。同一厂家同一批次的计量装置发生相同类型的故障可能性较大。运行时长的增加会使计量装置的运行可靠性下降,计量装置有一定的预期寿命,随着运行时长的增加,其零部件的老化和电池的损耗都会导致电表更容易发生故障。例如误差超差、电容损坏、电池损坏等硬件部分的损坏。工程质量问题也会影响电表的运行状况,工程质量问题分为电表质量和安装质量问题,包括虚焊、表内存在异物、焊点短路、参数设置不正确、进出线反接或相别错误。

基于上述对计量装置故障影响因素的分析,选择多种台账属性作为特征向量,数据由2类组成:1)用电环境与气象因素分量。针对用电环境与气象因素可选取使用地区、行业类别等维度;2)不同类型智能电表分量。对智能电表分量选择3个维度,分别为表记类别、生产厂商、标定电流。选取营销业务系统及用电信息采集系统的8000条智能电表数据作为样本,经过筛选后剩余6688条数据作为样本。其

中样本数据包括正常电表和故障电表的各项属性值,并对使用地区、表记类别、生产厂家、行业类别、标定电流等台账属性进行数据采集和整理。

2.2 实时用电曲线故障特征样本库

计量装置的历史运行数据反映了电表的运行健康程度,对于正常运行的智能电表,计量部门能够正常采集到其电流、电压、有功功率、无功功率、抄表电量等数据,这些运行数据反映了装置的运行状态。而具有较大潜在故障发生可能性的装置,往往处于较隐蔽的不正常运行状态,其运行数据与原本正常运行数据存在差异,但区别较小,不能达到计量中心故障诊断标准。因此,对实时运行数据的异常指标提取可以作为对计量装置运行状态健康程度的评价^[22]。

为了使模型有充足的样本进行学习,曲线长度取1个月,每日96点,曲线类别为电流、电压、功率。由于计量装置采集的电流、功率数据量大,若将曲线数据直接作为模型属性会导致维度过大,直接影响模型的运算时长与溯源精度。同时,短期曲线数据不能反映装置在较长时间内的运行状态。为了尽可能得到实时用电曲线数据的整体情况,便于模型对特征进行深度提取,采用卷积神经网络对电流与功率进行特征提取。

典型特征提取卷积神经网络如图2所示,其包含3层网络结构:输入层,卷积层和输出层^[23]。

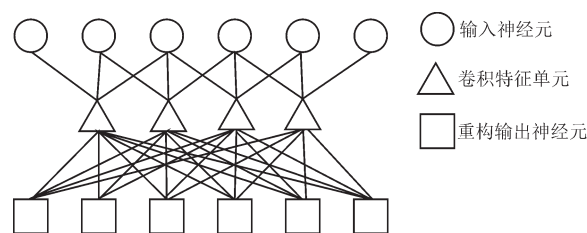


图2 卷积神经网络结构图

Fig.2 Structural diagram of convolutional neural network

其特征选取步骤如下:

(1)在输入层设置96个输入节点,并接收输入为1天96点的电流或功率序列。特别地,对于故障电表,输入的电流或功率为故障前1天的序列。

(2)在卷积层使用1个1维卷积核进行特征提取,提取后的特征为1条24点的特征序列。卷积核的特征提取具有局部稀疏连接及位移不变性的特点,可以很好地反映原始序列的关键特征。

(3)在输出层将上一层的特征序列通过全连接方式连接到输出层,并且输出1条96点的序列,以实现原始电流或功率序列的重构。如果卷积提

取的特征能够较好地反映原始电流或功率序列,则重构误差较小。因此,特征提取神经网络使用重构的均方误差作为目标函数,使用反向传播进行神经网络的参数训练,误差函数 E 为:

$$E = \frac{1}{96} \sum_{i=1}^{96} (y_i - o_i)^2 \tag{10}$$

式中: y_i 为实际标签,其表示原始的 96 点电流或功率序列; o_i 为模型输出,代表神经网络重构的电流或功率序列。

神经网络训练完毕后,重构误差较小,认为此时提取的特征具有典型性;将卷积层提取到的特征输出,即历史电流曲线、功率曲线浓缩成 2 条 24 点的特征序列。

通过卷积神经网络提取故障电表故障前的电流和功率特征,集中为两种类型,如图 3 所示的两类特征簇。其中一类簇在序列首段集中为负值,另一类在序列首段为较大的正值,提取出的两类特征簇为电表发生故障前的电流、功率特征。

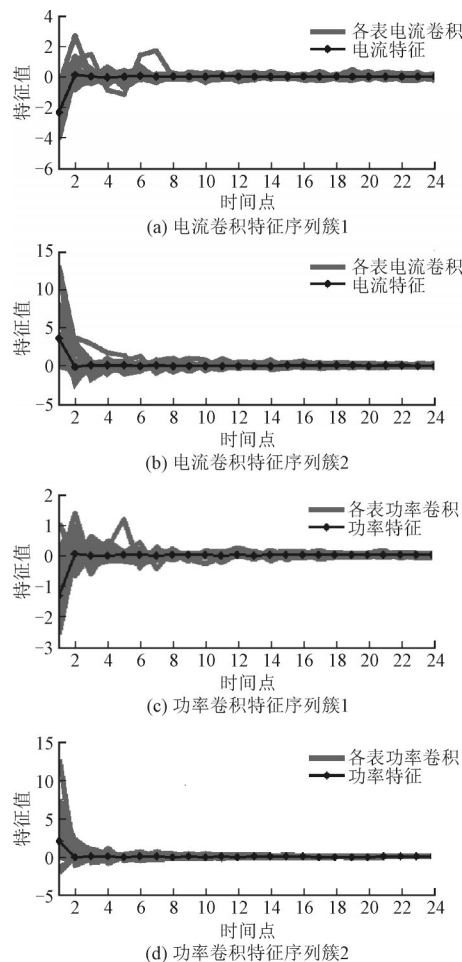


图3 卷积网络提取的特征结果

Fig.3 Feature result of convolutional network extraction
将故障电表的电流特征与功率特征分别归结

为两类,构建用电曲线故障特征样本库。将目标电表的用电曲线特征与故障特征进行比对时,使用余弦相似度对两种特征的相似性进行评价。余弦相似度 $D_{\cos i}$ 反映了用电曲线故障特征曲线 x 和目标电表用电曲线特征 y 的变化趋势间的相似性,其值越接近 1 表示越相似。

$$D_{\cos i} = \frac{\sum_{k=1}^m x_k y_k}{\sqrt{\sum_{k=1}^m x_k^2 \sum_{k=1}^m y_k^2}} \tag{11}$$

式中: m 为用电曲线特征向量长度。本文中 $m=24$; x_k, y_k 分别为用电故障特征曲线和目标电表用电特征曲线第 k 个点的特征分量。

3 计量装置故障溯源测试

3.1 实际电表测试

3.1.1 电压测试

首先测试了已知故障电表和未发现故障电表的一相电压数据。其中,选取 100 个电表进行测试,总测试点个数为 50 416。通过算法测试,最终寻找到电压异常点个数为 267 个,占测试点个数的比例为 0.530%。部分测试电压异常的结果如表 1 所示(若电压值为 NULL,属于采集缺失,不判断为异常点)。

表1 部分测试电压异常点列表
Table 1 List of abnormal points of partial test voltage

序号	数据日期	异常点的编号	异常数值/V	异常原因
1	2017-5-1	11	227.2	电压突变
2	2017-5-2	40	227.1	电压突变
3	2017-5-2	42	218.1	电压突变
4	2017-5-2	75	227.8	电压突变
5	2017-5-1	33	225.7	电压突变
6	2017-5-4	75	229.3	电压突变
7	2017-5-5	34	225.7	电压突变
8	2017-5-10	2	228.9	电压突变
9	2017-5-10	34	224.6	电压突变
10	2017-5-13	73	222.4	电压突变

3.1.2 电流与功率测试

对电流和功率数据进行异常值检测研究,表 2 和表 3 分别为部分测试电流和功率异常点列表,测试电表包括已知故障电表和未发现故障电表。

表2 部分测试电流异常点列表
Table 2 List of abnormal points of partial test current

序号	数据日期	异常点的编号	异常数值/A
1	2017-5-7	37	0.064
2	2017-5-14	37	0.068
3	2017-6-4	75	0.104
4	2017-5-1	35	2.136
5	2017-5-1	36	0.331
6	2017-5-2	55	2.631
7	2017-5-2	56	0.902
8	2017-5-6	36	3.107
9	2017-5-7	71	2.095
10	2017-5-8	29	2.657

表3 部分测试有功功率异常点列表
Table 3 List of abnormal points of partial test active power

序号	数据日期	异常点的编号	异常数值/kW
1	2017-5-4	3	0.191 6
2	2017-5-5	4	0.150 1
3	2017-5-5	5	0.044
4	2017-5-5	6	0.147 4
5	2017-5-7	7	0.044 4
6	2017-5-8	3	0.159 4
7	2017-5-8	4	0.066 5
8	2017-5-9	6	0.160 9
9	2017-6-4	4	0.185 5
10	2017-6-4	5	0.079 3

3.2 台账样本库溯源测试

测试分析使用 DBN 模型对台账样本库进行学习,输出测试集电表的运行状态,实现台账样本库溯源。将通过台账样本库判定为故障的电表输入故障曲线特征样本库,计算测试集电表的运行曲线特征与故障曲线特征的相似程度,根据结果判定电表是否发生故障,实现用电曲线故障特征溯源。根据最终溯源判定结果检测该方法的准确度。

3.2.1 台账样本库溯源

使用 DBN 模型学习台账样本库,按照十折交叉验证的方法随机抽取测试集进行测试。设置参数如下:建立由 RBM1 和 RBM2 组成的双隐含层 DBN 网络结构,如表 4 所示。

表4 DBN 参数
Table 4 DBN parameters

指标	参数
结构参数	5-2-2-1
RBM1 隐藏神经元	2
RBM2 隐藏神经元	2
DBN 输出层神经元	1
RBM 学习率	0.01
迭代次数	1200

采用相关指标比较 DBN 算法与常见分类算法贝叶斯 (Bayes)、决策树 (Decision Tree)、随机森林 (Random Forest)、决策树提升算法 (Adaboost1)、逻辑

分类 (Logistic) 进行比较,比较方法和指标。

针对分类验证方式采用十折交叉验证,其主要目的是为了防止模型出现过度拟合训练样本,而失去对其他数据的拟合和预测能力。通过交叉验证选出最好的分类模型,减少模型的过拟合,增强模型对于未知数据的适应能力。具体实现流程为:1) 将数据集随机平分为 10 份;2) 依次将其中 9 份作为训练数据,剩下的 1 份作为测试数据,进行试验;3) 求出 10 次结果的正确率;4) 进行多次十折交叉验证,以其结果的平均值作为对算法准确性的估计。

此外,其评价指标有分类正确率、kappa 系数、均方根误差、相对绝对误差以及模型训练时间等。

kappa 系数是一种衡量分类精度的指标,其计算公式为:

$$k = \frac{p_0 - p_e}{1 - p_e} \quad (12)$$

式中: p_0 为每一类正确分类的样本数值之和除以总样本数即总体分类精度; p_e 为参考系数。

假定各种真值样本数分别为 a_1, a_2, \dots, a_c , 预测出来的各种样本数分别为 b_1, b_2, \dots, b_c , 总样本数为 n , 则有:

$$p_e = \frac{a_1 \times b_1 + a_2 \times b_2 + \dots + a_c \times b_c}{n \times n} \quad (13)$$

均方根误差计算公式为:

$$E_{\text{RMS}} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (14)$$

式中: n 为测试电表总数目; y_i, \hat{y}_i 分别为第 i 个测试电表的真实分类情况和预测分类情况。 y_i 和 \hat{y}_i 的值为 0/1, 即电表故障/非故障。

观测者操作特性曲线 (Receiver Operating Characteristic Curve, ROC) 反映敏感性和特异性连续变量的综合指标, 曲线下面积越大, 诊断准确性越高。在 ROC 曲线上, 最靠近坐标图左上方的点为敏感性和特异性均较高的临界值。它的横坐标表示假正类率, 纵坐标表示真正类率^[24]。ROC 测试结果如表 5 所示。

表5 多种算法分类结果比较
Table 5 Comparison of classification results of various algorithms

使用方法	正确率/%	kappa 系数	均方根误差	ROC 值
DBN	87.78	0.754 3	0.123 9	0.942 5
Bayes	70.00	0.397	0.305 6	0.795
Decision Tree	80.00	0.599	0.227 5	0.912 5
Random Forest	81.67	0.625 9	0.230 4	0.913 3
Adaboost1	82.78	0.655 6	0.192 6	0.904 3
Logistic	84.45	0.686 6	0.154 3	0.904 2

在分类算法评价中, kappa 系数越接近 1 代表效果越好; 均方根误差越小, 代表分类结果更加精确; ROC 曲线值越大代表模型在不同数据下的表现更稳定。从表 5 可知, DBN 算法的正确率高于其余算法, 其 kappa 系数、均方根误差以及 ROC 曲线值的计算结果具有明显的优势。因此, DBN 算法与 Bayes、Decision Tree、Random Forest、Adaboost1、Logistic 算法相比, 在计量装置故障溯源方面更具有工程实用性。在此基础上, 将需要进行故障判定的电表台账数据输入到 DBN 模型中, 输出电表是否故障, 以此作为故障溯源的第一阶段。

3.2.2 用电曲线故障特征溯源

选取 40 个未发生故障的电表、40 个异常电表的台账数据作为测试集, 通过 DBN 对该测试集电表进行测试。提取判定为故障电表的用电曲线数据, 将其输入用电曲线故障特征样本库, 进行故障溯源的第二阶段。用电曲线属性为经过特征提取的电流、功率序列。使用余弦相似度进行相似度计算, 将用电特征曲线与故障特征曲线相似度高于 0.8 的电表确定为故障电表。测试电表 ID 及故障溯源结果如表 6 所示。

表 6 电表溯源测试结果

Table 6 Traceability test result of ammeter

电表ID	实际状态	电表溯源判断结果
2649932	已故障	故障
3029047	已故障	故障
3073657	已故障	故障
1466252	已故障	故障
4105960	已故障	故障
5685009	已故障	故障
4115661	已故障	正常
2745000	已故障	故障
2649932	已故障	故障
3029047	已故障	正常
3073657	已故障	故障
1466252	已故障	故障
4105960	已故障	故障
5685009	已故障	故障
2445834	已故障	正常
2553738	已故障	故障
2573443	已故障	故障
0405580	已故障	故障
9315635	已故障	故障
8959274	已故障	故障
1610308	未故障	故障
1061017	未故障	正常
1072716	未故障	正常
1110705	未故障	正常
1072722	未故障	正常
1006895	未故障	正常

由表 6 的测试结果可知, 对于 3.1 节中已经测试出有异常点的智能电表, 故障溯源测试正确率为 85.7%, 具有较高的检测精度。

4 结论

本文采用深度信念网络对计量装置故障溯源开展研究, 主要结论如下:

(1) 使用卷积神经网络对历史电流、功率曲线序列进行特征提取, 构建台账数据样本库与用电曲线故障特征样本库, 并使用深度信念网络算法学习台账样本库, 输出该计量装置的运行状况, 可实现异常电表故障溯源。

(2) 采用异常数据序列、异常故障样本、营销数据及用采系统数据训练深度信念网络结构参数, 基于训练完成的网络结构, 实现了计量装置状态判断, 从而可为计量装置轮换周期及检修策略提供科学合理的指导。结果表明, 对于台账样本库溯源测试正确率为 87.78%; 对于用电曲线故障特征溯源测试, 已测试出有异常点的智能电表的检测正确率为 85.7%。

参考文献

- [1] 耿俊成, 马文栋, 郭志民, 等. 基于智能电表大数据分析的台区变压器三相不平衡治理[J]. 智慧电力, 2018, 46(9): 103-108.
GENG Juncheng, MA Wendong, GUO Zhimin, et al. Three phase imbalance control of distribution transformer based on big data analysis of smart meters[J]. Smart Power, 2018, 46(9): 103-108.
- [2] 王赛一, 余建平, 孙丰杰, 等. 电力大数据的价值密度评价及结合改进 k-means 的提升方法研究[J]. 智慧电力, 2019, 47(3): 8-15.
WANG Saiyi, YU Jianping, SUN Fengjie, et al. Evaluation and promotion methods with improved k-means for value density of electric power big data[J]. Smart Power, 2019, 47(3): 8-15.
- [3] 潘森, 王栋霞, 乔俊峰, 等. 面向海量用电信息采集数据的质量分析方法研究[J]. 供用电, 2018, 35(3): 37-41.
PAN Sen, WANG Dong, QIAO Junfeng, et al. Research on quality analysis method for massive data from power consumption information collection system[J]. Distribution & utilization, 2018, 35(3): 37-41.
- [4] 赵永良, 付鑫. 大数据与智能客服的融合应用实例[J]. 供用电, 2018, 35(6): 72-76.
ZHAO Qingzhou, FU Xin, TIAN Shiming, et al. An example of integrated application of big data and intelligent customer service[J]. Distribution & utilization, 2018, 35(6): 72-76.
- [5] 赵一鸣, 谢炯, 陈晓宇, 等. 基于大数据的电力舆情管家系统的设计与实现[J]. 电力信息与通信技术, 2019, 17(8): 31-36.
ZHAO Yining, XIE Jiong, CHEN Xiaoyu, et al. Design and implementation of power public opinion butler system based on big data[J]. Electric Power Information and Communication Technology, 2019, 17(8): 31-36.

- [6] 王林信, 杨鹏, 江元, 等. 智能电网大数据隐私保护技术与实现[J]. 电力信息与通信技术, 2019, 17(12): 24-30.
WANG Linxin, YANG Peng, JIANG Yuan, et al. Research and implementation of privacy protection technology for big data in smart grid[J]. Electric Power Information and Communication Technology, 2019, 17(8): 31-36.
- [7] 陈钦柱, 符传福, 韩来君. 智能电网大数据分析决策系统的研究[J]. 电子设计工程, 2020, 28(5): 30-34.
CHEN Qinzhu, FU Chuanfu, HAN Laijun. Research on big data analysis and decision making system for smart grid[J]. Electronic Design Engineering, 2020, 28(5): 30-34.
- [8] BARNET V, LEWIS T. Outlier in statistical data[M]. New York: John Wiley & Sons, 1994.
- [9] RAMASWAMY S, RASTOGI R, SHIM K. Efficient algorithms for mining outlier from large data sets[C]//Proc of the ACM SIGMOD Int Conf on Management of Data. New York: ACM, 2000: 427-438.
- [10] MARKOU M, SINGH S. Novelty detection: A review—part 2: neural network based approaches[J]. Signal Processing, 2003, 83(12): 2499-2521.
- [11] MOURAO-MIRANDA J, HARDOON D R, HAHN T, et al. Patient classification as an outlier detection problem: An application of the one-class support vector machine[J]. Neuroimage, 2011, 58(3): 793-804.
- [12] WANG J S, CHIANG J C. A cluster validity measure with outlier detection for support vector clustering[J]. IEEE Transactions on System Man and Cybernetics, Part B—Cybernetics, 2008, 38(1): 78-89.
- [13] 陈墨, 金磊, 龚向阳, 等. 面向5G海量网管数据的故障溯源技术[J]. 北京邮电大学学报, 2018, 41(5): 96-100.
CHEN Mo, JIN Lei, GONG Xiangyang, et al. Research on fault tracing technology for 5G mass network management data[J]. Journal of Beijing University of Posts and Telecommunications, 2018, 41(5): 96-100.
- [14] 张根保, 张恒, 庞继红, 等. 数控机床液压系统故障溯源及分析[J]. 计算机应用研究, 2012, 29(4): 1356-1370.
ZHANG Genbao, ZHANG Heng, PANG Jihong, et al. Fault trace and analysis of hydraulic system attached to CNC machine tools[J]. Application Research of Computers, 2012, 29(4): 1356-1370.
- [15] 姜洪权, 王金字, 高建民, 等. 面向复杂系统故障溯源的SDG-FG模型建模方法[J]. 计算机集成制造系统, 2015, 21(3): 749-757.
JIANG Hongquan, WANG Jinyu, GAO Jianmin, et al. Modeling method for SDG-FG model oriented to fault root caused tracing analysis of complex electromechanical system[J]. Computer Integrated Manufacturing Systems, 2015, 21(3): 749-757.
- [16] 安军, 王孜航, 穆钢, 等. 基于WAMS测量和戴维南等值的电力系统动态仿真误差溯源及可信度验证方法[J]. 电网技术, 2013, 37(5): 1389-1394.
AN Jun, WANG Zihang, MU Gang, et al. Confidence level verification and tracking to error source for power grid dynamic simulation based on wams and thevenin equivalence[J]. Power System Technology, 2013, 37(5): 1389-1394.
- [17] 郭伟, 冯曙明, 汪伦. 基于大数据的电网设备供应链溯源管理评价研究[J]. 电力大数据, 2018, 21(9): 20-26.
GUO Wei, FENG Shuming, WANG Lun. Research on traceability management evaluation of power grid equipment supply chain based on big data[J]. Power Systems and Big Data, 2018, 21(9): 20-26.
- [18] 蓝天宝, 张志海, 蔡晓兰, 等. 基于不良数据挖掘的电力通信传输网实时态势感知与评估[J]. 电子设计工程, 2020, 28(4): 163-166, 171.
LAN Tianbao, ZHANG Zhihai, CAI Xiaolan, et al. Real-time situation awareness and assessment of power communication transmission network based on bad data recognition[J]. Electronic Design Engineering, 2020, 28(4): 163-166, 171.
- [19] 易灵芝, 常峰铭, 龙谷宗, 等. 基于进化深度学习短期负荷预测的应用研究[J]. 电力系统及其自动化学报, 2020, 32(3): 1-6, 13.
YI Lingzhi, CHANG Fengming, LONG Guzong, et al. Real-time situation awareness and assessment of power communication transmission network based on bad data recognition[J]. Proceedings of the CSU-EPSA, 2020, 32(3): 1-6, 13.
- [20] 陈伟, 何家欢, 裴喜平. 深度信念网络在电能质量复合扰动识别中的应用[J]. 电力系统及其自动化学报, 2018, 30(9): 75-82.
CHEN Wei, HE Jiahuan, PEI Xiping. Application of deep belief network in power quality compound disturbance identification[J]. Proceedings of the CSU-EPSA, 2018, 30(9): 75-82.
- [21] 石鑫, 朱永利, 萨初日拉, 等. 基于深度信念网络的电力变压器故障分类建模[J]. 电力系统保护与控制, 2016, 44(1): 71-76.
SHI Xin, ZHU Yongli, SA Churila, et al. Power transformer fault classifying model based on deep belief network[J]. Power System Protection and Control, 2016, 44(1): 71-76.
- [22] 肖桂雨, 向健平, 凌永志, 等. 基于小波神经网络的风力发电机故障预测方法[J]. 电力科学与技术学报, 2019, 34(2): 195-202.
XIAO Guiyu, XIANG Jianping, LING Yongzhi, et al. Prediction of wind turbine faults based on wavelet neural networks[J]. Journal of Electric Power Science and Technology, 2019, 34(2): 195-202.
- [23] 王宏刚, 田洪迅, 李浩松, 等. 考虑小样本统计的BP神经网络配电系统可靠性预测方法[J]. 电力科学与技术学报, 2019, 34(2): 40-46.
WANG Honggang, TIAN Hongxun, LI Haosong, et al. Reliability prediction method of power system based on a BP neural network considering small samples[J]. Journal of Electric Power Science and Technology, 2019, 34(2): 40-46.
- [24] 黄南天, 杨学航, 蔡国伟, 等. 采用非平衡小样本数据的风机主轴轴承故障深度对抗诊断[J]. 中国电机工程学报, 2020, 40(2): 563-573.
HUANG Nantian, YANG Xuehang, CAI Guowei, et al. A deep adversarial diagnosis method for wind turbine main bearing fault with imbalanced small sample scenarios[J]. Proceeding of the CSEE, 2020, 40(2): 563-573.

(责任编辑 徐秋芳)

收稿日期: 2019-03-25; 修改日期: 2020-05-20

作者简介: 李 宁(1968), 男, 新疆乌鲁木齐人, 高级工程师, 研究方向为电力营销、低压线损、电力计量、用电信息采集等。