

DOI:10.12158/j.2096-3203.2021.06.001

基于强化学习的异常用电判决方法

蔡云芹, 王非

(华中科技大学电信学院, 湖北 武汉 430074)

摘要:目前异常用电检测问题有许多基于分类的方法,但大多都是基于短期用电行为的判决来判断长期用电行为,判决阈值与比例难以确定,且在实际应用中,不同区域、时段的用户用电数据分布差异较大,比例与阈值也会有较大的不同,难以以固定的比例通用于所有的用户数据。针对此问题,文中提出一种基于强化学习的异常用电判决方法,创新地利用强化学习模型来动态生成阈值,以适应差异较大的不同数据集。首先获取分类器输出的数个用户短期行为的异常概率,然后输入到强化学习模型深度递归Q网络(DRQN)中,学习得到动态阈值即判决阈值与判决比例。试验结果表明,相比于人工调参的传统投票法,文中方法在评估指标上有明显提升,面对数据分布差异较大的数据集时也有较好的表现,说明文中方法具有较强的泛化能力,在数据类型复杂的现实环境中也有较好的应用场景。

关键词:智能电网;强化学习;神经网络;异常用电;动态阈值;深度Q网络;异常检测

中图分类号:TM933

文献标志码:A

文章编号:2096-3203(2021)06-0002-07

0 引言

随着电力系统信息化程度的不断提高^[1],通过用电数据来判断客户是否有异常用电行为已成为常用的方法。计量装置故障或用户窃电行为,均会导致采集到的数据是人为构造的虚假数据,这些虚假数据被称为异常用电数据。异常用电数据会对供电公司供电秩序造成严重干扰和影响,因此及时排查异常用电情况一直是电力公司关注的重点问题,异常用电的检测方法也受到广大学者关注。

目前学术界所研究的异常用电检测方法可以大致分为3类:基于系统状态的异常用电检测^[2-4],基于博弈论的异常用电检测^[5-7]和基于数据驱动的异常用电检测^[8]。其中,基于系统状态方法的理论基础为电网的物理本质决定了系统电压、注入功率等物理量的量测结果应具有一致性^[9]。该方法计算量大,所需数据多,且由于配电网的某些连接会发生变化,参数未必一成不变,存在收敛性问题。博弈论的方法是基于窃电用户的决策集与正常用户不同,最终影响双方在电量分布上的不同。但博弈论的方法只经过了理论推导和仿真,尚未经过实证检验。

在当前海量数据环境下,基于数据驱动的异常用电检测方法应用广泛,该方法基本可分为3类:分类、回归和聚类,其中分类方法应用最多。国内外学者广泛使用了许多基于机器学习的分类算法,主

要包括决策树^[10]、支持向量机^[11-13]、K近邻^[14]和人工神经网络^[15]等方法。文献[2,5,16]分别提出了一些基于深度学习的异常用电检测方法。这些研究成果的共同之处在于它们都是基于样本的判决,单样本覆盖的时间周期较短。然而,要对一个用户的用电行为是否存在异常进行判决,必须进行更长周期的分析。因此,上述方法面临巨大的挑战,要么由于时间周期过长,数据维度太大而无法完成分析,要么由于异常比例在用户间参差不齐,难以完成判决。

针对此问题,文中提出一种针对用户级别的基于强化学习的异常用电判决方法。文中专注于判决阈值与判决比例的设计,首先分析了基于分类器方法的用户级别判决的难点,然后提出了一种基于强化学习的异常用电判决模型,最后阐明试验结果与结论。

1 用户长周期用电数据判决问题分析

在实际工程应用中,用户级别的异常用电检测是非常重要的,而目前的研究往往都是在样本级别上的判决。通常单样本覆盖的时间周期较短,如果想要对用户的长周期数据进行判决,直接将长周期(一个月乃至一年)的用电数据输入到分类器中是不合适的。因为在长期的用电数据中,正常用电的数据比例较多,其中少部分的异常用电数据相当于噪声,会严重影响分类器输出的分类结果。且长期的用电数据直接输入到分类器中,意味着检测周期长,判决用户是否有异常用电行为的延迟长,不符合实际应用需求场景。

收稿日期:2021-05-31;修回日期:2021-07-21

基金项目:国家自然科学基金资助项目(62172169)

通常的方法是将长周期分成数个短周期,得到数个短周期的判决结果,再使用简单投票法,得到用户级别的判决结果。这种方法虽简单直观,却也存在着许多问题。

(1) 不同用户的有效用电数据可能长度差别较大,其异常比例会有较大区别。

例如一个新用户只有 20 d 的用电数据,而使用的投票阈值是针对老用户 100 d 的用电数据,新用户刚好这 20 d 数据出现了几天的异常,其异常比例相对较高,就超过了投票阈值,被判决为异常。但实际上通过后续观察发现,该用户后期的用电数据都是正常数据,并不是异常用电用户。这就是由于待测用户用电数据的有效长度差异较大,而投票阈值固定所引起的误判。

(2) 不同地区甚至是不同时间段的用户用电数据分布差别较大,异常比例在用户间参差不齐,较难以单一的固定票数比例去完成所有用户的判决。

例如用户数据显示,A 地异常用户的异常用电天数占总用电天数比例偏高,而 B 地异常用户的异常用电天数占用电天数比例偏低,则适用于 A 地区的投票阈值直接套用在 B 地区上,会有大量的异常用电用户检测不出,引起较大的误判。

为了解决简单投票法存在的上述各种问题,文中提出用动态的判决阈值与判决比例去做用户级别的异常用电判决。具体方法是建立一个基于深度递归 Q 网络(deep recurrent Q network, DRQN)的判决阈值优化模型,来动态优化判决阈值与判决比例。

2 基于 DRQN 的异常用电判决模型

2.1 强化学习算法 DRQN

Q 学习(Q-Learning)算法属于强化学习算法中的基础算法,深度 Q 网络(deep Q network, DQN)算法则是基于 Q-Learning 的一种流行的神经网络模型,DRQN 算法则是在 DQN 算法上做了改进。

在传统的 Q-Learning 算法中,智能体通过探索未知环境,得到每一个状态 s 和在这个 s 下每个行为 a 所拥有的 Q 值。Q 值记录了探索过程中连续动作的累积奖励,将其储存在 Q 表中,智能体依据 Q 表,选择使下一个状态 Q 值最大的动作采取行动。但在状态空间无限的问题中,Q 表根本无法在计算机中存储。因此在这种情况下,Q 表由 Q 值函数代替,而 DQN 正是用神经网络来逼近 Q 值函数,DQN 输入当前的状态值,输出 Q 值最大的动作。

DRQN 将 DQN 中的最后一层全连接层替换为

长短期记忆(long short-term memory, LSTM)网络,能够辅助记忆更多的回合信息。DRQN 的算法原理如图 1 所示。

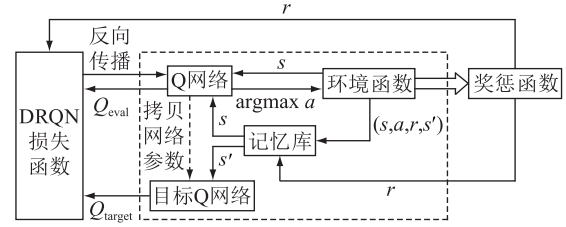


图 1 DRQN 算法原理

Fig.1 DRQN algorithm principle

在 DRQN 模型中,测试时由环境输入当前状态值 s 到 Q 网络,Q 网络的输出作为 a ,即所有动作的权值,对 a 进行 argmax 操作,得到所有动作的概率,然后按贪婪策略选择动作输入到环境函数中,得到下一步状态值 s' ,完成一个回合中一步的探索,循环多步,完成智能体一个回合的探索。

在训练 DRQN 的过程中,智能体首先需要探索一些回合,得到一些经验,存储在记忆库里,训练时只从记忆库中随机采样, s 输入到 Q 网络中得到 Q 估计值 Q_{eval} ,用这个估计值来逼近 Q 值, s' 输入到目标 Q 网络中,得到 Q 目标值 Q_{target} ,计算损失时用 Q_{target} 加上奖惩值 r ,再减去 Q_{eval} 。

2.2 异常用电判决模型

环境函数由状态与动作组成,输入当前状态值与选择的动作值,输出下一步的状态值。在文中的方法中,状态就是由判决阈值与判决比例组成的五维数组:

$$s = (x_1, x_2, x_3, x_4, x_5) \quad (1)$$

式中: $x_1 \in (0, E_{\text{max}})$ 为判决阈值 1, E_{max} 为用户用电数据理论上的最大值; $x_2 \in (0, E_{\text{max}})$ 为判决阈值 2; $x_3 \in (0, 1)$ 为超过判决阈值 1 的天数比例; $x_4 \in (0, 1)$ 为超过判决阈值 2 的天数比例。将待测用户数据超过阈值 1 的天数比例是否大于 x_3 记为 O_1 ,超过阈值 2 的天数比例是否大于 x_4 记为 O_2 , x_5 决定 O_1 和 O_2 之间取与操作还是或操作。也就是说, s 包含 2 个独立的判决阈值,2 个对应判决阈值的判决比例和最后 2 个结果 O_1 , O_2 之间是取与还是或,由 DRQN 模型生成。为在下文中方便解释,将判决阈值和判决天数比例合称为判决阈值。在 DRQN 模型中,首先要初始化状态即判决阈值,在网络迭代后会输出优化后的判决阈值,作为本次用来判断用户是否属于异常用电用户的指标。

动作 a 是指对状态 s 的五维数组分别进行加减,具体来说前四维分别是对应 s 的前四维进行加

减固定值的操作,第五维则是选择 s 的第五维是与操作还是或操作。

奖惩函数是由状态 s 得到的奖惩值, s 为当前状态,即网络目前输出的判决阈值,将此阈值与所有的用户概率序列进行对比得到每个用户的判决结果,在与正确用户标签进行比对后,得到当前阈值判决用户异常用电与否的正确率,根据该正确率输出奖惩值。

记忆库在每个回合结束后,记录下每步的当前状态 s , 选择动作值 a , 奖惩值 r 和下一步状态 s' , 组成四元组 (s, a, r, s') , 存入记忆库。当记忆库容量开始溢出时,删除最开始记录的部分四元组数据。

DRQN 包括 2 个结构一样的网络:Q 网络和目标 Q 网络。Q 网络由两层全连接网络和一层 LSTM 网络组成,第一层全连接网络的神经元数量为 10,第二层为 50, LSTM 网络层的神经元数量为 64,输入为 s , 输出为 Q_{eval} , 对 Q_{eval} 进行 argmax 操作,得到概率值最大的动作。Q 网络的结构如图 2 所示。

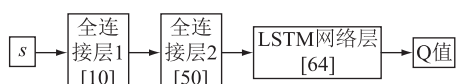


图 2 Q 网络结构

Fig.2 Q network structure

目标 Q 网络与 Q 网络结构完全一致,但是在训练时冻结网络参数,只有 Q 网络的网络参数在进行学习,每隔固定步数 k ,目标 Q 网络的网络参数与 Q 网络同步一次。

损失 l 由 Q 网络的输出 Q_{eval} , 目标 Q 网络的输出 Q_{target} 和 r 计算得到:

$$l = [Q_{target} - (r + \gamma Q_{eval})]^2 \quad (2)$$

式中: γ 为衰减系数,属于超参数。

3 异常用电判决模型的训练策略

在文中的模型中,智能体探索的一个回合,即是判决阈值寻优完成的一个回合。回合开始时状态初始化,智能体选择动作,得到下一步状态值,如此循环探索完固定步数后,智能体探索的一个回合结束,之后再开始新的探索回合。

DRQN 的训练是穿插在智能体的探索中的,具体来说,智能体每在环境里探索若干步,得到若干步的状态值与动作值,DRQN 就训练一次,并把这次探索的经验记录下来,训练流程如图 3 所示。

(1) 初始化 Q 网络、目标 Q 网络的模型参数,清零记忆库;

(2) 初始化状态,将当前状态值置为一个初始

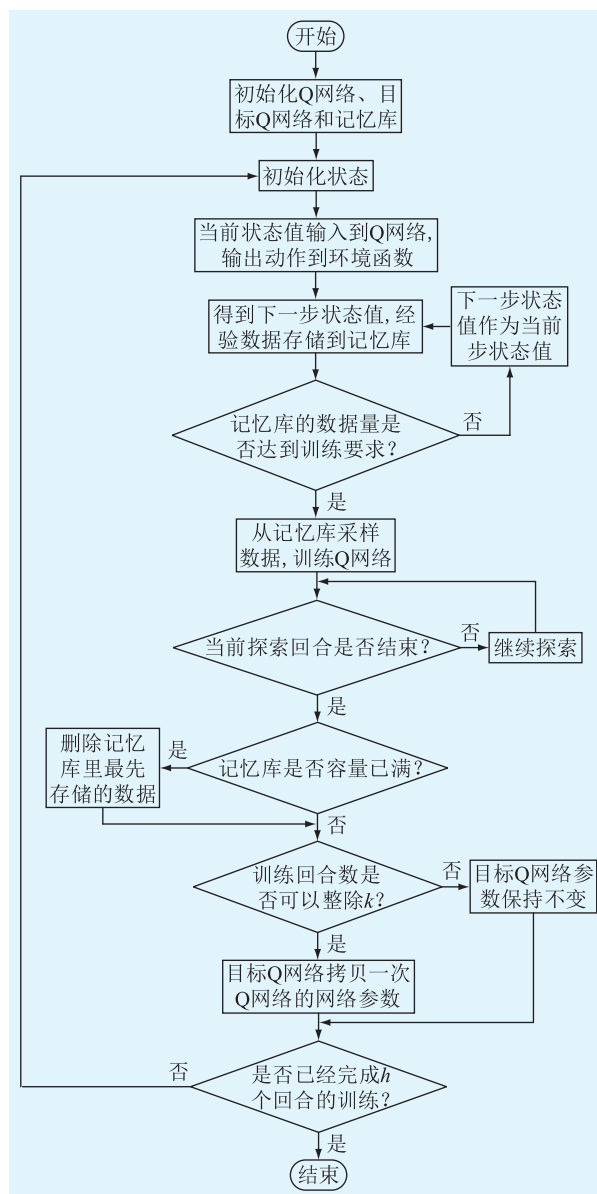


图 3 DRQN 训练流程

Fig.3 DRQN training process

状态值;

(3) 将当前状态输入到 Q 网络中, Q 网络输出所选择的动作,将动作输入到环境函数,得到下一步的状态值,完成智能体的一步探索,并将当前的经验数据即四元组 (s, a, r, s') 存入记忆库;

(4) 将(3)中的下一步状态值 s' 作为当前步的状态值,重复(3)若干步;

(5) 智能体每探索 n_s 步,判断一次当前记忆库所存储的数据量是否满足采样训练的要求,若不足则到(4)继续累积经验数据,若满足则到(6)开始训练;

(6) 从记忆库中采样若干步的经验数据, Q 网络开始学习,目标 Q 网络参数冻结,状态值输入到 Q 网络,下一步状态值输入到目标 Q 网络,计算出损

失,进行反向传播;

(7) 判断当前回合是否结束,若未结束,则继续探索,一个回合智能体固定探索 m 步;

(8) 判断记忆库的容量是否已满,若已满则删除最先存储的经验数据,继续添加当前经验数据;

(9) 判断 Q 网络的训练回合数是否可以整除 k ,若是,则目标 Q 网络拷贝一次 Q 网络的网络参数,否则目标 Q 网络的网络参数固定不变;

(10) 判断智能体是否已经完成了 h 个回合的探索,若是,则整个训练流程结束,否则回到(2)重新开始新的回合的智能体探索。

4 异常用电判决模型的判决流程

文中提出的基于 DRQN 的判决模型能够动态地对判决阈值进行优化,在现场测试中的具体判决流程如图 4 所示。

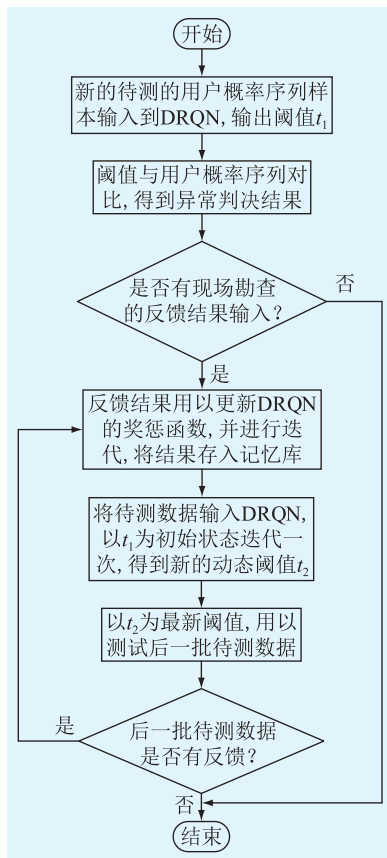


图 4 判决流程

Fig.4 Judgment process

(1) 将待测的长周期用户数据分为多个短周期数据,输入到分类器中,输出多个短周期对应的异常用电概率,组成概率序列;

(2) 训练好的 DRQN 模型输出得到判决阈值;

(3) 将(2)中的判决阈值和(1)中的概率序列进行比较,得出用户异常与否的判决结果;

(4) 工作人员依据判决结果去现场勘查一部分用户,得到真实的用户标签,即反馈;

(5) 将这部分反馈输入到 DRQN 模型中,即将反馈得到的真实用户标签和用户数据用来更新奖惩函数,并重新将待测数据输入到 DRQN 中,用新的奖惩函数进行迭代,将迭代结果存入到记忆库;

(6) 将(2)中得到的判决阈值作为初始状态进行一次迭代,输出得到新的判决阈值,阈值以此得到更新,因此称为动态阈值;

(7) 再将新的判决阈值用以测试后一批的待测数据,若之后的测试数据得到反馈,可以重复(5)、(6),使模型更加适用于新待测数据分布。

5 试验与分析

5.1 数据集和分类器介绍

所有训练和测试数据均为电力公司提供的真实案例数据,包含 331 户用户。其中 205 户来自同一个地区 A,每个用户的时间跨度为 1~300 d 不等,但大多都在 300 d 左右,这 205 户中又有 160 户是三相四线的数据,另外 45 户是三相三线的数据;另外 126 户来自另一个地区 B,时间跨度为 8~30 d 不等。每小时采样一次用户数据,将一天的 24 个用电数据作为分类器的样本,输入到分类器中,得到该样本的异常概率。一个用户有 n_d 天的用电数据(n_d 个分类器样本),则分类器输出长度为 n_d 的概率序列,该序列用于文中试验,序列中每个数字表示该用户在某天的用电数据为异常的概率。

将来自 A 地的 160 名三相四线用户的数据作为训练集,45 名三相三线用户数据作为测试集 1,用来测试模型面对不同种类数据的泛化能力,将来自 B 地的 126 名用户的数据作为测试集 2,用来测试模型面对不同地区、不同序列长度数据的泛化能力。数据集的概况如表 1 所示。

表 1 数据集概况

Table 1 Overview of data set

用户数据分类	训练集	测试集 1	测试集 2
正常用户数据	117	26	75
异常用户数据	43	19	51

文中试验所使用的分类器来自文献[17],该分类器基于知识嵌入和残差网络,在文中数据集的样本级上有较好的表现,将一天 24 个用电数据作为一个样本输入到分类器中, $F1$ 分数最高能达 0.73,受试者工作特征曲线的曲线下方面积 (area under curve, AUC) 分数最高能达 0.95。

5.2 评估指标

文中选择通用的查全率 R 、查准率 P 作为评估指标,具体定义如下:

$$P = \frac{N_{TP}}{N_{TP} + N_{FP}} \times 100\% \quad (3)$$

$$R = \frac{N_{TP}}{N_{TP} + N_{FN}} \times 100\% \quad (4)$$

式中: N_{TP} (真阳性)为正确检测到的异常用电用户数; N_{FP} (假阳性)为被归类为正常的异常用电用户数; N_{FN} (假阴性)为被归类为异常用电用户的正常用电用户数。

5.3 试验设置

文中的试验分为 2 组,一组为基准测试,即简单投票法,作为对比基准,没有测试结果的反馈;一组为动态测试,将测试集的数据分批输入进行测试,除了第一批次测试,后面每次测试都会将前一批次的试验数据的真实标签反馈给网络。

5.3.1 静态测试

基准测试为简单投票法,将模型在训练集上得到的阈值作为简单投票法的投票阈值,与用户的概率序列比较得到判决结果。在基准测试中,无论是对用电数据类型不同的测试集 1,还是对地域不同的测试集 2,都采用了固定的判决阈值与投票比例。

5.3.2 动态测试

动态测试是将待测数据分成 n_b 批,第一次测试用最初的静态阈值 t_1 测试全部数据,得到第一次测试的各项评估指标 r_1 ;然后得到第一批测试数据的测试反馈,即第一批数据的真实标签,再将第一批数据及标签输入到 DRQN 中,开始新的探索回合,取回合最后一步的输出作为动态阈值 t_2 ,用来测试除第一批以外的测试数据,得到第二次测试的各项评估指标 r_2 ;再得到第二批测试数据的测试反馈,将第二批数据及标签输入到 DRQN 中,开始新的探索回合,得到动态阈值 t_3 ,用来测试除第一、二批以外的测试数据,得到第三次测试的各项评估指标 r_3 ;以此类推,一直到得到倒数第二批的测试反馈,测试最后一批测试数据,整个动态测试完成。

DRQN 模型输出阈值,动态测试每批次测试完毕,都会得到不同的判决阈值,故而输出的阈值称为动态阈值,因为阈值会根据上一批次测试反馈的结果而改变。

这样设计动态试验是为了展现模型的学习能力。在持续得到所测试数据结果反馈的情况下,DRQN 的记忆库与状态空间会发生改变。试验并非是对样本级别的判决,而是对用户级别的判决,在

用户的用电数据发生变化时(无论是因为时间还是地域不同而引起的变化),模型输出的判决阈值随用户数据分布的改变而改变,从而能够提升模型在不同数据分布下的泛化能力。

5.3.3 模型实现与参数设置

Tensorflow 是由谷歌人工智能团队开发的开源软件库,通常应用于各种机器学习算法的实现,功能强大且在被广泛应用在各种工程中。因此,文中的网络模型是基于 Tensorflow 实现的。

文中的试验设置为 $n_d = 20, m = 400, k = 300, h = 1\ 000$ 。优化器选择为 RMSPropOptimizer,网络模型包含两层全连接网络和一层 LSTM 网络,神经元数量如图 2 所示。

5.4 试验结果与分析

5.4.1 基准测试结果

基准测试为一次性测试所有测试数据,直接把判决阈值与待测数据作比较,试验结果数据见表 2。

表 2 基准测试试验结果

评估指标	训练集	测试集 1	测试集 2
查准率	100.00	84.21	90.20
查全率	96.69	84.21	68.66

基准测试在 2 个测试集上的混淆矩阵见表 3。

表 3 基准测试在各个测试集上的混淆矩阵

Table 3 Confusion matrix of the benchmark on each test set

真实值	预测值			
	测试集 1		测试集 2	
	异常用户	正常用户	异常用户	正常用户
异常用户	16	3	46	5
正常用户	3	23	21	54

可以看出,模型得到的阈值在训练集上效果良好,在测试集 1 上表现也尚可,在数据分布差异较大的测试集 2 上则效果较差。

5.4.2 动态测试结果

文中将测试集 1 的 45 份待测试数据分为 3 份,每份 15 个待测序列数据;将测试集 2 的 126 份待测试数据分为 6 份,每份 21 个待测序列数据。在每个测试集上进行 3 次完整的动态测试,再将测试得到的评价指标结果进行平均,绘制成曲线图,比较每次反馈后,相应的评价指标是否有所提升。动态试验曲线如图 5、图 6 所示。

将每次动态测试的混淆矩阵叠加在一起,与基准测试的混淆矩阵进行对比,3 次动态测试各个批次在测试集上的混淆矩阵叠加后如表 4 所示。

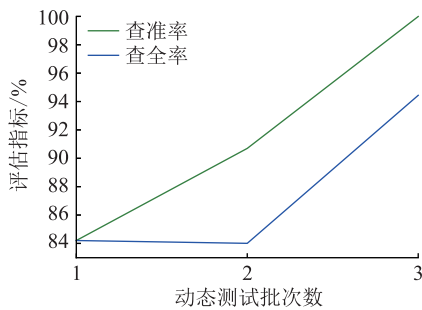


图5 测试集1上评价指标随动态测试批数的变化

Fig.5 Variation of evaluation index on test set 1 with the number of dynamic test batches

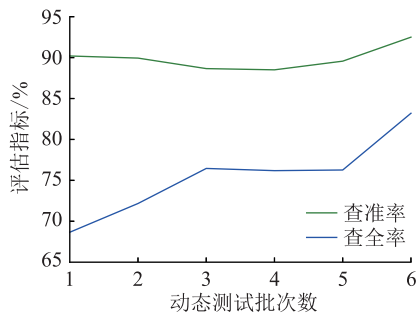


图6 测试集2上评价指标随动态测试批数的变化

Fig.6 Variation of evaluation index on test set 2 with the number of dynamic test batches

表4 动态测试在各个测试集上的混淆矩阵

Table 4 Confusion matrix of the dynamic test on each test set

真实值	预测值			
	测试集1		测试集2	
	异常用户	正常用户	异常用户	正常用户
异常用户	49	4	350	42
正常用户	8	74	115	438

由动态测试的结果可知,一旦引入被测数据的反馈,模型在新数据域上的适应能力会得到大幅度提升;而随着引入的被测数据反馈增加,模型在各项指标上都有所提升。具体表现在测试集1查准率上升趋势明显,查全率稍有波动,但整体还是上升趋势;在测试集2上,查全率的上升趋势明显,查准率则有一点波动,但最后还是呈现上升趋势。

因此可以说明文中的模型在引入待测数据的反馈后,有较强的学习能力,能够适应不同数据分布的待测数据,在实际工程应用中具有较强的灵活性。

6 结语

针对异常用电长周期检测所需的判决阈值问题,文中提出了一种新颖的基于强化学习的异常用电判决方法。试验结果说明,在基于分类器的异常用电检测中,通过强化学习生成的判决阈值可以有

效提升用户级别的判决准确率。其次,通过对比传统投票法和文中方法在2个测试集上的表现,得出文中方法表现均更加良好。

由此可以看出,文中模型能够解决传统投票法存在的一些问题,即不同用户的有效用电数据可能长度差别较大,不同地区甚至是不同时间段的用户用电数据分布差别较大,异常比例在用户间参差不齐,难以以单一固定的票数比例去完成所有用户的判决的问题。

文中提出的DRQN模型可以通过输入待测数据的测试反馈来提升模型的泛化能力,在工程应用上更加灵活有效,异常用电检测性能更加稳定、可靠,可有效支撑实际环境中的异常用电识别。

参考文献:

- [1] 苗新,张东霞,孙德栋. 在配电网中应用大数据的机遇与挑战[J]. 电网技术,2015,39(11):3122-3127.
MIAO Xin,ZHANG Dongxia,SUN Dedong. The opportunity and challenge of big data's application in power distribution networks[J]. Power System Technology, 2015, 39(11): 3122-3127.
- [2] NIKOVSKI D N,WANG Z. Method for detecting power theft in a power distribution system; US20140236506 A1 [P]. 2014-08-21.
- [3] LO Y L,HUANG S C,LU C N. Non-technical loss detection using smart distribution network measurement data [C]//IEEE PES Innovative Smart Grid Technologies. Tianjin,China. IEEE, 2012:1-5.
- [4] ARANHA NETO E A C,COELHO J. Probabilistic methodology for technical and non-technical losses estimation in distribution system[J]. Electric Power Systems Research,2013,97:93-99.
- [5] AMIN S,SCHWARTZ G A,TEMBINE H. Incentives and security in electric distribution networks[C]//International Conference on Decision and Game Theory for Security. Springer, Berlin,Heidelberg,2012.
- [6] CARDENAS A A,AMIN S,SCHWARTZ G,et al. A game theory model for electricity theft detection and privacy-aware control in AMI systems [C]//2012 50th Annual Allerton Conference on Communication,Control, and Computing (Allerton). Monticello,IL,USA. IEEE,2012:1830-1837.
- [7] AMIN S,SCHWARTZ G A,CARDENAS A A,et al. Game theoretic models of electricity theft detection in smart utility networks:providing new capabilities with advanced metering infrastructure[J]. IEEE Control Systems Magazine, 2015, 35(1): 66-81.
- [8] 钱忠敏,贺兴,葛晓军,等. 基于相关性分析的电力系统数据驱动异常早期检测[J]. 供用电,2019,36(10):30-35.
QIAN Zhongmin,HE Xing,GE Xiaojun,et al. Data-driven early anomaly detection in power system based on correlation analysis [J]. Distribution & Utilization,2019,36(10):30-35.
- [9] 陈启鑫,郑可迪,康重庆,等. 异常用电的检测方法:评述与

- 展望[J]. 电力系统自动化,2018,42(17):189-199.
- CHEN Qixin,ZHENG Kedi,KANG Chongqing, et al. Detection methods of abnormal electricity consumption behaviors: review and prospect[J]. Automation of Electric Power Systems,2018,42(17):189-199.
- [10] COMA-PUIG B,CARMONA J,GAVALDÀ R, et al. Fraud detection in energy consumption: a supervised approach [C]//2016 IEEE International Conference on Data Science and Advanced Analytics (DSAA). Montreal, QC, Canada. IEEE, 2016:120-129.
- [11] JIANG R,LU R X,WANG Y, et al. Energy-theft detection issues for advanced metering infrastructure in smart grid[J]. Tsinghua Science and Technology,2014,19(2):105-120.
- [12] JOKAR P,ARIANPOO N,LEUNG V C M. Electricity theft detection in AMI using customers' consumption patterns [J]. IEEE Transactions on Smart Grid,2016,7(1):216-226.
- [13] 胡天宇,郭庆来,孙宏斌. 基于堆叠去相关自编码器和支向量机的窃电检测[J]. 电力系统自动化,2019,43(1):119-125.
- HU Tianyu,GUO Qinglai,SUN Hongbin. Nontechnical loss detection based on stacked uncorrelating autoencoder and support vector machine [J]. Automation of Electric Power Systems, 2019,43(1):119-125.
- [14] RAMOS C C O,DE SOUZA A N,FALCAO A X, et al. New insights on nontechnical losses characterization through evolutionary-based feature selection[J]. IEEE Transactions on Power Delivery,2012,27(1):140-146.
- [15] ZHENG Z B,YANG Y T,NIU X D, et al. Wide and deep convolutional neural networks for electricity-theft detection to secure smart grids[J]. IEEE Transactions on Industrial Informatics,2018,14(4):1606-1615.
- [16] 赵文清,沈哲吉,李刚. 基于深度学习的用户异常用电模式检测[J]. 电力自动化设备,2018,38(9):34-38.
- ZHAO Wenqing,SHEN Zheji,LI Gang. Anomaly detection for power consumption pattern based on deep learning [J]. Electric Power Automation Equipment,2018,38(9):34-38.
- [17] 李江腾,王非. 基于知识嵌入和 DNN 的工商业用户异常用电检测[J]. 电力工程技术,2020,39(3):158-165.
- LI Jiangteng,WANG Fei. Non-technical loss detection based on energy measurement knowledge and deep neural network among industrial and commercial customers[J]. Electric Power Engineering Technology,2020,39(3):158-165.

作者简介:



蔡云芹

蔡云芹(1997),女,硕士在读,研究方向为深度学习(E-mail:caiyunqin@hust.edu.cn);

王非(1977),男,博士,副教授,研究方向为机器学习与电力数据分析。

Judgment method of abnormal electricity consumption based on reinforcement learning

CAI Yunqin, WANG Fei

(School of Electronic Information and Communications, Huazhong University of Science and Technology, Wuhan 430074, China)

Abstract: There are many classification-based methods for abnormal electricity consumption detection now, but most of them are based on short-term electricity consumption to judge long-term electricity consumption behavior. It is difficult to determine the threshold and ratio of these methods. In engineering application, the distribution of power consumption data in different regions and time periods is quite different, so the proportion and threshold value are quite different. It is difficult to apply the fixed proportion to all user data. To solve this problem, a method for judging abnormal electricity consumption based on reinforcement learning is proposed, which innovatively uses reinforcement learning model to dynamically generate threshold for different data sets. Firstly, the abnormal probability of the short-term behavior of several users output by the classifier is obtained. Then, the dynamic threshold is obtained by inputting the probability into the deep recurrent Q network (DRQN) of the enhanced learning model, where, the dynamic threshold can be Judgment threshold and judgment ratio as well. The experimental results show that, compared with the traditional voting method of manual parameter adjustment, this method has a significant improvement in the evaluation index, and also has a good performance in data sets with large differences in data distribution. It shows that this method has strong generalization ability in the real environment with complex data types.

Keywords: smart grid; reinforcement learning; neural network; abnormal electricity consumption; dynamic threshold; deep Q network; anomaly detection

(编辑 陆海霞)