引用格式:金保华,张明星,吴怀广,等. 一种基于电力大数据的反窃电预测方法[J]. 轻工学报,2020,35(4):81 - 87.
中图分类号:TM73 **文献标识码**:A
DOI:10.12187/2020.04.011
文章编号:2096 - 1553(2020)04 - 0081 - 07

一种基于电力大数据的反窃电预测方法

A prediction method of anti-electricity stealing based on big data of electric power

金保华,张明星,吴怀广,石永生
JIN Baohua,ZHANG Mingxing,WU Huaiguang,SHI Yongsheng

郑州轻工业大学 计算机与通信工程学院,河南 郑州 450001
College of Computer and Communication Engineering, Zhengzhou University of Light Industry, Zhengzhou 450001, *China*

关键词:
反窃电预测;异常规则分析;电力大数据;线损率增长率

Key words:
prediction of anti-electricity stealing; abnormal rule analysis; big data of electric power; growth rate of line loss rate

摘要:针对传统的反窃电预测方法准确度低的问题,提出了一种基于电力大数据的反窃电预测方法. 该方法根据异常规则构造窃电数据样本,引入线损率增长率这一约束条件,使用 4 种机器学习分类算法分别在电压、电流和功率因数数据集上构建预测模型,将其输出的数据异常用户与线损异常用户相结合,输出疑似窃电用户清单. 实验结果表明,该方法预测准确度令人满意,在疑似窃电用户识别方面是高效可行的.

收稿日期:2020 - 03 - 27
基金项目:国家自然科学基金项目(61672470,61802350);国家重点研发计划项目(2016YFE0100600,2016YFE0100300)
作者简介:金保华(1966—),男,河南省郑州市人,郑州轻工业大学教授,主要研究方向为人工智能、计算机决策支持系统、计算机软件和理论、应急管理.

Abstract: Aiming at the problem of low accuracy of traditional prediction methods of anti-electricity stealing, a prediction method of anti-electricity stealing based on big data of electric power was proposed. This method firstly constructed samples of electricity stealing data according to abnormal rules, and introduced the growth rate of line loss rate constraint conditions. Then it used four machine learning classification algorithms to build a prediction model on the voltage, current and power factor data sets respectively, combined the users with abnormal data output and users with abnormal line loss to output a list of users suspected of stealing electricity. Experimental results showed that the prediction accuracy of the method was satisfactory, and it was efficient and feasible in identifying users suspected of stealing electricity.

0　引言

麦肯锡研究所认为:大数据已经成为能够改变世界的第一科学技术[1-2]. 2008 年,Nature 在其杂志里提到“Big Data”这一概念,描述了现代互联网技术和计算机系统面对未来海量数据时将面临的种种挑战. 随后,Amazon、IBM、Google 等跨国公司纷纷加入到对大数据的研究队伍中,推动了大数据技术的快速发展,并将其普及至行业应用[3-9]. 大数据技术和智能电网的融合就是大数据理念和方法在电力行业的实践和应用. 目前,我国拥有丰富的数据资源、用户资源和应用市场优势,促使大数据的关键技术研发取得了重大突破. 随着国家电网智能化和信息化程度的加深,智能电网的管理也进入到了大数据时代.

近年来,电力与经济社会发展的关系密不可分,人们的生产生活对电能的依赖程度也越来越高[10-11],但窃电行为的存在使得世界各国都承受着巨大的损失. 我国每年因为窃电造成的经济损失高达 200 亿元左右[12-13],除此之外,由于窃电导致的各种事故所造成的间接损失更加巨大. 当前窃电行为的主要特点是窃电过程隐蔽化、窃电手段高科技化、窃电数量大额化. 目前,欧美等发达国家走在高科技反窃电技术的前沿,已经研发出了用于降低非技术损耗的反窃电装置,并在北美的各大电力部门进行使用,反窃电产品还被大量销往巴西、委内瑞拉、哥伦比亚等国家[14-17].

长期以来,我国在反窃电方面已取得了一些成效,但是传统的检测手段逐渐无法满足当前反窃电的需求. 随着信息科技的进步,很多学者开始结合数据挖掘、分析技术来预测可能存在的窃电行为. 窦健等[18]在异常检测模型里加入了线损异常这一约束条件,提高了异常检测的准确率. 吴迪等[19]利用鱼骨图表示特征参量,建立了多维度电参量的相关特征参量集合,并提出基于大数据的防窃电结构化模型. 庄池杰等[20]提出了基于无监督学习的异常用电检测模型,包括特征提取、主成分分析、网格处理等模块. 程超等[21]为了解决传统方法时效低的问题,由实际案例统计确定研究重点,根据计量装置的特点、电压电流值的规律和离群点理论,提出了基于距离的离群点检测法来判定是否存在窃电行为. 王新霞等[22]也提出根据离群点检测的方式来挖掘窃电用户. 以上方法均可在一定程度上识别窃电行为,但是时间复杂度较高,且仅对采集系统中的数据进行预处理,可能会遗漏很多有价值的信息,从而导致模型准确率不高. 当前电力系统建设尚不完善导致采集系统数据类型繁多、数据结构复杂,如果将异常规则分析与电压、电流、线损等数据相结合构建反窃电模型,将会覆盖更多的异常情况.

鉴于此,本文拟提出基于电力大数据的反窃电预测方法,该方法将异常识别规则和机器学习算法结合起来,并引入线损率增长率进行综合分析,以期更加高效地识别窃电用户.

1 反窃电预测方法设计

1.1 设计思路

基于电力大数据的反窃电预测方法设计思路如图 1 所示,具体可分为以下 4 个步骤.

步骤 1 确定能反映用户用电行为的特征,从采集系统中提取相关数据. 本文用于模型训练的数据是电压、电流和功率因数.

步骤 2 对数据进行预处理,查看数据中是否存在缺失值、异常值,以及数据类型不统一等问题,并对数据进行清洗,使用拉格朗日插值法填充缺失值,使用独热编码(one-hot)对不同特征中的数据格式进行转换,使之全部变成数值型数据. 根据异常规则分析构造窃电数据,非窃电数据从没有窃电记录的用户数据中获得,窃电数据和非窃电数据共同组成建模数据.

步骤 3 使用多种机器学习分类算法对数据进行训练以构建反窃电预警模型,计算各个线路前、后几天的线损率平均值,并计算两个平均值的增长率,如果线损率平均值增长率过大,即可认为该线路上的用户有可能发生窃电行为. 综合考虑算法模型输出的数据异常用户和线损异常用户,输出最终的疑似窃电用户清单.

步骤 4 对比分析多个模型的实验结果,分别从算法角度和数据角度对反窃电预警模型进行评估,最终实现对多种算法和多种数据建模效果的对比分析. 如果窃电用户的识别正确率比较低,则将实验分析结果反馈给异常规则、算法模型和线损异常判断模块,根据结果对异常规则中的阈值作相应的调整,并调节算法模型和线损异常判断中的参数,以达到不断优化模型的目的,直到取得较好的实验结果.

1.2 数据的获取

在步骤 3 中,构建反窃电算法模型需要正、负两类样本,其中正样本是指正常用电数据,负样本是指窃电数据. 但是由于前期反窃电工作的不完善,导致采集系统中无法直接根据窃电起始日期截取窃电数据,这也是模型建立过程中的难点之一. 考虑到目前反窃电研究中使用比较多的手段是异常规则分析[23],为了获取窃电数据,本文使用几种常见的异常规则对数据进行分析;非窃电数据就从没有窃电记录的用户数据中获得.

异常规则分析就是根据电能表中的电流数据、电压数据、功率因数数据对用电客户的用电情况进行检测分析. 在实际应用中,如果用户的用电数据符合相应的异常规则,那么就认为该用户符合窃电的异常情况. 常见的异常规则是电压断相异常、电压越限异常和三相不平衡异常,具体异常规则描述如表 1 所示,其中,U'为额定电压;规则中 K 的取值范围为 50% ~ 70%,$K_1=110\%$,$K_2=90\%$,$K_3=60\%$,这些取值

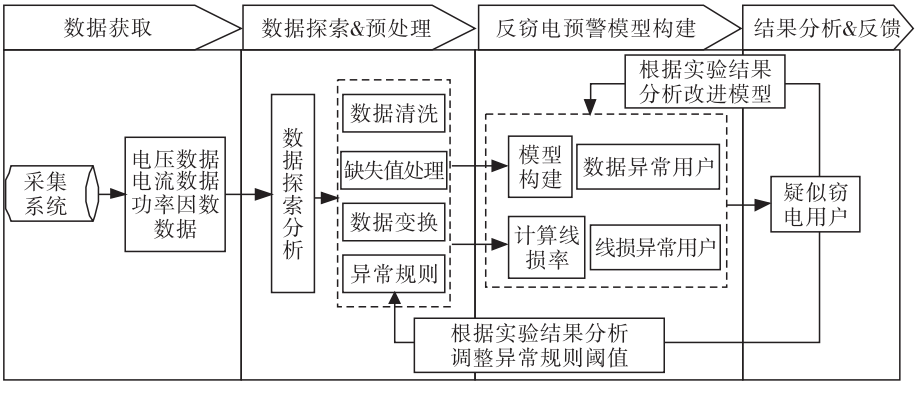

图 1 反窃电预测方法设计思路

Fig. 1 Design idea of the prediction method of anti-electricity stealing

参考了文献[24]在异常用电方面的研究结果；m、n、i 都是需要调节的整数型阈值，根据步骤 4 的分析结果反馈调整取值，调整时步长均为 1，直到达到较好的预测效果. 由于电能表计量方式和接线方式不同，表 1 中的额定电压取值也有所不同，其取值如表 2 所示.

1.3　线损率增长率

本文在使用机器学习算法进行模型构建时，引入了线损率增长率，综合考虑算法模型输出的数据异常用户和线损异常用户，输出最终的疑似窃电用户清单. 由于在确定疑似窃电用户清单时，增加了线损异常这一限制条件，故较单纯的算法模型准确率更高.

表 1　异常规则描述

Table 1　Description of the abnomal recognition rules

规则名称	规则描述	备注
电压断相	三相三线：$\min(U_a,U_c)<K\times U'$ & $U_b\geqslant K\times U'$ 三相四线：$\min(U_a,U_b,U_c)<K\times U'$ & $\max(U_a,U_b,U_c)\geqslant K\times U'$	如果 1 d 内有连续 m 个以上的数据点出现断相异常，且该状态持续 n d，即可确定该用户为异常用电
电压越限	三相三线：$\max(U_a,U_c)>K_1\times U'$ || $\max(U_a,U_c))\geqslant K_3\times U'$ & $\min(U_a,U_c)\leqslant K_2\times U'$ 三相四线：$\max(U_a,U_b,U_c)>K_1\times U'$ || $\max(U_a,U_b,U_c))\geqslant K_3\times U'$ & $\min(U_a,U_b,U_c)\leqslant K_2\times U'$	如果 1 d 内有连续 m 个以上的数据点符合越限异常，且该状态持续 n d，即可确定该用户为异常用电
三相不平衡	三相三线：$imbalance\_rate=\frac{\max(I_a,I_c)-\min(I_a,I_c)}{\max(I_a,I_c)}$ 三相四线：$imbalance\_rate=\frac{\max(I_a,I_b,I_c)-\min(I_a,I_b,I_c)}{\max(I_a,I_b,I_c)}$	$imbalance\_rate>i\%$ 即为异常点，如果 1 d 内的异常点大于 m 个，且该状态持续 n d，那么该用户的用电情况就被视为异常

表 2　不同电能表额定电压的取值

Table 2　Value of rated voltage of different vott-hour meters

计量方式	接线方式	U'/V
高供高计	三相三线	100.0
高供高计	三相四线	57.7
高供低计	三相四线	220.0

电网的线损率是一项重要技术指标，可以直接反映供电线路中的电能损耗. 虽然线损在电路输电过程中无法避免，但是线损率一般都会控制在合理范围内，如果线损率异常，即可视为该线路可能存在窃电用户. 线损率计算公式为

$$t_l=\frac{s_l-f_l}{s_l}\times 100\%$$

其中，s_l 是第 l d 某线路的总供电量，f_l 是该线路上所有用户的总用电量.

通常情况下，用户如果发生窃电行为，则其所在线路的线损率应该也会下降，但是由于用户每天的用电量都存在波动，因此单纯地以当天线损率的下降作为是否窃电的评判标准并不合适. 本文设置 5 d 的统计窗口期，用以计算线损率平均值，其中线损率增长率计算公式如下：

$$increase\_rate=\frac{V_1-V_2}{V_2}\times 100\%$$

其中，V_1 为前 5 d 的线损率平均值，V_2 为后 5 d 的线损率平均值. 若 $increase\_rate>1\%$，则认为该线路上可能存在窃电用户.

2　实验结果与分析

2.1　实验数据

本文使用的数据是某市高压用户的电压数据、电流数据、功率因数数据和线损数据. 其中，电流数据共 12 455 147 条，对应 19 049 个用户；电压数据共 10 525 026 条，对应 19 048 个用户；功率因数数据共 13 269 943 条，对应 19 042 个用户；线损数据共 3 328 392 条，对应 16 946 个线路的线损. 由于电压数据、电流数据、功率因数数据特征之间关系较为复杂，这些数据在进行空值填充、数据格式转换等预处理后，还需要进行数据降维和数据合并.

1）数据降维：在实际的反窃电工作中，窃电用户一旦开始实施窃电，一般都会维持至少

数个小时,因此只需要保留每天的24个整点时刻数据,即进行数据降维.

2)数据合并:已有的数据中,电能表每隔15 min记录一次数据,因此一天中每个电能表的每一相都会有96点数据.每个电能表有三相(*A*相、*B*相、*C*相),即每天对应有3条数据,通过计算三相不平衡率将每天的3条数据合并成一条,可以降低数据复杂度,提高程序的运算效率.

本文数据点数 m 的取值范围是1—24,对应一天中的24 h;天数 n 的取值范围是1—10,n 值过大会漏掉一些短期窃电用户,因此最大调整到10即可;i 的取值范围是1—100.它们所对应的最佳阈值分别为:$m=n=3$,$i=80$.

2.2 模型评估指标

本文引入了召回率(recall rate)和精准率(precision rate)两项指标对模型预测能力进行评估.召回率越高,代表预测正确的窃电用户数量在真实窃电用户中所占的比例越高,反映了正确识别窃电用户的能力.精准率越高,代表预测窃电用户的准确性越高.召回率和精准率的公式分别为

$$召回率=n_{pre\_real}/n_{real}$$

$$精准率=n_{pre\_real}/n_{pre}$$

其中,n_{real}是真实窃电用户的个数;n_{pre}是模型预测的窃电用户个数;n_{pre_real}是 n_{real}和 n_{pre}的交集,即预测结果与实际相符的窃电用户个数.

当召回率和精准率无法同时兼顾时,以度量值 F_1 来衡量模型的优劣:

$$F_1=2\times\frac{召回率\times精准率}{召回率+精准率}\times100\%$$

2.3 结果与分析

为了验证本文提出方法的有效性,使用随机森林、逻辑回归、决策树和支持向量机4种分类算法构建分类模型,将这4种算法模型分别应用在电流、电压、功率因数3种不同的数据集上,以此来比较不同算法模型在反窃电预测方法中的实践效果.为了提高模型的泛化性,实验均采用五折交叉验证,实验结果如表3—5所示.对比表3—5中引入线损率增长率前后的 F_1 值可以看出,引用线损率增长率之后,F_1 均有不同幅度的提高,这说明引入线损率增长率有助于提高疑似窃电用户识别的准确率,因此接下来的对比分析均采用引入线损率增长率的实验结果.

由表3可知,相比较其他3种算法模型,决策树分类模型表现最好,其精准率虽然不是最高,但是其 F_1 和召回率都达到最高值,分别为79.07%和85.00%,意味着该模型可以找到85.00%的窃电用户;逻辑回归分类模型表现最差,F_1 和召回率最低,只有57.97%和50.00%,意味着该模型只能够找到一半的窃电用户.因此,在使用电流数据建模时,使用决策树分类算法来训练数据,构建的反窃电模型最为有效.

由表4可知,4种模型的 F_1 差距不大,均在70.00%左右,说明使用电压数据进行模型构建,结果相对稳定.其中,随机森林分类模型表现最好,F_1 为72.00%,且召回率和精准率也都能达到70.00%以上,3个度量指标比较集中稳定;逻辑回归分类模型和支持向量机分类模型效果最差.因此在使用电压数据做训练时,采用随机森林算法进行建模效果最好,可以准确地识别出70.00%以上的窃电用户.

由表5可知,4种算法模型的 F_1 都低于70.00%,说明使用功率因数数据进行窃电用户识别的模型构建并不理想.对比这4种算法的表现,支持向量机分类模型相对于另外3种稍占优势,F_1 为68.60%,虽然该模型的精准率只有52.2%,但是其召回率可以达到100.00%.因此在实际应用中,如果需要尽可能覆盖所有窃电用户,可以考虑使用支持向量机进行模型的构建.

根据表3—5,计算每种算法的实验结果在

表 3　基于电流数据的多个算法模型评估结果

Table 3　The assessment results of multiple algorithm models based on the data of electric current　%

算法模型	精准率		召回率		F_1	
	引入线损率增长率前	引入线损率增长率后	引入线损率增长率前	引入线损率增长率后	引入线损率增长率前	引入线损率增长率后
随机森林	69.55	81.08	78.38	75.00	73.70	77.92
逻辑回归	52.98	68.97	60.04	50.00	56.29	57.97
决策树	72.33	73.90	85.67	85.00	78.44	79.07
支持向量机	43.45	58.70	78.77	67.50	56.01	62.79

表 4　基于电压数据的多个算法模型评估结果

Table 4　The assessment results of multiple algorithm models based on the data of voltage　%

算法模型	精准率		召回率		F_1	
	引入线损率增长率前	引入线损率增长率后	引入线损率增长率前	引入线损率增长率后	引入线损率增长率前	引入线损率增长率后
随机森林	59.66	72.80	75.23	71.30	66.55	72.00
逻辑回归	49.07	52.50	100.00	100.00	65.83	68.90
决策树	52.33	69.90	76.47	69.10	62.14	69.50
支持向量机	51.32	52.50	100.00	100.00	67.83	68.90

表 5　基于功率因数数据的多个算法模型评估结果

Table 5　The assessment results of multiple algorithm models based on power factor data　%

算法模型	精准率		召回率		F_1	
	引入线损率增长率前	引入线损率增长率后	引入线损率增长率前	引入线损率增长率后	引入线损率增长率前	引入线损率增长率后
随机森林	53.73	57.70	74.04	73.20	62.27	64.50
逻辑回归	35.02	55.70	92.06	89.00	50.74	68.50
决策树	57.01	61.40	66.34	65.90	61.32	63.50
支持向量机	43.02	52.20	100.00	100.00	60.16	68.60

不同数据集上的算术平均值,结果见表 6. 由表 6 可知,随机森林算法在 3 个数据集上的表现较优,精准率和 F_1 最高,分别为 70.53% 和 71.47%,召回率也达到 73.17%,因此从算法角度进行综合考量,可以认为随机森林是 4 种常见的分类算法中最为合适的模型构建算法.

本文用来训练的数据集主要是电流数据、电压数据和功率因数数据. 从数据的角度分析,可以了解使用哪个数据集进行模型构建更为准确高效. 根据表 3—5,计算每种数据集上 4 种算法实验结果的算术平均值,结果见表 7. 由表 7 可知,相对于电流数据和功率因数数据,使用电压数据的模型精准率虽然不高,但是其召回

表 6　使用不同算法模型的评估结果

Table 6　The assessment results of different algorithm models　%

算法模型	精准率	召回率	F_1
随机森林	70.53	73.17	71.47
逻辑回归	59.06	79.67	65.12
决策树	68.40	73.33	70.69
支持向量机	54.47	89.17	66.76

表 7　使用不同数据集构建模型的评估结果

Table 7　The assessment results of model construction with different dataset　%

数据集	精准率	召回率	F_1
电流	70.66	69.38	69.44
电压	61.93	85.10	69.83
功率因数	56.75	82.03	66.28

率和 F_1 都最高,说明可以覆盖更多的窃电用户,因此可以认为电压数据用于反窃电模型构建最合适;使用功率因数数据进行建模分析效果最差,精准率只有 56.75%,意味着实验中有近一半的疑似窃电用户识别错误,F_1 也只有 66.28%. 综合 F_1 分析,可以认为在实际应用中,使用电压数据构建的反窃电模型最为高效可行.

无论是从算法角度分析不同机器学习模型的综合表现,还是从数据角度对比不同数据进行模型训练的实验结果,引入线损率增长率后,3 个评估指标均可达到 70.00% 左右,因此可以认为本文提出的基于电力大数据的反窃电预测方法在疑似窃电用户识别方面高效可行.

3 结语

针对窃电事件屡禁不止、传统的反窃电手段准确度低的问题,提出了一种基于电力大数据的反窃电预测方法. 该方法融合异常规则分析和机器学习技术,首先根据异常规则构造窃电样本,引入线损率增长率,然后使用多种算法分别在电压、电流和功率因数数据集上构建预测模型,将其输出的数据异常用户与线损异常用户相结合,输出疑似窃电用户清单. 为了验证方法的可行性,分别从算法角度和数据角度对反窃电模型进行评估,结果表明,本文提出的基于电力大数据的反窃电预测方法在疑似窃电用户识别方面高效可行. 但是该方法目前只是针对高压用户构建模型,后续将会针对低压用户的数据进行研究分析.

参考文献:

[1] PAN W, YANG Q, AGGARWAL C, et al. Big data [J]. IEEE Intelligent Systems, 2007, 32(2): 7.

[2] 王红平,唐永锋. 大数据思维在高校学生信息化管理中的支撑作用[J]. 科技创新导报, 2018, 15(13): 231.

[3] JOSHI P, RAO P. Global pulses scenario: Status and outlook[J]. Annals of the New York Academy of Sciences, 2017, 1392(1): 6.

[4] SEVERIN A J. Dealing with data: Training new scientists[J]. Science, 2011, 331(6024): 1516.

[5] BIRNEY E. The making of ENCODE: Lessons for big-data projects [J]. Nature, 2012, 489 (7414): 49.

[6] 孙宏斌,郭庆来,潘昭光. 能源互联网:理念、架构与前沿展望[J]. 电力系统自动化, 2015, 39(19): 1.

[7] PULSE U. Big data for development: Challenges & opportunities [EB/OL]. (2012 - 11 - 05) [2018 - 09 - 10]. http://www.unglobalpulse.org/projects/BigdataDevelopment.

[8] AGRAWAL D, BERNSTEIN P, BERTINO E, et al. Challenges and opportunities with big data [EB/OL]. (2012 - 02 - 01) [2018 - 09 - 15]. http://www.cra.org/ccc/resources/ccc-led-white-papers.

[9] FANG X, MISRA S, XUE G, et al. Smart grid——The new and improved power grid: A survey [J]. IEEE Communications Surveys & Tutorials, 2012, 14(4): 944.

[10] CHEN L J, XU X H, WANG C M. Research on anti-electricity stealing method base on state estimation [C] // Proceedings of Power Engineering and Automation Conference (PEAM). Piscataway: IEEE, 2011.

[11] 李海. 用电监察面临的问题及反窃电措施[J]. 企业改革与管理, 2014(4): 119.

[12] 周瑾. 窃电与防窃电[J]. 电力与电工, 2004, 24(3): 73.

[13] 沈海泓. 远方电能计量运行监测系统研究[D]. 保定:华北电力大学(河北), 2004.

[14] 李小佳. 对反窃电技术研究及“零距离”复录系统的实现[D]. 广州:华南理工大学, 2011.

(下转第 95 页)

2011:2293.

[5] SUGANESHWAR G, IBRAHIM S P. A survey on collaborative filtering based recommendation system[C]// 3rd International Symposium on Big Data and Cloud Computing. Chengdu: Springer-Verlag, 2016:503.

[6] KOOHI H, KIANI K. User based collaborative filtering using fuzzy C-means[J]. Measurement, 2016, 91(1):134.

[7] SU X, KHOSHGOFTAAR T M. Collaborative filtering for multi-class data using belief nets algorithms[J]. International Journal on Artificial Intelligence Tools, 2008, 17(1):71.

[8] 孙小华,陈洪,孔繁胜. 在协同过滤中结合奇异值分解与最近邻方法[J]. 计算机应用研究,2006,23(9):206.

[9] 刘艺,冯钧,魏童童,等. 一种改进的协同过滤推荐算法[J]. 计算机与现代化,2017(1):1.

[10] SARWAR B, KARYPIS G, KONSTAN J, et al. Item-based collaborative filtering recommendation algorithms[C]// Proceedings of the 10th International Conference on World Wide Web. New York: ACM, 2001:285.

[11] PANDEY A, PANDEY R. Elective recommendation support through K-means clustering using R-tool[C]// International Conference on Computational Intelligence and Communication Networks. Kolkata: ICRCICN, 2016:851.

[12] HARPER F M, KONSTAN J A. The MovieLens datasets: History and context[J]. ACM Transactions on Interactive Intelligent Systems, 2015, 5(4):1.

[13] 陈清洁. 基于 SVD 的协同过滤推荐算法研究[D]. 成都:西南交通大学,2017.

(上接第 87 页)

[15] 李大勇,王瑜,黎灿兵,等. 基于无线射频技术的防窃电开箱记录仪设计[J]. 电测与仪表,2008,45(10):51.

[16] 孙凤杰,刘争芳,张永灿. 基于 GPRS 无线传输的防窃电系统[J]. 电力系统通信,2007,28(171):53.

[17] 余昌华,谢剑英. Winsocket 技术在电力远程监控系统中的应用[J]. 计算机工程,2000,26(10):81.

[18] 窦健,陈秀群,张海龙,等. 一种具有约束条件的用电异常检测模型:201711154836.7[P]. 2018－05－22.

[19] 吴迪,王学伟,窦健,等. 基于大数据的防窃电模型与方法[J]. 北京化工大学学报(自然科学版),2018,45(6):79.

[20] 庄池杰,张斌,胡军,等. 基于无监督学习的电力用户异常用电模式检测[J]. 中国电机工程学报,2016,36(2):379.

[21] 程超,张汉敬,景志敏,等. 基于离群点算法和用电信息采集系统的反窃电研究[J]. 电力系统保护与控制,2015,43(17):69.

[22] 王新霞,王珂,焦东翔,等. 基于正态分布离群点算法的反窃电研究[J]. 电气应用,2017,36(7):60.

[23] 窦健,刘宣,卢继哲,等. 基于用电信息采集大数据的防窃电方法研究[J]. 电测与仪表,2018,55(21):43.

[24] 任玮蒙,许庆,谢智奕,等. 基于电量、电压和电流异常分析的异常用电判断方法:201410706073.2[P]. 2015－03－11.