

基于平均影响值-启发式前向搜索的 异常光伏用户识别方法

陆双¹,彭曙蓉¹,杨云皓²,苏盛¹,刘登港¹,张恒¹,王书龙³

(1. 长沙理工大学 电气与信息工程学院,湖南 长沙 410114;2. 浙江大学
计算机科学与技术学院,浙江 杭州 310027;3. 武汉大学 电气与自动化学院,湖北 武汉 430072)

摘要:随着国家对光伏产业的大力推进与扶持,以及国家补贴政策具有长期性,出现了很多不法用户以虚假记录发电量骗取国家补贴的行为。针对现有的分布式光伏防窃电技术,提出了一种基于平均影响值(MIV)-启发式前向搜索的异常光伏用户识别方法。通过获取同一地区的标杆光伏用户及其他光伏用户在同一时段的发电数据,利用原始数据训练BP神经网络,再根据MIV的计算原理构造2组新的训练样本,用新样本的仿真结果计算各光伏用户的MIV,结合启发式前向搜索算法筛选得到与标杆光伏用户发电数据关联性大的用户,未被筛选的用户就是异常光伏用户。仿真结果验证了所提方法对异常光伏用户识别的有效性。

关键词:数据相关性;标杆光伏用户;异常光伏用户识别;MIV;启发式前向搜索;BP神经网络

中图分类号:TM 615

文献标志码:A

DOI:10.16081/j.epae.202112019

0 引言

自2013年3月国家电网公司出台《关于做好分布式电源并网服务工作的意见》以来,新能源发电产业不断兴起,大大小小的分布式光伏电站不断建成并网,极大地推进了我国光伏产业的飞速发展^[1],但也带来了不少的问题,分布式光伏电站一般仅向电网公司上传光伏发电系统的并网状态、发电量等信息,缺乏有效的辅助测量信息,存在用户通过一定的技术手段进行光伏窃电的行为。光伏窃电行为是在光伏发电系统中注入功率,但是注入的功率并不是来自光伏电池阵列输出的功率,而是来自电网的输出功率。光伏窃电用户将这些注入功率伪装成光伏电池阵列输出功率流过光伏电能计量表,使得光伏电能计量表虚高,进而获取高额补贴^[2-4]。

基于光伏特殊的发电方式,催生了很多新型的窃电手段,而光伏防窃电与传统防窃电有所不同。目前关于防窃电技术的研究主要针对传统窃电方式开展,对象侧重于电能计量表本体,每台电能计量表都需安装防窃电装置,工程量大且效果不好。文献[5]建立了分布式光伏并网信号的数学表征式,并基于时分割乘法器(TDM)电能计量模型,提出了适应分布式光伏动态特性的改进电能计量模型;文献[6]提出将关口电表数据随光伏计量表数据变化的斜率

作为虚增电量的检测指标,提高了检测便捷性。另一方面,对于光伏窃电现象的识别主要通过以下方法实现:基于线损率标杆值确定台区光伏窃电的严重程度,将发电量预测理论值与光伏计量值进行比较,再结合窃电嫌疑判别方法进行判定。文献[7]基于全天空云图,利用数字图像处理技术提取与辐射相关的图像特征,将大气层外辐射、大气质量、图像亮度和云量作为输入因子,较好地解决了云造成的影响;文献[8]采用欧氏距离对天气类型进行处理,将不同天气类型的平均发电功率之间的倍率关系映射为一个天气类型指数,该模型在各种天气类型下都有较准确的预测能力和较强的适用性;文献[9]提出了基于无监督学习的异常用电模式检测模型,适用于电力用户数据集缺乏训练样本的情况。但是影响光伏出力的因素众多^[10],预测模型难以进行全面考虑,导致预测精度不高以及识别准确率低。

本文针对现有的光伏窃电识别方法,发现电量数据相比气象数据更易获取齐全,且受影响较小,基于此提出了一种通过计算光伏用户的平均影响值MIV(Mean Impact Value)来表示同一地区其他光伏用户与当地标杆光伏用户发电量之间的关联程度,挖掘异常光伏用户的方法。MIV可反映神经网络中权重矩阵的变化情况,被认为是在神经网络中评价变量之间关联性最好的指标之一,可用于确定输入神经元对输出神经元的影响程度。该方法能够降低光伏用户发电量影响因素的干扰程度,MIV和前向搜索算法的组合运用提高了异常光伏用户的识别效率及精度^[11-13]。最后基于相关数据进行算例仿真,结果验证了所提识别方法的有效性。

收稿日期:2021-03-17;修回日期:2021-10-21

基金项目:国家自然科学基金资助项目(51777015);湖南省教育厅重点项目(20A021)

Project supported by the National Natural Science Foundation of China(51777015) and the Key Project of Hunan Education Department(20A021)

1 光伏电站地域相关性分析

1.1 所研究区域的光伏电站介绍

本文所研究区域为比利时某一包含11座光伏电站的区域,总面积为 $3.05 \times 10^4 \text{ km}^2$,地理位置图见附录A图A1。图中,地区5与地区11之间的直线距离约为220 km,地区8与地区9之间的直线距离约为140 km。各地区周边相邻基本为陆地(除了地区5、8附近),气候条件大致类似。分析各光伏电站的光伏出力相关性,根据相关性结果选定最终光伏电站的研究区域。

1.2 光伏电站出力相关性分析

为了说明本文所提异常光伏用户识别方法的可行性,需要考虑该区域同一时段内不同光伏电站之间的出力相关性,相关性反映了同一时段内不同光伏电站出力之间的关系。对于包含 N 座光伏电站的光伏基地而言,一般认为电能计量表会维持自身相应时间的读数,每15 min记录1次数据,那么不考虑延时的相关系数矩阵为:

$$\mathbf{R} = [\rho_{c,d}]_{N \times N} \quad (1)$$

$$\rho_{c,d} = \frac{\sum_{i=1}^n (P_i^c - \bar{P}^c)(P_i^d - \bar{P}^d)}{\sqrt{\sum_{i=1}^n (P_i^c - \bar{P}^c)^2 \sum_{i=1}^n (P_i^d - \bar{P}^d)^2}} \quad (2)$$

式中: $\mathbf{R}$ 为同一时段内 N 座光伏电站之间的出力相关系数矩阵; $\rho_{c,d}$ 为同一时段光伏电站 c 与光伏电站 d 之间的出力相关系数; P_i^c 、 P_i^d 分别为第 i 个时间点光伏电站 c 、 d 的发电量; $\bar{P}^c$ 、 $\bar{P}^d$ 分别为光伏电站 c 、 d 发电量的平均值; n 为时间点数量。

本文在比利时区域选取11座光伏电站(以地区编号标记光伏电站编号)进行出力相关性分析,根据式(2)分别讨论11座光伏电站在小时时间尺度和日时间尺度下的出力相关性,即分别取 $n=4$ 和 $n=49$ 计算各光伏电站两两之间的出力相关系数矩阵,从而选取地理位置合适的光伏电站。光伏电站小时时间尺度、日时间尺度的出力曲线分别见附录A图A2和图A3。小时时间尺度、日时间尺度出力相关性结果分别如附录A表A1和表A2所示(表中加粗数据为相关系数小于0.5的数据)。由表可知:在日时间尺度下,本文所研究比利时区域的11座光伏电站之间的出力相关性都比较高;但在小时时间尺度下,光伏电站2、5、11之间的出力相关性比较低,故本文最终选取研究其他8座光伏电站。其他8座光伏电站之间的出力相关系数较接近1,反映了光伏电站之间的出力相关性较高,为此能够对这些光伏电站统一进行光伏出力异常行为分析,为后文的分析提供可行性依据。

2 光伏用户行为分析

2.1 标杆光伏用户

分布式光伏发电是指在用户场地附近建设的光伏电站,运行方式以“用户侧自发自用、多余电量上网”为主。由于国家会给予光伏用户高额的补贴,不少用户利用技术手段使电能计量表数据虚高。

通过从电站发电能力、故障情况、安全文明生产、运行管理、班组建设、设备运行等多方面进行综合评估后得到标杆光伏用户,其能够代表该地区光伏电站的发电水平。通过计算同一地区其他光伏用户与当地标杆光伏用户发电量之间的相关性,可以挖掘异常光伏用户,有效识别光伏窃电行为。

2.2 光伏用户发电数据及数据处理

2.2.1 数据选取原则

光伏电站的有效发电时段为06:00—18:00,在此基础上获取比利时区域8座光伏电站连续多天且时间间隔为15 min的发电数据,具体天数需根据光伏电站的实际发电情况而定,本文选取1个月的发电量数据,且选定地区3(Flemish_Brabant)用户为标杆光伏用户。定义该地区的光伏用户集合 $X = \{X_1, X_2, \dots, X_N\}$ (N 为光伏用户的数量; X_1 为标杆光伏用户; X_i ($i=2, 3, \dots, N$)为其他光伏用户),根据上述数据计算得到光伏用户 X_i ($i=1, 2, \dots, N$)的日发电量数据序列 $X_i = \{x_{i1}, x_{i2}, \dots, x_{iT}\}$ (x_{ij} ($j=1, 2, \dots, T$)为光伏用户 X_i 在第 j 天的发电量, T 为所选取光伏用户发电数据的天数),并根据式(3)进行归一化处理,得到归一化发电量数据序列 $Y_i = \{y_{i1}, y_{i2}, \dots, y_{iT}\}$ 。

$$y_{ij} = \frac{x_{ij} - \min_{i'=1,2,\dots,N} \{X_{i'}\}}{\max_{i'=1,2,\dots,N} \{X_{i'}\} - \min_{i'=1,2,\dots,N} \{X_{i'}\}} \quad (3)$$

式中: $y_{ij} \in [0, 1]$ 。

标杆光伏用户 X_1 连续 T 天的发电量数据为:

$$[x_{11} \ x_{12} \ \dots \ x_{1T}] \quad (4)$$

其他光伏用户连续 T 天的发电量数据为:

$$\begin{bmatrix} x_{21} & x_{22} & \dots & x_{2T} \\ x_{31} & x_{32} & \dots & x_{3T} \\ \vdots & \vdots & & \vdots \\ x_{N1} & x_{N2} & \dots & x_{NT} \end{bmatrix} \quad (5)$$

2.2.2 BP神经网络模型训练

在人工神经网络中,BP神经网络是目前研究最为成熟、应用最为广泛的神经网络模型之一^[14-15],因此本文选用BP神经网络仿真得到与MIV相关的影响变化值(IV)。BP神经网络结构图如图1所示,其包含输入层、隐含层、输出层。在BP神经网络模型训练过程中,输入量为除标杆光伏用户外其他光伏用户的发电量数据,将标杆光伏用户的发电量数据作为网络输出量。BP神经网络输入层、隐含层、输

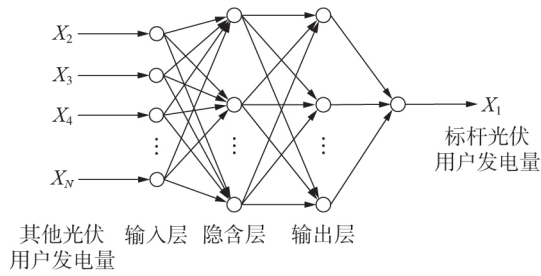


图1 BP神经网络结构图

Fig.1 Structure diagram of BP neural network

出层的定义式如式(6)~(8)所示。

$$\alpha_h = \sigma \left(\sum_{i=1}^D v_{ih} X_i \right) \quad (6)$$

$$\beta_j = \sigma \left(\sum_{h=1}^q w_{hj} \alpha_h \right) \quad (7)$$

$$\hat{y}_k = \sigma \left(\sum_{j=1}^p u_{jk} \beta_j \right) \quad (8)$$

式中: α_h 为隐含层的第 h 个神经元接收到的输入值; β_j 为隐含层的第 j 个神经元输入输出层的结果; $\hat{y}_k$ 为输出层第 k 个神经元的输出值; $\sigma(\cdot)$ 为激活函数,用于增加网络非线性拟合能力; v_{ih} 为输入层的第 i 个神经元与第1个隐含层的第 h 个神经元之间的连接权重; w_{hj} 为第1个隐含层的第 h 个神经元与第2个隐含层的第 j 个神经元之间的连接权重; u_{jk} 为第2个隐含层的第 j 个神经元与输出层的第 k 个神经元之间的连接权重; D 、 q 、 p 分别为输入层、隐含层、输出层的神经元个数。

当实际输出与期望输出不符时,进入误差的反向传播阶段。误差通过输出层,按误差梯度下降的方式修正各层权重值,向隐含层、输入层逐层反向传播。周而复始的信息正向传播和误差反向传播过程,是各层权重值不断调整的过程,也是神经网络学习训练的过程,重复该过程直至网络输出的误差减小到可以接受的程度。

3 基于MIV-启发式前向搜索的异常识别模型

本文所提异常识别方法能够降低光伏用户发电量影响因素的干扰程度,通过计算MIV和运用启发式前向搜索算法,提高了异常光伏用户的检测效率与精度。本文所提异常光伏用户的识别流程图如图2所示。

3.1 MIV的计算准则

MIV可反映BP神经网络中权重矩阵的变化情况,被认为是在神经网络中评价变量之间关联性最好的指标之一,可用于确定输入神经元对输出神经元的影响程度,其数值的正负性表示自变量对因变量产生影响的方向。

可以根据MIV计算得到当地其他光伏用户与标

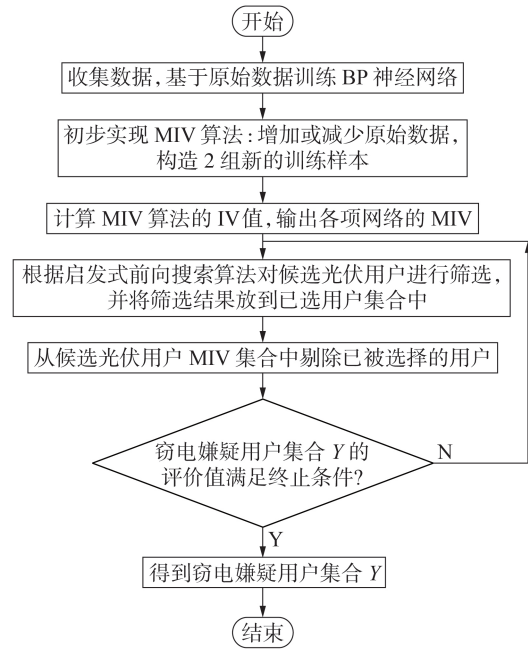


图2 异常光伏用户的识别流程图

Fig.2 Flowchart of identifying abnormal photovoltaic users

杆光伏用户发电量之间的相关性,具体计算过程如下:BP神经网络训练终止后,在原有发电量数据样本的基础上构建新的训练样本,将除标杆光伏用户之外的其他光伏用户的发电量数据在其原始数据基础上增加10%和降低10%构成2组新的训练样本,分别记为样本 S_1 和 S_2 ;然后将样本 S_1 和 S_2 分别作为网络的输入仿真样本,利用已建成的BP神经网络进行仿真,将标杆光伏用户的发电量作为网络输出,得到2组仿真结果,分别记为 A_1 和 A_2 ,它们之间的差值就是变动原始数据后对网络输出产生的IV;最后将IV除以天数 T 得到光伏用户与标杆光伏用户之间的MIV,即光伏用户的MIV集合 $M = \{M_{MIV2}, M_{MIV3}, \dots, M_{MIVN}\}$ (M_{MIVi} ($i=2, 3, \dots, N$)为光伏用户 X_i 的MIV), M_{MIVi} 的绝对值大小反映了光伏用户与当地标杆光伏用户发电量之间的相关性程度大小。

3.2 启发式前向搜索算法原理

采用启发式前向搜索算法逐个选取特征,即从空集开始,从当前候选光伏用户的MIV集合 $M = \{M_{MIV2}, M_{MIV3}, \dots, M_{MIVN}\}$ 中选取MIV最大的光伏用户放置到已选用户集中,并在候选光伏用户集合中将其剔除,直至已选用户集合的MIV达到要求,即所选光伏用户 X_j 的MIV为:

$$M_{MIVy} = \operatorname{argmax} \{ M_{MIVi} \mid i=2, 3, \dots, N \} \quad (9)$$

当筛选的光伏用户集合的评价值满足式(10)所示终止条件时,停止光伏用户的筛选,判定未被筛选的用户为异常光伏用户。

$$D' = \frac{\max\{M\}}{\sum M_{MIVy}} \leq 0.1 \quad (10)$$

式中: D' 为未被筛选的光伏用户MIV中的最大值与已筛选光伏用户MIV总和的比值,当其不大于所设定的终止阈值0.1时,停止对光伏用户的筛选。

上述方法能同时兼顾计算效率与选择效果,实现了以较小的计算量得到较好的筛选效果。

4 算例分析

本文从比利时区域选取8个光伏电站用户(地区1、3、4、6—10用户)连续1个月(5月份)06:00—18:00以15 min为时间间隔的发电量数据,其中地区3(Flemish_Brabant)用户为标杆光伏用户。为了使本文方法的有效性得到充分体现,特地人为构造了4组窃电用户数据,4组窃电用户数据的构造准则为:构造用户1,将光伏发电量按比例放大一定的倍数;构造用户2,在某特定时段将光伏发电量放大一定的倍数;构造用户3,在原始数据缩放的基础上随机增加某一数值,但一天中最大值出现时间点的数据保持不变;构造用户4,上述准则两两混合构造^[16-17]。

按照本文所提方法,采用原始发电量数据训练BP神经网络,结合MIV计算方法利用已建成的BP神经网络仿真得到除标杆光伏用户之外其他光伏用户的MIV,然后利用启发式前向搜索算法进行筛选,直到满足终止条件,未被筛选的用户则为异常光伏用户。采用MATLAB进行算例仿真,电脑配置为2.6 GHz,8 GB RAM。

在利用原始发电量数据训练BP神经网络时,将标杆光伏用户作为网络输出,其他11个光伏用户作为网络输入;以5 d为时间间隔对1个月分6次进行MIV计算,这样可以排除一些非窃电故障因素的影响,减小误报率。经过仿真计算得到各光伏用户的6次MIV结果如附录A表A3所示。然后采用启发式前向搜索算法对光伏用户进行筛选,所得结果如表1所示(表中加粗数据为未被筛选用户的MIV)。各用户某一周内的光伏发电量曲线如图3所示。

由表1可知,人为构造的窃电用户基本上每次都被筛选出来,结合图3所示发电量数据可以验证

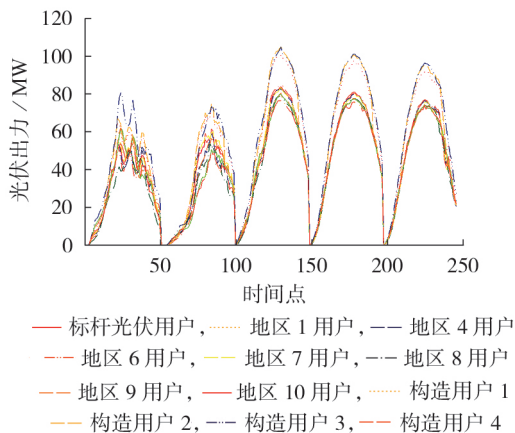


图3 各用户的发电量曲线

Fig.3 Power generating curve of each user

本文方法的有效性。与标杆光伏用户地区邻近的其他地区光伏用户被判定为正常用户,除了地区9和地区10用户分别有1次和2次未被筛选,有可能是因为这2个地区相较于其他地区离标杆光伏用户地区稍远,超出了适用范围(140 km),也有可能是因为光伏用户在那段时间内某个部件出现故障现象,此时先不判定窃电嫌疑,如果后续又出现未被筛选的情况,则判定其为异常光伏用户,而且可能是从首次未被筛选时开始窃电,需要对其重新进行检查。本文后续也进行了1月份及8月份相关数据的仿真,所得结果均能很好地识别窃电嫌疑用户,验证了所提方法具有较好的普适性。

为了进一步体现本文所提方法的有效性,将其与以下2种方法进行对比验证。

1)方法1,利用光伏发电量容量比及其偏离值分析识别异常光伏用户。首先计算每天各用户发电量容量比和地区光伏用户总发电量容量比,然后核查容量比偏离值 $K(K=(P_{av}-\bar{P}_{total,av})/\bar{P}_{total,av}, P_{av}$ 为光伏用户发电量容量比, $\bar{P}_{total,av}$ 为全地区平均发电量容量比),并计算30 d的K值之和 K_{total} 。若 $K_{total}>30$,则将该用户列为重点现场排查用户。上述11个光伏用户的容量比偏离值结果如表2所示。由表可知,方法1只适用于识别窃电非常严重的用户,对于其他情况无法进行有效排查。

表1 启发式前向搜索算法的筛选结果

Table 1 Filter results of heuristic forward searching algorithm

序号	MIV										
	地区1 用户	地区4 用户	地区6 用户	地区7 用户	地区8 用户	地区9 用户	地区10 用户	构造 用户1	构造 用户2	构造 用户3	构造 用户4
1	0.2380	0.2830	0.3143	0.2783	0.2651	0.2601	0.2019	0.1966	0.1719	0.1620	0.2107
2	0.1961	0.3706	0.2115	0.2092	0.1978	0.1310	0.1816	-0.1129	0.1186	0.1053	0.0980
3	0.4491	0.4344	0.4995	0.4156	0.4030	0.3784	0.3743	-0.2342	0.1958	0.3610	0.2006
4	0.3464	0.6172	0.3240	0.3094	0.3038	0.3008	-0.1950	0.1899	0.1857	-0.1690	0.2010
5	0.3549	0.3225	0.3139	0.2781	0.2675	0.2645	0.2022	-0.1340	-0.1367	0.0830	-0.1083
6	0.1710	0.1480	0.3410	0.1320	0.0900	0.1060	0.0650	0.0500	0.0610	0.0550	0.1130

表2 各用户的容量比偏离值

Table 2 Deviation values of users' capacity ratio

用户	容量比偏离值	用户	容量比偏离值
地区1用户	1.8309	地区10用户	4.2019
地区4用户	3.7969	构造用户1	15.8697
地区6用户	1.0908	构造用户2	22.2243
地区7用户	0.5033	构造用户3	29.0156
地区8用户	2.1664	构造用户4	16.7565
地区9用户	4.8907		

2)方法2,利用发电量频率分布相对熵进行异常光伏用户识别。将本文所用仿真数据按各自时段内的最大发电量数值进行标准化,并将其等分为10个频率区间,但由于数据量少且轻度异常时无明显频率特性特征,在频率特性的基础上运用相对熵进一步识别异常用户,但需要设置合理的时间窗口和阈值,比较复杂,该方法的实现要求较高。

综上可知,本文所提方法可以降低光伏用户发电量影响因素的干扰程度,同时系统性地提高异常光伏用户的识别效率及精度。

5 结论

本文首次提出采用MIV评估光伏用户与区域内标杆光伏用户之间的关联程度,并结合启发式前向搜索算法进行光伏用户MIV筛选,未被筛选的用户被判定为异常光伏用户。本文方法仿真所需数据为光伏发电量数据,不再受制于气象数据的准确度以及预测的精度。同时考虑以5d为时间间隔进行筛选,很好地排除了光伏出力的故障因素。经过对比验证可知,本文方法可以降低光伏用户发电量影响因素的干扰程度,同时系统性地提高异常光伏用户的识别效率及精度。

附录见本刊网络版(<http://www.epae.cn>)。

参考文献:

- [1] 黄爱颖,赵新,袁新润,等.光伏发电防作弊系统:CN1033-89397A[P]. 2013-11-13.
- [2] 王飞.并网型光伏电站发电功率预测方法与系统[D].北京:华北电力大学,2013.
WANG Fei. Power forecasting approach and system of grid-connected PV plant[D]. Beijing: North China Electric Power University, 2013.
- [3] 罗一凡,蒋传文,李春哲,等.考虑分布式电源入网的反窃电综合管理方法[J].电器与能效管理技术,2015(4):49-55,59.
LUO Yifan, JIANG Chuanwen, LI Chunzhe, et al. Integrated management method of anti-electricity stealing considering distributed generation[J]. Electrical & Energy Management Technology, 2015(4):49-55, 59.
- [4] 李燕青,杜莹莹.基于双维度顺序填补框架与改进Kohonen天气聚类的光伏发电短期预测[J].电力自动化设备,2019,39(1):60-65.
LI Yanqing, DU Yingying. Short-term photovoltaic power forecasting based on double-dimensional sequential imputation framework and improved Kohonen clustering[J]. Electric Po-

wer Automation Equipment, 2019, 39(1):60-65.

- [5] 窦晓波,杨磊,马建,等.适应分布式光伏动态特性的改进电能计量模型[J].电力自动化设备,2017,37(10):34-42.
DOU Xiaobo, YANG Lei, MA Jian, et al. Improved model for power metering based on dynamic characteristics of distributed photovoltaic power generation[J]. Electric Power Automation Equipment, 2017, 37(10):34-42.
- [6] 谢云云,袁晓冬,孙正洋,等.基于光伏模拟器的虚增电量特性分析与检测方法[J].电网技术,2016,40(6):1703-1708.
XIE Yunyun, YUAN Xiaodong, SUN Zhengyang, et al. Feature analysis and detection method research of inflated electricity using PV generation simulator[J]. Power System Technology, 2016, 40(6):1703-1708.
- [7] 陈志宝,丁杰,周海,等.地基云图结合径向基函数人工神经网络的光伏功率超短期预测模型[J].中国电机工程学报,2015,35(3):561-567.
CHEN Zhibao, DING Jie, ZHOU Hai, et al. A model of very short-term photovoltaic power forecasting based on ground-based cloud images and RBF neural network[J]. Proceedings of the CSEE, 2015, 35(3):561-567.
- [8] 袁晓玲,施俊华,徐杰彦.计及天气类型指数的光伏发电短期出力预测[J].中国电机工程学报,2013,33(34):57-64.
YUAN Xiaoling, SHI Junhua, XU Jieyan. Short-term power forecasting for photovoltaic generation considering weather type index[J]. Proceedings of the CSEE, 2013, 33(34):57-64.
- [9] 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):379-387.
ZHUANG Chijie, ZHANG Bin, HU Jun, et al. Anomaly detection for power consumption patterns based on unsupervised learning[J]. Proceedings of the CSEE, 2016, 36(2):379-387.
- [10] 李建文,焦衡,刘凤梧,等.基于相似时段的分时段光伏出力短期预测[J].电力自动化设备,2018,38(8):183-188.
LI Jianwen, JIAO Heng, LIU Fengwu, et al. Short-time segmented photovoltaic output forecasting based on similar period[J]. Electric Power Automation Equipment, 2018, 38(8):183-188.
- [11] 苏适,李康平,严玉廷,等.基于密度空间聚类和引力搜索算法的居民负荷用电模式分类模型[J].电力自动化设备,2018,38(1):129-136.
SU Shi, LI Kangping, YAN Yuting, et al. Classification model of residential power consumption mode based on DBSCAN and gravitational search algorithm[J]. Electric Power Automation Equipment, 2018, 38(1):129-136.
- [12] 陆俊,朱炎平,彭文昊,等.智能用电用户行为分析特征优选策略[J].电力系统自动化,2017,41(5):58-63,83.
LU Jun, ZHU Yanping, PENG Wenhao, et al. Feature selection strategy for electricity consumption behavior analysis in smart grid[J]. Automation of Electric Power Systems, 2017, 41(5):58-63, 83.
- [13] 王英立,陶帅,候晓晓,等.基于MIV分析的GA-BP神经网络光伏短期发电预测[J].太阳能学报,2020,41(8):236-242.
WANG Yingli, TAO Shuai, HOU Xiaoxiao, et al. GA-BP neural network photovoltaic power generation short-term forecast based on MIV analysis[J]. Acta Energiae Solaris Sinica, 2020, 41(8):236-242.
- [14] 陆俊,陈志敏,龚钢军,等.基于极限学习机的居民用电行为分类分析方法[J].电力系统自动化,2019,43(2):97-104.
LU Jun, CHEN Zhimin, GONG Gangjun, et al. Classification analysis method for electricity consumption behavior based on extreme learning machine algorithm[J]. Automation of Electric Power Systems, 2019, 43(2):97-104.
- [15] 杨志淳,靖晓平,乐健,等.基于MI-PSO-BP算法的配电设备状态实时评估方法[J].电力自动化设备,2019,39(12):108-113.
YANG Zhichun, JING Xiaoping, LE Jian, et al. Real-time condition assessment method based on MI-PSO-BP algorithm for

distribution equipment[J]. Electric Power Automation Equipment, 2019, 39(12): 108-113.

- [16] KRISHNA V B, GUNTER C A, SANDERS W H. Evaluating detectors on optimal attack vectors that enable electricity theft and DER fraud[J]. IEEE Journal of Selected Topics in Signal Processing, 2018, 12(4): 790-805.
- [17] ISMAIL M, SHAABAN M F, NAIDU M, et al. Deep learning detection of electricity theft cyber-attacks in renewable distributed generation[J]. IEEE Transactions on Smart Grid, 2020, 11(4): 3428-3437.

作者简介:



陆 双

陆 双(1997—),女,湖南衡南人,硕士研究生,主要研究方向为智能电网中的信息处理技术(E-mail: 1836996837@qq.com);

彭曙蓉(1975—),女,湖南长沙人,副教授,博士,主要研究方向为智能信息处理(E-mail: 173764138@qq.com)。

(编辑 陆丹)

Identification method of abnormal photovoltaic users based on mean impact value and heuristic forward searching

LU Shuang¹, PENG Shurong¹, YANG Yunhao², SU Sheng¹, LIU Denggang¹, ZHANG Heng¹, WANG Shulong³

(1. School of Electrical & Information Engineering, Changsha University of Science & Technology, Changsha 410114, China;

2. College of Computer Science and Technology, Zhejiang University, Hangzhou 310027, China;

3. School of Electrical Engineering and Automation, Wuhan University, Wuhan 430072, China)

Abstract: With the country's vigorous promotion and support of the photovoltaic industry and the long-term nature of the national subsidy policy, many illegal users cheat the state subsidy by falsely recording the electricity generation. Aiming at the existing distributed photovoltaic anti-theft technologies, an identification method of abnormal photovoltaic users based on MIV (Mean Impact Value) and heuristic forward searching is proposed. By obtaining the power generation data of benchmark photovoltaic users and other photovoltaic users in the same time segment and the same area, the original data are used to train the BP neural network. And then according to the MIV calculation principle, two sets of new training samples are constructed, and the simulative results of new samples are used to calculate the MIV of each photovoltaic user. Combining heuristic forward searching algorithm, the users with high power generation data correlation with the benchmark photovoltaic users are filtered, and the unfiltered users are abnormal photovoltaic users. The simulative results verify the effectiveness of the proposed method for identifying abnormal photovoltaic users.

Key words: data correlation; benchmark photovoltaic users; identification of abnormal photovoltaic users; mean impact value; heuristic forward searching; BP neural network

(上接第91页 continued from page 91)

Multi-stage planning of offshore wind farm considering wind power accommodation level and load increase

JIANG Yuewen^{1,2,3}, ZHANG Jinhui¹

(1. College of Electrical Engineering and Automation, Fuzhou University, Fuzhou 350108, China;

2. Fujian Province University Engineering Research Center of Smart Distribution Grid Equipment, Fuzhou 350108, China;

3. Research Center of Integrated Energy Planning and Optimal Operation, Fuzhou University, Fuzhou 350108, China)

Abstract: The current planning and construction of offshore wind farm rarely considers the construction time sequence, which mostly belongs to static single-stage planning and is easily lead to the lack of effective connection between the initial planning scheme and the operation at middle and late stages, for which, a multi-stage planning scheme for offshore wind farm considering wind power accommodation level is proposed. The day-ahead market trading model with the lowest electricity purchase cost as its objective is established to evaluate the maximum accommodation level of power grid to offshore wind power, and the maximum accommodation levels of regional power grid to offshore wind power are obtained under different load levels in the future years. On this basis, the planning period is divided into several stages, and a multi-stage planning model of offshore wind farm with the maximum net revenue of wind farm as its objective is built, which takes the investment year, number, location and the belonging investment stage of wind turbines as its optimization variables. Examples verify the effectiveness of multi-stage planning model in improving economic benefit of offshore wind farm compared with single-stage planning model.

Key words: offshore wind farm; multi-stage planning; wind power accommodation level; day-ahead market; installed capacity optimization

附录 A

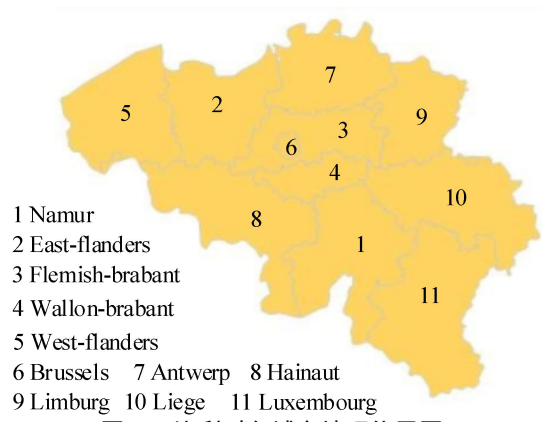


图 A1 比利时各城市地理位置图
Fig.A1 Location map of cities in Belgium

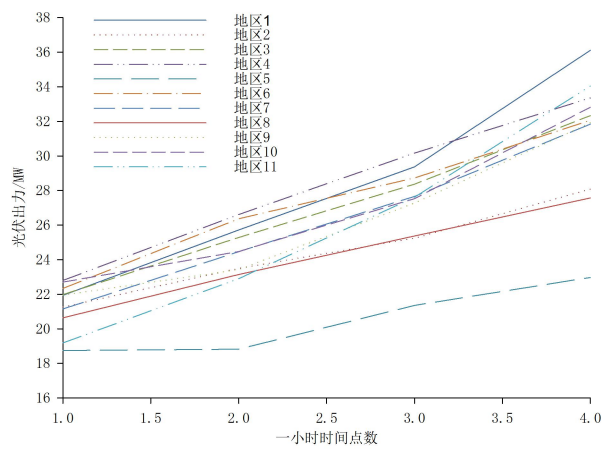


图 A2 各地区小时时间尺度的光伏出力曲线
Fig.A2 Hourly time scale photovoltaic output curves of various regions

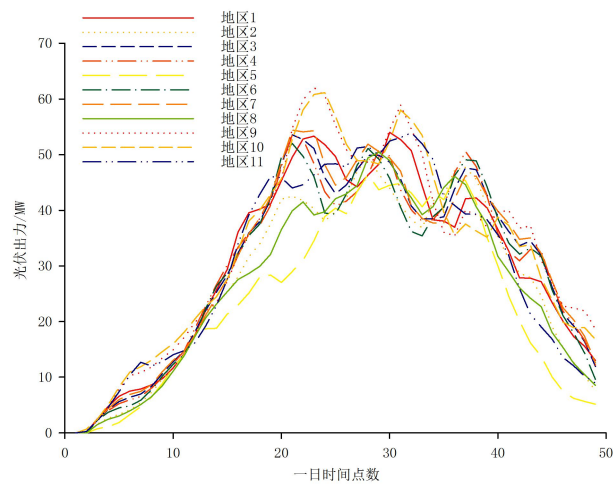


图 A3 各地区日时间尺度的光伏出力曲线
Fig.A3 Day time scale photovoltaic output curves of various regions

表 A1 光伏电站小时时间尺度的出力相关性
Table A1 Output correlation of photovoltaic power plants under hourly time scale

地区	出力相关系数										
	地区 1	地区 2	地区 3	地区 4	地区 5	地区 6	地区 7	地区 8	地区 9	地区 10	地区 11
1	1	0.58149	0.98644	0.96717	0.28047	0.91045	0.89832	0.77251	0.88415	0.98537	0.99530
2	0.58149	1	0.55049	0.57947	0.61378	0.55663	0.55602	0.67368	0.53971	0.47572	0.37162
3	0.98644	0.55049	1	0.91253	0.40137	0.98463	0.99363	0.78101	0.94402	0.95579	0.66432
4	0.96717	0.57947	0.91253	1	0.25375	0.96935	0.95238	0.73236	0.75425	0.98126	0.73802
5	0.28047	0.61378	0.40137	0.25375	1	0.31783	0.33335	0.67064	0.19264	0.28260	0.34775
6	0.91045	0.55663	0.98463	0.96935	0.31783	1	0.99803	0.79117	0.98710	0.90050	0.54978
7	0.89832	0.55602	0.99363	0.95238	0.33335	0.99803	1	0.79061	0.97515	0.91317	0.46654
8	0.77251	0.67368	0.78101	0.73236	0.67064	0.79117	0.79061	1	0.78469	0.83643	0.77501
9	0.88415	0.53971	0.94402	0.75425	0.19264	0.98710	0.97515	0.78469	1	0.99551	0.56321
10	0.98537	0.47572	0.95579	0.98126	0.28260	0.90050	0.91317	0.83643	0.99551	1	0.90899
11	0.99530	0.37162	0.66432	0.73802	0.34775	0.54978	0.46654	0.77501	0.56321	0.90899	1

表 A2 光伏电站日时间尺度的出力相关性
Table A2 Output correlation of photovoltaic power plants under day time scale

地区	出力相关系数										
	地区 1	地区 2	地区 3	地区 4	地区 5	地区 6	地区 7	地区 8	地区 9	地区 10	地区 11
1	1	0.98252	0.98268	0.98791	0.98305	0.97303	0.97891	0.99017	0.99197	0.98999	0.90488
2	0.98252	1	0.99482	0.99463	0.98380	0.99494	0.99118	0.99521	0.97517	0.96854	0.87919
3	0.98268	0.99482	1	0.99856	0.96919	0.99786	0.99921	0.98900	0.98192	0.96522	0.84557
4	0.98791	0.99463	0.99856	1	0.97543	0.99483	0.99649	0.99349	0.98449	0.96901	0.85590
5	0.98305	0.98380	0.96919	0.97543	1	0.96250	0.96221	0.99119	0.97371	0.97971	0.93657
6	0.97303	0.99494	0.99786	0.99483	0.96250	1	0.99677	0.98536	0.96859	0.95227	0.83292
7	0.97891	0.99118	0.99921	0.99649	0.96221	0.99677	1	0.98388	0.98117	0.96130	0.83389
8	0.99017	0.99521	0.98900	0.99349	0.99119	0.98536	0.98388	1	0.97964	0.97392	0.89500
9	0.99197	0.97517	0.98192	0.98449	0.97371	0.96859	0.98117	0.97964	1	0.99022	0.88754
10	0.98999	0.96854	0.96522	0.96901	0.97971	0.95227	0.96130	0.97392	0.99022	1	0.93487
11	0.90488	0.87919	0.84557	0.85590	0.93657	0.83292	0.83389	0.89500	0.88754	0.93487	1

表 A3 除标杆用户外其他光伏用户的 MIV
Table A3 MIVs of other photovoltaic users besides benchmark user

序号	MIV										
	地区 1 用户	地区 4 用户	地区 6 用户	地区 7 用户	地区 8 用户	地区 9 用户	地区 10 用户	构造用户 1	构造用户 2	构造用户 3	构造用户 4
1	0.2380	0.2830	0.3143	0.2783	0.2651	0.2601	0.2019	0.1966	0.1719	0.1620	0.2107
2	0.1961	0.3706	0.2115	0.2092	0.1978	0.1310	0.1816	-0.1129	0.1186	0.1053	0.0980
3	0.4491	0.4344	0.4995	0.4156	0.4030	0.3784	0.3743	-0.2342	0.1958	0.3610	0.2006
4	0.3464	0.6172	0.3240	0.3094	0.3038	0.3008	-0.1950	0.1899	0.1857	-0.1690	0.2010
5	0.3549	0.3225	0.3139	0.2781	0.2675	0.2645	0.2022	-0.1340	-0.1367	0.0830	-0.1083
6	0.1710	0.1480	0.3410	0.1320	0.0900	0.1060	0.0650	0.0500	0.0610	0.0550	0.1130