



(12) 发明专利

(10) 授权公告号 CN 108593990 B

(45) 授权公告日 2021.06.01

(21) 申请号 201810562347.3

审查员 汤莎亮

(22) 申请日 2018.06.04

(65) 同一申请的已公布的文献号  
申请公布号 CN 108593990 A

(43) 申请公布日 2018.09.28

(73) 专利权人 国网天津市电力公司  
地址 300010 天津市河北区五经路39号  
专利权人 国家电网公司

(72) 发明人 冯瑛敏 任国岐 赵晶 黄丽妍  
刘瑾 毛华 赵新 谢秦

(74) 专利代理机构 天津才智专利商标代理有限公司 12108  
代理人 王颢

(51) Int. Cl.

G01R 11/24 (2006.01)

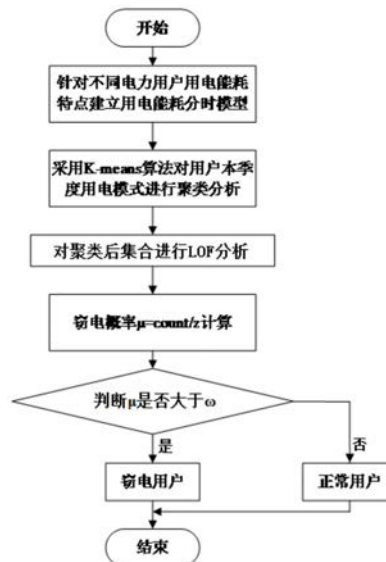
权利要求书1页 说明书6页 附图6页

(54) 发明名称

一种基于电能用户用电行为模式的窃电检测方法和应用

(57) 摘要

本发明属于电力检测技术领域,特别涉及电力窃电检测技术领域,尤其涉及一种基于电能用户用电行为模式的窃电检测方法和应用。本发明的检测方法从用户的用电行为特性出发,依据不同电力用户用电能耗特点,考虑行为习惯、气候、季节等因素,建立不同用户能耗分时模型。基于K-means聚类算法与LOF算法实现用户电能计量数据异常检测后,考虑到电能用户存在合理的用电行为改变,并不能准确判定用户存在窃电行为。本发明结合用电信息采集系统计量电能损耗,根据台区下能量消耗公式,引入窃电概率,对电能窃电概率进行计算,实现电能用户不同类型窃电行为的可靠监测。



1. 一种基于电能用户用电行为模式的窃电检测方法,其特征在于,所述检测方法至少包括以下步骤:

(1) 建立能耗分时模型:基于用电信息采集系统中的用户电能计量信息,建立不同类型用户的能耗分时模型;

所述用户的不同类型包括大型专变用户、中小型专变用户、工行业用户和城镇居民用户;

(2) 将不同类型用户的用电行为数据进行聚类,得到聚类后集合;

(3) 对聚类后集合进行分析,定义局部离群因子 $LOF_{k(p)}$ ,并对疑似窃电行为进行筛选;

(4) 结合用电信息采集系统计量电能损耗,根据台区下能量消耗公式,进行用户诚信度评价,实现窃电用户监测;

设某一台区下电能用户 $m$ 的第 $j$ 个计量时刻的数值为 $x_{mj}$ ,电网线损为 $E_{TL}$ ,为避免因网络中设备计量以及温度对线损带来的影响,设置补偿阈值为 $\delta$ ,其取值与台区的规模有关,则某一台区下能耗关系为:

$$E_{delivered} - (x_{1j} + x_{2j} + \dots + x_{mj} + E_{TL}) \leq \delta \quad (4)$$

$E_{delivered}$  表示总表的电量;

若用户数据异常时,台区下能耗满足公式(4),则认为该数据为正常数据;反之,则认为该用户存在窃电可能;

设检测的用户集合数据条数为 $z$ ,count为数据集中不满足上述公式的采集点量,则窃电概率表示为:

$$\mu = \frac{\text{count}}{z} \quad (5)$$

设 $\omega$ 为窃电概率检测阈值,若 $\mu > \omega$ 则判定该用户为窃电用户,相反则为正常用户。

2. 根据权利要求1所述的检测方法,其特征在于,

考虑行为习惯、气候、季节因素中的至少一种,建立能耗分时模型。

3. 根据权利要求1-2中任一项所述的检测方法在电能用户窃电行为监测方面的应用。

## 一种基于电能用户用电行为模式的窃电检测方法和应用

### 技术领域

[0001] 本发明属于电力检测技术领域,特别涉及电力窃电检测技术领域,尤其涉及一种基于电能用户用电行为模式的窃电检测方法和应用。

### 背景技术

[0002] 随着智能电网建设的不断推进以及能源格局的变化,电能在经济社会中扮演着越来越重要的角色。用电信息采集系统作为对用户电能信息采集、分析处理和 数据应用的平台,在现代化电力营销及智能电网双向互动方面发挥着重要作用, 余额报警、远程充值等业务的应用,为人们的生活提供了更多的便利。然而,半 开放网络结构以及智能电表硬件资源限制,同样带来了用电信息采集系统安全问题。如何在电网线路损耗等技术性损耗存在的前提下,合理高效的对窃电等非技术性损耗进行识别,避免因非法窃电对供电企业造成的经济损失,是用电信息采集系统需要解决的重要任务之一。

[0003] 目前,在实际情况中,不同用户在工作日模式和非工作日模式下的用电能耗不同;因此,在窃电检测中考虑电能用户用电行为特点,才能够提升用电异常数据检测的准确性,实现窃电用户的可靠检测。

### 发明内容

[0004] 本发明的目的是提供一种基于电能用户用电行为模式的窃电检测方法和应用,本发明的方法在窃电检测中考虑电能用户用电行为特点,提升用电异常数据检测的准确性,实现窃电用户的可靠检测。

[0005] 为此,本发明提供的技术方案如下:

[0006] 第一方面,本发明提供一种基于电能用户用电行为模式的窃电检测方法,所述检测方法至少包括以下步骤:

[0007] (1) 建立能耗分时模型:基于用电信息采集系统中的用户电能计量信息,建立不同类型用户的能耗分时模型;

[0008] (2) 将不同类型用户的用电行为数据进行聚类,得到聚类后集合;

[0009] (3) 对聚类后集合进行分析,定义局部离群因子 $LOF_{k(p)}$ ,并对疑似窃电行为进行筛选;

[0010] (4) 结合用电信息采集系统计量电能损耗,根据台区下能量消耗公式,进行用户诚信度评价,实现窃电用户监测。

[0011] 优选地,在步骤(1)中,所述用户的不同类型包括大型专变用户、中小型专变用户、工行业用户和城镇居民用户;

[0012] 优选地,考虑行为习惯、气候、季节因素中的至少一种或至少两种因素的组合,建立能耗分时模型。

[0013] 优选地,在步骤(2)中,基于K-means算法将不同类型用户的用电行为数据进行聚类,得到聚类后集合,为用户用电行为数据异常分析奠定前期基础;

[0014] 优选地,所述聚类的个数为 $k$ ,聚类中心为 $C_k$ ;

[0015] 优选地,步骤(2)包括以下步骤:

[0016] (a) 聚类初始阶段:在整体的 $n$ 个数据样本中随机选取 $k$ 个样本, $C_1, C_2, \dots, C_k$ 作为初始聚类中心;

$$[0017] \quad (b) \text{ 令 } dis(n) = \sqrt{\sum_{q=1}^{96} (x_{iq}^2 - C_{nq}^2)} \quad (1)$$

[0018]  $dis(n)$  表示用户的第 $i$ 个计量日的数据与聚类中心 $C_n$ 的几何距离, $x_i$ 表示用户的第 $i$ 个计量日的数据; $q$ 为迭代次数;

[0019] 按照最小距离原则,将 $x_i$ 划分到 $\min\{dis(n) | n=1, 2, \dots, k\}$ 对应的聚类中;

$$[0020] \quad (c) \text{ 更新聚类中心为 } C_n = \frac{\sum_{i \in n} x_i}{f} \quad (2)$$

[0021]  $f$ 表示步骤(b)中获得的该聚类中心 $C_n$ 中的元素个数;

[0022] (d) 迭代步骤(b)和步骤(c),以误差平方 $E$ 作为聚类性能判定条件,当 $E < \varepsilon$ 时,近似表示为用电模型聚类中心 $C_n$ 不再发生变化,转到步骤(e);其中, $\varepsilon$ 在极限讨论中代表的是一个大于0的很小的数,可以任意小,只要不等于零;

[0023] (e) 算法结束,输出 $k$ 个聚类中心数值以及数据样本 $x$ 的聚类结果,形成聚类后集合。

[0024] 优选地,在步骤(3)中,运用LOF算法对聚类后集合进行分析,定义局部 离群因子 $LOF_k(p)$ ,并对疑似窃电行为进行筛选;

$$[0025] \quad LOF_k(p) = \frac{\sum_{o \in N_k(p)} \frac{lrd_k(o)}{lrd_k(p)}}{|N_k(p)|} \quad (3)$$

[0026]  $LOF_k(p)$ 数值大小反应了 $p$ 点的异常程度,其值越大异常程度越高。若在某一段时间存在窃电行为,则该用户在此时段内的数据聚类后 $LOF$ 数值将明显增大。

[0027] 设聚类后集合 $Y = \{y_1, y_2, \dots, y_j\}$ ,各局部离群因子表示为 $LOF_{(y,j)}$ ,令 $\eta$ 为 离群检测阈值,若 $LOF_{(y,j)} > \eta$ 则认为数据存在异常,转窃电行为检测;否则,认为是正常数据;

[0028] 优选地, $\eta > 1$ ,具体取值以实际检测精度而定,若取值过小则异常数据变多,取值过大则造成某些异常数据的遗漏; $\eta$ 的取值例如可以是2、2.5、3、3.5、4、4.5、5、5.5、6、6.5、7、7.5、8、8.5、9、9.5、10、10.5、11或12及所述范围内的所有取值,由于篇幅的限制,不再一一列举;更优选地, $\eta = 1.5$ 。

[0029] 优选地,在步骤(4)中,设某一台区下电能用户 $m$ 的第 $j$ 个计量时刻的数值为 $x_{mj}$ ,电网线损为 $E_{TL}$ ,为避免因网络中设备计量以及温度对线损带来的影响,设置补偿阈值为 $\delta$ ,其取值与台区的规模有关,则某一台区下能耗关系为:

$$[0030] \quad E_{\text{delivered}} - (x_{1j} + x_{2j} + \dots + x_{mj} + E_{TL}) \leq \delta \quad (4)$$

[0031] 其中, $E_{\text{delivered}}$ 表示总表的电量;若用户数据异常时,台区下能耗满足公式(4),则认为该数据为正常数据;反之,则认为该用户存在窃电可能;

[0032] 优选地,设检测的用户集合数据条数为 $z$ ,count为数据集中不满足上述公式的采

集点量,则窃电概率表示为:

$$[0033] \quad \mu = \frac{\text{count}}{z} \quad (5)$$

[0034] 设  $\omega$  为窃电概率检测阈值,若  $\mu > \omega$  则判定该用户为窃电用户,相反则为正常用户。

[0035] 作为优选技术方案,所述检测方法包括以下步骤:

[0036] (1) 建立能耗分时模型:基于用电信息采集系统中的用户电能计量信息,考虑行为习惯、气候、季节因素中的至少一种或至少两种因素的组合,建立不同类型用户的能耗分时模型;所述用户的不同类型包括大型专变用户、中小型专变用户、工行业用户和城镇居民用户;

[0037] (2) 基于K-means算法将不同类型用户的用电行为数据进行聚类,得到聚类后集合;所述聚类的个数为k,聚类中心为  $C_k$ ;包括以下步骤:

[0038] (a) 聚类初始阶段:在整体的n个数据样本中随机选取k个样本,  $C_1, C_2, \dots, C_k$  作为初始聚类中心;

$$[0039] \quad (b) \text{ 令 } dis(n) = \sqrt{\sum_{q=1}^{96} (x_{iq}^2 - C_{nq}^2)} \quad (1)$$

[0040]  $dis(n)$  表示用户的第i个计量日的数据与聚类中心  $C_n$  的几何距离,  $x_i$  表示用户的第i个计量日的数据;  $q$  为迭代次数,表示每次将用户的第i个计量日的数据与聚类中心  $C_n$  的几何距离进行计算时都迭代96次;

[0041] 按照最小距离原则,将  $x_i$  划分到  $\min\{dis(n) | n=1, 2, \dots, k\}$  对应的聚类中;

$$[0042] \quad (c) \text{ 更新聚类中心为 } C_n = \frac{\sum_{i \in n} x_i}{f} \quad (2)$$

[0043]  $f$  表示步骤 (b) 中获得的该聚类中心  $C_n$  中的元素个数;

[0044] (d) 迭代步骤 (b) 和步骤 (c),以误差平方  $E$  作为聚类性能判定条件,当  $E < \varepsilon$  时,近似表示为用电模型聚类中心  $C_n$  不再发生变化,转到步骤 (e);

[0045] (e) 算法结束,输出k个聚类中心数值以及数据样本  $x$  的聚类结果,形成聚类后集合;

[0046] (3) 对聚类后集合进行分析,定义局部离群因子  $LOF_{k(p)}$ ,并对疑似窃电行为进行筛选;运用LOF算法对聚类后集合进行分析,定义局部离群因子  $LOF_k(p)$ ,并对疑似窃电行为进行筛选;

$$[0047] \quad LOF_k(p) = \frac{\sum_{o \in N_k(p)} lrd_k(o)}{|N_k(p)|} \quad (3)$$

[0048] 设聚类后集合  $Y = \{y_1, y_2, \dots, y_j\}$ ,各局部离群因子表示为  $LOF_{(y_j)}$ ,  $\eta$  为离群检测阈值,若  $LOF_{(y_j)} > \eta$  则认为数据存在异常,转窃电行为检测;否则,认为是正常数据;其中,  $\eta = 1.5$ ;

[0049] (4) 结合用电信息采集系统计量电能损耗,根据台区下能量消耗公式,进行用户

诚信度评价,实现窃电用户监测;

[0050] 具体地,设某一台区下电能用户 $m$ 的第 $j$ 个计量时刻的数值为 $x_{mj}$ ,电网线损为 $E_{TL}$ ,为避免因网络中设备计量以及温度对线损带来的影响,设置补偿阈值为 $\delta$ ,其取值与台区的规模有关,则某一台区下能耗关系为:

$$E_{\text{delivered}} - (x_{1j} + x_{2j} + \dots + x_{mj} + E_{TL}) \leq \delta \quad (4)$$

[0052] 若用户数据异常时,台区下能耗满足公式(4),则认为该数据为正常数据;反之,则认为该用户存在窃电可能;

[0053] 设检测的用户集合数据条数为 $z$ ,count为数据集中不满足上述公式的采集点量,则窃电概率表示为:

$$\mu = \frac{\text{count}}{z} \quad (5)$$

[0055] 设 $\omega$ 为窃电概率检测阈值,若 $\mu > \omega$ 则判定该用户为窃电用户,相反则为正常用户。

[0056] 第二方面,本发明提供第一方面所述的检测方法在电能用户窃电行为监测方面的应用。

[0057] 本发明提供的基于电能用户用电行为模式的窃电检测方法和应用首先依据不同用户的用电行为特点,建立不同用户能耗分时模型,基于K-means聚类算法与LOF算法实现用户电能计量数据异常检测后,考虑到用户合理的用电行为改变情况,结合用电信息采集系统计量电能损耗,根据台区下能量消耗公式,确定用户数据是否异常,并引入窃电概率 $\mu$ ,最终实现窃电用户的可靠检测;与现有技术相比,本发明至少具有以下有益效果:

[0058] 1) 与传统窃电检测方法相比,本发明的检测方法从用户的用电行为特性出发,依据不同电力用户用电能耗特点,考虑行为习惯、气候、季节等因素,建立不同用户能耗分时模型。

[0059] 2) 本发明的方法基于K-means聚类算法与LOF算法实现用户电能计量数据异常检测后,考虑到电能用户存在合理的用电行为改变,并不能准确判定用户存在窃电行为,随后结合用电信息采集系统计量电能损耗,根据台区下能量消耗公式,引入窃电概率,对电能窃电概率进行计算,实现电能用户不同类型窃电行为的可靠监测。

## 附图说明

[0060] 图1为本发明的窃电检测方法的流程示意图;

[0061] 图2为不同类型用户日负荷曲线图,其中,A,B,C,D为选取的四个用户,没有特殊含义;

[0062] 图3为D用户三月份的日电力负荷曲线图;

[0063] 图4为K-means算法聚类后用户D的用电模型图;

[0064] 图5为D用户三月份用电数据LOF值图;

[0065] 图6为仿真实验台区图;

[0066] 图7为D用户外出情况下的电能计量数据分析图;

[0067] 图8是窃电数据分析图。

## 具体实施方式

[0068] 下面结合附图1-8及具体实施例对本发明做进一步的说明,但下述实施例绝非对本发明有任何限制。

### [0069] 实施例1

[0070] (1) 建立能耗分时模型:基于用电信息采集系统中的用户电能计量信息,考虑行为习惯、气候、季节因素中的至少一种或至少两种因素的组合,建立不同类型用户的能耗分时模型;所述用户的不同类型包括大型专变用户、中小型专变用户、工行业用户和城镇居民用户;

[0071] (2) 以图3所示D用户3月份的日电力负荷曲线图为例,运用K-means算法将不同类型用户用电行为数据进行聚类,获得k个聚类和聚类中心 $C_k$ 。考虑居民用户电能消耗模型可分为工作日模型与非工作日模型两种,得出不同工作模式下D用户用电模型,如图4所示;

[0072] 图3中用电信息采集系统以15分钟/次的负荷计量频率实现用电负荷的计量,也可根据实际情况选择不同的时间间隔,在此不做过多限制,但从实际情况来看,10-30分钟/次的频率为最佳(例如可以是10分钟/次、11分钟/次、12分钟/次、13分钟/次、14分钟/次、15分钟/次、16分钟/次、17分钟/次、18分钟/次、19分钟/次、20分钟/次、21分钟/次、22分钟/次、23分钟/次、24分钟/次、25分钟/次、26分钟/次、27分钟/次、28分钟/次、29分钟/次、30分钟/次)。

[0073] (3) 运用LOF算法对电力用户计量数据聚类后集合进行分析,定义局部离群因子 $LOF_k(p)$ ,并对疑似窃电行为进行筛选:

$$[0074] \quad LOF_k(p) = \frac{\sum_{o \in N_k(p)} lrd_k(o)}{|N_k(p)|} \quad (3)$$

[0075]  $LOF_k(p)$  数值大小反应了p点的异常程度,其值越大异常程度越高;若在某段时间存在窃电行为,则该用户在此时段内的数据聚类后LOF数值将明显增大。

[0076] 图5中选取D用户3月份用电数据进行分析,其离群检测阈值 $\eta$ 的选取以实际检测精度而定,若取值过小则异常数据变多,取值过大则造成某些异常数据的遗漏,在本文的实验中选取 $\eta=1.5$ 。

[0077] (4) 结合用电信息采集系统计量电能损耗,根据台区下能量消耗公式,进行用户诚信度评价,实现窃电用户监测;

[0078] 设台区下电能用户m的第j个计量时刻的数值为 $x_{mj}$ ,电网线损为 $E_{TL}$ ,为避免因网络中设备计量以及温度对线损带来的影响,设置补偿阈值为 $\delta$ ,其取值与台区规模有关,则台区下能耗关系可表示为:

$$[0079] \quad E_{delivered} - (x_{1j} + x_{2j} + \dots + x_{mj} + E_{TL}) \leq \delta \quad (4)$$

[0080] 若用户数据异常时刻台区下能耗满足上式,则认为该数据为正常数据。反之,则认为该用户存在窃电可能。进一步引入用户窃电概率,更为准确的判定用户是否存在窃电行为。设检测的用户集合数据条数为z,count为数据集中不满足上述公式的采集点量,则窃电概率表示为:

$$[0081] \quad \mu = \frac{\text{count}}{z} \quad (5)$$

[0082] 令 $\omega$ 为窃电概率检测阈值( $\omega$ 具体取值以实际检测精度而定),若 $\mu > \omega$ 则判定该用户为窃电用户,相反则为正常用户。

[0083] 图6代表本实施例选取的某个20用户组成的台区,在电力系统中,台区是指(一台)变压器的供电范围或区域。其中,T表示变压器,S表示集中器,K表示考核表, $n_i$ ( $i=1-20$ )表示用户侧智能电表。

[0084] 图7中选取图6的台区下12号用户( $n_{12}$ )4月份电能计量数据集计算LOF值,8日、9日两天LOF数值分别等于26.9与27.3,明显高于其他时段,但不能据此判定该用户存在窃电行为;进一步计算台区电能损失值和窃电概率 $\mu$ ,得8、9两日台区下并无电能的非技术性损失,且 $\mu=0$ ,因此通过数值分析可排除该用户窃电行为,为正常用户;数值判定与电力工作人员实际入户排查结果相吻合;说明本发明的检测方法得出的结果与实际相一致,消除了因方法局限造成的误判行为。

[0085] 图8中选取图6台区下11号用户( $n_{11}$ )4月份用电数据为分析对象,在11号、12号以及21号的电能计量数据LOF值分别为3.18与3.12和3.19,远高于正常值1,被判定为异常数据。进一步计算窃电概率 $\mu$ ,分别为0.66、0.68与0.52,远大于0,表明在上述时间内存在窃电行为,数值判定与电力工作人员实际入户排查结果相吻合;说明本发明的检测方法得出的结果与实际相一致。与传统窃电检测方法相比,本发明的检测方法从用户的用电行为特性出发,依据不同电力用户用电能耗特点,考虑行为习惯、气候、季节等因素,建立不同用户能耗分时模型。基于K-means聚类算法与LOF算法实现用户电能计量数据异常检测后,考虑到电能用户存在合理的用电行为改变,并不能准确判定用户存在窃电行为。本发明结合用电信息采集系统计量电能损耗,根据台区下能量消耗公式,引入窃电概率,对电能窃电概率进行计算,实现电能用户不同类型窃电行为的可靠监测。

[0086] 应该注意到并理解,在不脱离后附的权利要求所要求的本发明的精神和范围的情况下,能够对上述详细描述的本发明做出各种修改和改进。因此,要求保护的技术方案的范围不受所给出的任何特定示范教导的限制。

[0087] 申请人声明,以上内容是结合具体的优选实施方式对本发明所作的进一步详细说明,不能认定本发明的具体实施只局限于这些说明。对于本发明所属技术领域的普通技术人员来说,在不脱离本发明构思的前提下,还可以做出若干简单推演或替换,都应当视为属于本发明的保护范围。



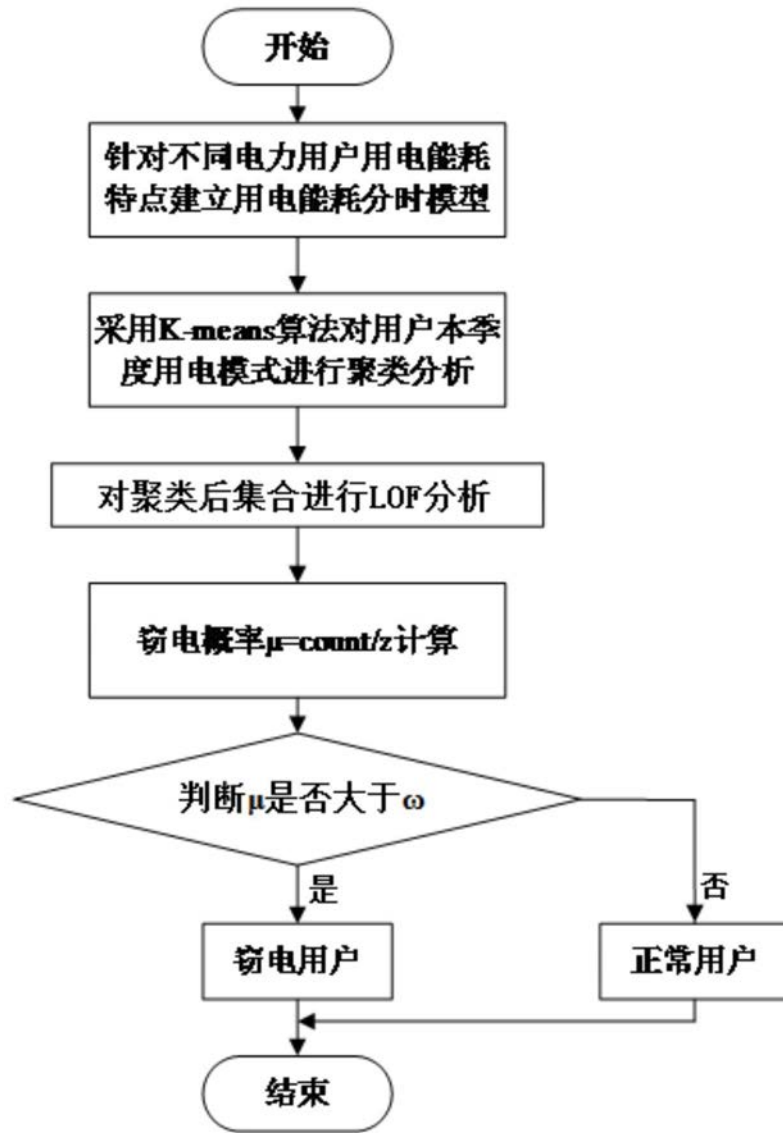


图1

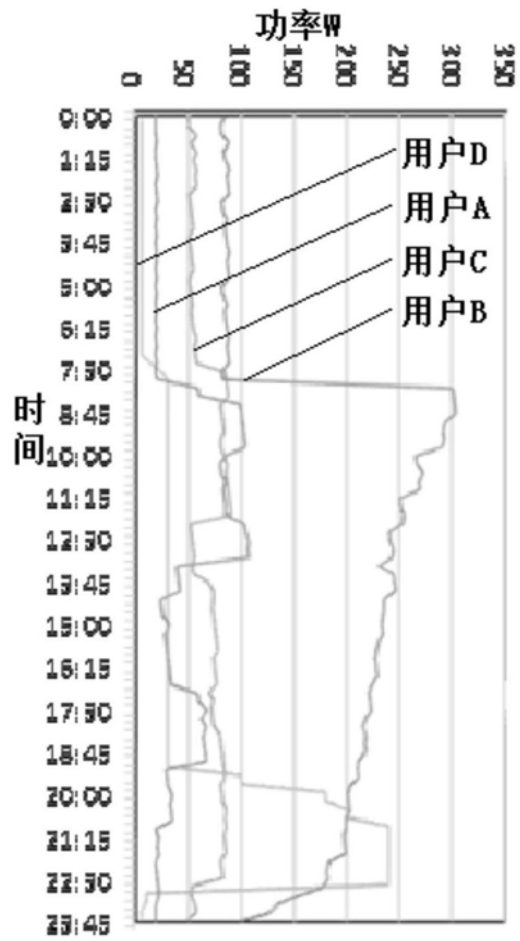


图2

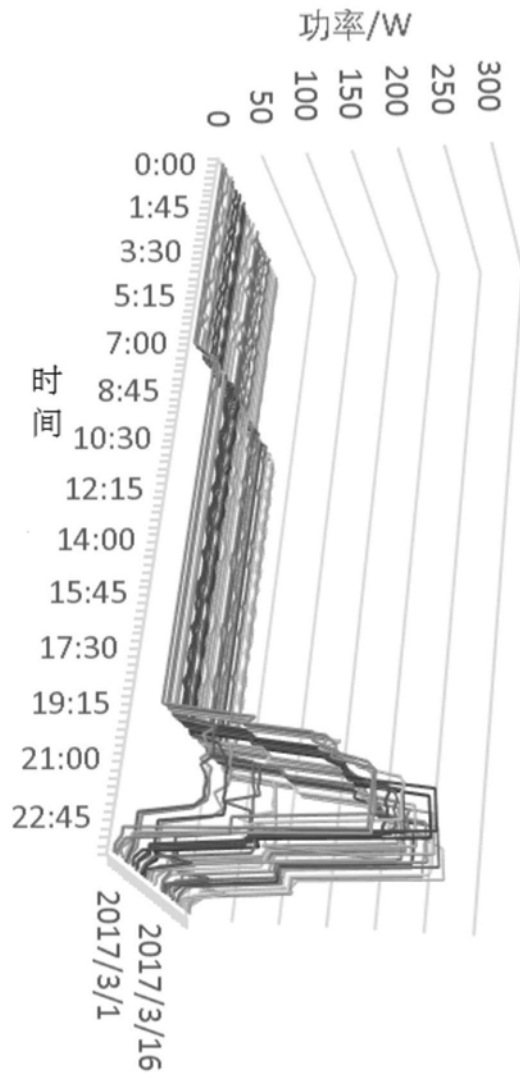


图3

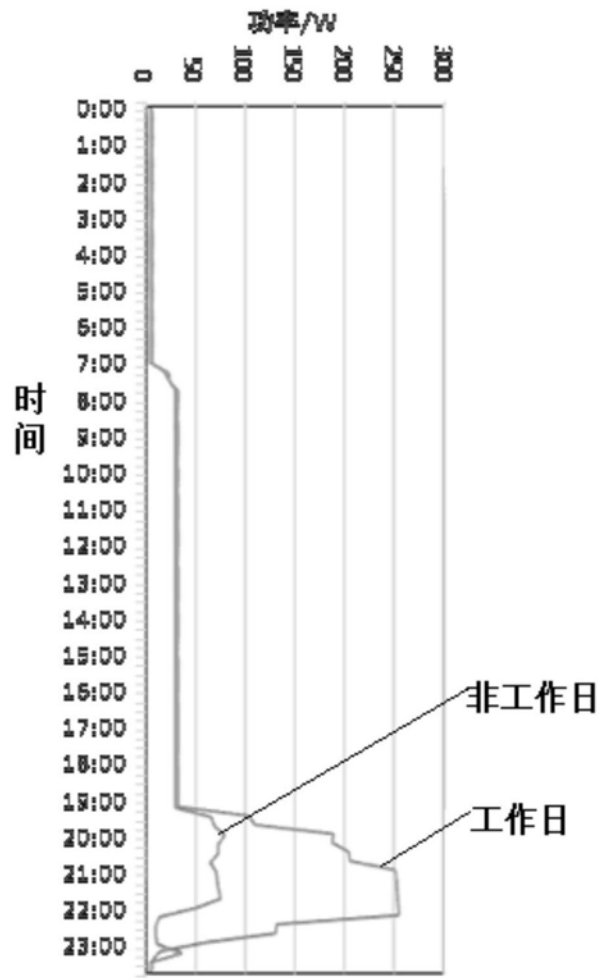


图4

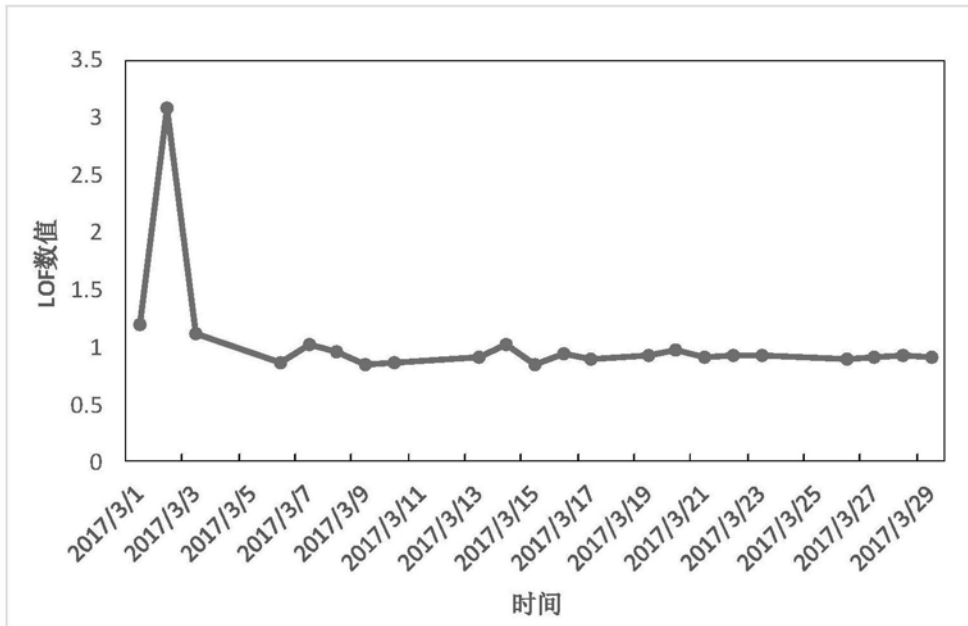


图5

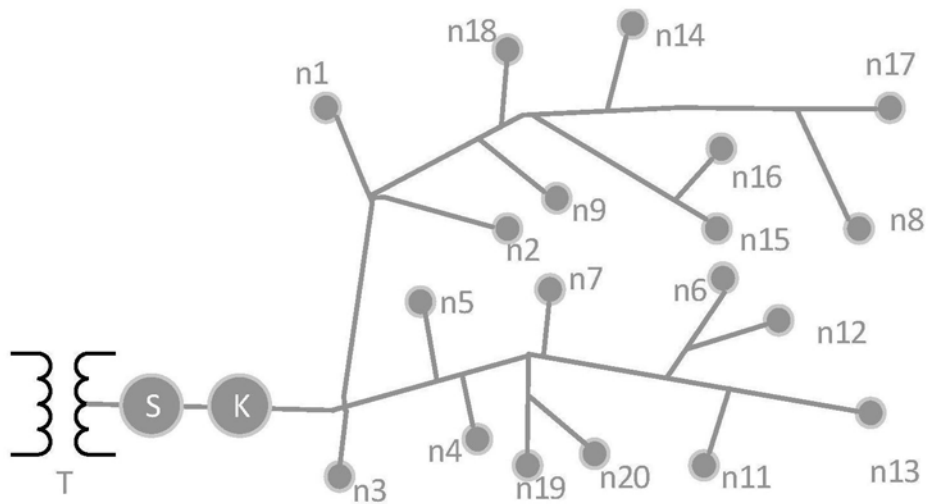


图6

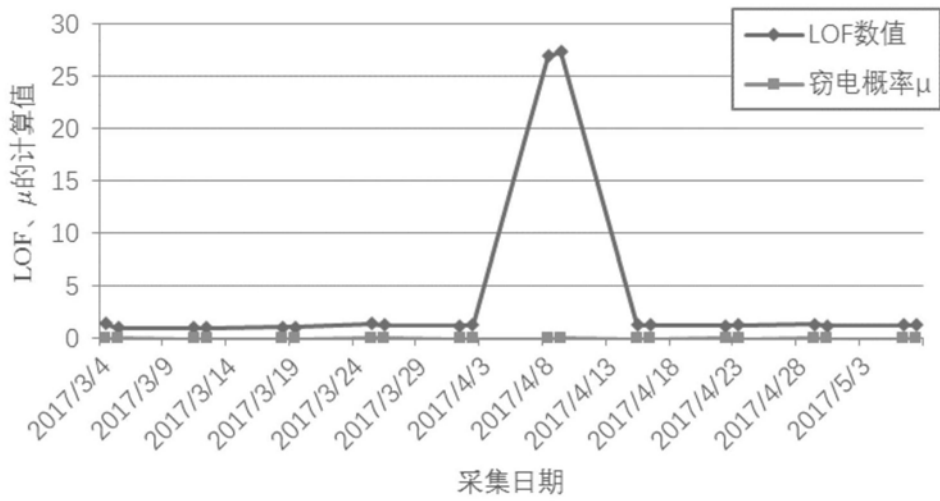


图7



图8