



(12)发明专利申请

(10)申请公布号 CN 110097261 A

(43)申请公布日 2019.08.06

(21)申请号 201910309828.8

(22)申请日 2019.04.17

(71)申请人 三峡大学

地址 443002 湖北省宜昌市西陵区大学路8号

(72)发明人 熊炜 马玉婷

(74)专利代理机构 宜昌市三峡专利事务所
42103

代理人 余山

(51) Int. Cl.

G06Q 10/06(2012.01)

G06Q 50/06(2012.01)

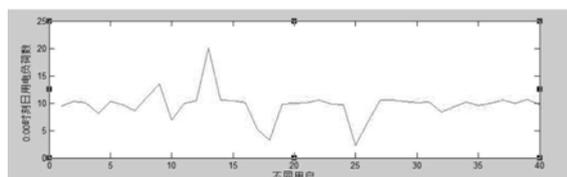
权利要求书2页 说明书5页 附图2页

(54)发明名称

一种判断用户用电异常的方法

(57)摘要

一种判断用户用电异常的方法,它包括以下步骤:1)根据用户电负荷曲线,构造隶属函数,利用隶属函数求出负荷曲线上各负荷数据的隶属度值;2)再结合模糊类聚的区间调整方法,采用聚类思想,选取划分三类区间(谷、正常、峰)的最优阈值;3)根据三类区间判断负荷曲线上的各负荷数据是否异常。本发明目的是为了提供一种能很好辨识异常用电行为的方法,用于解决传统异常用电检测效率低下、覆盖范围小、监测精度低的技术问题。



1. 一种判断用户用电异常的方法,其特征在於,它包括以下步骤:

1) 根据用户电负荷曲线,构造高斯型隶属函数,利用高斯型隶属函数求出负荷曲线上各负荷数据的隶属度值;

2) 再结合模糊类聚的区间调整方法,采用聚类思想,选取划分三类区间的最优阈值;

3) 根据三类区间判断负荷曲线上的各负荷数据是否异常。

2. 根据权利要求1所述的一种判断用户用电异常的方法,其特征在於:在步骤1)中,根据用户电负荷曲线,构造高斯型隶属函数,利用高斯型隶属函数计算不同用户用电量的隶属度,电力负荷曲线的横坐标由不同用户集合 $S = \{a_1, a_2, a_3, \dots, a_m\}$ 组成,纵坐标代表负荷值,由集合 $Q = \{q_1, q_2, q_3, \dots, q_m\}$ 组成,利用隶属度函数求出负荷曲线上各用户同一时刻点对应的负电荷数的隶属度 $x_i = (x_1, x_2, x_3, \dots, x_m)$ 。

3. 根据权利要求1或2所述的一种判断用户用电异常的方法,其特征在於:在步骤2)中,使区间(偏高,正常,偏低)内各隶属度的方差最小,区间间距离最大,并找到最优阈值。

4. 根据权利要求3所述的一种判断用户用电异常的方法,其特征在於:在步骤3)中,根据最优阈值,将负荷曲线合理的划分为偏大、正常、偏小三个区域。

5. 根据权利要求3所述的一种判断用户用电异常的方法,其特征在於,在步骤2)中,在找最优阈值时,具体包括以下步骤:

(1) 设定合适步长 ε 及初始阈值 λ ,划分区间为偏大、正常、偏小;

(2) 根据阈值 λ ,分别求出不同区间内各用户隶属度的方差:

$$M = \frac{x_1 + x_2 + x_m}{m}, \quad s^2 = \frac{(M - x_1)^2 + (M - x_2)^2 + \dots + (M - x_m)^2}{m};$$

(3) 建立模糊相似矩阵,每个用户的特征均由隶属度来表示,然后得到初始的数据矩阵:

$$x_i = \begin{pmatrix} x_{i1} & \dots & x_{im} \\ \vdots & \ddots & \vdots \\ x_{1m} & \dots & x_{im} \end{pmatrix},$$

再采用距离法确定相似系数,其公式如下:

$$r_{mn} = 1 - c \sum_{i=1}^J |x_{mi} - x_{ni}|$$

其中, c 为选取适当参数,它能保证 $r_{mn} \in [0, 1]$, $\sum_{i=1}^J |x_{mi} - x_{ni}|$ 表示 x_{mi} 和 x_{ni} 之间的距离;

(4) 则令 $\lambda = \lambda + \varepsilon$,

重复步骤三四五,直到区间(偏高,正常,偏低)内各隶属度的方差最小,区间间距离最大。

6. 一种判断用户用电异常的方法,其特征在於,它包括以下步骤:

1) 根据用户电负荷曲线,构造正态分布型隶属函数,利用正态分布型隶属函数求出负荷曲线上各负荷数据的隶属度值;

2) 再结合模糊类聚的区间调整方法,采用聚类思想,选取划分三类区间(谷、正常、峰)的最优阈值;

3) 根据三类区间判断负荷曲线上的各负荷数据是否异常。

7. 根据权利要求6所述的一种判断用户用电异常的方法,其特征在於:在步骤1)中,根

据用户电负荷曲线,构造正态分布型隶属函数,利用正态分布型隶属函数计算不同用户用电量的隶属度,电力负荷曲线的横坐标由不同用户集合 $S = \{a_1, a_2, a_3 \dots a_m\}$ 组成,纵坐标代表负荷值,由集合 $Q = \{q_1, q_2, q_3 \dots q_m\}$ 组成,利用隶属度函数求出负荷曲线上各用户同一时刻点对应的负电荷数的隶属度 $x_i = (x_1, x_2, x_3 \dots x_m)$ 。

8. 根据权利要求6或7所述的一种判断用户用电异常的方法,其特征在于:在步骤2)中,使区间(偏高,正常,偏低)内各隶属度的方差最小,区间间距离最大,并找到最优阈值。

9. 根据权利要求8所述的一种判断用户用电异常的方法,其特征在于:在步骤3)中,根据最优阈值,将负荷曲线合理的划分为偏大、正常、偏小三个区域。

10. 根据权利要求8所述的一种判断用户用电异常的方法,其特征在于,在步骤2)中,在找最优阈值时,具体包括以下步骤:

(1) 设定合适步长 ϵ 及初始阈值 λ ,划分区间为偏大、正常、偏小;

(2) 根据阈值 λ ,分别求出不同区间内各用户隶属度的方差:

$$M = \frac{x_1 + x_2 + x_m}{m}, \quad s^2 = \frac{(M - x_1)^2 + (M - x_2)^2 + \dots + (M - x_m)^2}{m};$$

(3) 建立模糊相似矩阵,每个用户的特征均由隶属度来表示,然后得到初始的数据矩

阵: $x_i = \begin{pmatrix} x_{i1} & \dots & x_{im} \\ \vdots & \ddots & \vdots \\ x_{1m} & \dots & x_{im} \end{pmatrix},$

再采用距离法确定相似系数,其公式如下:

$$r_{mn} = 1 - c \sum_{i=1}^j |x_{mi} - x_{ni}|$$

其中, c 为选取适当参数,它能保证 $r_{mn} \in [0, 1]$, $\sum_{i=1}^j |x_{mi} - x_{ni}|$ 表示 x_{mi} 和 x_{ni} 之间的距离;

(4) 则令 $\lambda = \lambda + \epsilon$,

重复步骤三四五,直到区间(偏高,正常,偏低)内各隶属度的方差最小,区间间距离最大。

一种判断用户用电异常的方法

技术领域

[0001] 本发明属于数据处理技术领域,具体涉及一种判断用户用电异常的方法。

背景技术

[0002] 用户异常用电行为不仅造成电力公司巨大的经济损失,还危害到了电网的安全运行,所以一直以来是电力稽查的重点。传统的异常用电检测方法过于依赖人力排查,效率低下,而后发展的基于计量系统的异常检测方法则误报太多,实用性不足,需要发展新的技术。随着用电信息数据越来越多地汇集到中心站,使得利用数据挖掘来辨识异常用电行为成为可能。

[0003] 由于电能的不可储存性,异常用电检测一直是电力公司的难点。居民区用户用电异常主要包括电力系统故障和用户窃电两方面原因。电力系统故障现象有农村区域电缆过多穿插建立导致电缆之间互相叠加在一起,使得用户用电负荷异常变大;或电线表面包扎的绝缘带用明显烧焦痕迹,出现漏电行为;或供电线路三相负荷分布严重不均匀等;或者以反窃电技术为主,从源头上防止异常用电,然后现场检测为辅。用户的用电行为可以使用负荷数据来描述,异常用电行为则表现为异常的用电负荷数据。传统的检测异常用电方法有定期巡查,定期校验电表,用户举报窃电等手段来发现窃电或计量装置故障等,存在的不足有:

[0004] 1、巡视检查法效率底且无法覆盖全部用户;

[0005] 2、计量自动化系统远程在线监测功能经常误报,存在过多的无用信息,造成大量人力物力资源的浪费;

[0006] 经典的SVDD模型在异常用电辨识的应用中,由于用户用电数据具有内部分类的特性,导致其构建的超球紧度特性不佳,影响了监测精度。

发明内容

[0007] 本发明目的是为了提供一种基于高斯型隶属函数的能很好辨识异常用电行为的方法,用于解决传统异常用电检测效率低下、覆盖范围小、监测精度低的技术问题。

[0008] 为了解决上述技术问题,本发明提供了如下的技术方案:

[0009] 一种判断用户用电异常的方法,它包括以下步骤:

[0010] 1) 根据用户电负荷曲线,构造高斯型隶属函数,利用高斯型隶属函数求出负荷曲线上各负荷数据的隶属度值;

[0011] 2) 再结合模糊类聚的区间调整方法,采用聚类思想,选取划分三类区间(谷、正常、峰)的最优阈值;

[0012] 3) 根据三类区间判断负荷曲线上的各负荷数据是否异常。

[0013] 在步骤1)中,根据用户电负荷曲线,构造高斯型隶属函数,利用高斯型隶属函数计算不同用户用电量的隶属度,电力负荷曲线的横坐标由不同用户集合 $S = \{a_1, a_2, a_3 \dots a_m\}$ 组成,纵坐标代表负荷值,由集合 $Q = \{q_1, q_2, q_3 \dots q_m\}$ 组成,利用隶属度函数求出负荷曲线

上各用户同一时刻点对应的负电荷数的隶属度 $x_i = (x_1, x_2, x_3 \dots x_m)$ 。

[0014] 在步骤2)中,使区间(偏高,正常,偏低)内各隶属度的方差最小,区间间距离最大,并找到最优阈值。

[0015] 在步骤3)中,根据最优阈值,将负荷曲线合理的划分为偏大、正常、偏小三个区域。

[0016] 在步骤2)中,在找最优阈值时,具体包括以下步骤:

[0017] (1) 设定合适步长 ε 及初始阈值 λ ,划分区间为偏大、正常、偏小;

[0018] (2) 根据阈值 λ ,分别求出不同区间内各用户隶属度的方差:

$$[0019] \quad M = \frac{x_1 + x_2 + x_m}{m}, \quad s^2 = \frac{(M - x_1)^2 + (M - x_2)^2 + \dots + (M - x_m)^2}{m};$$

[0020] (3) 建立模糊相似矩阵,每个用户的特征均由隶属度来表示,然后得到初始的数据

$$\text{矩阵: } x_i = \begin{pmatrix} x_{11} & \dots & x_{i1} \\ \vdots & \ddots & \vdots \\ x_{1m} & \dots & x_{im} \end{pmatrix},$$

[0021] 再采用距离法确定相似系数,其公式如下:

$$[0022] \quad r_{mn} = 1 - c \sum_{i=1}^J |x_{mi} - x_{ni}|$$

[0023] 其中, c 为选取适当参数,它能保证 $r_{mn} \in [0, 1]$, $\sum_{i=1}^J |x_{mi} - x_{ni}|$ 表示 x_{mi} 和 x_{ni} 之间的距离;

[0024] (4) 则令 $\lambda = \lambda + \varepsilon$,

[0025] 重复步骤(2)(3)(4),直到区间(偏高,正常,偏低)内各隶属度的方差最小,区间间距离最大。

[0026] 一种判断用户用电异常的方法,它包括以下步骤:

[0027] 1) 根据用户电负荷曲线,构造正态分布型隶属函数,利用正态分布型隶属函数求出负荷曲线上各负荷数据的隶属度值;

[0028] 2) 再结合模糊类聚的区间调整方法,采用聚类思想,选取划分三类区间(谷、正常、峰)的最优阈值;

[0029] 3) 根据三类区间判断负荷曲线上的各负荷数据是否异常。

[0030] 在步骤1)中,根据用户电负荷曲线,构造正态分布型隶属函数,利用正态分布型隶属函数计算不同用户用电量的隶属度,电力负荷曲线的横坐标由不同用户集合 $S = \{a_1, a_2, a_3 \dots a_m\}$ 组成,纵坐标代表负荷值,由集合 $Q = \{q_1, q_2, q_3 \dots q_m\}$ 组成,利用隶属度函数求出负荷曲线上各用户同一时刻点对应的负电荷数的隶属度 $x_i = (x_1, x_2, x_3 \dots x_m)$ 。

[0031] 在步骤2)中,使区间(偏高,正常,偏低)内各隶属度的方差最小,区间间距离最大,并找到最优阈值。

[0032] 在步骤3)中,根据最优阈值,将负荷曲线合理的划分为偏大、正常、偏小三个区域。

[0033] 在步骤2)中,在找最优阈值时,具体包括以下步骤:

[0034] (1) 设定合适步长 ε 及初始阈值 λ ,划分区间为偏大、正常、偏小;

[0035] (2) 根据阈值 λ ,分别求出不同区间内各用户隶属度的方差:

$$[0036] \quad M = \frac{x_1 + x_2 + \dots + x_m}{m}, \quad s^2 = \frac{(M - x_1)^2 + (M - x_2)^2 + \dots + (M - x_m)^2}{m};$$

[0037] (3) 建立模糊相似矩阵, 每个用户的特征均由隶属度来表示, 然后得到初始的数据

$$\text{矩阵: } x_i = \begin{pmatrix} x_{i1} & \dots & x_{im} \\ \vdots & \ddots & \vdots \\ x_{1m} & \dots & x_{im} \end{pmatrix},$$

[0038] 再采用距离法确定相似系数, 其公式如下:

$$[0039] \quad r_{mn} = 1 - c \sum_{i=1}^J |x_{mi} - x_{ni}|$$

[0040] 其中, c 为选取适当参数, 它能保证 $r_{mn} \in [0, 1]$, $\sum_{i=1}^J |x_{mi} - x_{ni}|$ 表示 x_{mi} 和 x_{ni} 之间的距离;

[0041] (4) 则令 $\lambda = \lambda + \varepsilon$,

[0042] 重复步骤三四五, 直到区间 (偏高, 正常, 偏低) 内各隶属度的方差最小, 区间间距离最大。

[0043] 本发明的有益效果是:

[0044] 本发明利用数据挖掘技术相关算法、基于隶属函数的区间划分方法、以及隶属函数求出负荷曲线上各负荷数据的隶属度值, 再结合模糊类聚的区间调整方法, 采用聚类思想, 选取划分三类区间 (谷、正常、峰) 的最优阈值, 提高各区间划分精度, 最后根据三类区间判断负荷曲线上的各负荷数据是否异常, 能高效、准确的判断用户用电的异常问题。

附图说明

[0045] 图1是本发明实施例1中40个不同用户日用电负荷数曲线采集图;

[0046] 图2是本发明实施例1中日用电负荷模型还原图;

[0047] 图3为本发明实施例2中40个不同用户日用电负荷数曲线采集图;

[0048] 图4为本发明实施例2中日用电负荷模型还原图。

具体实施方式

[0049] 实施例1: 一种基于高斯型隶属函数判断用户用电异常的方法, 它包括以下步骤:

[0050] 1) 根据用户电负荷曲线, 构造高斯型隶属函数, 利用高斯型隶属函数求出负荷曲线上各负荷数据的隶属度值;

[0051] 其中, 构造的高斯型隶属函数为: $f(x, \sigma, c) = \exp\left[-\frac{(x - c)^2}{2\sigma^2}\right]$;

[0052] 2) 再结合模糊类聚的区间调整方法, 采用聚类思想, 选取划分三类区间 (谷、正常、峰) 的最优阈值;

[0053] 3) 根据三类区间判断负荷曲线上的各负荷数据是否异常。

[0054] 在步骤1) 中, 根据用户电负荷曲线, 构造高斯型隶属函数, 利用高斯型隶属函数计算不同用户用电量的隶属度, 电力负荷曲线的横坐标由不同用户集合 $S = \{a_1, a_2, a_3 \dots a_m\}$ 组成, 纵坐标代表负荷值, 由集合 $Q = \{q_1, q_2, q_3 \dots q_m\}$ 组成, 利用隶属度函数求出负荷曲线

上各用户同一时刻点对应的负电荷数的隶属度 $x_i = (x_1, x_2, x_3 \dots x_m)$ 。

[0055] 在步骤2)中,使区间(偏高,正常,偏低)内各隶属度的方差最小,区间间距离最大,并找到最优阈值。

[0056] 在步骤3)中,根据最优阈值,将负荷曲线合理的划分为偏大、正常、偏小三个区域。

[0057] 在步骤2)中,在找最优阈值时,具体包括以下步骤:

[0058] (1) 设定合适步长 ε 及初始阈值 λ ,划分区间为偏大、正常、偏小;

[0059] (2) 根据阈值 λ ,分别求出不同区间内各用户隶属度的方差:

$$[0060] \quad M = \frac{x_1 + x_2 + x_m}{m}, \quad s^2 = \frac{(M - x_1)^2 + (M - x_2)^2 + \dots + (M - x_m)^2}{m};$$

[0061] (3) 建立模糊相似矩阵,每个用户的特征均由隶属度来表示,然后得到初始的数据

$$\text{矩阵: } x_i = \begin{pmatrix} x_{11} & \dots & x_{i1} \\ \vdots & \ddots & \vdots \\ x_{1m} & \dots & x_{im} \end{pmatrix},$$

[0062] 再采用距离法确定相似系数,其公式如下:

$$[0063] \quad r_{mn} = 1 - c \sum_{i=1}^J |x_{mi} - x_{ni}|$$

[0064] 其中, c 为选取适当参数,它能保证 $r_{mn} \in [0, 1]$, $\sum_{i=1}^J |x_{mi} - x_{ni}|$ 表示 x_{mi} 和 x_{ni} 之间的距离;

[0065] (4) 则令 $\lambda = \lambda + \varepsilon$,

[0066] 重复步骤三四五,直到区间(偏高,正常,偏低)内各隶属度的方差最小,区间间距离最大。

[0067] 在一个实施例中,根据居民用户相关历史数据,采用多元线性回归,支持向量回归的变权组合模型,预测出未来一段时间居民平均日用电负荷数据,并借助MATLAB进行数据清洗等预处理画出如图1所示40个不同用户日用电负荷数曲线。

[0068] 在该实施例中,提出了利用高斯型隶属函数求出居民平均日用电负荷数据的隶属度值,借助MATLAB画出用电负荷数隶属度的云图像并根据图像基本划分阈值区间,如图2所示。再采用模糊类聚思想,根据隶属度构造模糊相似矩阵,根据隶属度不同区间的方差及矩阵数据之间的距离对隶属函数的划分结果进行优化,并确定最优阈值合理划分区间,根据各个区间来判断居民日用电负荷数是否正常。

[0069] 实施例2:一种基于正态分布型隶属函数判断用户用电异常的方法,它包括以下步骤:

[0070] 1) 根据用户电负荷曲线,构造正态分布型隶属函数,利用正态分布型隶属函数求出负荷曲线上各负荷数据的隶属度值;

$$[0071] \quad \text{其中,构造正态分布型隶属函数: } A(x) = \begin{cases} e^{-k(x-a)} & x \ll a \\ e^{-k(x-a)} & x > a \end{cases};$$

[0072] 2) 再结合模糊类聚的区间调整方法,采用聚类思想,选取划分三类区间(谷、正常、峰)的最优阈值;

[0073] 3) 根据三类区间判断负荷曲线上的各负荷数据是否异常。

[0074] 在步骤1)中,根据用户电负荷曲线,构造正态分布型隶属函数,利用正态分布型隶属函数计算不同用户用电量的隶属度,电力负荷曲线的横坐标由不同用户集合 $S = \{a_1, a_2, a_3 \dots a_m\}$ 组成,纵坐标代表负荷值,由集合 $Q = \{q_1, q_2, q_3 \dots q_m\}$ 组成,利用隶属度函数求出负荷曲线上各用户同一时刻点对应的负电荷数的隶属度 $x_i = (x_1, x_2, x_3 \dots x_m)$ 。

[0075] 在步骤2)中,使区间(偏高,正常,偏低)内各隶属度的方差最小,区间间距离最大,并找到最优阈值。

[0076] 在步骤3)中,根据最优阈值,将负荷曲线合理的划分为偏大、正常、偏小三个区域。

[0077] 在步骤2)中,在找最优阈值时,具体包括以下步骤:

[0078] (1) 设定合适步长 ε 及初始阈值 λ ,划分区间为偏大、正常、偏小;

[0079] (2) 根据阈值 λ ,分别求出不同区间内各用户隶属度的方差:

$$[0080] \quad M = \frac{x_1 + x_2 + x_m}{m}, \quad s^2 = \frac{(M - x_1)^2 + (M - x_2)^2 + \dots + (M - x_m)^2}{m};$$

[0081] (3) 建立模糊相似矩阵,每个用户的特征均由隶属度来表示,然后得到初始的数据

$$\text{矩阵: } x_i = \begin{pmatrix} x_{11} & \dots & x_{i1} \\ \vdots & \ddots & \vdots \\ x_{1m} & \dots & x_{im} \end{pmatrix},$$

[0082] 再采用距离法确定相似系数,其公式如下:

$$[0083] \quad r_{mn} = 1 - c \sum_{i=1}^J |x_{mi} - x_{ni}|$$

[0084] 其中, c 为选取适当参数,它能保证 $r_{mn} \in [0, 1]$, $\sum_{i=1}^J |x_{mi} - x_{ni}|$ 表示 x_{mi} 和 x_{ni} 之间的距离;

[0085] (4) 则令 $\lambda = \lambda + \varepsilon$,

[0086] 重复步骤三四五,直到区间(偏高,正常,偏低)内各隶属度的方差最小,区间间距离最大。

[0087] 在一个实施例中,根据居民用户相关历史数据,采用多元线性回归,支持向量回归的变权组合模型,预测出未来一段时间居民平均日用电负荷数据,并借助MATLAB进行数据清洗等预处理画出如图3所示40个不同用户日用电负荷数曲线。

[0088] 在该实施例中,提出了利用正态分布型隶属函数求出居民平均日用电负荷数据的隶属度值,借助MATLAB画出用电负荷数隶属度的云图像并根据图像基本划分阈值区间,如图4所示。再采用模糊类聚思想,根据隶属度构造模糊相似矩阵,根据隶属度不同区间的方差及矩阵数据之间的距离对隶属函数的划分结果进行优化,并确定最优阈值合理划分区间。根据各个区间来判断居民日用电负荷数是否正常。

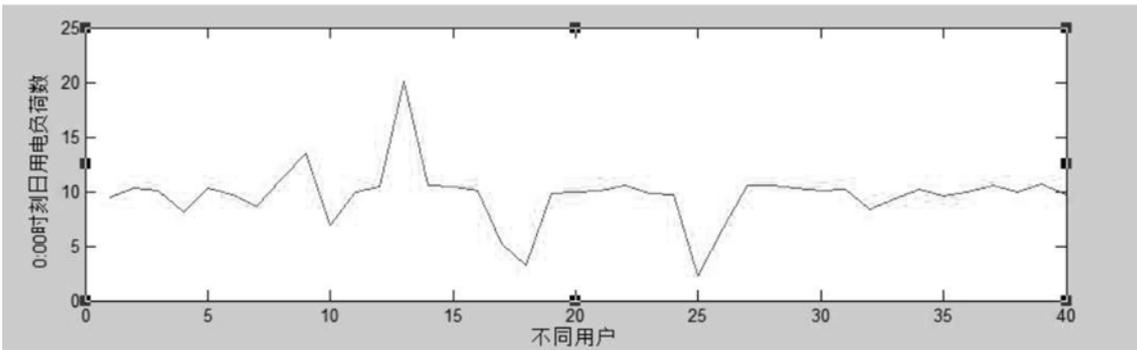


图1

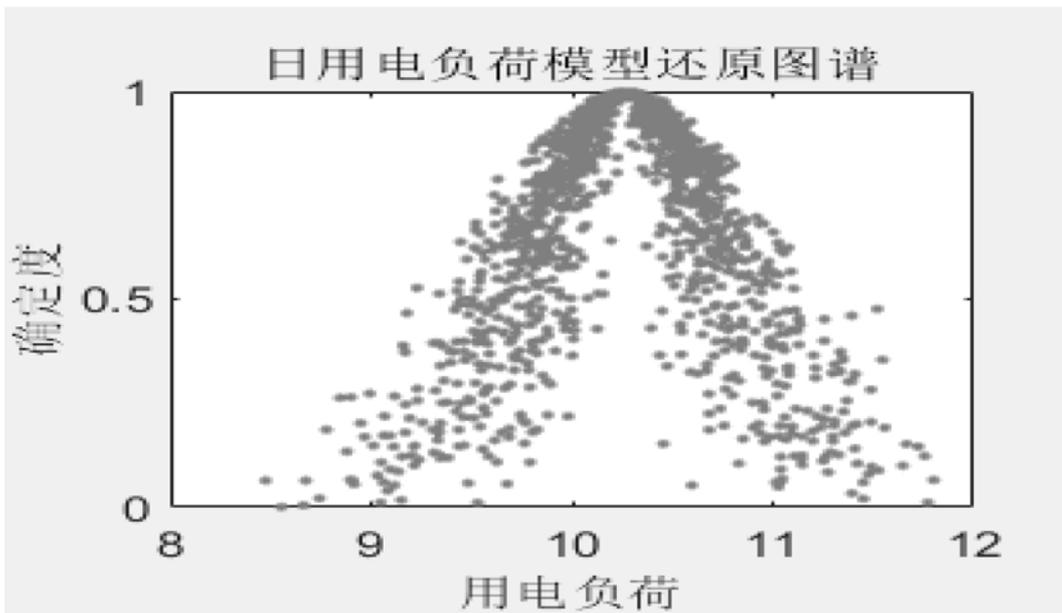


图2

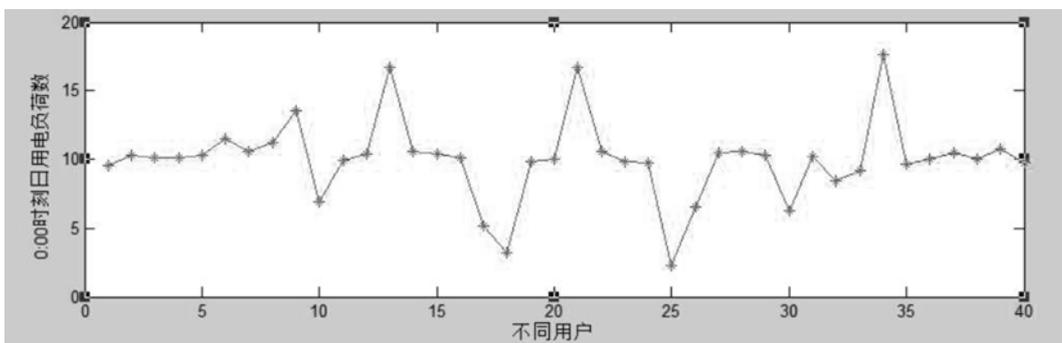


图3

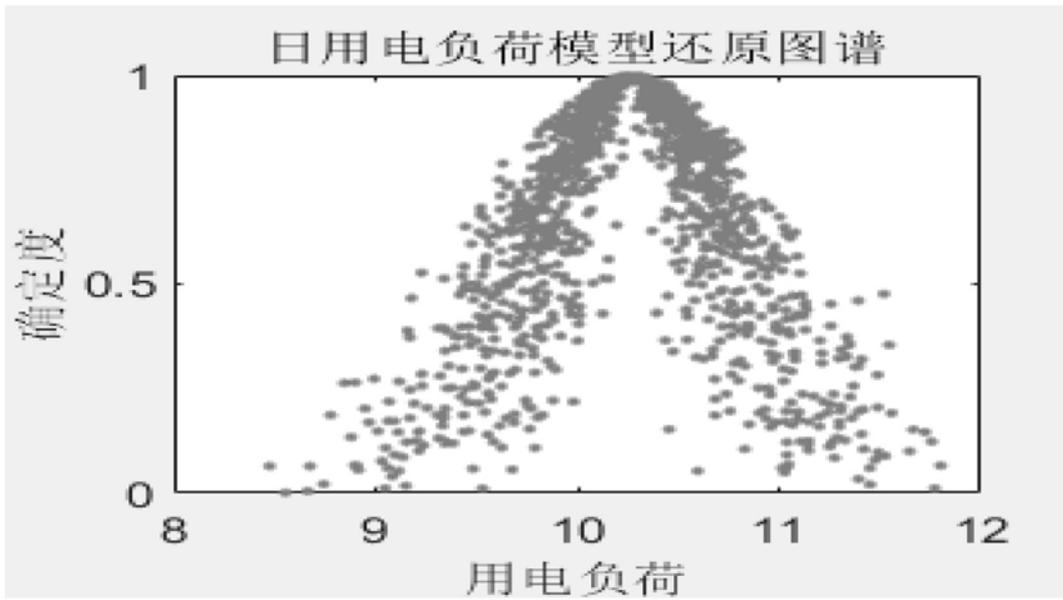


图4