

doi:10.11835/j.issn.1674-4764.2018.02.016



建筑用户行为节能潜力评估新方法

胡宾, 俞准, 李郡, 张国强

(湖南大学 土木工程学院, 长沙 410082)

摘要: 现有建筑用户行为节能潜力评估方法未能充分考虑不同用户之间的差异性, 从而显著降低了评估结果的准确性。针对此种不足, 提出一种行为节能潜力评估的新方法。通过主成分分析对用户行为影响因素进行降维处理, 在此基础上, 采用数据挖掘方法中的聚类分析技术对样本用户进行合理分类, 并针对不同类用户特征分别进行节能潜力评估。由于该方法既全面考虑了同一用户不同因素的影响, 又充分体现了不同用户之间的差异性, 可显著提高评估结果的准确性。通过将其应用于湖南省 3 所高校 100 名研究生用户电脑待机行为的节能潜力评估表明, 该方法有效可行。

关键词: 用户行为; 节能潜力评估; 数据挖掘; 聚类分析; 主成分分析

中图分类号: TU244.3 **文献标志码:** A **文章编号:** 1674-4764(2018)02-0103-06

A new energy-saving potential evaluation method for building occupant behavior improvement

Hu Bin, Yu Zhun, Li Jun, Zhang Guoqiang

(College of Civil Engineering, Hunan University, Changsha 410082, P. R. China)

Abstract: Existing energy-saving potential evaluation methods for building occupant behavior improvement do not take occupants' diversity into consideration. This significantly decreases the accuracy of evaluation results. To address this issue, a new energy-saving potential evaluation method has been established. First, principal component analysis is used to reduce the number of influencing factors of occupant behavior. Then, a data mining technique, clustering analysis is used to classify sample occupants into different groups. At last, energy-saving potentials are evaluated in terms of different groups characteristics. This method takes both the impact of various influencing factors and the diversity of occupants into account, and thus could improve the evaluation accuracy significantly. The method was applied to evaluate the energy-saving potential of 100 graduate students' computer standby usage behavior in 3 universities in Hunan province. The results demonstrate its effectiveness and feasibility.

Keywords: occupant behavior; energy-saving potential evaluation; data mining; clustering analysis; principal component analysis

收稿日期: 2017-06-10

基金项目: 国家自然科学基金(51408205)

作者简介: 胡宾(1991-), 女, 主要从事建筑节能及行为节能研究, E-mail: binh8284@foxmail.com。

俞准(通信作者), 男, 副教授, 博士生导师, E-mail: zhunyu@hnu.edu.cn。

Received: 2017-06-10

Foundation item: National Natural Science Foundation of China (No. 51408205)

Author brief: Hu Bin (1991-), main research interests: building energy saving and standby energy, E-mail: binh8284@foxmail.com.

Yu Zhun (corresponding author), associate professor, doctoral supervisor, E-mail: zhunyu@hnu.edu.cn.

研究表明,建筑用户行为对建筑总能耗有显著的影响,且通过改变用户行为减少能耗是实现建筑节能的有效手段之一^[1-2]。为深入分析用户行为对建筑能耗的影响和有效降低建筑能耗,有必要对行为节能的潜力进行量化分析和评估。当前对行为节能潜力的研究涵盖不同行为类型^[3-4](如正常使用行为和待机行为)、不同建筑类型^[5-7](如住宅建筑、办公建筑和学校建筑)和不同电器种类^[8-9](如空调和电脑),所采用的方法主要是:收集一定数量的样本用户数据并建立数据库,通过将样本用户设备或电器的实际功率及使用时长降低至某一较低的预设值,以进行相应节能潜力评估。该方法的局限性主要表现在未能充分考虑不同用户之间的差异性,对所有用户均采用相同的功率及使用时长预设值,这在实际生活中往往难以实现,导致过高或过低估计了部分用户的行为节能潜力,从而显著降低了节能潜力评估结果的准确性。为克服上述局限,一个可行的方法是:先全面采集用户行为潜在影响因素相关数据,再根据这些因素将样本用户合理分类,然后,分别针对每一类用户特征进行进一步的节能潜力评估。实现该方法的关键障碍在于:1)用户行为的影响因素往往数量较多,如仅选择部分因素进行分类易造成信息缺失导致分类不合理,而选择所有因素进行分类又易因数量过多难以进一步处理;2)在对样本用户进行分类时难以同时考虑不同因素的影响。例如,将用户按某一种因素(如年龄)进行分类时,通常无法同时考虑到其他因素(如学历和生活习惯)的影响。

为克服上述障碍,提高评估准确性,提出一种建筑用户行为节能潜力评估新方法。该方法首先采用主成分分析技术分析用户行为的潜在影响因素,该技术可在保持原有影响因素绝大部分信息的同时进行降维处理,避免了选择部分因素易造成信息缺失,选择所有因素又导致数量过多难以进一步处理的缺点。采用数据挖掘方法中聚类分析技术对样本用户进行分类,该技术可在考虑不同影响因素的同时将样本用户合理分类,最后,进一步针对不同类别用户特征进行节能潜力分析。此外,以高校建筑电脑待机行为节能潜力为例,分析该方法的有效性。

1 研究方法

1.1 数据收集

待机状态是指电器连接电源时没有执行其主要功能但仍消耗电能的状态^[3]。电器在待机状态下所消耗的电能即为待机能耗。就电脑而言,其待机状

态主要包括屏保、锁屏、睡眠和关机 4 种模式。以湖南地区 3 所高校 100 名研究生用户为例,采用型号为 GPM-8212 的功率计(在 23 ± 5 °C 的环境下,测量范围为 0.32~13.10 kW,精度为 $\pm 2\%$)测量其电脑在各待机模式下的功率值,并通过问卷形式调查了各待机模式的使用时长及待机行为的潜在影响因素,其中,潜在影响因素的种类及部分因素取值如表 1 所示。

表 1 电脑待机行为潜在影响因素

Table 1 The potential influencing factors of standby behavior of computer

种类	具体因素及部分因素取值范围
个人信息	性别(男/女)、年龄、是否已婚、兴趣爱好(电脑相关/无关)
专业相关信息	专业(文科/理工科)、学历(硕士/博士)、入学年数、专业成绩(优/良/中/一般)
社会经济因素	生活费用、是否定期接受节能教育、是否有上网流量限制
生活习惯	是否午休、是否住校、是否打卡出勤

1.2 电脑待机能耗计算

基于所采集数据可获得电脑各待机模式下的功率 P 以及相应待机时长 T ,从而可根据式(1)计算单台电脑年待机能耗。

$$E_i = \sum_{j=1}^4 \frac{P_{ij} \times T_{ij} \times M}{1\ 000} \quad (1)$$

式中: E_i 为第 i 台电脑的年待机能耗, $\text{kW} \cdot \text{h}$; P_{ij} 、 T_{ij} 分别为第 i 台电脑在第 j 种待机模式下的功率值和一周内的使用时长($j=1, 2, 3, 4$; 分别对应于电脑的屏保、锁屏、睡眠和关机模式); M 为 1 a 内学生在校周数(去除寒暑假), 本文取 44。

1.3 用户行为节能潜力评估

首先,采用主成分分析技术与聚类分析技术相结合的方法对样本用户进行分类,再针对每类用户采用行为节能方案分别进行行为节能潜力分析。主成分分析、聚类分析及所采用的行为节能方案介绍如下。

1.3.1 主成分分析 主成分分析是将多个可能存在相关性的变量,通过正交变换转化为少数几个互相无关的综合变量的一种统计分析方法。这些综合变量保持了原始变量的大部分信息。

1)数据标准化 由于数据库中的参数包含分类属性参数和数值属性参数,如用户性别为分类属性(男或女),年龄为数值属性(20岁或25岁),难以直接对所有数据进行直接比较和分析。此外,数值属

性参数的取值范围差异较大,而取值范围大的参数影响将覆盖取值范围小的参数影响。因此,需将数据库中所有参数均转换为数值属性参数进行统一比较分析,并对所有数值属性参数进行标准化处理以使其结果落到同一区间,取 $[0, 1]$ 区间。

对于分类属性参数,应首先按其不同状态排序,如专业成绩可分4个等级[优,良,中,一般],再根据式(2)计算。

$$z_{ij} = \frac{\text{参数 } j \text{ 第 } i \text{ 个状态的排序值} - 1}{\text{参数 } j \text{ 所有状态的最大排序值} - 1} \quad (2)$$

计算后上述4个等级可转换为 $[0, 1/3, 2/3, 1]$ 。

对于数值属性参数,可根据式(3)计算。

$$z_{ij} = \frac{x_{ij} - x_{j\min}}{x_{j\max} - x_{j\min}} \quad (3)$$

式中: z_{ij} 、 x_{ij} 、 $x_{j\min}$ 和 $x_{j\max}$ 分别表示参数 j 的转换值、原始值、最小值和最大值。

采用上述方法,原始数据矩阵 \mathbf{X} 变换为标准化矩阵 \mathbf{Z} 。

$$\mathbf{X} = \begin{bmatrix} x_{12} & x_{13} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \cdots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix} \rightarrow \mathbf{Z} = \begin{bmatrix} z_{12} & z_{13} & \cdots & z_{1p} \\ z_{21} & z_{22} & \cdots & z_{2p} \\ \vdots & \vdots & \cdots & \vdots \\ z_{n1} & z_{n2} & \cdots & z_{np} \end{bmatrix}$$

2) 计算标准化矩阵 \mathbf{Z} 的样本相关系数矩阵 \mathbf{R} , 样本相关系数矩阵 \mathbf{R} 的表达式为

$$\mathbf{R} = \begin{bmatrix} r_{12} & r_{13} & \cdots & r_{1p} \\ r_{21} & r_{22} & \cdots & r_{2p} \\ \vdots & \vdots & \cdots & \vdots \\ r_{p1} & r_{p2} & \cdots & r_{pp} \end{bmatrix}$$

式中: r_{uv} ($u, v = 1, 2, \dots, P$) 为参数 Z_u 、 Z_v 的相关系数,且 $r_{uv} = r_{vu}$,计算式为

$$r_{uv} = \frac{\sum_{k=1}^n (z_{ku} - \bar{z}_u)(z_{kv} - \bar{z}_v)}{\sqrt{\sum_{k=1}^n (z_{ku} - \bar{z}_u)^2 \sum_{k=1}^n (z_{kv} - \bar{z}_v)^2}} \quad (4)$$

式中: z_{ku} 、 z_{kv} 分别为参数 Z_u 、 Z_v 的第 k 个值; \bar{z}_u 、 \bar{z}_v 分别为参数 Z_u 、 Z_v 的平均值。

3) 求特征值和特征向量 解特征方程 $|R - \lambda I_P| = 0$ 得 P 个特征值 $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_P \geq 0$ 。分别求出对应特征值 λ_i 的特征向量 e_i ($i = 1, 2, \dots$), 且 $e_i = 1$ 。

4) 确定主成分个数 m 主成分个数 m 的确定

方法: $\lambda_1 \geq 1$ 且 $\frac{\sum_{j=1}^m \lambda_j}{\sum_{j=1}^p \lambda_j} \geq 0.8$ 。

5) 求出 $z_i = (z_{i1}, z_{i2}, \dots, z_{ip})^T$ 的 m 个主成分

$$u_{is} = z_i^T e_i \quad (s = 1, 2, \dots, m) \rightarrow$$

$$U = \begin{bmatrix} u_{12} & u_{13} & \cdots & u_{1m} \\ u_{21} & u_{22} & \cdots & u_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ u_{n1} & u_{n2} & \cdots & u_{nm} \end{bmatrix}$$

1.3.2 聚类分析 聚类分析是采用数学方法按照某种相似性或差异性参数定量地确定样本之间的亲疏关系,再根据亲疏关系对样本进行分类,使同类中的样本对象具有最高相似性,不同类间的样本具有最低相似性。该方法的一个主要优点在于可在考虑不同参数影响的同时进行合理分类^[12-13]。

现有的聚类分析典型算法包括 k 均值算法、BIRCH算法和CLIQUE算法等^[12-13]。其中, k 均值算法由于其原理简单和计算速度快等优点得到广泛应用。采用该算法进行聚类分析,其一般步骤为

1) 从 n 个样本对象中任意选择 k 个对象作为初始聚类中心 O_1, O_2, \dots, O_k ;

2) 根据式(6)计算样本对象与各聚类中心的欧式距离,并根据最小欧式距离重新划分相应对象;

$$E(a, b) =$$

$$\sqrt{(x_{a1} - x_{b1})^2 + (x_{a2} - x_{b2})^2 + \cdots + (x_{am} - x_{bm})^2} \quad (5)$$

3) 重新计算各聚类特征参数的平均值,即新的聚类中心;

4) 循环迭代步骤2)和3),直到各聚类中心不再发生变化。

1.3.3 行为节能方案 考虑到用户待机行为主要表现为其采用不同待机模式及待机时长,为全面评估相应行为节能潜力,采用3个行为节能方案:1)为保持待机时长不变而仅调整待机模式,即用功率值较低的待机模式替换功率值较高的待机模式;2)为保持待机模式不变而仅调整待机时长,即将各类用户不同待机模式使用时长降低到每一类用户相应待机模式使用时长平均值;3)将上述两种方案相结合以评估最大行为节能潜力。

2 研究结果及分析

2.1 主成分分析

对采集得到的样本数据矩阵进行主成分分析得出特征值和方差贡献率,如表2所示。考虑到特征值的大小与主成分对原始变量解释力度的大小有关,且当主成分的特征值小于1时,其对原始变量的解释力度不如原始变量自身的解释力度^[14]。因此,通常可选取特征值大于1的主成分为综合变量进行后续分析。此外,为保证所提取的主成分能够反映

原始变量的绝大部分信息,其累积方差贡献率通常应大于 80%(方差贡献率表示该主成分包含原始数据的信息量占全部信息量的比重)^[15]。由表 2 可知,前 4 个主成分的特征值均大于 1,且累积贡献率为 80.2%,因此,可将这 4 个主成分作为综合变量评估用户待机行为节能潜力。

表 2 特征值及方差贡献率

Table 2 Eigenvalue and contribution rate of variance

序号	特征值	方差贡献率/%	累积方差贡献率/%
1	6.9	49.3	49.3
2	1.8	12.9	62.2
3	1.4	10.0	72.2
4	1.1	8.0	80.2
5	0.7	5.0	85.3
6	0.6	4.2	89.4
7	0.4	2.8	92.3
8	0.3	2.5	94.7
9	0.3	2.1	96.8
10	0.4	1.4	98.2
11	0.1	1.0	99.2
12	0.1	0.6	99.8
13	0.0	0.2	100
14	9.021×10^{-16}	6.443×10^{-15}	100

为进一步理解各主成分与原始变量之间联系的紧密程度,计算了因子载荷矩阵,如表 3 所示。因子载荷是各主成分特征值的平方根与相应特征向量的乘积,实质上是各主成分与原始变量之间的相关系数,且因子载荷值较大(通常取大于 0.5)的变量被认为是该主成分的主要相关变量^[16]。根据该标准,由表 3 可知,主成分 1 主要与用户性别、婚姻状况、兴趣爱好、专业、学历、是否定期接受节能教育、是否有上网流量限制、是否午休、是否住校、是否打卡出勤相关;主成分 2 主要与用户年龄、入学年数、专业成绩相关;主成分 3 主要与年龄、生活费用有关;主成分 4 主要与用户生活费用相关。

表 3 因子载荷矩阵

Table 3 The factor loading matrix

变量	主成分 1	主成分 2	主成分 3	主成分 4
性别	0.8	0.0	-0.3	0.4
年龄	0.1	0.6	0.6	0.1
婚姻状况	0.8	0.2	-0.4	0.3
兴趣爱好	0.9	-0.1	0.1	-0.2
专业	0.7	-0.2	0.1	-0.1
学历	0.9	-0.1	0.0	-0.2

续表 3

变量	主成分 1	主成分 2	主成分 3	主成分 4
入学年数	0.1	0.8	-0.0	-0.4
专业成绩	0.1	0.7	-0.5	-0.2
生活费用	0.0	0.2	0.6	0.5
节能教育	0.9	-0.1	0.1	-0.2
流量限制	0.8	0.2	-0.2	0.2
是否午休	0.9	-0.1	0.2	-0.2
是否住校	0.8	0.2	-0.2	0.3
是否打卡	0.8	-0.1	0.2	-0.2

2.2 聚类分析

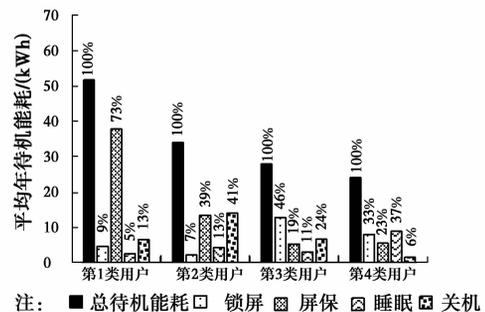
基于所得主成分对 100 名样本用户进行聚类分析并将其分为 4 类,每一类聚类中心(即各主成分的平均值)及用户数量如表 4 所示。

表 4 聚类分析结果

Table 4 The results of clustering analysis

聚类	主成分 1	主成分 2	主成分 3	主成分 4	用户数量
聚类 1	1.5	0.5	0.2	0.0	23
聚类 2	0.2	0.6	0.1	0.1	28
聚类 3	3.1	0.5	0.0	0.1	36
聚类 4	1.5	1.1	-0.7	1.0	13

由表 4 可知,不同类用户之间的聚类中心差异较大,表明各类用户之间的特征差异显著。为进一步分析各类用户待机能耗特征,图 1 给出了 4 类用户年总待机能耗及各待机模式下的年待机能耗。由图 1 可知,不同类用户间的能耗分布差异非常显著,这表明不同用户特征会导致不同用户待机行为和能耗,从而节能潜力也有所不同。因此,有必要先将样本用户分类,然后,针对用户特征分别计算节能潜力。此外,同类用户之间不同待机模式下的能耗差异较大,其中,锁屏和屏保模式待机能耗在 4 类用户中占总能耗比例均超过 45%,最高达到 82%,应重点进行考虑。



注: ■ 总待机能耗 □ 锁屏 ▨ 屏保 ▩ 睡眠 ▪ 关机

图 1 4 类用户的分项能耗特征

Fig. 1 The energy consumption of standby modes of four categories users

2.3 节能潜力分析

2.3.1 方案 1 节能潜力分析 方案 1 为保持待机时长不变而仅调整待机模式。为深入理解各待机模式及各类用户对其使用情况,表 5 给出了电脑各待机模式功率值及各类用户相应待机模式的使用率(表中使用率为使用某待机模式用户数量与该类用户数量的比值)。由表 5 可知,电脑各待机模式的功率值差异较大,且屏保与锁屏模式的功率值远高于睡眠和关机模式,因此,调整待机模式时,应用睡眠或关机模式下的功率值替换屏保和锁屏模式下的功率值。就这 4 种模式的使用率而言,各类用户之间差距较大,进一步表明方案中应分别计算其节能潜力。

表 5 各待机模式功率值及各类用户相应待机模式的使用率

Table 5 The power and utilization rate of standby modes of four categories users

待机模式	功率值/W				使用率/%			
	最小值	平均值	最大值	标准差	第 1 类用户	第 2 类用户	第 3 类用户	第 4 类用户
屏保	12	46.8	112	25.8	18	19	14	10
锁屏	10	41.1	111	25.3	9	11	10	0
睡眠	0.4	6.4	49	11.1	46	51	50	54
关机	0.2	3.3	49	6.7	27	18	26	35

除了关机待机模式外,当电脑处于屏保、锁屏、睡眠待机模式时,电脑均能快速从待机状态进入工作状态,因此,尽管关机待机模式下功率值相对较低,本方案仍从方便用户使用角度出发,采用睡眠模式下的功率值代替屏保和锁屏模式下的功率值,同时,保持其使用时间不变。结果表明,各类用户的年平均待机能耗分别从 52、34、28、24 kWh 减少到 4、6、7、9 kWh。平均节能潜力为 48、28、21、15 kWh。

2.3.2 方案 2 节能潜力分析 方案 2 为保持待机模式不变而仅调整待机时长。为深入理解各类用户待机时长分布特征,图 2 给出了各类用户不同待机模式下的日平均使用时的最小值、平均值及最大值。由图可知,4 类用户各待机模式的时长分布范围较广,不同类用户间时长分布不均且显著不同,进一步表明在本方案中应分别计算其节能潜力。

本方案的节能潜力计算是通过将各类用户不同待机模式时长降低为其相应待机模式的使用时长平均值。结果显示,各类用户的年平均待机能耗分别从 52、34、28、24 kWh 减少到了 10、13、10、10 kWh。平均节能潜力为 42、21、18、14 kWh。

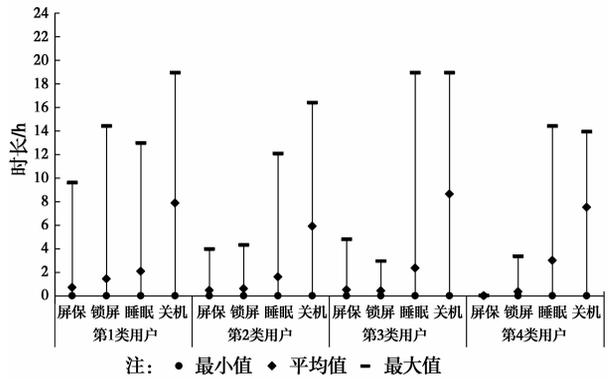


图 2 4 类用户不同待机模式下使用时长分布

Fig. 2 The distribution of standby modes of four categories users

2.3.3 方案 3 节能潜力分析 方案 3 为将上述两种方法相结合。结果表明,各类用户的年平均待机能耗分别从 52、34、28、24 kWh 减少到了 4、5、6、8 kWh。平均节能潜力为 48、29、22、16 kWh。

表 6 为 4 类用户不同方案下行为节能潜力。此外,表中同时给出了采用传统方法(即不对用户进行分类)后不同方案下的行为节能潜力作为参考。

表 6 4 类用户 3 种方案下行为节能潜力

Table 6 The energy-saving potential by taking three scenarios in four categories users kWh

用户类别	平均年待机能耗			年节能潜力			
	参考值	方案 1	方案 2	方案 3	方案 1	方案 2	方案 3
第 1 类用户	52	4	10	4	48	42	48
第 2 类用户	34	6	13	5	28	21	29
第 3 类用户	28	7	10	6	21	18	22
第 4 类用户	24	9	10	8	15	14	16
传统方法 (全部用户)	35	6	12	6	29	23	30

3 结论

为提高用户行为节能潜力评估的准确度,首先,采用主成分分析技术对用户行为的潜在影响因素进行降维处理,在此基础上,采用聚类分析技术对用户进行合理分类,并针对不同类用户特征分别进行节能潜力评估。将该方法应用于湖南省 3 所高校 100 名研究生样本用户电脑待机行为节能潜力的分析和评估,结果验证了该方法的有效性和可行性。该方法可应用于不同类型建筑(如住宅建筑、办公建筑和学校建筑)中不同种类行为(如正常使用行为和待机行为)的节能潜力评估,进而帮助深入分析用户行为对建筑能耗的影响和有效降低建筑能耗,促进我国建筑节能工作的有效开展。

参考文献:

- [1] JIA M, SRINIVASAN R S, RAHEEM A A. From occupancy to occupant behavior: An analytical survey of data acquisition technologies, modeling methodologies and simulation coupling mechanisms for building energy efficiency [J]. *Renewable & Sustainable Energy Reviews*, 2017, 68:525-540.
- [2] HONG T Z, TAYLOR L S C, D'OCA S, et al. Advances in research and applications of energy-related occupant behavior in buildings [J]. *Energy & Buildings*, 2015, 116:694-702.
- [3] SAHIN M C, KOKSAL M A. Standby electricity consumption and saving potentials of Turkish households [J]. *Applied Energy*, 2014, 114 (2): 531-538.
- [4] CARMENATE T, INYIM P, PACHEKAR N, et al. Modeling occupant-building-appliance interaction for energy waste analysis [J]. *Procedia Engineering*, 2016, 145:42-49.
- [5] AZAR E, ANSARI H. Framework to investigate energy conservation motivation and actions of building occupants: The case of a green campus in Abu Dhabi, UAE [J]. *Applied Energy*, 2017, 190:563-573.
- [6] CHEN S, YANG W, YOSHINO H, et al. Definition of occupant behavior in residential buildings and its application to behavior analysis in case studies [J]. *Energy & Buildings*, 2015, 104:1-13.
- [7] WANG C, YAN D, REN X. Modeling individual's light switching behavior to understand lighting energy use of office building [J]. *Energy Procedia*, 2016, 88: 781-787.
- [8] YUN G Y, STEEMERS K. Behavioural, physical and socio-economic factors in household cooling energy consumption [J]. *Applied Energy*, 2011, 88 (6): 2191-2200.
- [9] KAMILARIS A, NEOVINO J, KONDEPUDI S, et al. A case study on the individual energy use of personal computers in an office setting and assessment of various feedback types toward energy savings [J]. *Energy & Buildings*, 2015, 104(1):73-86.
- [10] SMITH L I. A Tutorial on principal components analysis [J]. *Information Fusion*, 2002, 51(3):219-226.
- [11] 韩小孩, 张耀辉, 孙福军, 等. 基于主成分分析的指标权重确定方法[J]. *四川兵工学报*, 2012, 33(10): 124-126.
- HAN X H, ZHANG Y H, SUN F J, et al. Weight determination of indicators based on principal component analysis [J]. *Journal of Sichuan Ordnance*, 2012, 33(10):124-126. (in Chinese)
- [12] 孙吉贵, 刘杰, 赵连宇. 聚类算法研究[J]. *软件学报*, 2008, 19(1):48-61.
- SUN J G, LIU J, ZHAO L Y. Clustering algorithms research [J]. *Journal of Software*, 2008, 19(1):48-61. (in Chinese)
- [13] HAN J W, KAMBER M, PEI J. Data mining: concepts and techniques [M]. 3rd ed. Beijing: China Machine Press, 2012.
- [14] 陈欢, 曹承富, 张存岭, 等. 基于主成分-聚类分析评价长期施肥对砂姜黑土肥力的影响[J]. *土壤学报*, 2014 (3):609-617.
- CHEN H, CAO C F, ZHANG C L, et al. Princl component-cluster analysis of effects of long-term fertilization on fertility of lime concretion black soil [J] *Acta Pedologica Sinica*, 2014(3):609-617. (in Chinese)
- [15] 张鹏. 基于主成分分析的综合评价研究[D]. 南京:南京理工大学, 2004.
- ZHANG P. Research on comprehensive evaluation based on principal component analysis [D]. Nanjing: Nanjing University of Science & Technology, 2004. (in Chinese)
- [16] 顾文, 庄智, 谭洪卫. 基于聚类分析的酒店建筑分类与用能评价[J]. *建筑技术*, 2014, 45(11):1026-1029.
- GU W, ZHUANG Z, TAN H W. Hotel classification based on cluster analysis and evaluation of energy consumption [J]. *Architecture Techology*, 2014, 45 (11):1026-1029. (in Chinese)
- [17] 曾亮, 傅丽亚, 罗理勇, 等. 不同品种和花期茶树花挥发性物质的主成分和聚类分析[J]. *食品科学*, 2015, 36(16):88-93.
- ZENG L, FU L Y, LUO L Y, et al. Principal component and cluster analyses of volatile components in tea flowers from different cultivars at different stages of bloom [J]. *Food Science*, 2015, 36(16):88-93. (in Chinese)
- [18] 公丽艳, 孟宪军, 刘乃侨, 等. 基于主成分与聚类分析的苹果加工品质评价[J]. *农业工程学报*, 2014, 30 (13):276-285.
- GONG L Y, MENG X J, LIU N Q, et al. Evaluation of apple quality based on principal component and hierarchical cluster analysis [J]. *Transactions of the Chinese Society of Agricultural Engineering*, 2014, 30 (13):276-285. (in Chinese)
- [19] 李郡, 俞准, 刘政轩, 等. 住宅建筑能耗基准确定及用能评价新方法[J]. *土木建筑与环境工程*, 2016, 38(2): 75-83.
- LI J, YU Z, LIU Z X, et al. A method for residential building energy benchmarking and energy use evaluation [J]. *Journal of Civil, Architectural and Environmental Engineering*, 2016, 38(2):75-83. (in Chinese)