

基于大数据的物联网用户行为模式挖掘

陆兴华, 林佳聪, 谢欣殷, 林家豪
(广东工业大学华立学院, 广东 广州 511325)

摘要:对智能家居物联网用户的行为模式进行准确有效的挖掘,提高智能家居物联网的优化组网能力,实现智能家居的优化控制,提出一种基于大数据的智能家居物联网用户的行为模式挖掘方法。构建智能家居物联网用户行为模式的大数据分析模型,采用模糊调度方法对用户行为特征进行关键行为特征点定位,采用资源标识方法进行用户行为模式自适应标定和状态重组,建立用户行为模式的大数据分类模型。根据用户行为特征的聚类性实现智能家居物联网用户行为特征挖掘和自适应聚类,采用极限机学习算法进行智能家居物联网用户行为模式挖掘的收敛性控制,提高用户行为模式挖掘的自适应性。仿真结果表明,采用该方法进行智能家居物联网用户的行为模式挖掘的准确性较高,挖掘过程的收敛性较好。

关键词:极限机学习;物联网;行为模式;挖掘;大数据分析

中图分类号:TP391

文献标识码:A

文章编号:1673-629X(2019)12-0099-05

doi:10.3969/j.issn.1673-629X.2019.12.018

Mining of User Behavior Pattern in Internet of Things Based on Big Data

LU Xing-hua, LIN Jia-cong, XIE Xin-yin, LIN Jia-hao
(Huali College Guangdong University of Technology, Guangzhou 511325, China)

Abstract: The behavior pattern of smart home users are accurately and effectively excavated to improve the networking capability of the intelligent home Internet of things and to realize the optimal control of intelligent home. A method of mining the behavior pattern of smart home users based on big data is proposed. We construct a big data analysis model of smart home user behavior pattern, use fuzzy scheduling method to locate the key behavior feature points of user behavior characteristics, and adopt resource identification method to self-adaptively calibrate and reorganize user behavior mode. The big data classification model of user behavior pattern is established, and user behavior feature mining and adaptive clustering are realized according to the clustering of user behavior characteristics. The limit machine learning algorithm is used to control the convergence of user behavior pattern mining in the Internet of things in smart home, and the self-adaptability of user behavior pattern mining is improved. The simulation shows that the proposed method is accurate and convergent.

Key words: limit machine learning; Internet of things; behavior pattern; mining; big data analysis

0 引言

随着人工智能控制技术和物联网技术的快速发展,智能家居技术取得了较快的发展并不断成熟,为用户提供了现代化的家居智能控制体验。智能家居是建立在物联网基础上,通过物联网技术将家中的各种设备(如音视频设备、照明系统、窗帘控制、空调控制等)连接到一起,提供家电控制、照明控制、室内外遥控、防盗报警、环境监测、暖通控制、红外转发以及可编程定时控制等多种功能和手段。在智能家居设计中,每个

用户都具有自身偏好的特征,需要对用户进行针对性的特征分析,对智能家居物联网用户的行为模式进行挖掘和特征分析,制定符合用户行为特征的智能家居控制模型,提高智能家居的用户价值体验^[1]。

当前,对智能家居物联网用户行为模式数据挖掘主要采用的是多源数据的信息数据库构建方法,结合 QoS 预测实现智能家居服务的评估和数据挖掘^[2],提高智能家居的用户行为模式挖掘性能。典型的用户行为模式挖掘方法主要有主成分分析方法、支

收稿日期:2018-12-18

修回日期:2019-04-22

网络出版时间:2019-06-27

基金项目:2019 年“攀登计划”广东大学生科技创新培育专项资金立项项目(pdjh2019b0617)

作者简介:陆兴华(1981-),男,硕士,副教授,通讯作者,研究方向为嵌入式技术、无人机飞行稳定性控制方法、机器人运动控制方法;林佳聪(1996-),男,研究方向为计算机控制技术、传感器应用。

网络出版地址: <http://kns.cnki.net/kcms/detail/61.1450.TP.20190627.1105.048.html>

持向量机算法、模糊 C 均值算法、语义指向性特征提取算法等^[3-4]。通过对用户行为模式的数据分析和大数据数据库模型的构建,结合特提取方法进行智能家居物联网用户行为模式挖掘和特征分析,取得了较好的用户行为模式挖掘效果。其中,文献[5]提出一种基于模糊 C 均值聚类算法的智能家居物联网用户行为模式挖掘方法,对智能家居物联网用户行为特征分布大数据进行 FCM 聚类预处理,对聚类输出的数据进行时频分析和特征点检测,实现用户行为特征的准确挖掘;文献[6]提出一种基于语义本体模型和关联指向性特征提取的智能家居物联网用户行为挖掘算法,采用语义特征信息提取和关联规则挖掘方法,实现对用户行为特征的模糊指向性挖掘,实现对用户行为的动态特征挖掘模型优化,但该方法在用户行为挖掘中的计算开销较大,挖掘精度不高。

针对上述问题,文中提出一种基于大数据的智能家居物联网用户的行为模式挖掘方法。构建智能家居物联网用户行为模式的特征数据模型,采用关联规则特征分解方法进行用户行为模式的大数据分析和信息重构。根据用户行为模式大数据之间的差异性进行指向性行为特征分析,根据用户的行为偏好进行特征分类和信息融合处理,建立用户行为模式的大数据分类模型,根据用户的行为特征实现智能决策和判断,采用极限机器学习算法进行收敛性控制,提高用户行为模式挖掘的自适应性。最后进行实验分析,验证该方法在提高智能家居物联网用户行为模式挖掘准确性和收敛性方面的优越性能。

1 智能家居物联网用户行为模式的大数据分布模型及特征提取

1.1 智能家居物联网用户行为模式的大数据分布模型构建

首先需要构建智能家居物联网用户行为模式的大数据分布模型。智能家居物联网用户行为的特征向量集 $S = \{s_1, s_2, \dots, s_k\}$, 采用有向图 $G_1 = (M_1^\alpha, M_1^\beta, Y_1)$, $G_2 = (M_2^\alpha, M_2^\beta, Y_2)$ 表示智能家居物联网用户行为模式的大数据节点分布,在给定的行为模式下,用向量 $x =$

$[x_1, x_2, \dots, x_k]$ 表示特征分量,用户行为模式的信息特征分量的 a_i 的属性值为 $\{c_1, c_2, \dots, c_k\}$ 。在用户行为特征采样信息交换中心,满足 $G_1 \subseteq G_2 \Leftrightarrow Y_1 \subseteq Y_2$, 输出的智能家居物联网用户行为特征分布满足 $X \sim S_\alpha(\sigma, \beta, \mu)$, $0 < \alpha < 2$ 。采用三元组形式表示大规模的智能家居物联网用户行为特征分布集:

$$V_0(k) = \begin{cases} \gamma(1)\gamma^T(1), k = 1 \\ \frac{[\rho V_0(k-1) + \gamma(k)\gamma^T(k)]}{1 + \rho}, k > 1 \end{cases} \quad (1)$$

$$\gamma(k) = \tilde{M}(k) - \tilde{H}(k)\hat{x}(k|k-1) \quad (2)$$

在关联约束下,用户行为模式满足 Wigner-Ville 分布:

$$cx + b \sim S^\alpha(|c|, \sigma, \text{sgn}(c), \mu + b) \quad (3)$$

在 Wigner-Ville 分布空间进行大规模异构数据重组^[7],建立智能家居物联网用户行为模式的检测模型为:

$$\tilde{u}_{elv,k} = \tilde{u}_{e,k} + \tilde{\Sigma}_{ee,k}^{-1} \tilde{\Sigma}_{ev,k}^T (v_k - \tilde{u}_{v,k}) \quad (4)$$

$$\begin{cases} v_k \sim t_{v_k}(\tilde{u}_{v,k}, \tilde{\Sigma}_{vv,k}) \\ e_k \sim t_{e_k}(\tilde{u}_{e,k}, \tilde{\Sigma}_{ee,k}) \end{cases} \quad (5)$$

使用 E_i, e_s^t 表示用户行为模式的频率分量和时域分量,用户行为模式特征分布的多维向量模型表示为:

$$X_n = \{X_n, X_{n-\tau}, \dots, X_{n-(d-1)\tau}\} \quad (6)$$

采用模糊调度方法对用户行为特征进行关键行为特征点定位,得到的定位结果为:

$$CW_{\min}^{l_{i+1}} = \begin{cases} CW_{\min}^0 \times (1 + \bar{D}_{l_0})^x, i = 0 \\ CW_{\min}^{l_i} \times (1 + \bar{D}_{l_i})^x, 0 < i < M \end{cases} \quad (7)$$

对用户行为模式进行 QoS 控制,得到 QoS 控制加权的闭频繁项集合为:

$$\chi = \log_{(1+\bar{D})} A / CW_{\min}^0 \quad (8)$$

其中, CW_{\min}^0 为协同过滤模糊预测初始化值; $CW_{\min}^{l_i}$ 为第 W_{\min}^0 步的语义相似性尺度。

由此构建智能家居物联网用户行为模式的大数据分布模型,如图 1 所示。

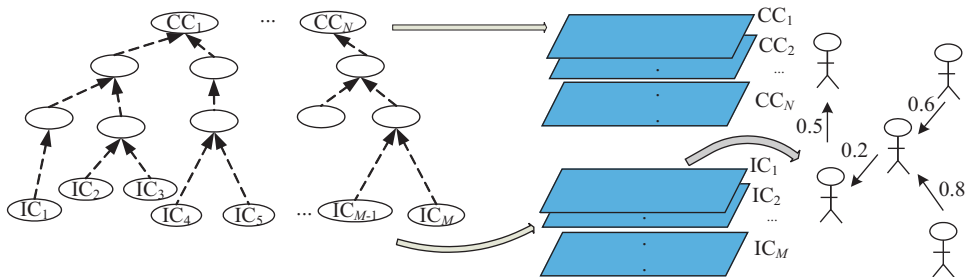


图 1 智能家居物联网用户行为模式的分布模型

1.2 智能家居物联网用户行为特征提取

采用相空间重构方法对智能家居物联网用户行为模式数据信息进行非线性特征分解,在 $l(n)$ 层的语义本体模型中,得到智能家居物联网用户行为模式特征集合的平均子节点为 $\bar{D}_{l(n)}$;从源节点至目标节点的拓扑树的总平均子节点数为 M ,采用资源标识方法进行用户行为模式自适应标定和状态重组,用户行为特征点的总数满足:

$$\bar{D} = \sum_{l_i=1}^{M-1} |D_{l_i}| / \sum_{j=1}^{M-1} |L_j| \quad (9)$$

采用关联规则特征分解方法进行用户行为模式的大数据分析和信息重构,重构结果为:

$$\begin{aligned} CW_{\min}^{l_u} &= CW_{\min}^{l_{u-1}} \times (1 + \bar{D}_{l_{u-1}})^x = CW_{\min}^0 \times (1 + \bar{D}_{l_0})^x \times (1 + \bar{D}_{l_1})^x \times \cdots \times (1 + \bar{D}_{l_{u-1}})^x = \\ &= CW_{\min}^0 \times [(1 + \bar{D}_{l_0}) \times (1 + \bar{D}_{l_1}) \times \cdots \times (1 + \bar{D}_{l_{u-1}})]^x \end{aligned} \quad (10)$$

其中

$$\begin{aligned} [(1 + \bar{D}_{l_0}) \times (1 + \bar{D}_{l_1}) \times \cdots \times (1 + \bar{D}_{l_{u-1}})] &\leq \\ [(1 + \bar{D}_{l_0} + 1 + \bar{D}_{l_1} + \cdots + 1 + \bar{D}_{l_{u-1}})/M]^M &= \\ (1 + \bar{D})^M \end{aligned} \quad (11)$$

考虑存在不确定时延下的用户信任权重,得到智能家居物联网用户行为特征分布的特征值满足约束条件:

$$\begin{aligned} CW_{\min}^{l_u} &\leq CW_{\min}^0 \times [(1 + \bar{D})^M]^x = \\ CW_{\min}^0 \times A / CW_{\min}^0 &= A \end{aligned} \quad (12)$$

由此构建智能家居物联网用户行为模式数据信息流模型,通过特征分解实现用户行为模式挖掘。

2 用户行为模式挖掘算法优化

2.1 用户行为模式大数据分析

在构建智能家居物联网用户行为模式的特征数据模型,采用关联规则特征分解方法进行用户行为模式的大数据分析和信息重构的基础上,进行智能家居物联网用户的行为模式挖掘的优化设计。文中提出一种基于大数据的智能家居物联网用户的行为模式挖掘方法。根据用户行为模式大数据之间的差异性进行指向性行为特征分析^[8],得到用户行为特征的分布模型为:

$$\begin{aligned} CW_{\min}^n &= \\ \left\{ \begin{aligned} &[(1 - B_{l(n)})e^{1-\alpha_n} + B_{l(n)}] \times CW_{\min}^{l(n)}, \alpha_n > 1 \\ &CW_{\min}^{l(n)}, 0 \leq \alpha_n \leq 1 \end{aligned} \right. \end{aligned} \quad (13)$$

$$B_{l(n)} = (1/(1 + \bar{D}_{l(n)-1}))^x \quad (14)$$

$$\alpha_n = \begin{cases} |D_n|/\bar{D}_{l(n)}, \bar{D}_{l(n)} \neq 0 \\ 0, \bar{D}_{l(n)} = 0 \end{cases} \quad (15)$$

其中, α_n 为用户 A、B 之间的信任权重; $B_{l(n)}$ 为信任度评价渐进系数。

提取智能家居物联网用户行为模式的 QoS 信息特征为:

$$F_i = \frac{C_s + \frac{C_1}{bw_s}}{SIZE_i} \times (C_i)^{c_2} \quad (16)$$

采用多层时态属性分箱处理,建立 QoS 映射关系的动态平衡模型,进行智能家居物联网用户行为特征的分类处理,根据用户的行为偏好进行特征分类和信息融合。

2.2 极限机学习算法及挖掘优化

采用极限机学习算法进行收敛性控制,根据智能家居物联网用户行为模式的支持度单调递增原理,得到用户行为模式的信息融合输出为:

$$\text{Bel}(x_t) = p(x_t | z_t, u_{t-1}, z_{t-1}, \cdots, u_0, z_0) \quad (17)$$

采用如下极限机学习算法进行回归分析:

$$\begin{aligned} v_t &= wv_{t-1} + c_1 \text{rand}_1() \cdot (\text{pbest} - x_{t-1}) + \\ &+ c_2 \text{rand}_2() \cdot (\text{gbest} - x_{t-1}) \end{aligned} \quad (18)$$

$$x_t = x_{t-1} + v_t \quad (19)$$

其中, v_t 是极限机学习过程中的迭代速度,表示智能家居物联网用户行为模式挖掘全局寻优速度; x_t 是极限机学习的适应度值; c_1 和 c_2 是最优学习算子,一般取 $c_1 = c_2 = 2$; $\text{rand}_1()$ 和 $\text{rand}_2()$ 是 $[0, 1]$ 之间的随机数。

通过极限机学习,得到智能家居物联网用户行为大数据特征分布 X_i 对应的一个函数。构建智能家居物联网用户行为模式挖掘的统计决策目标函数为:

$$\begin{cases} x_{id}^{t+1} = \omega x_{id}^t + c_1 r_1 (p_{ad} - x_{id}^t) + c_2 r_2 (p_{gd} - x_{id}^t) \\ p_{ad} = \frac{1}{m} \sum_{i=1}^m p_{id} \end{cases} \quad (20)$$

其中, p_{ad} 为用户行为模式数据挖掘寻优过程中的最大适应度值。

采用动态惯性权重加权方法得到个体极值 p_{best} ,在收敛条件约束下,根据如下两个公式来更新极限机学习的交叉寻优过程:

$$\begin{cases} v_{id}^{t+1} = \omega v_{id}^t + c_1 r_1 (p_{id} - x_{id}^t) + c_2 r_2 (p_{gd} - x_{id}^t) \\ x_{id}^{t+1} = x_{id}^t + v_{id}^{t+1} \end{cases} \quad (21)$$

根据挖掘目标函数的参数寻优,实现对智能家居物联网用户行为大数据的准确挖掘和特征分析^[9-11],

得到特征函数为：

$$r = \frac{M^{-1} \sum_i j_i k_i - [M^{-1} \sum_i \frac{1}{2}(j_i + k_i)]^2}{M^{-1} \sum_i \frac{1}{2}(j_i^2 + k_i^2) - [M^{-1} \sum_i \frac{1}{2}(j_i + k_i)]^2} \tag{22}$$

分析智能家居物联网用户行为特征数据挖掘模板集的静态特性^[12],根据极限机学习算法进行收敛性控制,得到优化挖掘结果为:

$$\begin{aligned} \min(w, \xi, \xi^*) &= \dot{X} \frac{1}{2} \|w\|^2 + \dot{Y} C \sum_{i=1}^n (\xi_i + \xi_i^*) \\ \text{s. t. } H_i(z) &\begin{cases} y_i - w^T \Phi(x) - b \leq \varepsilon + \xi_i^* \\ -y_i + w^T \Phi(x) + b \leq \varepsilon + \xi_i \\ \xi_i^*, \xi_i \geq 0 \end{cases} \end{aligned} \tag{23}$$

其中, \dot{X} 为平衡因子; \dot{Y} 为数据挖掘的负载; $H_i(z)$

为智能家居物联网用户行为特征数据分布的模糊决策系统函数^[13]。

综上处理,实现了智能家居物联网用户行为模式挖掘^[14],提高了用户行为模式挖掘的自适应性。

3 仿真实验与结果分析

对文中方法在实现智能家居物联网用户行为模式挖掘中的性能进行仿真实验。惯性权重为 0.12,相关系数 $R = 0.22$,均方误差 $MSE = 0.087$,智能家居物联网用户行为数据样本集为 3 000,测试集为 200,用户规模为 4 000,对用户行为特征大数据采集的归一化初始频率 $f_1 = 0.3$,归一化终止频率 $f_2 = 0.05$ 。根据上述仿真环境和参数设定,进行智能家居物联网用户的行为模式挖掘,得到的挖掘输出的散点图如图 2 所示。

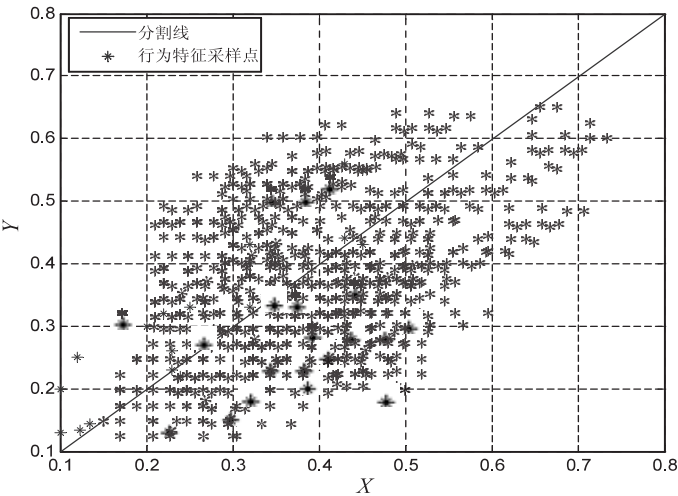


图 2 用户行为模式数据挖掘输出的散点图

从图 2 得知,通过对智能家居物联网用户行为模式的挖掘和预测,提高了对用户行为模式的智能分析能力。采用模糊调度方法对用户行为特征进行关键行

为特征点定位,采用资源标识方法进行用户行为模式自适应标定和状态重组,实现对用户行为特征提取,提取结果如图 3 所示。

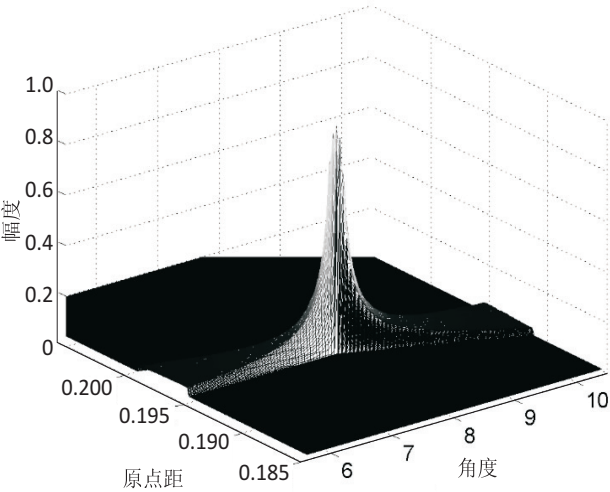


图 3 智能家居物联网用户行为模式提取定位结果

分析图 3 得知,采用文中方法进行智能家居物联网用户行为模式提取的准确性较高,抗干扰性能较强。采用不同方法进行挖掘的精度对比,得到的挖掘误差

对比结果如图 4 所示。分析得知,利用文中方法进行智能家居物联网用户行为模式挖掘的精度较高,准确性较好。

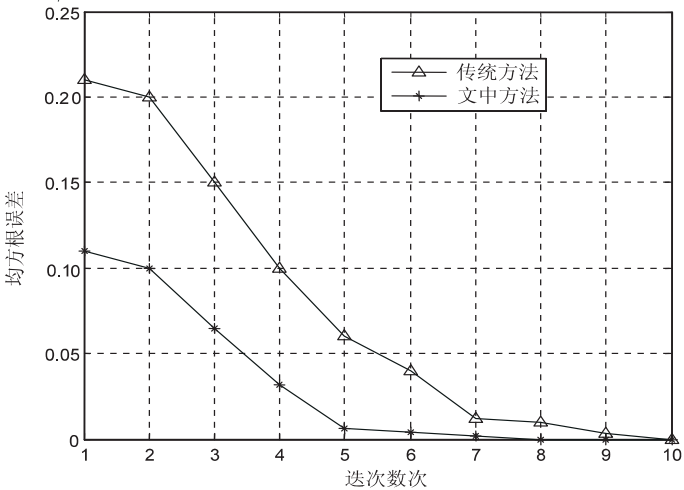


图 4 挖掘性能对比

4 结束语

文中提出一种基于大数据的智能家居物联网用户的行为模式挖掘方法。构建智能家居物联网用户行为模式的大数据分析模型,采用模糊调度方法对用户行为特征进行关键行为特征点定位,建立用户行为模式的大数据分类模型。根据用户的行为特征的聚类性实现智能家居物联网用户行为特征挖掘和自适应聚类,采用极限机器学习算法进行智能家居物联网用户行为模式挖掘的收敛性控制,提高用户行为模式挖掘的自适应性。实验结果表明,采用该方法进行智能家居物联网用户的行为模式挖掘的准确性较高,挖掘过程的收敛性较好,在智能家居物联网控制设计中具有很好的应用价值。

参考文献:

[1] 陆兴华,郑昶宏,吴焕彬,等. 加入前馈补偿的时滞系统鲁棒性跟踪控制器[J]. 计算机技术与发展,2017,27(6):136-141.

[2] 陆兴华,吴恩桑,黄冠华. 基于 Android 的智能家居控制系统软件设计研究[J]. 物联网技术,2015,5(11):14-16.

[3] 文天柱,许爱强,程 恭. 基于改进 ENN2 聚类算法的多故障诊断方法[J]. 控制与决策,2015,30(6):1021-1026.

[4] XU Yinyin, TONG Shaocheng, LI Yongming. Prescribed performance fuzzy adaptive fault-tolerant control of non-linear systems with actuator faults[J]. IET Control Theory and Ap-

plications,2014,8(6):420-431.

[5] HUANG Xia, WANG Zhen, LI Yuxia, et al. Design of fuzzy state feedback controller for robust stabilization of uncertain fractional-order chaotic systems[J]. Journal of the Franklin Institute,2015,351(12):5480-5493.

[6] XIONG He, GUO Youqiang, ZHU Honghao, et al. Robust non-negative matrix factorization on manifold via projected gradient method[J]. Information and Control,2018,47(2):166-175.

[7] 邢长征,刘 剑. 基于近邻传播与密度相融合的进化数据流聚类算法[J]. 计算机应用,2015,35(7):1927-1932.

[8] 陶新民,宋少宇,曹盼东,等. 一种基于流形距离核的谱聚类算法[J]. 信息与控制,2012,41(3):307-313.

[9] 田 刚,何克清,王 健,等. 面向领域标签辅助的服务聚类方法[J]. 电子学报,2015,43(7):1266-1274.

[10] 周于皓,张红玲,李芳菲,等. 局部关注支持向量机算法[J]. 计算机应用,2018,38(4):945-948.

[11] 刘雪晴. 面向科研社交网络的小同行双向推荐算法[J]. 计算机应用与软件,2018,35(10):171-180.

[12] 吴 磊,原 鹏,丁维龙. 智能家居网关与云服务器数据同步协议的研究[J]. 计算机技术与发展,2018,28(9):151-155.

[13] 李国良,楚娅萍,冯建华,等. 多社交网络的影响力最大化分析[J]. 计算机学报,2016,39(4):643-656.

[14] ULUKUS S, YENER A, ERKIP E, et al. Energy harvesting wireless communications: a review of recent advances [J]. IEEE Journal on Selected Areas in Communications,2015,33(3):360-381.