

基于马尔科夫决策过程的家庭能量管理智能优化策略

傅质馨^{1,2}, 李潇逸¹, 朱俊澎^{1,2}, 袁越^{1,2}

(1. 河海大学 能源与电气学院, 江苏 南京 211100;

2. 河海大学 可再生能源发电技术教育部工程研究中心, 江苏 南京 211100)

摘要:在迅速发展的通信技术和泛在电力物联网建设的背景下,结合多种信息交互方式和人工智能技术可为提高家庭能量管理的智能化程度提供新的思路。提出一种结合实时信息交互的家庭能量管理智能优化策略。首先,给出了以用户用能费用为基础的马尔科夫决策过程模型,采用动态规划方法求解模型,重点在家庭用电设备调度过程中考虑实时电价信息和用户的随机行为等不确定因素的影响;在此基础上,结合事件触发机制有效提高家庭能量管理系统的运行效率,进而给出从家庭能量管理控制中心到用电设备的智能优化调度方法;最后,通过仿真算例证实了所提方法的有效性,结果表明其能在减少用户用电费用的同时给出满足用户用电需求的优化用电策略。

关键词:家庭能量管理系统;马尔科夫决策过程;随机动态规划;实时电价;泛在电力物联网;智能优化

中图分类号:TM 732

文献标志码:A

DOI:10.16081/j.epae.202006029

0 引言

泛在电力物联网建设要求构建更高效、更安全、更可靠、更具弹性和可持续性的智能网络化电力能源系统^[1-2]。能量管理系统作为电网稳定、高效、经济运行不可或缺的一部分,其在需求侧和供给侧都起着至关重要的作用^[3]。而家庭能量管理系统 HEMS(Home Energy Management System)作为能量管理技术在用户侧的体现,可给出满足用户需求的各类负荷及分布式能源的优化控制策略,为需求响应和可再生能源的接入提供支持^[4-5],并对电网调峰做出显著贡献^[6-7]。

现阶段,国内外学者在家庭能量管理系统的用电模型、调度策略及优化算法等研究方面都取得了显著的成果^[8-9]。然而,家庭能量管理系统由多种负荷以及分布式电源构成,并且具有面向用户需求多样性的特性,这就决定了其必然存在大量的不可控因素与不确定性,主要包括对可再生能源预测的不确定性、与电力市场信息交互过程中实时电价 RTP(Real-Time electricity Price)的不确定性以及作为控制核心的管理系统在与用户交互过程中产生的多种不可预知的用户随机行为,这些都给家庭能量管理控制策略的制定带来了挑战。

目前处理上述多种不确定因素的方法主要分为在调度前处理或在调度中将不确定性与模型相结合 2 种。前者主要是采用智能优化算法处理不确定信息,如文献[10]提出了一种神经网络和遗传算法相

结合的方法,用于减小能源消耗并提高了可再生能源的利用率。但是,由于信息具有随机性和高复杂度,其未给出有效的调度模型。有些学者采用日前调度实时调整的方法,如文献[11]给出了一种基于事件驱动的在线能量管理策略,用一种实时的状态监测系统代替了传统的预测机制。然而,对随机信息的处理未做出详细设定。基于上述研究中存在的问题,有些学者采用将不确定性信息与模型相结合的方法来提高调度效率,主要方法包括鲁棒优化、机会约束优化、随机规划、随机动态规划 SDP(Stochastic Dynamic Programming)等。文献[12]给出了一种采用机会约束规划的家庭设备调度方法来应对用户的随机行为。文献[13]给出了一种采用随机优化和鲁棒优化的方法来响应实时电价的能量管理系统,所提算法能够应对电价的不确定性,但对 2 种方法的协同控制未给出详细的解决方案。文献[14]给出了一种基于建筑热力学模型的结合混合整数二次规划和模型预测控制的家庭能量管理系统,有效减少了系统的储能容量。文献[15]建立了综合的家庭能源局域网,通过模型预测控制方法提高系统面对预测误差的鲁棒性。文献[16]给出了一种以用户净支出与购电功率波动最小为目标的日内家庭用电双层优化策略及实时用电调整策略,但针对突发情况的预期风险估计略有不足。

随机动态规划是马尔科夫链的扩展,常将其建模为马尔科夫决策过程 MDP(Markov Decision Process)。在家庭能量管理系统的环境下,已有一些文献给出了采用 MDP 的控制框架,例如:文献[17]利用 k -means 方法对历史数据特征进行整合,建立了考虑光伏发电和设备用电需求的 MDP 模型;文献[18]将系统分为实时监控、随机调度和实时控制阶段,能

收稿日期:2019-12-23;修回日期:2020-05-06

基金项目:国家重点研发计划项目(2016YFB0900100)

Project supported by the National Key Research and Development Program of China(2016YFB0900100)

够结合实时电价信息给出较为合理的调度方案;文献[19]基于排队理论搭建了家庭恒功率设备模型,并采用部分可见的马尔科夫决策过程 POMDP (Partially Observable Markov Decision Process) 方法对模型进行求解;文献[20]给出了一种基于马尔科夫链的能量管理策略,在状态设置中加入了激活时间的约束,使整体模型更加灵活,但对设备的约束在实际中会降低系统的灵活性;文献[21]基于 MDP 给出了一种结合需求侧和供给侧整体福利最大化的模型,建立了多种设备的转移概率矩阵。

综上所述,现有文献在系统的控制模型和优化算法方面已取得了较多的成果,但在外部环境信息的反馈与控制系统的智能性方面还有待进一步研究。因此,本文拟给出一种结合 MDP 模型的家庭能量管理智能优化策略。通过实时的策略调整与判断机制,考虑光伏预测误差、实时电价和用户随机行为等不确定因素的影响,在有效减少系统误差和预测误差的同时提高系统的计算效率。首先,对家庭能量管理系统的框架进行说明;然后,结合家庭用电设备的负荷特性建立 MDP 模型,并给出基于动态规划的模型求解算法以及家庭能量管理系统的优化调度过程;最后,通过仿真分析验证所提方法的有效性,算例结果表明其能在减少用户用电花费的同时给出满足用户用电需求的优化用电策略。

1 家庭能量管理系统框架

1.1 系统整体框架

本文给出了家庭能量管理系统的整体框架,如图1所示,其包括电冰箱、热水器、空调、洗碗机、洗衣机、干衣机、电饭煲、热水壶等日常家庭用电设备,以及电动汽车和家用光伏系统。

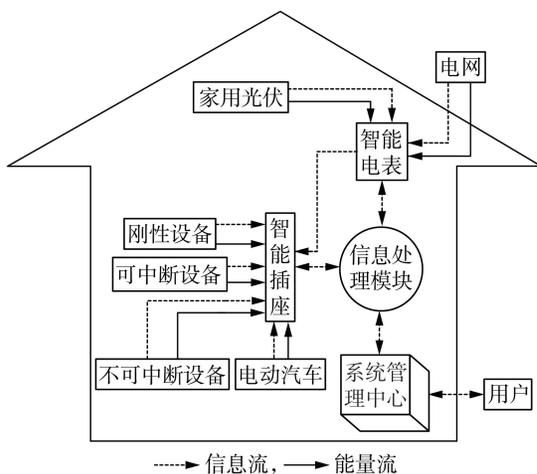


图1 家庭能量管理系统框图
Fig.1 Block diagram of HEMS

本文给出的家庭能量管理系统主要包括以下三部分:①信息感知部分,用于实时监控各类用电设

备、光伏系统、电动汽车的运行状态和电能变换情况以及电网的实时电价信息;②信息存储与传输部分,用于存储感知信息,并根据应用需求将必要的信息传递给管理中心;③系统管理中心,用于信息的接收与融合、智能优化策略的制订、故障的诊断与处理。这3个部分相互联系,即管理中心对获得的感知信息进行进一步融合处理,结合用户需求制订智能化的调度策略,在此过程中若发现设备发生故障,则发出报警信息。用户作为系统面向对象的主体,可根据自身需求对管理中心进行操作,这是自主控制部分。同时,管理中心可在用户预先设定的权限下控制设备,这是智能自动控制部分。2种控制方式相结合有利于实现智能用电。

家庭能量管理系统的运行过程中存在信息流和能量流的交互。能量的流动是通过管理中心控制日常用电设备、光伏系统和电动汽车的运行状态来进行彼此间及与上级电网之间的交互。信息的流动是通过智能电表、无线传感设备等物联网技术对设备信息进行实时感知,并将其传递至管理中心,同时管理中心还将通过通信设备及时获取实时电价信息。

综上,家庭能量管理系统即为一个用户侧的家庭物联网系统,各类设备是基本单元,其中用电设备、光伏系统、电动汽车也是被监控的对象,通信设备是连接这些监控对象与管理中心的桥梁,管理中心为核心部分,承载着整个系统智能优化运行的重要职责,也是能量流与信息流交互融合,最终形成具体应用价值的重要环节。

1.2 用电设备模型

结合现有研究的成果^[8],本节给出所涉及的家庭光伏系统与日常家庭用电设备的具体模型。

家用光伏系统是光伏板,其发电功率由光照强度和光伏板面积决定,可根据式(1)计算得到。

$$P_{pv} = r \eta_{pv} A_{pv} B \quad (1)$$

其中, P_{pv} 为光伏发电功率; r 为光照强度; η_{pv} 为光电转换效率; A_{pv} 为有效光照面积; B 为光伏板块数。

刚性负荷主要包括电视机、电脑、电灯等,其使用情况客观上反映了用户的使用习惯和生活方式。由于这些设备的启停会直接影响用户的居家体验,故将其视为基线负荷,在没有用户允许的情况下不对其进行调度。这些设备可通过响应设置在用户需要时直接启动,其模型如式(2)所示。

$$\begin{cases} P_{\min}^d \leq P_t^d \leq P_{\max}^d \\ \sigma^t = 0 \quad t \notin [t_{\text{start}}, t_{\text{end}}] \\ \sum_{t=\tau+1}^{\tau+T} \sigma^t \geq t_{\text{total}} (\sigma^{\tau+1} - \sigma^{\tau}) \quad \tau \in [t_{\text{start}} - 1, t_{\text{end}} - t_{\text{total}}] \\ \sum_{\tau=t_{\text{start}}}^{t_{\text{start}}+\lambda} \sigma^{\tau} \geq 1, \lambda = 0 \end{cases} \quad (2)$$

其中, $d \in D_{\text{gx}}, D_{\text{gx}}$ 为刚性负荷设备集合; P_t^d 为设备 d 在 t 时刻的功率; $P_{\text{max}}^d, P_{\text{min}}^d$ 分别为设备 d 功率的上、下限; σ^t 为设备在 t 时刻的开关状态量, 取值为 0 表示关闭, 取值为 1 表示开启; $t_{\text{start}}, t_{\text{end}}$ 分别为期望的开启、关闭时间, 由用户设定; t_{total} 为设备的额定持续运行时间; λ 为设备的最大延迟时间, $\lambda=0$ 表示刚性负荷不可延迟; T 为调度周期。

柔性负荷是指可控制的设备, 其特点是运行时间具有较高的可调性, 在允许运行的时间范围内对其进行控制时, 对居家用户使用体验的影响很小。柔性负荷分为可中断负荷、可转移负荷这 2 种。可中断负荷包括家用加湿器、干衣机、水泵等, 在其运行期间可以自由关断且不会对用户生活造成显著的影响, 模型如式(3)所示。

$$\begin{cases} P_{\text{min}}^d \leq P_t^d \leq P_{\text{max}}^d \\ \sigma^t = 0 \quad t \notin [t_{\text{start}}, t_{\text{end}}] \\ \sum_{t=\tau+1}^{\tau+T} \sigma^t \geq t_{\text{total}} \quad \tau \in [t_{\text{start}} - 1, t_{\text{end}} - t_{\text{total}}] \\ \sum_{\tau=t_{\text{start}}}^{t_{\text{start}}+\lambda} \sigma^\tau \geq 1, 0 \leq \lambda \leq t_{\text{end}} - t_{\text{start}} - t_{\text{total}} \end{cases} \quad (3)$$

其中, $d \in D_{\text{rx}}, D_{\text{rx}}$ 为可中断负荷设备集合。

可转移负荷主要包括洗衣机、电饭煲、洗碗机等设备, 其运行期间不可中断, 但是其启动与运行时间可以根据用户需求灵活设定, 模型如式(4)所示。

$$\begin{cases} P_{\text{min}}^d \leq P_t^d \leq P_{\text{max}}^d \\ \sigma^t = 0 \quad t \notin [t_{\text{start}}, t_{\text{end}}] \\ \sum_{t=\tau+1}^{\tau+T} \sigma^t \geq t_{\text{total}} (\sigma^{\tau+1} - \sigma^\tau) \quad \tau \in [t_{\text{start}} - 1, t_{\text{end}} - t_{\text{total}}] \\ \sum_{\tau=t_{\text{start}}}^{t_{\text{start}}+\lambda} \sigma^\tau \geq 1, 0 \leq \lambda \leq t_{\text{end}} - t_{\text{start}} - t_{\text{total}} \end{cases} \quad (4)$$

其中, $d \in D_{\text{zy}}, D_{\text{zy}}$ 为可转移负荷设备集合。

对于温控类设备, 例如热水器, 其能够将电能转化为热能存储。对这类设备采用温度变化来描述, 模型如式(5)所示。

$$\begin{cases} P_{\text{h, min}} \leq P_{\text{h}} \leq P_{\text{h, max}} \\ \sigma^t = 0 \quad t \notin [t_{\text{start}}, t_{\text{end}}] \\ \theta_{t+1} = \theta_{\text{en}} + P_{\text{h}} D_{\text{h}} - (\theta_{\text{en}} + P_{\text{h}} D_{\text{h}} - \theta_t) e^{-\Delta t / (D_{\text{h}} G_{\text{h}})} \\ \theta_{\text{min}} \leq \theta_t \leq \theta_{\text{max}} \end{cases} \quad (5)$$

其中, θ_t, θ_{t+1} 分别为 $t, t+1$ 时刻的水温; θ_{en} 为环境温度; P_{h} 为热水器功率; $P_{\text{h, max}}, P_{\text{h, min}}$ 分别为热水器功率的上、下限; D_{h} 为热水器的热阻系数; Δt 为时间间隔; G_{h} 为热水器热容; $\theta_{\text{max}}, \theta_{\text{min}}$ 分别为设定温度的上、下限。

2 基于 MDP 的优化模型

2.1 MDP 模型的建立

虽然实时的信息感知能够较为准确地获取家庭

设备的使用情况, 但是多种随机状况的发生(如电价发生变化、光伏发电功率变化、用户随机行为等)都可能给最优策略的制订带来一定的偏差, 因此本文将建立 MDP 模型, 在每个时段内都将考虑不确定性因素。建立离散时间分布的随机过程, 并给出时段集合 $Y = \{1, 2, \dots, n, \dots, N\}$ 。对于大部分经济型用户而言, 其主要驱动因素是用能费用。因此, 本文将用户的用能费用表示如下:

$$C_n = (c_n + \Delta c) P_n^{\text{total}} \Delta t \quad (6)$$

$$P_n^{\text{total}} = P_{\text{must}} + \sum_{n=1}^N P_n^{\text{con}} - (P_{\text{pv}} + \Delta P_{\text{pv}}) \quad (7)$$

其中, C_n 为 n 时刻的用电费用; c_n 为 n 时刻的实时电价信息; Δc 为电价信息偏差; P_n^{total} 为 n 时刻设备的总耗电量; P_{must} 为刚性负荷功率; P_n^{con} 为 n 时刻连接负荷功率; ΔP_{pv} 为光伏功率预测偏差; N 为时段总数。用户的行为习惯包含在 P_{must} 内; 式(7)中的 P_n^{total} 能够直观地反映家庭中的能源消耗情况, 同时通过式(6)将其与具体的用电费用 C_n 直接相联系。

基于上述整体的用能费用, 结合现有文献的成果^[18], 本文给出下述 MDP 模型。当将整体用能费用视为离散费用的集合, 即整体用能费用如式(8)所示时, 由此会产生一序列的费用变量, 在一定的范围内费用变量可看作一个状态, 形成状态值和用能费用的映射。

$$C^{\text{total}} = \{C_1, C_2, \dots, C_N\} \quad (8)$$

未来的用能费用状态只与现在的用能费用有关, 而与之之前的用能费用保持独立, 此时费用序列即可看作马尔科夫链, 如式(9)所示。

$$\mathbf{T}_{\text{rob}}^{12}(n+1) = \mathbf{T}_{\text{rob}}(C_{n+1} \rightarrow S_2 | C_n \rightarrow S_1) \quad (9)$$

其中, $\mathbf{T}_{\text{rob}}^{12}(n+1)$ 为系统从状态 S_1 转移至状态 S_2 时的状态转移矩阵。

由此, 进一步建立 MDP 模型, 主要包括状态空间(S)、动作集合(A)、转移概率矩阵(T)以及回报函数(R)的定义。MDP 问题的核心是在最大化奖励函数的条件下给出最优策略, 映射在本文中的具体目的是在用户用能费用最小的前提下给出最优的设备控制策略。

状态空间(S): 在考虑最小化用电费用和电量消耗的优化目标下, 利用用能费用的一定范围表征状态空间, 采用概率密度函数来划分状态边界, 如式(10)所示。

$$S_k = [B_k, B_{k+1}) = \{C_n | B_k \leq C_n < B_{k+1}\} \quad (10)$$

其中, B_{k+1}, B_k 分别为状态边界的上、下限, 由用能费用概率密度函数求得。

动作集合(A): 动作集合由一系列动作变量表征, 不同变量表示不同设备的开关状态。由信息控

制阶段给出设备的运行状态,不同运行状态的组合构成控制设备集合, t 时刻的动作集合为可控集合的子集,由此给出动作集合,如式(11)所示。

$$A = \{A_m, m = 1, 2, \dots, m_{\max}\} \quad (11)$$

$$A_m^q(S_k) \in \{0, 1, \dots, x_c, 12, \dots, 12, \dots, x_c\} \quad (12)$$

其中, m 为不同的可控设备; m_{\max} 为最大可控设备数; x_c 为最大可控设备序号; $A_m^q(S_k)$ 为在状态 S_k 下的动作集合。每一个动作表征不同设备的开关组合。

转移概率矩阵(T):表征了状态之间变化的概率。本文采用蒙特卡洛模拟法对状态转移概率进行逼近。由一个季度的用户用电数据进行模拟,能够有效表示用户行为变化和用电费用间的潜在关系。当考虑动作变量后,概率表征在当前状态下采用所给动作转移至下一状态的概率,如式(13)所示。

$$T_{rob}^{ij}(S_j|S_i, A_m) = N_{ij}^{A_m} / \sum_{k=1}^{k_{\max}} N_{ik}^{A_m} \quad (13)$$

其中, $T_{rob}^{ij}(S_j|S_i, A_m)$ 为在动作 A_m 下由状态 S_i 转移至状态 S_j 的转移概率; k_{\max} 为计算转移次数的最大值; $N_{ij}^{A_m}$ 为动作 A_m 下由状态 S_i 转移至状态 S_j 时的次数。同时,给出在动作变量下状态变化的定义为:

$$(S_i, S_j)_{A_m} \rightarrow C_1 + \sum_{k=1}^m \delta_k P_k = C_2 \quad (14)$$

其中, $C_1 \in S_i$; $C_2 \in S_j$; P_k 为在动作集合中包涵动作设备的运行功率; δ_k 为设备的运行状态量,以正负值的形式对整体用能费用产生影响。

回报函数(R):回报函数的定义是MDP中重要的一部分,直接影响整体最优策略集合的选取。本文给出的回报函数采用功率、电价、能耗以及设备状态时间来表征整体目标,如式(15)所示。

$$R = F(P_n^d, M_n^d, C_n, W_n) \quad (15)$$

其中, M_n^d 为表征 n 时刻设备 d 运行状态的0-1变量,处于运行状态取值为1,不处于运行状态则取值为0; W_n 为 n 时刻设备的可延时间。

整体回报采用用能费用和满意程度的加权来表征:第一部分用能费用变化由状态的变化决定,第二部分整体舒适度由动作给出的设备状态及设备的运行时间决定。

$$R_{ij}^{A_c}(n) = \alpha \frac{S_i}{S_j} + \beta \left(\sum_{d=1}^{x_c} M_n^d P_n^d \frac{W_n}{W_{set}} \right) \quad (16)$$

其中, $R_{ij}^{A_c}(n)$ 为 n 时刻由状态 S_i 在动作 A_c 下转移至状态 S_j 时的回报值; α 、 β 为比例系数; W_{set} 为设备设定的剩余时间。回报函数分为三部分:①表征能耗变化趋势的状态比例;②表征用户设备功率缩减的程度;③表征设备等待时间限制的充分延迟控制,用时间延迟比例表示。

2.2 模型求解

上述MDP模型可用值迭代和 q 学习的方法进行求解。本文中求解的目标是在最大化奖励函数的条件下给出优化控制策略,即在满足用户用电费用最少的条件下给出设备的最优控制策略。模型求解的具体步骤见附录A。在后续仿真过程中将对上述2种迭代方法的最终优化策略求解结果进行比较。

2.3 家庭能量管理系统优化调度过程

整体优化调度过程主要包含数据预处理、条件判断、设备状态监测、实时优化控制4个阶段。传统的家庭能量管理系统一般都着眼于用电设备运行状态的优化,而本文将在优化用电设备运行状态的同时更加关注光伏功率的预测误差、实时电价和用户随机行为这些不确定因素对优化结果的影响。在优化过程的4个阶段中,每个时段下都依据实时电价信息、光伏状态信息进行判断,并给出相应的计算结果。

(1)数据预处理阶段:包括光伏系统发电状态、实时电价信息、家用电器状态实时信息的获取。由于传感器的传输时间间隔和实时电价信息间隔为5 min,所以本文给定时间间隔为5 min。

(2)条件判断:由给定信息的判断机制触发控制系统进入循环。为了提高系统有效运算的次数和效率,在进入控制中心前进行设备控制机制判断,当满足判断条件时,则允许进入,否则按正常运行。由此利用事件触发机制来有效提高家庭能量管理系统的运行效率。本文给出了以下4种判断条件(即触发机制)。

a. 当需求能耗小于光伏产出时,将余电回馈电网以赚取利润。

$$P_{pv} - P_D > \partial_p \quad (17)$$

其中, P_D 为需求光伏功率; ∂_p 为设定的阈值。

需要说明的是,为了提高可再生能源的利用率和消纳程度,本文采取优先利用光伏出力应对负荷用电的策略。当光伏出力不足时,从电网购电以保证负荷正常运行。

b. 当电力市场给出电价波动过大时,重新给出调度策略。

$$|\pi_t - \pi_{t+1}| \geq \partial_\pi \quad (18)$$

其中, π_t 、 π_{t+1} 分别为 t 、 $t+1$ 时刻的实时电价; ∂_π 为实时电价的限制值。

c. 设备的整体运行功率需在家庭负荷的允许范围内。

$$P_{total} < P_{limit} \quad (19)$$

其中, P_{total} 为家庭设备的整体运行功率; P_{limit} 为家庭负荷功率限制值。

d. 当用户产生随机行为或负荷出现规定以外

的运行情况时,需重新调度设备。

$$t^d \notin Y_{\text{set}} \quad (20)$$

其中, t^d 为设备 d 的启动时间; Y_{set} 为设定的设备运行时间集合。

(3)设备状态监测阶段:该阶段给出设备的可控信息,包括设备运行状态、可连接设备集合、可控设备集合。

(4)实时优化控制阶段:首先通过迭代求解给出最优控制策略,再由所处状态给出实际控制集合,最后将控制集合转化为设备控制信息,并实时传输给设备进行调控。由于在控制阶段存在多种设备的状态判断,定义以下4种设备状态集合,即计划外设备集合、运行设备集合、可控制设备集合、可延迟设备集合。

a. 计划外设备集合表征在设备允许运行时间外的设备启停,是用户的随机用电行为,在常规情况下不进行控制,当超出预设标准或得到用户许可后可进行控制。此外,每个控制阶段都会调整设备的调度策略,从而能够很好地结合实时电价进行设备的用能分配。

b. 运行设备集合:在设备允许的运行时段内,最大等待时间为0。对运行设备集合给出现阶段的控制策略,强制转换已连接的设备的运行状态。

$$\begin{cases} M_t^{\text{con}} > 0 \\ \delta = 1 \\ W_t^d = 0 \end{cases} \quad (21)$$

c. 可控制设备集合:表示已连接的设备,并且在设备允许的运行时段内有等待时间的设备集合,设备状态可由控制信息控制并进行转换。

$$\begin{cases} M_t^{\text{con}} > 0 \\ \delta = 1 \\ W_t^d \in W_{\text{max}} \end{cases} \quad (22)$$

其中, M_t^{con} 为 t 时刻设备的连接状态,可由后续控制过程表详细观察; W_t^d 为设备 d 在 t 时刻的允许等待时间; δ 为设备的运行状态; W_{max} 为设备的最大允许等待时间。

d. 可延迟设备集合:在设备允许的运行时段内,已连接设备,并且运行时间小于最大等待时间。这类设备是优先控制的设备。根据其运行时间与最大用户等待时间决定其状态的转换。

$$\begin{cases} M_t^{\text{con}} \geq 0 \\ \delta = 0 \\ W_t^d \in W_{\text{max}} \end{cases} \quad (23)$$

整体优化调度控制的流程见附录B中图B1。

3 仿真结果分析

3.1 算例参数设置

为了验证本文所提家庭能量管理系统框架和优

化调度策略的有效性,在 MATLAB 中搭建仿真模型进行分析验证。各设备的用电时段及具体信息见附录C中表C1。其中,温控设备的温度范围如下:空调设备需保持室温 24°C ,热水器保持温度在 $40\sim 60^\circ\text{C}$ 。

本文采用澳洲能源网站的实时电价信息^[22],见附录C中图C1。光伏系统的发电功率由仿真系统产生,光伏功率的上网电价为 0.4096 元 / (kW·h),基线负荷功率见附录C中图C2。其他计算参数的具体设置见附录C中表C2。

3.2 仿真结果分析

本节给出并分析了家庭能量管理系统的调度策略与调度结果。图2为系统在某时刻的回报函数值。可以看出,整体的状态转移倾向于往用能费用较小的状态进行转移,符合目标特征趋势。优化策略表示设备的组合编号,将全部设备进行排列组合并按顺序编号后得到策略组合1—256。

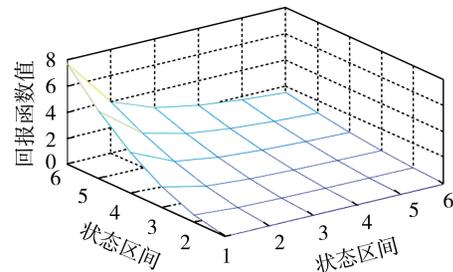


图2 回报函数值

Fig.2 Reward function values

由此可得到设备的最优策略曲线如图3所示。由图3可看出,在晚间负荷高峰期(18:00—21:00),会频繁触发控制机制进行调度策略的调整。在白天尤其是中午时段(11:00—13:00),当光伏电量充足时,由于采取优先消纳光伏出力应对负荷用电的方式,因此较少地调整调度策略,由此减少了控制机制的触发次数,提高了优化调度的效率。在无负荷需要控制且电价较低的时段(00:00—04:00和22:00—24:00)进行电动汽车充电。需要说明的是,由于电动汽车的储能特性需要结合用户的用电行为进行分析,且需要在电动汽车在家的情况下进行充电,因此

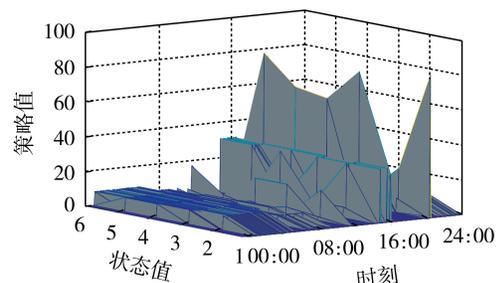


图3 最优策略曲线

Fig.3 Optimal strategy curves

本文暂未考虑电动汽车的储能特性。当负荷波动较小时,对负荷的调度着重于向费用低的状态即用户费用降低的方向转移。

在给出设备的调度策略后,结合实时信息控制系统给出当前的设备状态量,由系统控制中心输出调度策略。表1给出了用户某日20:30时的控制过程。对于未连接的设备(如电水壶、增湿器)而言,不对其进行控制,若用户在此时启动这些设备,会直接改变设备的连接状态,将其加入调度计划中。对于已连接的设备而言,由动作集合和状态集合共同决定其运行状态。表1中状态集合的取值0—3表征了设备的状态,取值为0表示计划外设备,取值为1表示运行中设备,取值为2表示可控设备,取值为3表示设备可延。

表1 20:30时的设备控制过程

Table 1 Device control process at 20:30

设备编号	用电设备	可控制设备集合	动作设备集合	状态设备集合	控制前运行设备集合	控制后运行设备集合
1	洗碗机	1	0	0	0	0
2	电饭煲	0	0	0	0	0
3	洗衣机	1	1	3	1	1
4	电水壶	0	0	0	0	0
5	增湿器	0	0	0	0	0
6	烘干机	0	0	0	0	0
7	热水器	1	1	2	1	1
8	电动汽车	1	0	2	0	0

在说明设备的调度过程后,给出采用优化调度策略前、后的对比结果,如图4所示。由图4可见,在06:00—08:00时段,由于电价上浮,采用调度策略后

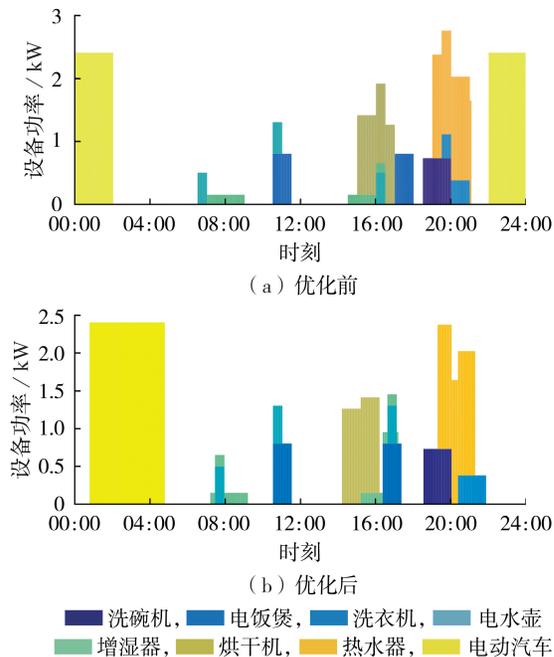


图4 优化前、后设备的调度策略对比

Fig.4 Comparison of equipment dispatch strategies between before and after optimization

电水壶和增湿器的启动时间向后推移,降低了用电费用;在11:00—13:00时段,虽然刚性负荷的用能增加,但光伏电量充足,使得负荷用电量不至于过高,即使在出现用户随机行为时也能够保证正常用电;在18:00—21:00时段,由于用电高峰导致电价波动较大,采用调度策略后使负荷运行时段往后适当推移,在满足用电需求的情况下最大限度地降低了用电费用;在00:00—04:00、22:00—24:00时段,电动汽车作为可控负荷,无出行计划,采用调度策略后其运行时间向午夜移动,进一步降低了用电费用。整体而言,优化调度策略能够在满足用电需求的同时给出合理的控制方案。

为了更为直观地反映调度策略的有效性,表2给出了5种家庭能量管理场景的优化结果。场景1—5分别代表未接入光伏的居民用电场景、未接入光伏的居民优化调度用电场景、接入光伏的居民用电场景、享受光伏上网补贴的居民用电场景、接入光伏的居民优化调度用电场景。由表2可知,在未进行优化调度的场景下用户的费用较高,进行优化调度后能够有效地减少用户费用;在接入家庭光伏后,优化调度策略可以在提高光伏消纳的同时,有效地减少用户费用,提高系统的优化率。

表2 5种场景的优化结果对比

Table 2 Comparison of optimal results among five scenes

场景	用户费用 / 元	上网收益 / 元	上网电量 / (kW·h)	优化率 / %
1	30.06	0	0	0
2	27.55	0	0	8.35
3	23.98	0	0	0
4	19.93	4.05	9.89	0
5	16.38	4.26	10.40	17.81

为了验证采用不同求解算法计算的多样性,给出了值迭代和 q 学习这2种迭代方法所求得的最优策略结果,如表3所示。由表3可以看出,2种迭代方法所求得的优化结果相差较小,但各有特点:值迭代方法能够给出偏向于贴合用户用电习惯的、较为稳定的控制方案,更加符合经济型用户的需求; q 学习方法能够给出更多变的策略集合,更多地探索动作空间,从而降低了用户的用能费用,更加符合体验型用户的需求。

表3 2种迭代方法所得优化结果对比

Table 3 Comparison of optimal results solved by two iterative methods

迭代方法	用能费用 / 元		优化率 / %	上网收益 / 元
	优化调度前	优化调度后		
值迭代	19.93	16.38	17.81	4.26
q 学习	19.93	17.22	13.60	4.20

为了验证存在多种不确定因素时系统的决策结果是否仍能够满足用户需求,给出了包含预测误差和随机用户行为情况下的优化效果,如表4所示。表中分析了光伏预测误差、实时电价预测误差、用户随机行为以及综合考虑上述不确定因素带来的影响。其中预测值均由传统预测方法给出,误差均在可接受的范围内。由最终结果可以看出,关于光伏预测误差的影响,在白天时段,正常的用户用电习惯下,用电负荷较少,而光伏系统的出力较为集中,因此较多的光伏出力会在白天时段回馈给电网,在光伏上网电价不变的情况下,光伏预测误差对整体调度策略结果的影响较小;实时电价预测误差和用户的随机行为会对整体的用电费用产生较大的影响,本文给出的调度策略能有效减少上述不确定因素变化所带来的影响。对于单个家庭用户而言,虽然优化前、后的整体误差相差不大,但是优化后长期带来的用电收益将会十分可观且会使用户获得更为舒适的用电体验。

表4 存在多种不确定因素时的优化结果对比

Table 4 Comparison of optimal results with many uncertainties

不确定因素	预测值 均值	实际值 均值	误差率/ %	误差 / %	
				均值 误差	平均绝 对误差
光伏预测误差 / kW	1.0191	0.8320	3.89	14.21	16.51
实时电价预测 误差 / 元	0.5676	0.6082	6.33	6.20	16.03
用户随机行为 / kW	0.5250	0.4545	5.11	15.51	22.49
综合上述3种因素	—	—	5.06	—	—

4 结论

本文基于泛在物联环境给出了家庭能量管理系统的框架,在此基础上,通过建立MDP模型构建了包含事件触发机制的家庭能量管理系统优化调度策略,并通过动态规划方法对调度策略进行迭代求解。仿真结果表明,本文所提方法对包含实时电价信息和用户随机行为的不确定性因素有很好的处理效果,能够在满足用户日常用电需求的基础上给出节省用电费用的用电设备最优控制方案。后续将进一步考虑电动汽车充放电能力和不确定因素,研究其与其他用电设备的协调优化运行策略。

附录见本刊网络版(<http://www.epae.cn>)。

参考文献:

[1] 杨挺,翟峰,赵英杰,等. 泛在电力物联网释义与研究展望[J]. 电力系统自动化,2019,43(13):9-20,53.
YANG Ting,ZHAI Feng,ZHAO Yingjie,et al. Explanation and prospect of ubiquitous electric power Internet of Things[J]. Automation of Electric Power Systems,2019,43(13):9-20,53.

[2] 傅质馨,李潇逸,袁越. 泛在电力物联网关键技术探讨[J]. 电

力建设,2019,40(5):1-12.

FU Zhixin,LI Xiaoyi,YUAN Yue. Research on key technologies of ubiquitous power Internet of Things[J]. Electric Power Construction,2019,40(5):1-12.

- [3] 张涛,张福兴,张彦. 面向能源互联网的能源管理系统研究[J]. 电网技术,2016,40(1):146-155.
ZHANG Tao,ZHANG Fuxing,ZHANG Yan. Study on energy management system of energy Internet[J]. Power System Technology,2016,40(1):146-155.
- [4] 曾鸣,杨雍琦,向红伟,等. 兼容需求侧资源的“源-网-荷-储”协调优化调度模型[J]. 电力自动化设备,2016,36(2):102-111.
ZENG Ming,YANG Yongqi,XIANG Hongwei,et al. Optimal dispatch model based on coordination between “generation-grid-load-energy storage” and demand-side resource[J]. Electric Power Automation Equipment,2016,36(2):102-111.
- [5] 高乾恒,黄帅飞,李二超,等. 市场环境下的含氢储能的售电公司优化调度[J]. 电力建设,2019,40(4):42-48.
GAO Qianheng,HUANG Shuaifei,LI Erchao,et al. Optimal dispatching of electricity retailers considering hydrogen storage in the electricity market[J]. Electric Power Construction,2019,40(4):42-48.
- [6] SHAREEF H,AHMED M S,MOHAMED A,et al. Review on home energy management system considering demand responses, smart technologies,and intelligent controllers[J]. IEEE Access,2018,6:24498-24509.
- [7] 王德志,张孝顺,余涛,等. 基于帕累托纳什均衡博弈的电网/多元家庭用户互动多目标优化算法[J]. 电力自动化设备,2017,37(5):114-121,128.
WANG Dezhi,ZHANG Xiaoshun,YU Tao,et al. Multi-objective optimization algorithm based on Pareto curve and Nash equilibrium game for interaction between grid and multi-home user[J]. Electric Power Automation Equipment,2017,37(5):114-121,128.
- [8] 周磊,李扬. 分时电价环境下基于家居能量管理系统的家居负荷建模与优化运行[J]. 电网技术,2015,39(2):367-374.
ZHOU Lei,LI Yang. Modeling and optimal dispatch for residential load based on home energy management system under time-of-use pricing[J]. Power System Technology,2015,39(2):367-374.
- [9] 葛少云,李吉峰,刘洪,等. 考虑物理特征与行为因素的家庭用能特性建模[J]. 电力自动化设备,2019,39(3):36-44.
GE Shaoyun,LI Jifeng,LIU Hong,et al. Modelling of household energy consumption characteristics considering physical features and behavior factors[J]. Electric Power Automation Equipment,2019,39(3):36-44.
- [10] YUCE B,REZGUI Y,MOURSHED M. ANN-GA smart appliance scheduling for optimised energy management in the domestic sector[J]. Energy and Buildings,2016,111:311-325.
- [11] 樊玮,刘念,张建华. 事件驱动的智能家庭在线能量管理算法[J]. 电工技术学报,2016,31(13):130-140.
FAN Wei,LIU Nian,ZHANG Jianhua. An online algorithm based on event triggering for energy management of smart home[J]. Transactions of China Electrotechnical Society,2016,31(13):130-140.
- [12] 曾博,蒋雯倩,杨舟,等. 基于机会约束规划的家庭用电设备负荷优化调度方法[J]. 电力自动化设备,2018,38(9):27-33.
ZENG Bo,JIANG Wenqian,YANG Zhou,et al. Optimal load dispatching method based on chance-constrained programming for household electrical equipment[J]. Electric Power Automation Equipment,2018,38(9):27-33.
- [13] CHEN Z,WU L,FU Y. Real-time price-based demand response management for residential appliances via stochastic optimi-

- zation and robust optimization[J]. IEEE Transactions on Smart Grid, 2012, 3(4): 1822-1831.
- [14] KILLIAN M, ZAUNER M, KOZEK M. Comprehensive smart home energy management system using mixed-integer quadratic programming[J]. Applied Energy, 2018, 222: 662-672.
- [15] 张彦, 张涛, 刘亚杰, 等. 基于模型预测控制的家居能源局域网最优能量管理研究[J]. 中国电机工程学报, 2015, 35(14): 3656-3666.
ZHANG Yan, ZHANG Tao, LIU Yajie, et al. Optimal energy management of a residential local energy network based on model predictive control[J]. Proceedings of the CSEE, 2015, 35(14): 3656-3666.
- [16] 张禹森, 孔祥玉, 孙博伟, 等. 基于电力需求响应的多时间尺度家庭能量管理优化策略[J]. 电网技术, 2018, 42(6): 1811-1819.
ZHANG Yusen, KONG Xiangyu, SUN Bowei, et al. Multi-time scale home energy management strategy based on electricity demand response[J]. Power System Technology, 2018, 42(6): 1811-1819.
- [17] KEERTHISINGHE C, CHAPMAN A C, VERBIC G. PV and demand models for a Markov decision process formulation of the home energy management problem[J]. IEEE Transactions on Industrial Electronics, 2019, 66(2): 1424-1433.
- [18] VIVEKANANTHAN C, MISHRA Y, LI F X. Real-time price based home energy management scheduler[J]. IEEE Transactions on Power Systems, 2015, 30(4): 2149-2159.
- [19] HANSEN T M, CHONG E K P, SURYANARAYANAN S, et al. A partially observable Markov decision process approach to residential home energy management[J]. IEEE Transactions on Smart Grid, 2018, 9(2): 1271-1281.
- [20] NISTOR M, ANTUNES C H. Integrated management of energy resources in residential buildings: a Markovian approach[J]. IEEE Transactions on Smart Grid, 2018, 9(1): 240-251.
- [21] ZHU Hongbo, GAO Yan, HOU Yong, et al. Real-time pricing considering type of smart home appliances based on Markov decision process[J]. International Journal of Electrical Power & Energy Systems, 2019, 107: 486-495.
- [22] AEMO. Price and demand data[EB/OL]. (2018-10-08)[2019-09-06]. <http://www.aemo.com.au/Electricity/National-Electricity-Market-NEM/Data-dashboard>.

作者简介:



傅质馨

傅质馨(1983—),女,江苏徐州人,副教授,博士,研究方向为物联网技术在可再生能源发电中的应用(**E-mail**: zhixinfu@hhu.edu.cn);

李潇逸(1995—),男,河南南阳人,硕士研究生,主要研究方向为基于物联网技术的能量管理策略(**E-mail**: 171306030033@hhu.edu.cn);

朱俊澎(1990—),男,江苏徐州人,讲师,博士,主要研究方向为主动配电网规划、运行与控制;

袁越(1966—),男,陕西西安人,教授,博士研究生导师,博士,主要研究方向为电力系统运行与分析、可再生能源发电技术等。

(编辑 陆丹)

Intelligent optimization strategy of home energy management based on Markov decision process

FU Zhixin^{1,2}, LI Xiaoyi¹, ZHU Junpeng^{1,2}, YUAN Yue^{1,2}

(1. College of Energy and Electrical Engineering, Hohai University, Nanjing 211100, China;

2. Research Center for Renewable Energy Generation Engineering of Ministry of Education, Hohai University, Nanjing 211100, China)

Abstract: Under the background of the rapid development of communication technology and the construction of Ubiquitous Power Internet of Things, the combination of multiple information interaction modes and artificial intelligence technologies can provide a new idea for improving the intelligence degree of home energy management. An intelligent optimization strategy of home energy management considering real-time information interaction is proposed. Firstly, the Markov decision process model based on user's energy consumption cost is given and the dynamic programming method is used to solve the model. The influence of uncertain factors, such as real-time electricity price information, user's random behavior, and so on, is considered in the scheduling process of household electrical equipment. On this basis, combining with the event trigger mechanism, the operation efficiency of home energy management system is effectively improved, and then the intelligent optimization scheduling method from the home energy management control center to the electrical equipment is proposed. Finally, the effectiveness of the proposed method is proved by a simulation example, and the results show that it can provide the optimal power consumption strategy to meet the user's demand while reducing the user's power consumption cost.

Key words: home energy management system; Markov decision process; stochastic dynamic planning; real-time electricity price; Ubiquitous Power Internet of Things; intelligent optimization

附录 A

MDP 模型的求解步骤具体如下。

控制策略是表征从状态集合到动作集合的映射，如式 (A1) 所示。

$$A_n = \pi_n(S_n) \quad (\text{A1})$$

值函数是状态集合到整体回报的映射，由此采用贝尔曼方程能够很好地表征当前状态的值函数与下一个状态的值函数之间的关系。式 (A2) 和式 (A3) 分别给出了贝尔曼期望公式和考虑状态动作对的贝尔曼期望公式。

$$v_\pi(s) = \sum_{a \in A} \pi(a|s) \left(R_s^a + \gamma \sum_{a' \in A} T_{ss'}^a v_\pi(s') \right) \quad (\text{A2})$$

$$q_\pi(s, a) = R_s^a + \gamma \sum_{s' \in S} T_{ss'}^a \sum_{a' \in A} \pi(a'|s') q_\pi(s', a') \quad (\text{A3})$$

由最优性原理：

$$v_* = \max v_\pi(s) \quad (\text{A4})$$

$$q_* = \max q_\pi(s, a) \quad (\text{A5})$$

给出最优贝尔曼期望方程：

$$v_*(s) = \max \left(R_s^a + \gamma \sum_{s' \in S} T_{ss'}^a v_*(s') \right) \quad (\text{A6})$$

$$q_*(s, a) = R_s^a + \gamma \sum_{s' \in S} T_{ss'}^a \max_{a'} q_*(s', a') \quad (\text{A7})$$

基于贝尔曼方程和贝尔曼最优性方程，即可采用动态规划求解 MDP 模型。本文分别采用值迭代 (式 (A8)) 和 q 学习 (式 (A9)) 迭代方法进行求解，从而得到最终的优化策略。

$$v_{k+1}(s) = \max \left(R_s^a + \gamma \sum_{s' \in S} T_{ss'}^a v_k(s') \right) \quad (\text{A8})$$

$$q_{k+1}(s, a) = q_k(s, a) + \alpha_q [R_s^a + \gamma \max_{a' \in A} q_k(s', a') - q_k(s, a)] \quad (\text{A9})$$

最终的优化策略分别由式 (A10)、式 (A11) 的最优方程给出。

$$\pi_* = \arg \max \left(R_s^a + \gamma \sum_{s' \in S} T_{ss'}^a v_k(s') \right) \quad (\text{A10})$$

$$\pi_q = \arg \max_{a \in A} q(s, a) \quad (\text{A11})$$

附录 B

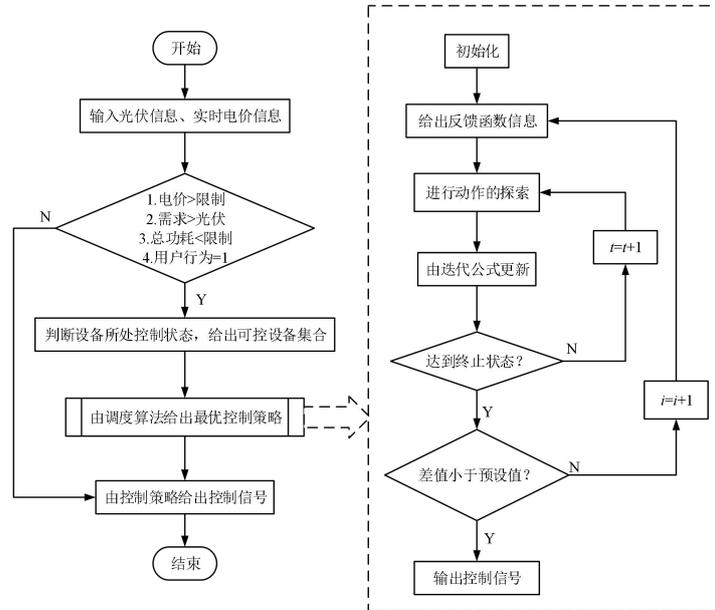


图 B1 控制流程图
Fig.B1 Control flowchart

附录 C

表 C1 设备的工作状态
Table C1 Operating status of devices

设备编号	用电设备	启动时间	停止时间	运行时间/h	设备功率/kW
1	洗碗机	18:30	21:00	1.5	0.73
2	电饭煲	10:00	12:00	1	0.8
		16:00	18:00		
3	洗衣机	19:00	22:00	1	0.38
		6:30	8:00		
4	电水壶	10:30	12:00	0.5	0.5
		16:00	17:30		
5	空气增湿器	7:00	10:00	2	0.15
6	烘干机	14:00	17:00	2	1.26
7	热水器	19:00	0:00	2	1.64
8	电动车	20:00	6:00	4	2.4

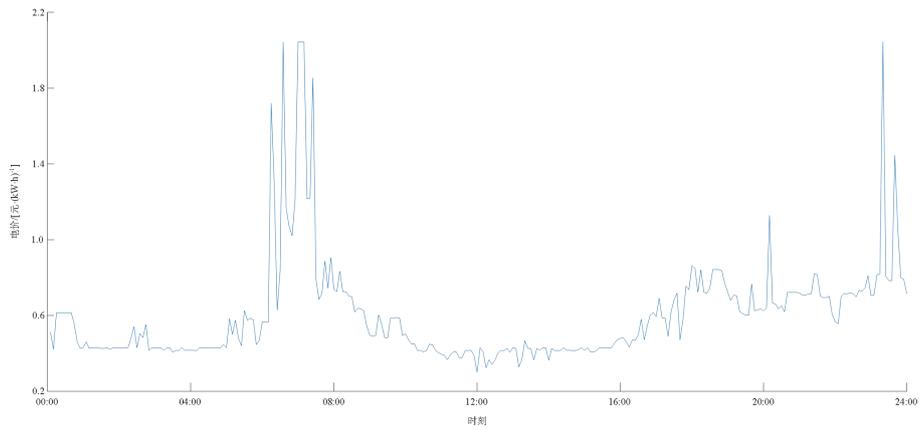


图 C1 实时电价
Fig.C1 Real-time price

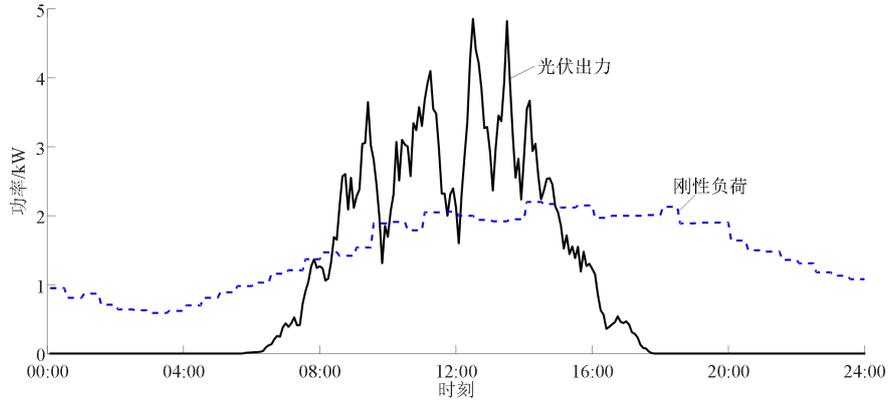


图 C2 光伏出力和基线负荷曲线
Fig.C2 Photovoltaic output and baseline load curves

表 C2 参数设定
Table C2 Parameter setting

参数	数值	参数	数值
α	0.5	整体迭代次数 N_{iter}	1000
β	0.5	q 学习方法的探索概率 ϵ	0.01
MDP 模型的折扣因子 γ	0.9	q 学习方法的学习率 α_q	0.1