

智能电网弹性响应时间业务需求的接入控制

江琦, 路改香, 唐昊, 谭琦

(合肥工业大学 电气与自动化工程学院, 合肥 230009)

摘要: 考虑智能电网多种类型业务需求的接入控制, 通过利用响应时间的弹性来平缓业务负荷的波动, 使得电网运行的长期平均代价最小. 针对业务需求和用户行为的随机分布特性, 建立连续时间 Markov 控制过程的系统分析模型; 结合性能势基于样本轨道的估计, 提出一种基于仿真的策略迭代优化算法, 有效缓解了系统大状态空间导致的维数灾问题, 具有较快的收敛速度和良好的应用效果. 仿真实验结果验证了所提出方法的有效性.

关键词: 智能电网; 需求负荷控制; Markov 控制过程; 策略迭代; 接入控制

中图分类号: TP202

文献标志码: A

Access control of demand requests with response time flexibility in smart grids

JIANG Qi, LU Gai-xiang, TANG Hao, TAN Qi

(School of Electrical Engineering and Automation, Hefei University of Technology, Hefei 230009, China.

Correspondent: JIANG Qi, E-mail: jiangqi@ustc.edu.cn)

Abstract: An access control strategy of multi-type service demands is proposed to minimize the long-run average cost of the operation in smart grids, which exploits the flexibility of response time to relieve the fluctuation of demand loads. By capturing the characters of stochastic distribution of service demands and users' behaviors, an analytical model based on continue-time Markov control processes is introduced. Combined with performance potentials estimate on sample path, a simulation based policy iteration algorithm is presented, which can alleviate the curse of dimensions induced by the large system state space and find a near-optimal policy with the rapid convergence rate. Simulation results show the effectiveness of the proposed method.

Key words: smart grid; demand load control; Markov control processes; policy iteration; access control

0 引言

智能电网以包含各种发电设备、输配电网络、用电设备和储能设备的物理电网为基础, 集成先进的传感测量、信息处理、网络通信与智能控制技术, 通过对全网资源和用户需求的实时观测、调度和配置, 实现系统运行、维护和管理动态优化, 以达到高效、经济、安全、可靠的运行目标.

需求负荷控制(DLC)是智能电网的基本功能构成, 旨在通过调节用户需求与电能供应相匹配, 提高电网运行的经济性和可靠性^[1]. 电网用电负荷在时间上具有显著的波动性. 在用电高峰期, 大量用户集中用电, 发电设备满负荷工作, 在电能供应不能满足用户需求时, 需要启动后备的辅助发电/储能逆变设备

或从外网输入电能来补充, 这便提高了电能供应的成本代价, 同时重负荷的输配电也对电网运行的可靠性带来不良影响; 在用电低谷期, 用户需求负荷较轻, 电网电能供应冗余造成较大的资源浪费, 在此状态下电网的单位电能供应成本也较高. DLC通过用户需求管理, 将高峰期的部分电能需求转移到非高峰期, 以平缓电网负荷的波动, 使得需求负荷与电能供应相匹配, 实现电网运行的经济性和可靠性, 同时也增大了电网容纳用户需求的能力.

DLC利用需求功率的可调节或需求响应时间的弹性来实现需求负荷与电能供应的匹配, 相关工作包括在用户端进行的用户参与决策^[2-4], 或从电网运行角度进行的集中控制^[5-6]. 文献[2]中运行者通过发布

收稿日期: 2013-04-12; 修回日期: 2013-09-24.

基金项目: 国家自然科学基金项目(61374158, 61233003, 61174186, 61074033); 国家国际科技合作项目(2011DFA10440); 安徽省自然科学基金项目(11040606M142).

作者简介: 江琦(1967-), 男, 副教授, 从事随机动态系统的控制与优化的研究; 路改香(1986-), 女, 硕士生, 从事智能电网的性能分析与优化的研究.

分时差异电价, 引导用户需求的选择性适时接入, 构成一个多用户参与的二级决策过程, 并提出一种启发式的注水算法求解最优的定价和接入策略. 文献[3]考虑未来一定时间段内分时电价已知情况下, 配有再生发电装置和储能设备用户的容许延迟响应需求的负荷控制, 提出一种基于 Lyapunov 函数的策略优化算法. 文献[4]基于实时电价的预测, 运用动态规划方法求解具有功率可调、时间弹性的电动汽车充放电的用户需求管理问题. 动态分时定价引导多用户参与决策的 DLC 受限于不同用户及用户与运行者之间协同的复杂性. 智能电网依托集成的高速双向通信网络, 使得集中式的实时 DLC 成为可能^[7]. 文献[5]中运行者通过控制大容量储能设备的充放电, 将其作为一个可逆的负荷来平缓业务需求的波动, 在稳态性能分析的基础上提出一种基于负荷阈值的充放电控制策略. 文献[6]考虑具有弹性延迟响应时间用户需求的接入控制问题, 采用一致化链的方法建立基于离散时间 Markov 决策过程的系统分析模型, 提出一种缓存队列长度相关的负荷阈值接入控制策略. 该研究仅考虑单种类型业务需求的接入, 基于即时接入与缓存请求对应状态的值函数之差关于电网负荷和缓存队列长度的严格单调性, 证明了基于负荷阈值的接入控制策略是最优接入控制策略, 但未能导出关于缓存队列长度的阈值函数的具体形式. 此分析方法和结论难以直接推广至多种类型业务需求的情况: 1) 上述单调性在多种类型业务需求情况下不复存在; 2) 对于多类型业务需求的情况, 关联队列长度的电网负荷阈值不能等效表示系统状态, 基于阈值的控制策略所构成的不再是等价于基于状态反馈的闭环控制, 不能达到最优的接入控制效果.

本文考虑多种类型的用户业务需求的接入控制问题, 不同类型业务的功率需求不同, 服务持续时间和容许响应延迟时间的分布参数也各异. 首先, 针对用户不同类型业务的请求到达、服务持续时间和容许响应延迟时间的随机分布特性, 采用连续时间 Markov 控制过程来描述智能电网 DLC 的动态特性, 建立系统性能分析与策略优化框架; 然后, 以电网运行的长期平均代价最小为目标, 提出一种策略迭代优化算法; 最后, 结合性能势基于样本轨道的估计, 提出该策略迭代算法基于仿真的实现. 本文算法有效地缓解了系统大状态空间导致的维数灾问题, 具有较快的收敛速度和良好的应用效果.

1 问题描述

考虑为多类具有弹性响应时间业务需求的用户提供服务的智能电网系统, 电网运行控制中心 (OCC) 和安装于用户端的智能测量装置 (SMD) 之间

通过双向通信实时传输需求和控制信息. 用户的用电需求通过 SMD 发送给 OCC, 各类业务请求到达的时间间隔、服务持续时间、最大延迟响应时间具有相互独立的随机性. 各类业务的随机特征可以用不同参数的概率分布来近似, 并由业务服务实测数据的统计分析获得.

OCC 接收 SMD 传送的用户请求, 根据当前电网的运行情况决定是否即时接入用电需求, 或推迟在需求的最大容许响应时间之前的某个时刻接入, 并通过 SMD 控制用户需求的接入, 如图 1 所示. 系统设置一个缓存区用以缓存未被即时接入的请求, 当系统中有服务完成导致负荷变化时, OCC 决定是否接入缓存中等待的服务请求. 若请求在缓存中等待时间达到其最大延迟响应时间, 则自动接入电网服务.

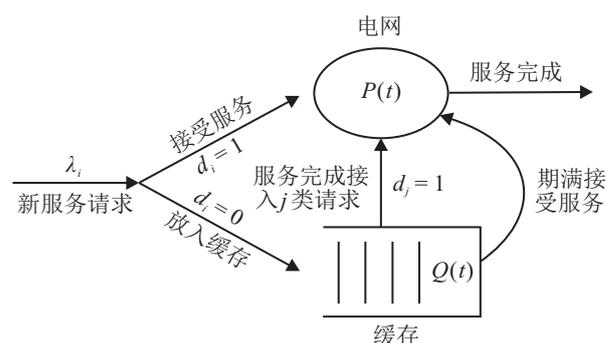


图 1 弹性响应时间业务需求的接入控制

鉴于在负荷较轻或较重时电网运行的经济性较差, 而以中等负荷运行电网的经济性良好, 电网运行的成本函数 $C(P)$ 是关于负载功率 P 的一个分段可微的凸函数, 如图 2 所示. 接入控制的实质是利用需求响应时间的弹性来平缓电网负荷的波动, 控制目标是使系统的长期运行成本最低.

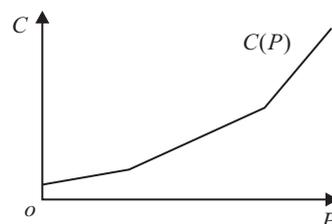


图 2 运行成本与负荷的函数关系

2 系统分析模型

假设电网用户的业务需求有 K 类, 第 $i \in I \triangleq \{1, 2, \dots, K\}$ 类需求的用电功率为 p_i , 其请求到达 OCC 服从参数为 λ_i 的 Poisson 分布, 用电持续时间满足均值为 $1/\mu_i$ 的指数分布, 最大容许延迟响应时间服从均值为 $1/\nu_i$ 的指数分布.

以电网中正在服务的各类业务的个数及缓存中各类业务请求的个数描述系统状态, 即 $s = (n_i,$

$m_i, i \in I$). 其中: n_i 表示电网中正在接受服务的第 i 类业务的个数, m_i 表示缓存中第 i 类业务请求的个数. 系统的状态空间为

$$S = \left\{ (n_i, m_i, i \in I) \mid \sum_{i \in I} p_i n_i \leq P_{\max}, \sum_{i \in I} m_i \leq M_{\max} \right\},$$

其中 P_{\max} 和 M_{\max} 分别表示电网的极限负荷和缓存的最大容量.

控制行动 $d = (d_i, i \in I)$, 其中 d_i 取 0 或 1: $d_i = 1$ 表示接入新到达或缓存的第 i 类业务请求; $d_i = 0$ 表示将新到达的第 i 类业务请求放入缓存中或不接入缓存中的第 i 类业务请求. 所有可行的行动构成行动集 D . 定义事件 e_i 为一个 i 类业务新请求到达, 事件 e'_i 为电网中一个 i 类业务完成服务离开, 事件集合 $E = \{e_i, e'_i, i \in I\}$.

控制行动由事件触发. 确定型策略 \mathcal{L} 是状态空间 S 与事件集合 E 至行动集 D 的映射, $\mathcal{L}: S \times E \rightarrow D$, 即 $\mathcal{L} = (\mathcal{L}_i(s, e), i \in I, s \in S, e \in E)$, 其中 $\mathcal{L}_i(s, e)$ 表示处于状态 s 发生事件 e 时选取的控制行动 d_i . 确定型策略空间记为 Π .

在策略 $\mathcal{L} \in \Pi$ 控制下系统的状态演化可以用一个连续时间 Markov 控制过程 $\{X_t, t \geq 0; S, A^{\mathcal{L}}\}$ 表示, 其中 $A^{\mathcal{L}}$ 为状态转移速率矩阵, 其元素

$$a_{ss'}^{\mathcal{L}} = \begin{cases} \lambda_i \mathcal{L}_i(s, e_i), s' = s^{(n_i+1)}; \\ \lambda_i I(\mathcal{L}_i(s, e_i) = 0), s' = s^{(m_i+1)}; \\ n_i \mu_i \mathcal{L}_j(s, e'_i), s' = s^{(n_i-1, n_j+1, m_j-1)}; \\ n_i \mu_i I(\mathcal{L}_j(s, e'_i) = 0), s' = s^{(n_i-1)}; \\ m_i \nu_i I(p_i + P(s) \leq P_{\max}), s' = s^{(n_i+1, m_i-1)}; \\ m_i \nu_i I(p_i + P(s) > P_{\max}), s' = s^{(m_i-1)}; \\ - \sum_{s' \neq s} a_{ss'}^{\mathcal{L}}, s' = s; \\ 0, \text{others.} \end{cases}$$

这里: $s, s' \in S$, $s^{(\cdot)}$ 表示其状态分量相对于 $s = (n_1, n_2, \dots, n_K, m_1, m_2, \dots, m_K)$ 发生括号标注的变化, $P(s) = \sum_{j=1}^K n_j p_j$, $I(\cdot)$ 为示性函数.

在策略 $\forall \mathcal{L} \in \Pi$ 下, 该 Markov 过程不可约, 其稳态概率 $q^{\mathcal{L}} = (q_s^{\mathcal{L}}, s \in S)$ 存在且唯一, 满足平衡方程

$$q^{\mathcal{L}} A^{\mathcal{L}} = 0, q^{\mathcal{L}} e = 1, A^{\mathcal{L}} e = 0,$$

其中 $e = (1, 1, \dots, 1)^T$.

设代价函数 $f: S \rightarrow R$, $f(s) = C[P(s)]$, 写成向量形式 $f = (f(s), s \in S)^T$. 系统的长期平均代价为

$$\eta^{\mathcal{L}} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T f(X_t) dt = q^{\mathcal{L}} f.$$

优化目标为寻求最优的控制策略 $\mathcal{L}^* \in \Pi$, 使得系统运行的长期平均代价最小, 即

$$\mathcal{L}^* \in \arg \min_{\mathcal{L} \in \Pi} \eta^{\mathcal{L}}.$$

3 策略优化算法

上述策略优化是一个无约束优化问题, 其最优策略可在确定型策略空间取得, 适用于采用策略迭代的方法进行策略寻优^[8].

设 $\{X_t, D_t, t \geq 0\}$ 是策略 \mathcal{L} 驱动下上述连续时间 Markov 控制过程的一条样本轨道, 对应于代价函数 f . 定义状态 $s \in S$ 的性能势^[9]为

$$g_s^{\mathcal{L}} = E \left[\int_0^{\infty} (f(X_t) - \eta^{\mathcal{L}}) dt \mid X_0 = s \right], \quad (1)$$

性能势向量 $g^{\mathcal{L}} = (g_s^{\mathcal{L}}, s \in S)^T$ 满足 Poisson 方程

$$-f + \eta^{\mathcal{L}} e = A^{\mathcal{L}} g^{\mathcal{L}}. \quad (2)$$

由式 (2) 可得性能势的理论计算公式

$$g^{\mathcal{L}} = -((A^{\mathcal{L}} - eq)^{-1} - eq)f. \quad (3)$$

设 $A^{\mathcal{L}'}$ 为策略 $\mathcal{L}' \in \Pi$ 诱导的另一个转移速率矩阵; $q^{\mathcal{L}'}, \eta^{\mathcal{L}'}$ 为对应的稳态概率和平均代价测度. 将式 (2) 两边同时乘以 $q^{\mathcal{L}'}$, 并应用 $q^{\mathcal{L}'} A^{\mathcal{L}'} = 0$ 得到性能差公式

$$\eta^{\mathcal{L}'} - \eta^{\mathcal{L}} = q^{\mathcal{L}'} ((A^{\mathcal{L}'} - A^{\mathcal{L}}) g^{\mathcal{L}} + (f^{\mathcal{L}'} - f^{\mathcal{L}})). \quad (4)$$

由性能差公式, 基于一个策略下的性能势比较两个策略的优劣, 有如下引理.

引理 1 比较定理和最优不等式.

1) 如果 $A^{\mathcal{L}'} g^{\mathcal{L}} + f^{\mathcal{L}'} \preceq A^{\mathcal{L}} g^{\mathcal{L}} + f^{\mathcal{L}}$ (符号 \preceq 表示各分量不大于且至少有一个分量小于), 则有 $\eta^{\mathcal{L}'} < \eta^{\mathcal{L}}$.

2) 当且仅当 $\forall \mathcal{L}' \in \Pi$ 时, 策略 \mathcal{L} 是最优策略, 有

$$A^{\mathcal{L}} g^{\mathcal{L}} + f^{\mathcal{L}} \leq A^{\mathcal{L}'} g^{\mathcal{L}} + f^{\mathcal{L}'}. \quad (5)$$

基于引理 1, 有如下策略迭代算法.

算法 1 基于数值计算的策略迭代算法.

Step 1: 给定初始策略 \mathcal{L}^0 , 置 $k = 0$.

Step 2 (策略评估): 根据式 (3) 计算 $g^{\mathcal{L}^k}$.

Step 3 (策略改进): 逐个分量选择

$$\mathcal{L}^{k+1} = \arg \min_{\mathcal{L} \in \Pi} (A^{\mathcal{L}} g^{\mathcal{L}^k} + f^{\mathcal{L}}). \quad (6)$$

若在状态 $s \in S$, 行动 $\mathcal{L}^k(s, e)$ 取到式 (6) 最小, 则置 $\mathcal{L}^{k+1}(s, e) = \mathcal{L}^k(s, e)$.

Step 4: 若 $\mathcal{L}^{k+1} = \mathcal{L}^k$, 则算法停止; 否则, 置 $k = k + 1$, 转入 Step 2.

由 Step 3, 如果 $\mathcal{L}^{k+1} \neq \mathcal{L}^k$, 则

$$f^{\mathcal{L}^{k+1}} + A^{\mathcal{L}^{k+1}} g^{\mathcal{L}^k} \preceq f^{\mathcal{L}^k} + A^{\mathcal{L}^k} g^{\mathcal{L}^k}.$$

根据引理 1, 则有 $\eta^{\mathcal{L}^{k+1}} \leq \eta^{\mathcal{L}^k}$, 即每一次策略迭代, 代价都有减少. 由于策略空间有限, 迭代将在有限步内停止. 当迭代在第 k 步停止时, 设 $\mathcal{L}^* = \mathcal{L}^k$, 由式 (6) 有

$$\mathcal{L}^* = \arg \min_{\mathcal{L} \in \mathcal{H}} (f^{\mathcal{L}} + A^{\mathcal{L}} g^{\mathcal{L}^*}),$$

或者

$$A^{\mathcal{L}^*} g^{\mathcal{L}^*} + f^{\mathcal{L}^*} \leq A^{\mathcal{L}} g^{\mathcal{L}^*} + f^{\mathcal{L}}.$$

由引理 1 可知, 策略 \mathcal{L}^* 是最优策略.

上述算法中, 策略评估步骤中的性能势数值计算涉及矩阵求逆运算. 随着系统规模的扩大, 转移速率矩阵的阶数急剧增长, 求逆运算带来很大的计算负担, 可能由于计算机硬件限制难以得出计算结果. 为了解决系统大状态空间导致的维数灾问题, 可设计基于仿真的策略迭代算法, 其主要思想是: 在每次迭代的策略评估步骤, 仿真系统在当前策略 \mathcal{L}^k 下运行一条样本轨道, 并基于样本轨道的状态转移信息计算性能势 $g^{\mathcal{L}^k}$ 的估计值 $\hat{g}^{\mathcal{L}^k}$; 在策略改进步骤中, 以 $\hat{g}^{\mathcal{L}^k}$ 替代 $g^{\mathcal{L}^k}$ 来计算改进策略. 具体算法如下:

算法 2 基于仿真的策略迭代算法.

Step 1: 选择一个较大的整数 N 和初始策略 \mathcal{L}^0 , 置 $k = 0$.

Step 2 (策略评估): 仿真在策略 \mathcal{L}^k 控制下的系统运行, 得到一条具有 N 次状态转移的样本轨道 $\{X_t, 0 \leq t \leq T_N\}$, 根据式 (7) 计算性能势估计值

$$\begin{cases} \hat{g}_s^{\mathcal{L}^k} = \int_{t_s}^{T_N} (f(X_t) - \hat{\eta}^{\mathcal{L}^k}) dt, \\ \hat{\eta}^{\mathcal{L}^k} = \frac{1}{T_N} \int_0^{T_N} f(X_t) dt. \end{cases} \quad (7)$$

其中: T_N 是样本轨道第 N 状态转移的时间, t_s 是样本轨道第一次转移到状态 s 的时间.

Step 3 (策略改进): 逐个分量选择

$$\mathcal{L}^{k+1} = \arg \min_{\mathcal{L} \in \mathcal{H}} (f^{\mathcal{L}} + A^{\mathcal{L}} \hat{g}^{\mathcal{L}^k}). \quad (8)$$

若在状态 $s \in S$, 行动 $\mathcal{L}^k(s, e), e \in E$ 取到式 (8) 最小, 则置 $\mathcal{L}^{k+1}(s, e) = \mathcal{L}^k(s, e), e \in E$.

Step 4: 若 $\mathcal{L}^{k+1} = \mathcal{L}^k$, 则算法停止; 否则, 置 $k = k + 1$, 转入 Step 2.

基于仿真的策略迭代的收敛性既依赖于性能势的估计误差, 又依赖于应用于估计的样本轨道长度. 算法 2 采用固定数量 N 次的状态转移来估计性能势, 当 N 足够大时, 性能势的估计值与真值间的误差满足一定的条件, 可以保证该算法是收敛的.

4 仿真结果

通过数值仿真实验评估算法的收敛性和有效性. 仿真参数选取如下: 电网的极限负荷 $P_{\max} = 4$, 最大缓存容量 $M_{\max} = 4$, 为两种类型的业务需求提供服务 ($K = 2$), 各类业务需求的具体参数如表 1 所示.

表 1 仿真参数

业务类型 i	p_i	μ_i	ν_i
1	1	2	4
2	2	3	8

电网的成本函数 $C(P)$ 是关于功率负载 P 的一个分段凸函数, 表示为

$$C(P) = \begin{cases} 30P + 40, & 0 \leq P \leq 2; \\ 60(P - 2) + 100, & 2 < P \leq 3; \\ 100(P - 3) + 160, & P > 3. \end{cases}$$

用不同的业务需求到达率模拟不同的应用环境, 评估算法 1 和算法 2 的收敛情况. 实验中, 基于理论计算的策略迭代算法以较快的速度收敛到最优控制策略 (通常 $k < 10$). 而基于仿真的策略迭代由于样本轨道 N 的有限取值 (实验中取 $N = 10\,000$), 性能势的估计存在一定的随机误差, 表现在算法 2 以较快速度收敛于最优策略附近之后有一个小幅振荡过程. 这时可以通过附加一个停止准则, 使得算法产生一个具有满意性能的近似最优策略. 也可以通过增大 N 的取值来提高性能势估计的精度, 以抑制振荡使得算法收敛到最优策略, 这会增加仿真计算量. 实际应用中, N 值需权衡计算代价与优化性能进行折中选取. 图 3 和图 4 为在两种不同的需求到达率下, 算法 1 和算法 2 的收敛过程.

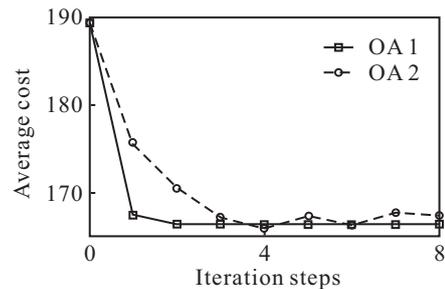


图 3 算法 1 (OA1) 和算法 2 (OA2) 的执行情况 ($\lambda_1 = 5, \lambda_2 = 3$)

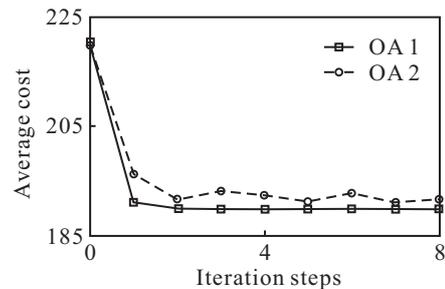


图 4 算法 1 (OA1) 和算法 2 (OA2) 的执行情况 ($\lambda_1 = 8, \lambda_2 = 6$)

通过比较实验评估算法的有效性. 图 5 为在不同业务需求到达率下 ($\lambda = \lambda_1 + \lambda_2, \lambda_1/\lambda_2 = 4/3$), 算法 2 产生的最优策略 (OA2)、即时接入控制策略 (AAI) 和基于阈值控制策略 (SHB)^[6] 的应用效果. 即时接入控制策略是当业务需求到达时即予以接入. 基于

阈值策略是设定一个与缓存队列长度相关的负荷阈值, 当电网实际负荷大于此阈值时, 不接入到达的业务需求; 反之则接入一个新的业务需求或激活一个缓存的业务请求. 通过比较表明, 算法2产生策略的应用效果优于基于阈值的接入策略, 而即时接入策略的性能最低. 这是由于算法2产生的是一种状态相关的最优策略, 基于阈值的策略在多类型业务的应用中不再等同于基于状态的策略(此时相同的缓存队列长度和电网负荷可对应多个不同的系统状态), 而即时接入策略实质上不具有利用响应时间弹性进行电网负荷调节的功能.

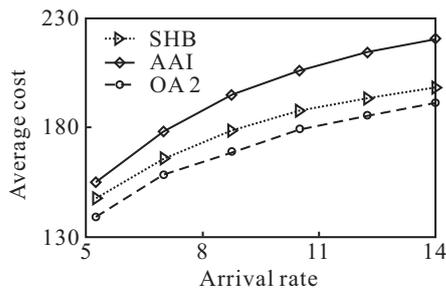


图5 应用效果比较

5 结 论

本文考虑智能电网多种类型业务需求的接入控制问题, 利用延迟响应时间的弹性来平缓需求负荷的高波动, 以达成电网的业务需求与电能供应的有效匹配. 针对业务需求和用户行为的随机分布特性, 建立基于连续时间 Markov 控制过程的系统性能分析模型, 进而提出基于仿真的策略迭代优化算法. 该算法有效地缓解了系统大状态空间带来的维数灾问题, 能以较快速度收敛到近似理论最优值. 后续工作将进一步考虑在系统参数未知情况下接入控制策略的在线自适应优化问题.

参考文献(References)

- [1] Moslehi K, Kumar R. A reliability perspective of the smart grid[J]. IEEE Trans on Smart Grid, 2010, 1(1): 57-64.
- [2] Gkatzikis L, Salonidis T, Hegde N, et al. Electricity markets meet the home through demand response[C]. Proc of the 51st IEEE Conf on Decision & Control. Maui HI, 2012: 5846-5851.
- [3] Chen S, Sinha P, Shroff N B. Scheduling heterogeneous delay tolerant tasks in smart grid with renewable energy[C]. Proc of the 51st IEEE Conf on Decision & Control. Maui HI, 2012: 1130-1135.
- [4] Rotering N, Ilic M. Optimal charge control of plug-in hybrid electric vehicles in deregulated electricity markets[J]. IEEE Trans on Power Systems, 2011, 26(3): 1021-1029.
- [5] Koutsooulos I, Hatzi V, Tassiulas L. Optimal energy storage control policies for the smart power grid[C]. Proc of the 2011 IEEE Int Conf on Smart Grid Communications. Brussels, 2011: 475-480.
- [6] Koutsooulos I, Tassiulas L. Optimal control policies for power demand scheduling in the smart grid[J]. IEEE J on Selected Areas in Communications, 2012, 30(6): 1049-1060.
- [7] Gungor V C, Sahin D, Kocak T. A survey on smart grid potential applications and communication requirements[J]. IEEE Trans on Industrial Informatics, 2013, 9(1): 28-42.
- [8] Puterman M L. Markov decision processes: Discrete stochastic dynamic programming[M]. New York: John Wiley & Sons, 1994: 377-385.
- [9] Cao X R. The potential structure of sample paths and performance sensitivities of Markov systems[J]. IEEE Trans on Automatic Control, 2004, 49(12): 2129-2142.

(责任编辑: 齐 霖)