考虑市场因素的电力系统供需互动混合博弈强化学习算法

包 涛1, 李昊飞2, 余 涛2†, 张孝顺3

(1. 广东电网责任有限公司 广州供电局, 广东 广州 510620; 2. 华南理工大学 电力学院, 广东 广州 510640;

3. 汕头大学 工学院, 广东 汕头 515063)

摘要:为对电力市场环境下电力系统供需互动问题更精确地建模,使其更好地与未来电力市场环境下需求侧负 荷聚合商之间多变的关系和复杂的通信拓扑结构相匹配,本文将电力系统供需互动的Stackelberg博弈与复杂网络 上反映需求侧负荷聚合商互动的演化博弈相结合,搭建考虑市场因素的电力系统供需互动混合博弈模型.并提出 混合博弈强化学习算法求解相应的非凸非连续优化问题,该算法以Q学习为载体,通过引入博弈论和图论的思想, 把分块协同和演化博弈的方法相结合,充分地利用博弈者之间互动博弈关系所形成的知识矩阵信息,高质量地求解 考虑复杂网络上多智能体系统的非凸优化问题.基于复杂网络理论搭建的四类3机-6负荷系统和南方某一线城市电 网的仿真结果表明:混合博弈强化学习算法的寻优性能比大多数集中式的智能算法好,且在不同网络下均可以保 证较好的寻优结果,具有很强的适应性和稳定性.

关键词: 混合博弈强化学习算法; 供需互动; Stackelberg博弈; 演化博弈; 复杂网络

引用格式:包涛,李昊飞,余涛,等.考虑市场因素的电力系统供需互动混合博弈强化学习算法.控制理论与应用, 2020, 37(4):907 – 917

DOI: 10.7641/CTA.2019.80814

Mixed game reinforcement learning of supply-demand interaction in power system dis-patch on electricity market

BAO Tao¹, LI Hao-fei², YU Tao^{2†}, ZHANG Xiao-shun³

(1. Guangzhou Power Supply Bureau of Guangdong Power Grid Co., Ltd, Guangzhou Guangdong 510620, China;

2. College of Electric Power, South China University of Technology, Guangzhou Guangdong 510640, China;

3. College of Engineering, Shantou China, Shantou Guangdong 515063, China)

Abstract: In order to solve the supply and demand interaction problem in electricity market more accurately, this paper builds a mixed game model of supply and demand interaction in power system considering electricity market factors, and proposes a mixed game reinforcement learning algorithm. Considering the ideas of game theory and graph theory, the algorithm combines block cooperation and evolutionary game methods to fully utilize the interaction of knowledge matrix information formed by interactive game relationships between players based on Q-learning. The corresponding non-convex optimization problem under complex networks can be solved efficiently. Finally, the simulation results of two test systems indicate that the optimization performance of the mixed game reinforcement learning algorithm is better than that of most centralized intelligent algorithms. Comparing with the existing center-based algorithms, this mixed game reinforcement learning algorithm has better search results, strong adaptability and stability under different networks.

Key words: mixed game reinforcement learning; supply and demand interaction; Stackelberg game; evolutionary game; complex network

Citation: BAO Tao, LI Haofei, YU Tao, et al. Mixed game reinforcement learning of supply-demand interaction in power system dis-patch on electricity market. *Control Theory & Applications*, 2020, 37(4): 907 – 917

1 引言

在电力系统的供需平衡中,需求侧通常只是供给 侧电价的被动接收者,其主动作用经常被忽视.近年 来,智能电网、能源互联网、综合能源系统等概念的提 出和大量分布式电源、电动汽车等新兴事物的涌现, 为需求侧参与供需互动提供了强大的技术支持^[1-3]. 目前综合考虑供给侧的经济调度和需求侧响应的电 力系统供需互动问题也受到了国内外学者的广泛重

收稿日期: 2018-10-22; 录用日期: 2019-09-09.

[†]通信作者. E-mail: taoyu1@scut.edu.cn; Tel.: +86 13002088518.

本文责任编委: 梅生伟.

国家自然科学基金项目(51477055)资助.

Supported by the National Natural Science Foundation of China (51477055).

视^[4-7]. 伴随着电力市场改革的不断深入, 多方利益实体的竞争化模式使其不断地寻求自身利润的最大化, 而需求侧的负荷聚合商之间有合作也有竞争, 信息网络拓扑复杂多变. 电力市场参与主体逐渐增多, 电力系统供需互动问题也日益复杂, 因此需要提出新的方法对其实现精确建模和高效求解.

近年来,越来越多的专家学者认识到博弈论在解 决拥有多主体特征的电力系统供需互动问题中的高 效性[8]. 文献[9]利用主从博弈的思想, 搭建电动汽车 车主与极大化的小区代理商的互动博弈模型. 文 献[10]研究了电力零售商在现货市场、双边合同市场 的购电策略,建立了发电商与大用户之间的主从博弈 模型. 文献[11]针对含有多个售电商的智能住宅小区, 建立了售电商和用户之间的主从博弈模型.因此本文 也采用Stackelberg博弈的框架建立考虑市场因素的供 需互动模型. 但现有研究一般认为需求侧博弈主体的 关系是全耦合或完全随机的,但实际并非如此,因此 如何描述需求侧主体之间的不同博弈关系也是一大 难点.目前复杂网络被广泛应用于电网问题的研究, 其可将主体的相互作用关系转换为网络拓扑的描述 方式[12-17]. 但目前鲜少有文献利用复杂网络探究需 求侧个体的相互作用.因此,本文将复杂网络与演化 博弈结合用于描述需求侧不同主体的互动关系.

对于此类供给侧与需求侧互动问题的求解方法有 解析法和智能算法两种. 解析法通常是将需求侧的优 化问题转化为等价的KKT条件^[9-10],从而将双层优化 问题转化为单层优化问题. 但解析法对于模型的要求 较高,求解非凸非连续的模型十分困难.而常用的智 能算法例如遗传算法(genetic algorithm, GA)、粒子群 搜索算法(particle swarm optimization, PSO)和人工蜂 群算法(artificial bee colony, ABC)等均为集中式算法, 无法保证互动主体信息的私密性.并且随着电力系统 供给侧经济调度考虑的约束逐渐复杂,且在供需互动 中引入市场价格后,优化问题的可行域缩小,亟待提 出一种新的分布式优化算法用于求解考虑市场因素 的电力系统供需互动问题. 而强化学习具有收敛性好、 收敛速度快和对模型的依赖性低的优势,并且能够存 储学习的知识[18],本文正是利用该类算法的优点实现 对非凸非连续的电力系统供需互动问题的求解.

基于上述理论研究,本文将电力系统供需互动的 Stackelberg博弈与复杂网络上反映需求侧负荷聚合商 互动的演化博弈相结合,搭建基于博弈论的电力系统 供需互动混合博弈模型.并提出了全新的分布式算 法—混合博弈强化学习算法(mixed game reinforcement learning, MGRL)对含有需求侧复杂网络的供需互 动模型进行求解,该方法可以有效求解非凸非连续优 化问题,并适用于分布式优化框架,保证信息私密性. 最后,以3机-6负荷系统和南方某一线城市为算例,对 本文所提模型及算法进行验证.

2 电力系统供需互动模型

2.1 考虑市场因素的供需互动模型

2.1.1 供给侧模型

在供给侧, 描述发电机的最经典数学模型是线性 函数和二次函数, 模型较为简单, 且凸函数的求解难 度较低, 但此类模型均没有考虑阀点效应, 导致了模 型的精度受到了影响. 本文考虑发电机组的"阀点效 应", 在二次煤耗特性曲线上叠加正弦修正分量^[19], 故发电成本函数变为

$$f_{i}^{s}(P_{Gi}) = a_{i}P_{Gi}^{2} + b_{i}P_{Gi} + c_{i} + |d_{i}\sin(e_{i}(P_{Gi}^{\min} - P_{Gi}))|, \qquad (1)$$

式中: $a_i, b_i, c_i, d_i n e_i$ 是模型中第i台发电机的煤耗特性参数; P_{Gi} 是第i台发电机的输出功率; P_{Gi}^{\min} 为第i台发电机输出功率下限.

约束条件方面, 输出功率应在禁止运行区域(prohibited operating zones, POZs)^[20]之外, P_{Gi} 应满足下述 不等式约束:

$$\begin{cases} P_{\text{G}in}^{\min} \leqslant P_{\text{G}i} \leqslant P_{\text{G}i,1}^{1}, \\ P_{\text{G}i,z-1}^{\text{u}} \leqslant P_{\text{G}i} \leqslant P_{\text{G}i,z}^{1}, z = 2, \cdots, Z_{i}, \\ P_{\text{G}i,Z_{i}}^{\text{u}} \leqslant P_{\text{G}i} \leqslant P_{\text{G}i}^{\max}, \quad \forall i \in \Psi, \end{cases}$$

$$(2)$$

式中: Ψ 是考虑POZs的发电机集合; $P_{G_{i,z}}^1 和 P_{G_{i,z}}^u \beta$ 别是第i台发电机第z个POZs的下限和上限; Z_i 是第i台发电机的POZs数量; $P_{G_i}^{\max}$ 为第i台发电机有功功率上限.

本文考虑市场因素作用,引入发电商报价,供给侧 的效益函数为式(3).发电商为追求其自身利益最大 化,将调整其输出功率直到边际成本等于发电商报价, 如式(4)所示:

$$\max U_{s,i} = r_i P_{Gi} - f_i^{s}(P_{Gi}), \qquad (3)$$

$$\frac{\mathrm{d}U_{s,i}}{\mathrm{d}P_{\mathrm{G}i}} = r_i - \frac{\mathrm{d}f_i^{\mathrm{s}}(P_{\mathrm{G}i})}{\mathrm{d}P_{\mathrm{G}i}} = 0, \tag{4}$$

式中: $U_{s,i}$ 为供给侧效益函数; r_i 为发电商报价; $f_i^s(P_{Gi})$ 即为式(1)中发电机i的发电成本.

2.1.2 需求侧模型

在需求侧,引入负荷聚合商,作为一个整体来整合 与调整区域内的负荷,通过评估该区域可协调负荷的 最大值与最小值,参与到系统的供需互动优化中.为 了与发电侧的发电成本函数相匹配,本文采用了同样 能够以金钱来衡量的效用函数来表示需求侧用电客 户消耗电量后所获取的满意度水平,数学表达式如下:

$$f_j^{\rm d}(P_{\rm Dj}) = \frac{1}{2}\alpha_j P_{\rm Dj}^2 + \omega_j P_{\rm Dj},\tag{5}$$

式中: P_{D_j} 是第j个负荷聚合商有功功率; α_j 和 ω_j 是模型中第j个柔性负荷效用参数.

在电力市场下,引入需求侧的用户报价,则用户侧 的效益变为式(6).用户为了追求其用电满意度达到最 大值,将持续不断购电.按照微观经济学理论,最大值 发生在边际效用等于用户报价的时候,该模式可用式 (7)来描述.

$$\max U_{d,j} = f_j^{\rm d}(P_{{\rm D}j}) - r_j P_{{\rm D}j}, \tag{6}$$

$$\frac{\mathrm{d}U_{d,i}}{\mathrm{d}P_{\mathrm{D}j}} = \frac{\mathrm{d}f_j^{\mathrm{d}}(P_{\mathrm{D}j})}{\mathrm{d}P_{\mathrm{D}j}} - r_j = 0, \tag{7}$$

式中: $U_{d,i}$ 为需求侧效益函数; r_j 是负荷聚合商j的用 户报价; $f_i^d(P_{Di})$ 为负荷聚合商j的效用函数.

2.1.3 社会效益模型

电力市场调度模式下,在供给侧和需求侧引入电价信息,整个电力系统供需互动采用社会效益最大化作为目标函数,则模型可表示为

$$\max \sum_{i \in \Omega_{s}} U_{s,i} + \sum_{j \in \Omega_{d}} U_{d,j}$$

s.t.
$$\begin{cases} \vec{\pi}(2), \\ P_{\mathrm{D}j}^{\min} \leqslant P_{\mathrm{D}j} \leqslant P_{\mathrm{D}j}^{\max}, \ j = 1, 2, \cdots, N_{\mathrm{D}}, \\ \sum_{i=1}^{N_{\mathrm{G}}} P_{\mathrm{G},i} - \sum_{j=1}^{N_{\mathrm{D}}} P_{\mathrm{D},j} = 0, \end{cases}$$

(8)

式中 $P_{D_j}^{\max}$ 和 $P_{D_j}^{\min}$ 为第j个负荷聚合商有功功率上下限.

若发电商报价r_i固定不变,发电机i将改变其输出 功率,使得其边际成本与r_i相等.类似地,用户也将持 续购电,使得其边际效用与用户报价r_j相等.同时,必 须根据功率约束和供需平衡约束条件来调整r_i和r_j. 最终,r_i和r_i将相等并固定下来.该均衡电价为^[21]

$$r_i \!=\! r_j \!=\! r^*, \tag{9}$$

式中: $i=1,2,\cdots,N_{\rm G}, j=1,2,\cdots,N_{\rm D}, r^*$ 为当供需 平衡时市场的出清电价.

2.2 复杂网络下的需求侧演化博弈模型

2.2.1 演化博弈

演化博弈是在经典博弈理论的基础上提出来的, 本文所描述的博弈者均是指理性的智能体Agent.在 博弈的过程中, Agent会以一定的概率学习其他Agent 的策略. 一般来说, 适应度越大, 被模仿学习的机会就 越多^[22]. 本文采用Fermi-Dirac函数^[23]来制定博弈者 之间的模仿学习规则. 用 $\Omega_i(s)(\Omega_i(s) \in \Omega_i)$ 代表Agent *i*的邻居中采用策略*s*的Agent, 则Agent *i*的策略从*s*变 为*s*'的概率:

$$\eta_i(s \to s') = \frac{\varphi}{|\boldsymbol{\Omega}_i|} \sum_{j \in \boldsymbol{\Omega}_i(s')} \frac{1}{1 + \exp(-\frac{(f_j - f_i)}{T})},$$
(10)

式中T代表了博弈者的理性程度,是一个常数.

2.2.2 复杂网络

复杂网络描述的是自然、社会和人工系统等中的 复杂连接关系.不同的统计特征表示不同的节点及节 点间连接特性和网络结构.根据不同平均路径长度、 聚类系数、度和度分布可搭建出最经典的4复杂网络, 分别是规则网络、随机网络、小世界网络和无标度网 络,如图1所示.



图 1 4种典型复杂网络结构

Fig. 1 Four networks structures

2.2.3 需求侧复杂网络演化博弈

在复杂网络的演化博弈中,节点表示Agent,边表示节点间的博弈关系.在每次迭代过程中,Agent根据边的关系与邻居展开博弈,再根据模仿规则更新其策略.

在本文讨论的电力系统中,以网络拓扑无向图*G* 来代表负荷聚合商间的博弈关系或通过通讯线路进 行信息交互的路径. 定义*G* = (*V*, *E*, *A*), *V* = { v_1 , v_2 , …, v_n }为顶点集合,为有限的非空集合,也称为顶点 集. *E*为*V*中元素构成的无序二元组的集合,称为边 集. 图*G*的顶点代表博弈者,边代表博弈者间的信息 传递关系和博弈关系. *A* = [a_{ih}] $\in \mathbb{R}^{NB \times NB}$ 为该图 的邻接矩阵, N_B 表示系统中博弈者的个数. 元素 a_{ih} 表示顶点i, h之间的边的权重. 如果顶点i能收到顶点 h的信息,则称顶点h为顶点i的邻居. N_i 表示顶点i的 邻居的集合(包括顶点i), 基数 d_i = $|N_i|$. 本文的分块 是指,与节点 v_i 相连接的节点构成的一个集合,表示 为 Ω_i = { $v_h \in V$ |(v_i , v_h) $\in E$ }, v_h 称为节点 v_i 的邻 居. 假设通信拓扑图的边满足双向并且等权重, 定义 邻接矩阵A的元素 a_{ih} 为

$$a_{ih} = \begin{cases} \frac{1}{d_i}, \ h \in N_i, \\ 0, \ h \notin N_i. \end{cases}$$
(11)

2.3 电力系统供需互动混合博弈模型

本文搭建的混合博弈模型,上层是以Stackelberg 博弈为框架,下层需求侧内部参与复杂网络上的演化 博弈,如图2所示.





Fig. 2 Interaction framework of supply-demand in power system

在模型的上层,发电机组与负荷聚合商在领导 者平衡机组的带领下,参与电力系统供需互动Stackelberg博弈.平衡机组一般会挑选调节范围较大的 发电机组,它可以作为主从博弈中的Leader,在迭代 过程中,与其余Agent交换功率信息,并只选取最优 策略.在模型的下层的需求侧,因为复杂网络的存 在,负荷聚合商将参与分块协同演化博弈,并通过 MGRL算法得到最优策略.

每个Follower在观察到Leader和其余Follower的 当前最优策略后,将追求全局最优来改变自身策略:

$$\max \left[U_{\text{leader}}(\bar{\boldsymbol{x}}^{*}), \ U_{\text{leader}}(\bar{\boldsymbol{x}}^{*}_{\text{best}}) \right]$$

s.t.
$$\begin{cases} U_{\text{leader}}(\bar{x}^{*}) = \sum_{i=1}^{n} U_{i}(\bar{x}^{*}), \\ \bar{x}^{*} = (\bar{x}_{1}^{*} \cdots \bar{x}_{i}^{*} \cdots \bar{x}_{n}^{*}), \\ \bar{x}_{i}^{*} = \arg \max_{x_{i} \in A_{i}} U_{i}(\bar{x}_{p}^{*}, x_{i}), \end{cases}$$
(12)

式中: $i=1,2,\dots,n,p$ 为领导者的可选策略; U_{leader} 为领导者的社会效益函数; U_i 为第i个跟随者的社会 效益函数; \bar{X}_p^* 为领导者的最优策略; \bar{x}^* 为在领导者 采用策略p后跟随者的最优策略集合; \bar{X}_i 为第i个跟 随者的最优策略; \bar{x}_{best}^* 为所有跟随者的最优策略; A_i 为跟随者的策略域.

其中供给侧跟随者的决策变量即为 P_{G_i} , 需满足 约束式(2). 而需求侧的跟随者在Stackelberg博弈的 框架下, 内部进行演化博弈的策略更新, 式为

 $\max [U_{\text{leader}}(\bar{\boldsymbol{x}}^*), U_{\text{leader}}(\bar{\boldsymbol{x}}_{\text{best}}^*)],$

s.t.
$$\begin{cases} U_{\text{leader}}(\bar{\boldsymbol{x}}^*) = \sum_{i=1}^n U_i(\bar{\boldsymbol{x}}^*), \\ \bar{\boldsymbol{x}}^* = (\bar{\boldsymbol{x}}_1^* \cdots \bar{\boldsymbol{x}}_i^* \cdots \bar{\boldsymbol{x}}_n^*), \\ \bar{\boldsymbol{x}}_i^* = \arg\max_{\boldsymbol{x}_i \in A_i} U_i(\bar{\boldsymbol{x}}_p^*, \boldsymbol{x}_i), i = 1, 2, \cdots, n, \\ \eta_i(s_i \to s_h) = \\ \frac{1}{1 + \exp[\frac{(U_i - U_h)}{T}]}, \quad \overleftarrow{\Box} a_{ih} = \frac{1}{d_i}, h \in \mathbb{N}_i, \end{cases}$$
(13)

式中: $s_i \rightarrow s_h$ 表示当需求侧的第h个Agent与第i个 Agent之间有拓扑关联时, 第i个Agent的策略从 s_i 变 为 s_h 的概率.

在目标函数的设计中,参与市场的每个博弈者 均须加上电价因素.第t个跟随者的目标函数为

$$F_{t} = \begin{cases} f_{i}^{s}(P_{Gi}) + f_{t}^{slack}(P_{G}^{slack}) + f_{t}^{other} - \\ r_{i}P_{Gi} + PF_{t}, \ t \in \Omega_{s}, \\ r_{j}P_{Dj} - f_{j}^{d}(P_{Dj}) + f_{t}^{slack}(P_{G}^{slack}) + \\ f_{t}^{other} + PF_{t}, \ t \in \Omega_{d}. \end{cases}$$
(14)

除去平衡机组和第*t*个跟随者后,其余发电机的 发电成本为

$$f_{t}^{\text{other}} = \begin{cases} \sum_{i \in \Omega_{s,-t}} (f_{i}^{s}(\bar{P}_{Gi}) - r_{i}\bar{P}_{Gi}) - \\ \sum_{j \in \Omega_{d}} (f_{j}^{d}(\bar{P}_{Dj}) - r_{j}\bar{P}_{Dj}), & t \in \Omega_{s}, \\ \sum_{i \in \Omega_{S}} (f_{i}^{s}(\bar{P}_{Gi}) - r_{i}\bar{P}_{Gi}) - \\ \sum_{i \in \Omega_{d}} (f_{j}^{d}(\bar{P}_{Dj}) - r_{j}\bar{P}_{Dj}), & t \in \Omega_{d}. \end{cases}$$
(15)

在该模型中, Leader平衡机组的目标函数为

$$F_{\text{leader}} = -f(\bar{\boldsymbol{x}}) + \sum_{i \in \Omega_{s}} PF_{i}(\bar{P}_{\text{G}i}) + \sum_{i \in \Omega_{d}} PF_{i}(\bar{P}_{\text{D}i}) + r_{j}P_{\text{D}j} - r_{i}P_{\text{G}i}.$$
(16)

同时,除了满足容量与平衡约束外,涵盖电价因素的供需两侧效益必须大于零,如下式所示:

$$U_{s,i} = r_i P_{Gi} - f_i^s(P_{Gi}) > 0, \qquad (17)$$

$$U_{d,i} = f_j^{\rm d}(P_{\rm Gi}) - r_j P_{\rm Dj} > 0.$$
 (18)

当达到博弈均衡时,同时满足约束条件,此时罚 函数为0,用户报价和发电商报价稳定在出清价格, *F*_{leader}达到最小值,即系统社会总效益*U*取得最大 值.

对于考虑市场因素的电力系统供需互动模型, 不仅仅有考虑发电机阀点效应和机组禁止运行区域 第4期

后导致模型的非凸特性,并且由于考虑了电价因素, 导致求解的可行域极大地缩小了,增加了运算收敛 的难度.因此亟待提出一种能够快速求解该非凸优 化问题的分布式优化算法.

3 混合博弈强化学习算法

本文基于Q学习算法,在Stackelberg博弈的框架 下,在需求侧实行分块协同演化博弈机制,提出了 全新的分布式算法—混合博弈强化学习算法.在该 算法中,各发电机组和负荷聚合商的有功功率策略 将汇集到Leader中集中发布.在每次迭代前,各Follower都能获知在前一次博弈场中其他博弈者的策 略信息;迭代结束时,各Follower都要向Leader上传 对应的Agent *i*的策略.Leader在得知各发电机的出 力策略或负荷聚合商的负荷策略后,可以得到本次 迭代的社会效益,并通过判断优化结果是否比上一 次迭代的好,来选择是否更新发布的出力或者负荷 策略,最终收敛到最优值.

3.1 学习算法

Q学习算法关键是通过对状态-动作对的值函数 Q(s,a)矩阵进行优化迭代获得最优解^[24].本文首先 对Q矩阵进行二进制编码, $Q_{im}(2@2)$ 是Agent *i*的 第m个二进制位的Q矩阵,该Q矩阵的动作空间 A_{im} 与下一个Q矩阵的状态空间 $S_i, m + 1$ 相等, Agent *i* 的Q矩阵 Q_{im} 更新式如下所示:

$$\begin{cases} Q_{im}^{k+1}(s_{im}^{kj}, a_{im}^{kj}) = Q_{im}^{k}(s_{im}^{kj}, a_{im}^{kj}) + \mu \Delta Q_{im}^{k}, \\ \Delta Q_{im}^{k} = R_{im}^{j}(s_{im}^{kj}, s_{im}^{k+1,j}, a_{im}^{kj}) + \\ \eta \max_{a_{in} \in A_{im}} Q_{im}^{k}(s_{im}^{k+1,j}, a_{im}) - \\ Q_{im}^{k}(s_{im}^{kj}, a_{im}^{kj}), \end{cases}$$

$$(19)$$

式中: $m=1, 2, \dots, M; j=1, 2, \dots, J; \mu$ 是学习因 子; η 是折扣因子; 上标k和j分别表示第k次迭代和 第j个个体; 下标i和m分别表示第i个智能体和第m个二进制位; M是二进制串的长度; ΔQ 是知识增 量; J为一次迭代中的种群大小; (s, a)表示状态-动 作对; $R(s_w^k, s^{k+1}, a^k)$ 是在动作 a^k 发生时从状态 s^k 转换到状态 a^{k+1} 的奖励函数; a^{im} 是可选动作.

本文根据蚁群算法的合作机制^[25-27]设计奖励 函数*R*,用于对*Q*矩阵的策略更新,并控制迭代的收 敛速度,如下所示:

$$R_{im}^{j}(s_{im}^{lj}, s_{im}^{k+1,j}, a_{im}^{kj}) = \begin{cases} \frac{p_{m}}{F_{i}^{\text{Best}}}, & (s_{im}^{kj}, a_{im}^{kj}) \in \text{SA}_{i}^{\text{Best}}, \\ 0, & \text{ 其他}, \end{cases}$$
(20)

$$F_i^{\text{Best}} = \min_{j=1,2,\cdots,J} F_i^{kj},\tag{21}$$

式中: F_i^{Best} 代表在第i个智能体的第k次迭代中, 种 群中最优个体的目标函数最小值; F_i^{kj} 代表了第i个 智能体的第k次迭代中的目标函数; p_m 是一个为正 的倍数; SA_i^{\text{Best}}表示第i个智能体的第k次迭代中最 优个体的状态–动作对集合.

3.2 Q矩阵的分块协同演化博弈

在需求侧,分块 Ω_i 可表示与Agent *i*有边的邻居 和Agent *i*本身. Ω_i 储存了对应Agent *i*的 Q_{ii} 矩阵及 同一个分块内邻居Agent *h*的 Q_{ih} 矩阵($h \in \Omega_i$),相当 于Q算法中的智能体.它在每次迭代的过程中,都会 对内部的知识矩阵及其对应的策略进行优化;同时, 因为Agent *i*有可能分布在不同的分块中,此时, Agent *i*的知识矩阵将被不同的分块同时优化,其实 也就是一个分块同时优化了多个知识矩阵.

各发电机组和负荷聚合商的有功功率策略将汇 集到Leader中集中发布.在迭代前,分块 Ω_i 将获知 前一轮在博弈场上的策略信息,并且获取分块 Ω_i 内 的邻居Agent h的 Q_{ih} 矩阵($h \in \Omega_i$)在其他分块中的 Q矩阵和社会效益信息;每轮迭代结束后,分块 Ω_i 需告知Leader对应的Agent i的负荷策略. Leader在 获知各负荷聚合商的策略后,再结合供给侧各发电 机组的策略,便可求取本轮的社会效益.如果此次 的结果比上一次好,则更新本轮的优化策略,否则 就选择保留上一轮的更新结果,并体现在奖励函 数R中,并改变下一轮的更新策略方向,最终收敛到 最优值.以"1–2–3"的三智能体简单拓扑为例,图3 为需求侧Q矩阵的分块协同演化博弈框架.



图 3 Q矩阵的分块协同演化博弈框架



而对于每一个Agent来说,均有可能与其他的 Agent产生联系,在拓扑中体现为有边的连接,此时 该Agent的分块中就含有不止一个知识矩阵. 所以 Agent *i*的*Q_{ii}*迭代更新变成如下所示:

$$\begin{cases} Q_{iim}^{k+1}(\boldsymbol{S}_{iim}^{kj}, a_{iim}^{kj}) = Q_{iim}^{k}(\boldsymbol{S}_{iim}^{kj}, a_{iim}^{kj}) + \alpha \Delta Q_{iim}^{k}, \\ \Delta Q_{iim}^{k} = R_{iim}^{j}(s_{iim}^{kj}, s_{iim}^{k+1,j}, a_{iim}^{kj}) + \\ \gamma \max_{a_{iim} \in A_{iim}} Q_{iim}^{k}(s_{iim}^{k+1,j}, a_{iim}) - \\ Q_{im}^{k}(s_{iim}^{kj}, a_{iim}^{kj}), \\ m = 1, 2, \cdots, M, \ j = 1, 2, \cdots, J. \end{cases}$$
(22)

第*i*个Agent的第*i*个知识矩阵的奖励函数的设计 如下所示:

$$R_{iim}^{j}(s_{iim}^{kj}, s_{iim}^{k+1,j}, a_{iim}^{kj}) = \begin{cases} \frac{p_{m}}{F_{i}^{\text{Best}} + c_{f}}, & (s_{iim}^{kj}, a_{iim}^{kj}) \in \text{SA}_{i}^{\text{Best}}, \\ 0, & \text{ 其他.} \end{cases}$$
(23)

在本文算法中,节点与其所有邻居在每次迭代 过程中进行博弈,积累博弈得到的收益,再依据策 略演化更新规则实施策略更新.采用式(10)的策略 更新规则,Agent *i*模仿邻居Agent *h*策略的概率为

$$\eta_{i}(Q_{iim}^{k}(s_{iim}^{kj}, a_{iim}) \to Q_{ihm}^{k}(s_{ihm}^{kj}, a_{ihm})) = \frac{1}{1 + \exp[\frac{(U_{i} - U_{h})}{T}]}.$$
(24)

算法在每次迭代过程中,会在分块i中寻找与i有 连接的获得最大社会效益的智能体h. 当 $U_h > U_i$ 时, 则以概率 ηi 直接复制分块h内博弈者i的知识矩阵 Q_{ihm}^k 矩阵.

传统的Q学习是智能体通过相应变量的Q矩阵 来寻优的,与此不同的是,在本文的分块协同中,不 同分块储存的对应于同一个节点i的 $Q_{hi(h\in\Omega_i)}(h \in \Omega_i)$ 矩阵,会进行交互学习,得到分块 Ω_i 内节点i对 应的协同 Q_{ii} 矩阵,用于指导动作选择.按照二进制 位进行迭代的公式具体如下:

$$C_{iim}^{k}(s_{iim}^{kj}, a_{iim}) = Q_{iim}^{k}(s_{iim}^{kj}, a_{iim}) - \boldsymbol{\theta} \times \sum_{h \in \Omega_{l}(k)} Q_{iim}^{k} \times (s_{iim}^{kj}, a_{iim} - Q_{ihm}^{k}(s_{ihm}^{kj}, a_{ihm})), \qquad (25)$$

式中*θ*为分块协同的权重值,不同网络拓扑下的权 重值不同,反映了拓扑的结构和性质.

上面步骤得到的协同矩阵*C^kiim*将最终决定动作选择. 而在动作选择过程中, 动作选择*a*的方式如下:

式中: q_0 是0到1之间的随机数; ε 是采取贪婪寻优法则策略的概率; a_{rand} 是采取随机寻优策略的概率.

而对于供给侧的发电机组,由于拓扑是固定的, 每一台均是独立且相互博弈的个体,所以此时的分 块只含有本身智能体的知识矩阵,即相当于没有演 化博弈的过程.

3.3 混合博弈强化学习算法流程图

本文所提的混合博弈强化学习算法的求解流程 如图4所示.其中kmax是最大迭代次数.



Fig. 4 Flow chart of MGRL

由图4流程图可知,每次迭代前,各智能体都会获取上一次所有博弈者的策略信息;每次迭代结束后,各智能体都要向领导者上传对应的博弈者*i*的策略.领导者在得知各智能体的功率策略后,可以得到本次迭代的社会效益,并通过判断优化结果是否比上一次迭代的好,来选择是否更新发布的功率策略,最终收敛到了最优值.

本文所提的混合博弈强化学习算法的优点主要 有以下3个方面: 1)该算法可以有效求解非凸优化问题.混合博 弈强化学习算法基于Q学习算法,以Q矩阵来描述 策略,对模型依赖程度低,可以有效求解本文所研究 的含阀点效应多智能体系统的非凸供需互动模型.

2) 该算法适应于分布式优化计算框架. 混合博 弈强化学习算法是在Stackelberg博弈与演化博弈的 混合博弈框架下提出来的, 能够充分调动多智能体 的互动, 并把复杂的总任务分配到下层每个智能体 的子任务中, 既保证了私密性和可靠性, 也极大地 提高了优化速度.

3) 该算法能够充分利用不同形式的网络拓扑结构进行有效求解. 混合博弈强化学习算法利用了图论来研究智能体直接的关系,并利用了分块协同和演化博弈来进行优化,充分实现并利用了网络拓扑中各智能体之间所形成的知识矩阵信息互动,通过多个智能体的协同优化明显提升了算法的寻优性能,并且能够在任何形式的网络拓扑结构下进行有效求解,具有很强的适应性和推广性.

4 算例分析

为了验证MGRL算法的可行性与优越性,本文 搭建了基于上文所描述的4种经典复杂网络的3机-6负荷系统用于仿真,同时把仿真扩大到在需求侧含 有无标度网络的南方某一线城市电网算例中.MG-RL算法的参数如表1所示.

表 1 混合博弈强化学习算法参数 Table 1 Parameters of MGRL

		3机6负荷系统				
参数	范围	规则	随机	小世界	无标度	南方某一线
		网络	网络	网络	网络	城市电网
α	$0\!<\!\alpha\!<\!1$	0.8	0.8	0.7	0.8	0.8
γ	$0\!<\!\gamma\!<\!1$	0.1	0.1	0.1	0.1	0.1
J	$J\!\geqslant\!1$	50	50	50	50	50
M	$M\!\geqslant\!1$	25	25	25	25	25
p_{m}	$p_{\rm m} \! \geqslant \! 0$	1000	1000	1000	1000	1000
$c_{\rm f}$	$c_{\rm f}\!\geqslant\!0$	0	0	0	0	0
ε	$0\!<\!\varepsilon\!<\!1$	0.5	0.8	0.8	0.8	0.8
θ		0.5	0.5	0.5	0.5	0.5
T	$T\!\geqslant\!0$	100	100	100	100	1000
k_{\max}	$k_{\max} \ge 2$	300	300	300	300	300

因为电力市场环境下考虑需求侧复杂网络的电 力系统供需互动问题的非凸性,本文分别采用了人 工蜂群算法(ABC)、遗传算法(GA)、生物地理学优 化算法(biogeography-based optimization, BBO)、群 搜索算法(group search optimizer, GSO)、磷虾群算 法(krill herd, KH)及粒子群算法(PSO)算法共6种集 中式人工智能算法与混合博弈强化学习算法进行对 比,这些算法均可以求解非凸模型.

为了兼顾算法最佳性能和公平性,本文经过大量测试,确定了最优的参数配置,如表2所示.所有算法的迭代步数均设置为300,种群规模均设置为50. 仿真计算均在CPU为英特尔i7-4710MQ、主频2.5 GHz、内存(RAM)8 GB的计算机运行.

表 2 对比算法最优参数配置

Table 2 Optimal parameters of the comparative algorithms

		取值			
算法	参数	40机54负荷	南方某一线		
		系统	城市电网		
	变异概率	0.8	0.8		
GA	交叉概率	0.1	0.05		
	遗传代沟	0.8	0.8		
	最小旋转速度	-5	-5		
	最大旋转速度	5	5		
PSO	加速系数 <u>c1</u>	1.50	1.75		
		1.50	1.75		
	最小惯性糸数	0.4	0.4		
	最大惯性系数	0.9	0.9		
	采蜜蜂	75	70		
	侦察蜂	50	45		
ABC	观察蜂	25	35		
	限制次数	800	800		
	游荡者比例	30%	25%		
GSO	最大搜索角	$\frac{\pi}{4}$	$\frac{\pi}{4}$		
	最大搜索转角	$\frac{\frac{4}{\pi}}{8}$	$\frac{\frac{4}{\pi}}{8}$		
DDO	迁入概率边界	[0, 1]	[0, 1]		
RRO	突变概率	0.015	0.009		
	最大诱导速度	0.01	0.009		
KH	觅食速度个体	0.025	0.02		
	最大扩散速度	0.005	0.01		

4.1 3机-6负荷系统算例

4.1.1 仿真模型

表3-4分别给出了3机-6负荷系统的燃煤发电机 的主要参数和负荷参数.

表 3 3机-6负荷系统燃煤发电机的主要参数

Table 3 Main parameters of coal-fired generators of3-generator 6-load system

宮早			发电成本参数				
1	Gi /MIW	Gi /INI W	а	b	С	d	е
G_1	60	120	0.00608	2.121	92.86	30	0.084
G_2	47	97	0.00342	1.605	44.66	36	0.077
G_3	25	110	0.00483	1.764	92.23	24	0.098

914

Table 4 Load parameters of 3-generator 6-load system

序号	$lpha_i$	ω_i	$P_{\mathrm{D}j}^{\mathrm{min}}/\mathrm{MW}$	$P_{\mathrm{D}j}^{\mathrm{max}}/\mathrm{MW}$
D_1	-0.055	5.7	22	44
D_2	-0.0595	6.2	23.51	47.02
D_3	-0.0487	5.1	20	40
D_4	-0.08	6.15	20	40
D_5	-0.137	7.25	16.5	33
D_6	-0.0912	5.25	12.5	25

表3-4中,机组3的容量范围最大,选为平衡机 组.图5根据上文给出的经典复杂网络模型搭建 了6负荷规则网络、6负荷随机网络、6负荷小世界网 络和6负荷无标度网络.分别对基于复杂网络的4 类3机-6负荷系统使用MGRL算法和其他6种智能 算法分别独立计算100次.



Fig. 5 Four networks of 6 load

4.1.2 仿真结果分析

利用混合博弈强化学习算法计算的仿真收敛图 如图6所示.寻优结果,即社会效益最优值、出清价 格及其对应的机组出力和负荷策略,如表5所示.



图 6 MGRL算法在3机--6负荷系统的仿真收敛图



表 5 混合博弈强化学习算法在3机-6负荷系统的 优化结果

Table 5	Optimization results of MGRL algorithm i	n
	3-generator 6-load system	

序号	功率/MW	序号	功率/MW	
G_1	60	D_3	35.0768	
G_2	87.8205	D_4	33.7921	
G_3	57.0495	D_5	28.3146	
D_1	41.5319	D_6	19.8152	
D_2	46.3392		_	
$r*/(\$ \cdot MW^{-1} \cdot h)$		4.0937		
U/	$(\$ \cdot h^{-1})$	297.5986		

从图6和表5可以明显地看出,混合博弈强化学 习算法在该模型中的第50次迭代步数之前就已经收 敛到了最优值297.5986 \$/h,收敛速度较快,而且收 敛效果较好.

为了验证混合博弈强化学习算法的性能,本文 把混合博弈强化学习算法(MGRL)、人工蜂群算 法(ABC)、遗传算法(GA)、生物地理学优化算法 (biogeography-based optimization, BBO)、群搜索算 法(group search optimizer, GSO)、磷虾群算法(krill herd, KH)及粒子群算法(PSO)在3机-6负荷系统中 计算100次后得到的社会效益最大值、平均值和最 小值放在了表6进行了比较.其中,除了混合博弈强 化学习算法是分布式算法之外,其余的6种算法都 是集中式的算法.

- 表 6 7种算法在3机-6负荷系统计算100次后的结 果比较
- Table 6 Comparison of results after calculating one hundred times by 7 different algorithms in 3generator 6-load system

笛辻		优化	社会效益/(\$ · h ⁻¹)			
	异位		最小值	平均值	最大值	
GA		集中式	250.69	290.94	299.63	
	PSO	集中式	267.31	290.49	298.82	
BBO		集中式	254.47	288.58	297.13	
ABC		集中式	254.31	287.22	296.98	
KH		集中式	251.25	271.92	296.51	
	GSO	集中式	265.12	280.54	295.60	
	无标度网络	分布式	280.29	291.83	297.60	
MGRL	规则网络	分布式	289.52	295.96	297.57	
	随机网络	分布式	280.24	292.95	297.43	
	小世界网络	分布式	284.02	295.23	297.32	

从表6可以明显看出, MGRL算法的社会效益最 大值仅比在这方面表现最优的GA算法大约低了 0.7%,几乎达到了最优,排名可达前3名.而MGRL 算法的社会效益最小值和平均值均排名第一.另一 方面,若只观察MGRL算法在4种网络拓扑中的寻 优结果,可以明显看出,该算法在这4类复杂网络中 的仿真结果非常相近.

4.2 南方某一线城市电网算例

4.2.1 仿真模型

在南方某一线城市电网中,一共含有110 kV及 以上的节点279个,发电厂一共有12个,支路有375 条,装机容量一共为13120 MW.在该城市电网中、 核电站、天然气发电厂和垃圾发电厂共6台由于被 用于满足刚性负荷需求中,必须实时满足,所以不 参与到本文讨论的实时供需互动优化中.而其他 6台燃煤发电机则用来平衡需求侧的柔性负载.同 样地,机组6由于容量约束范围最大,被挑选为平衡 机组.另外,该电网含有110 kV及以上的柔性负荷 共71个,它们都参与实时供需互动优化.

同时,由第1个算例可知,不同的复杂网络几乎 不影响MGRL算法的性能.而由上文复杂网络理论 可知,大多数大规模真实网络通常使用幕率分布来 描述度分布更为准确,幕率分布的节点基本上只有 少量的连接,而少量的节点则有着大量的连接,而 由于无标度网络的度分布呈现幂律分布的特性,这 与未来电力市场环境下,需求侧负荷聚合商最终演 化成的多寡头少联系的关系相适应.所以,在本小 节的算例中,本文利用该南方某一线城市电网中 的71负荷,构建了无标度网络,如图7所示.为了验 证MGRL算法在较大规模的推广性,本文利用上述 的7种算法对该大型电网算例进行了仿真.同样的, 为了排除偶然性,本文的每类算法均分别计算100次.



图 7 71负荷无标度网络 Fig. 7 Scale-free network of 71 load

4.2.2 仿真结果分析

混合博弈强化学习算法的仿真收敛过程如图8 所示.观察图8可知,在约第40次迭代步数之前,该 算法就已经收敛了.从收敛所耗费的时间来看是比 较短的,而从效果来看也是比较好的.最后收敛达 到的总经济效益是213912 \$/h. 与算例1类似的,7种 算法的寻优结果如表7所示.



图 8 MGRL算法在南方某一线城市电网的仿真收敛图

Fig. 8 Simulation convergence diagram of MGRL algorithm on a first-tier city power grid in the southern China

从表7可以明显看出,从表格的最大值一列可看出,MGRL算法在社会效益最大值方面仅仅比PSO 差了0.02%,几乎已经是最优值,排名第3名.而在社会效益的最小值和平均值方面,MGRL算法的表现也不差,分别为第4名和第3名.而且从绝对值来看,与最好的收敛值也是非常相近的.从表7的仿真结果来看,该算例充分验证了MGRL算法对实际的大规模智能电网供需互动模型优化的可行性和高效性.

- 表77种算法在南方某一线城市电网计算100次后 的结果比较
- Table 7 Comparison of results after calculating one
hundred times by 7 different algorithms on a
first-tier city power grid in the southern Chi-
na

算法	优化 方式	社会 最小值	☆益/(\$ · ŀ 平均值	n ⁻¹) 最大值
GSO	集中式	213564	213775	213961
GA	集中式	213491	213771	213925
MGRL	分布式	213446	213742	213912
BBO	集中式	212742	213383	213809
PSO	集中式	210511	213200	213780
ABC	集中式	213461	213608	213755
KH	集中式	212102	212957	213605

4.3 算法寻优性能分析

从前面两个算例的仿真结果可以得出,分布式 的混合博弈强化学习算法的寻优性能比大部分的集 中式智能算法优越,并且在4类复杂网络中,该算法 的收敛结果非常接近.由于本文研究的是电力市场 环境下的电力系统供需互动问题,集中式的智能算 法由于无法保证参与博弈的市场主体的私密性,所 以不适用于该框架.而本文提出的全新的混合博弈 强化学习算法属于分布式算法,算法性能非常接近 集中式智能算法的最优值,表明了该算法可以有效 地求解决考虑市场因素的电力系统供需互动问题.

另一方面,由第1个算例可知,混合博弈强化学 习算法在不同的复杂网络中都可以保证良好的仿真 结果,而这4种典型的复杂网络基本涵盖了所有类 型的网络拓扑,因此说明了混合博弈强化学习算法 在不同的网络中具有较强的适应性和稳定性.

显然, 混合博弈强化学习算法优异的寻优性能 是由其算法机理决定的. 正是因为该算法在图论的 基础上, 引入了分块协同和演化博弈机制. 这种机 制使得邻居之间互动博弈关系所形成的知识矩阵信 息互动得到了充分的利用, 再通过多个智能体的协 同优化明显提升了算法的寻优性能.

5 结论

本文提出一种基于Q学习理论的混合博弈强化 学习算法,用于求解考虑机组阀点效应,禁止运行 区域、考虑电价因素及需求侧复杂网络的非凸电力 系统供需互动模型.所得结论如下:

 本文所搭建得考虑需求侧复杂网络的电力系 统供需互动混合博弈模型可与未来电力市场环境下 需求侧负荷聚合商之间多变的关系和复杂的信息网 络拓扑相匹配.

2) 本文所提出的分布式MGRL算法相较于集中 式的智能算法更是适应于电力市场环境下多主体的 供需互动框架,并且相较与多数集中式智能算法, 其寻优性能更好.

3) 本文所提出的MGRL算法由于在图论的基础 上引入分块协同和演化博弈机制,使得知识矩阵信 息可以充分利用,因此对于不同的复杂网络拓扑均 显示了良好的寻优效果.3机--6负荷系统算例和南 方某一线城市电网算例两个仿真结果表明算法具有 较好的适应性和稳定性.

在下一步的研究中,笔者将对加入需求侧负荷 的精细化建模,并进行多时段耦合仿真分析,进一 步完善混合博弈强化学习算法的计算性能.

参考文献:

 MHANNA S, VERBIČ G, CHAPMAN A C. A faithful distributed mechanism for demand response aggregation. *IEEE Transactions on Smart Grid*, 2016, 7(3): 1743 – 1753.

- [2] LIU Guojing, HAN Xueshan, YANG Ming. Synergetic eco-nomic dispatch in power system operation. *Proceedings of the CSEE*, 2014, 34(16): 2668 2675.
 (刘国静, 韩学山, 杨明. 电力系统运行协同的经济调度. 中国电机工程学报, 2014, 34(16): 2668 2675.)
- [3] MA Z, CALLAWAY D, HISKENS I. Decentralized charging control for large populations of plug-in electric vehicles: Application of the Nash certainty equivalence principle. *IEEE International Conference* on Control Applications. Yokohama, Japan: IEEE, 2010: 206 – 212.
- [4] YAO Jianguo, YANG Shengchun, WANG Ke, et al. Frame-work and strategy design of demand response scheduling for balancing wind power fluctuation. *Automation of Electric Power Systems*, 2014, 38(9): 85 – 92.
 (姚建国,杨胜春,王珂,等. 平衡风功率波动的需求响应调度框架与

策略设计. 电力系统自动化, 2014, 38(9): 85 – 92.)

- [5] DENG R, YANG Z, CHOW M Y, et al. A survey on demand response in smart grids: Mathematical models and approaches. *IEEE Transactions on Industrial Informatics*, 2015, 11(3): 570 – 582.
- [6] WANG Xifan, XIAO Yunpeng, WANG Xiuli. Study and analysis on supply-demand interaction of power systems under new circumstances. *Proceedings of the CSEE*, 2014, 34(29): 5018 5028.
 (王锡凡,肖云鹏,王秀丽.新形势下电力系统供需互动问题研究及分析.中国电机工程学报, 2014, 34(29): 5018 5028.)
- [7] SAMADI P, MOHSENIAN-RAD A H, SCHOBER R, et al. Optimal real-time pricing algorithm based on utility maximization for smart grid. *IEEE International Conference on Smart Grid Communications*. Gaithersburg, MD, USA: IEEE, 2010: 415 – 420.
- [8] MEI Shengwei, LIU Feng, WEI Wei. Engineering Game Theory and Power System Application. Beijing: Science Press, 2016. (梅生伟, 刘锋, 魏韡. 工程博弈论及电力系统应用. 北京: 科学出版 社, 2016.)
- [9] WEI Wei, CHEN Yue, LIU Feng, et al. Stackelberg game based retailer pricing scheme and EV charging management in smart residential area. *Power System Technology*, 2015, 39(4): 939 945.
 (魏韓, 陈玥, 刘锋, 等. 基于主从博弈的智能小区代理商定价策略及 电动汽车充电管理. 电网技术, 2015, 39(4): 939 945.)
- [10] WU Cheng, GAO Bingtuan, TANG Yi, et al. Master-slave game based bilateral contract transaction model for generation companies and large consumers. *Automation of Electric Power Systems*, 2016, 40(22): 56 62.
 (吴诚,高丙团,汤奕,等. 基于主从博弈的发电商与大用户双边合同 交易模型. 电力系统自动化, 2016, 40(22): 56 62.)
- [11] DAI Yeming, GAO Yan, GAO Hongwei, et al. Leader-follower game model for demand response in smart residential grid. Automation of Electric Power Systems, 2017, 41(15): 88 – 94.
 (代业明,高岩,高红伟,等. 智能住宅小区的需求响应主从博弈模型. 电力系统自动化, 2017, 41(15): 88 – 94.)
- [12] LU Z X, MENG Z W, ZHOU S X. Cascading failure analysis of bulk power system using small-world network model. 2004 International Conference on Probabilistic Methods Applied to Power Systems. Ames, IA, USA: IEEE, 2004: 635 – 640.
- [13] DOROGOVTSEV S N, MENDERS J F F. Scaling properties of scalefree evolving networks: Continuous approach. *Physical Review E*, 2001, 63(5): 056125/1 – 056125/19.
- [14] WANG Xianpei, ZHU Guowei, HE Ruijuan, et al. Survey of cascading failures in cyber physical power system based on complex network theory. *Power System Technology*, 2017, 41(9): 2947 – 2956. (王先培,朱国威,贺瑞娟,等. 复杂网络理论在电力CPS连锁故障研 究中的应用综述. 电网技术, 2017, 41(9): 2947 – 2956.)
- [15] WEI Zhenbo, GOU Jing. An overview on application of complex network theory in power system analysis. *Power System Technology*, 2015, 39(1): 279 – 287.
 (魏震波, 苟竞. 复杂网络理论在电网分析中的应用与探讨. 电网技 术, 2015, 39(1): 279 – 287.)

[16] WANG Yunan, YANG Jingfei, HE Yeshuai, et al. Rapid assessment of multistage operation risk of active distribution network based on complex network theory. *Automation of Electric Power Systems*, 2016, 40(14): 65 - 71.
(王钰楠,杨镜非,何也帅,等. 基于复杂网络理论的主动配电网多级

运行风险快速评估. 电力系统自动化, 2016, 40(14): 65 – 71.)

- [17] PENG J, WILLIAMS R J. Incremental multi-step Q-learning. Machine Learning, 1996, 22(1/3): 283 – 290.
- [18] GONG C, WANG X, XU W, et al. A distributed real-time energy scheduling in smart grid: Stochastic model and fast optimization. *IEEE Transactions on Smart Grid*, 2013, 4(3): 1476 – 1489.
- [19] PARK J B, JEONG Y W, SHIN J R, et al. An improved particle swarm optimization for nonconvex economic dispatch problems. *IEEE Transactions on Power Systems*, 2010, 25(1): 156 – 166.
- [20] ZHANG W, XU Y, LIU W, et al. Distributed online optimal energy management for smart grids. *IEEE Transactions on Industrial Informatics*, 2017, 11(3): 717 – 727.
- [21] LI Yixiao. Research on modeling socio-economic systems based on theories of complex network and theories of complex network and evolutionary game. Hangzhou: Zhejiang University, 2010. (李一啸. 基于复杂网络和演化博弈理论的社会. 杭州: 浙江大学, 2010.)
- [22] REN J, WANG W X, QI F. Randomness enhances cooperation coherence resonance in evolutionary game. eprint ar Xiv.physics/0607457.
- [23] SHI Zhentao. Research on evolutionary agent-based artificial social system modeling methods and application. Lanzhou: Lanzhou University of Technology, 2013.
 (时振涛. 基于演化Agent的人工社会系统建模方法及其应用研究. 兰州: 兰州理工大学, 2013.)

- [24] WATKINS C J C H, DAYAN P. Q-learning. *Machine Learning*, 1992, 8(3/4): 279 – 292.
- [25] DORIGO M, GAMBARDELLA L M. A study of some properties of Ant-Q. International Conference on Parallel Problem Solving from Nature. Berlin: Springer, 1996: 656 – 665.
- [26] KRYNICKI K, HOULE M E, JAEN J. An efficient ant colony optimization strategy for the resolution of multi-class queries. *Knowledge-Based Systems*, 2016, 105(C): 96 – 106.
- [27] YU Tao, ZHANG Xiaoshun. A fast dynamic optimization algorithm with memory and self-learning and its application on reactive power optimization. *Scientia Sinica Technologica*, 2016, 46(3): 256 267. (余涛,张孝顺. 一种具有记忆自学习能力的快速动态寻优算法及其无功优化求解. 中国科学: 技术科学, 2016, 46(3): 256 267.)

作者简介:

包 涛 硕士研究生,主要研究方向为电力系统优化运行与控制, E-mail: baotaowork@foxmail.com;

李昊飞 硕士研究生,主要研究方向为电力系统及综合能源系统 优化调度, E-mail: li_haofei@sina.com;

余 涛 博士,教授,博士生导师,主要研究方向为复杂电力系统的非线性控制理论和仿真以及人工智能在电力系统中的应用, E-mail: taoyu1@scut.edu.cn;

张孝顺 博士, 副教授, 硕士生导师, 主要研究方向为电力系统优 化控制与启发式算法, E-mail: xszhang1990@sina.cn.