

基于上下文学习的电力物联网接入控制方法

周振宇^{1,2}, 贾泽晗^{1,2}, 廖海君^{1,2}, 赵雄文^{1,2}, 张磊³

(1. 华北电力大学新能源电力系统国家重点实验室, 北京 102206; 2. 东南大学移动通信国家重点实验室, 江苏 南京 210096;
3. 国网山东省电力公司电力科学研究院, 山东 济南 250003)

摘 要: 针对 6G 时代电力物联网海量终端接入冲突严重、队列积压大、能量效率低等问题, 提出了一种基于上下文学习的接入控制算法。所提算法基于强化学习和快速上行链路授权技术, 考虑终端活跃与休眠 2 种状态, 优化目标为在终端接入服务质量需求的长期约束下最大化网络总能量效率。利用李雅普诺夫优化对长期优化目标与约束进行解耦, 将长期优化问题转化为单时隙独立的确定性子问题, 并利用基于终端状态感知的上置信界算法进行求解。仿真结果表明, 所提算法能够在满足终端接入服务质量需求的同时, 有效提高网络总能量效率。相较于传统快速上行链路授权算法, 所提算法可提高平均能量效率 48.11%, 提高满足接入服务质量需求的终端比例 54.95%, 降低平均队列积压 83.83%。

关键词: 6G; 电力物联网; 海量终端接入; 上下文学习; 快速上行链路授权

中图分类号: TN915

文献标识码: A

DOI: 10.11959/j.issn.1000-436x.2021062

Context-aware learning-based access control method for power IoT

ZHOU Zhenyu^{1,2}, JIA Zehan^{1,2}, LIAO Haijun^{1,2}, ZHAO Xiongwen^{1,2}, ZHANG Lei³

1. State Key Laboratory of Alternate Electrical Power System with Renewable Energy Sources, North China Electric Power University, Beijing 102206, China
2. National Mobile Communications Research Laboratory, Southeast University, Nanjing 210096, China
3. State Grid Shandong Electric Power Research Institute, Ji'nan 250003, China

Abstract: In view of the problems of severe access conflicts, high queue backlog, and low energy efficiency in the massive terminal access scenario of the power Internet of things (power IoT) in 6G era, a context-aware learning-based access control (CLAC) algorithm was proposed. The proposed algorithm was based on reinforcement learning and fast uplink grant technology, considering active state and dormant state of terminals, and the optimization objective was to maximize the total network energy efficiency under the long-term constraint of terminal access service quality requirements. Lyapunov optimization was used to decouple the long-term optimization objective and constraint, and the long-term optimization problem was transformed into a series of single time-slot independent deterministic sub-problems, which could be solved by the terminal state-aware upper confidence bound algorithm. The simulation results show that CLAC can improve the network energy efficiency while meeting the terminal access service quality requirements. Compared with the traditional fast uplink grant, CLAC can improve the average energy efficiency by 48.11%, increase the proportion of terminals meeting access service quality requirements by 54.95%, and reduce the average queue backlog by 83.83%.

Keywords: 6G, power IoT, massive terminal access, context-aware learning, fast uplink grant

收稿日期: 2020-08-19; 修回日期: 2020-11-16

基金项目: 国家电网有限公司科技基金资助项目 (No.SGSDDK00KJJS1900405); 东南大学移动通信国家重点实验室开放研究基金资助项目 (No.2021D12)

Foundation Items: Science and Technology Project of State Grid Corporation of China (No.SGSDDK00KJJS1900405), The Open Research Fund of National Mobile Communications Research Laboratory, Southeast University (No.2021D12)

1 引言

电力物联网 (PIoT, power Internet of things) 是实现电力系统各环节万物互联、人机交互的工业级物联网, 基于深度的感知能力和先进的信息通信技术提高电网精准控制、智能调度的水平, 推动传统电力系统向能源互联网转变^[1]。配用电是电网的重要场景之一, 配用电电力物联网的建设对分布式能源友好接入、新型负载弹性承载以及用户多样性用能需求保障具有重要作用。然而, 由于无线资源和计算资源的限制, 配用电电力物联网中海量终端的并发接入极大地增加了接入网的承载压力, 导致网络拥塞、过载等问题, 严重威胁电网的安全稳定运行。因此, 海量终端的接入控制问题一直是备受关注的研究热点。

6G 系统将在 5G 系统“万物互联”的基础上进一步拓展至“万物智联”^[2], 为解决配用电电力物联网海量终端的接入控制问题提供了可能。一方面, 与 5G 相比, 6G 时代终端连接密度将可提高 100 倍, 达到每平方千米 1 亿。另一方面, 人工智能 (AI, artificial intelligence) 技术将进一步提高接入控制和资源管理的智能化水平^[3]。现有的接入控制技术可分为竞争和非竞争两大类。针对竞争接入场景, 文献[4]提出了一种自适应调整接入类别限制 (ACB, access class barring) 因子的方法, 基于最小均方算法优化 ACB 因子的选择策略, 提高了高拥塞期间的接入成功概率, 降低了接入时延。文献[5]提出了一种分组群呼预退避方案, 有效提高了终端资源共享的公平性。文献[6]提出了一种动态随机接入资源分配机制, 通过引入终端序号对终端加以区分, 进而降低前导码冲突概率。然而, 竞争接入控制技术依赖于基站与终端之间的频繁信令交互, 只适用于连接密度不高的场景, 难以满足爆炸式增长的终端接入需求^[7]。与竞争接入控制技术相比, 非竞争接入控制技术具有信令开销低、资源利用率高、承载能力大等优点。第三代合作伙伴计划 (3GPP, the third generation partnership project) 在 Release 14 中引入了快速上行链路授权技术^[8], 允许获得上行链路许可的终端在没有发送任何调度请求的情况下直接在基站预分配的信道上进行数据传输, 降低信令开销和接入冲突发生的概率。文献[9]介绍了快速上行链路授权技术的基本原理, 探讨了海量终端接入场景下快速上行链路授权面临的机遇与挑战。文

献[10]提出了一种基于压缩感知技术的快速授权信令, 允许时延敏感型终端在接收授权信令后立即进行上行链路访问, 降低了接入时延。然而, 现有技术大多基于全局信息已知的假设, 在实际应用中具有局限性。由于网络资源与信令开销的限制, 基站无法准确获得海量终端的全部信息, 包括终端状态、信道增益、队列积压等。综上所述, 面向高密度、万物智联的 6G 电力物联网接入控制研究在国际上仍处于起步阶段, 迫切需要对全局信息不确定场景下的接入模型、接入控制算法、仿真验证等方面进行深入研究。

强化学习是解决全局信息不确定场景下连续决策问题的有效方法, 已被广泛应用于终端接入控制方面的相关研究。文献[11]介绍了 2 种基于强化学习的无线接入技术: 基于毫米波的智能切换技术和基于多主体强化学习的多无线接入技术 (Multi-RAT, multi-radio access technology)。文献[12]提出了一种基于强化学习的基站选择算法, 终端通过自组织的方式选择负载较轻的基站接入, 有效解决了接入点选择问题, 缓解了网络拥塞。文献[13]利用强化学习解决了非授权用户信道选择问题, 通过减少信道转换过程, 降低了接入时延, 提高了系统容量。文献[14]提出了一种基于强化学习的路由选择方法, 通过利用与基站交互获得的反馈信息, 动态优化路由选择策略, 降低了路由开销。然而, 以上研究均将终端作为策略学习优化主体, 要求终端具有较强的算力和能量资源, 在计算资源受限、低功耗的电力物联网中不再适用。此外, 上述研究均假设所有终端在每个时隙都处于活跃状态, 而电力物联网终端仅在数据传输时处于活跃状态, 其余时隙则处于休眠状态。文献[15]提出了一种基于强化学习的快速上行链路授权方案, 通过优化信息价值、最大可容忍接入时延与传输速率的加权和, 降低接入时延, 提高网络吞吐量。然而, 该方案忽略了海量终端接入场景下终端能量效率对网络长期性能的影响, 以及不同终端的差异化接入需求, 容易导致信息价值低、最大可容忍接入时延高与传输速率低的终端产生严重的数据队列积压, 影响数据的及时传输。

为了解决电力物联网海量终端接入冲突严重、队列积压大、能量效率低等问题, 本文采用快速上行链路授权技术和强化学习, 提出了基于上下文学习的接入控制 (CLAC, context-aware learning-based

access control) 算法, 在保障终端接入服务质量需求的同时, 提升网络总能量效率。首先, 基于强化学习理论, 将基站建模为智能体, 进行上行链路授权策略和收益的设计, 并利用李雅普诺夫优化对长期的优化目标与约束进行解耦。其次, 基站在每个时隙初预测活跃终端的集合, 并根据终端性能上界选择终端分发授权, 获得授权的终端进行数据传输, 计算能量效率与接入性能的加权和并反馈给基站。最后, 基站依据终端反馈更新所执行策略的性能上界, 通过不断迭代实现策略优化。

2 系统模型

2.1 接入模型

基于快速上行链路授权的配用电力物联网多小区海量终端接入模型如图 1 所示。

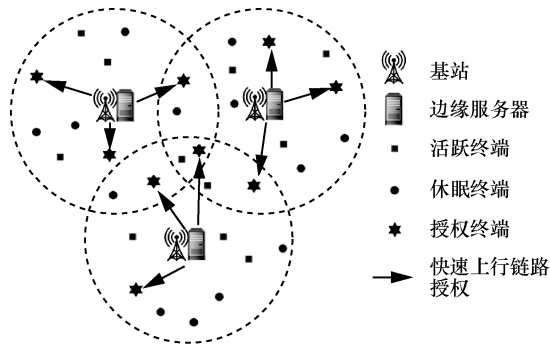


图 1 基于快速上行链路授权的终端接入模型

整个网络由 J 个小区和 K ($K \gg J$) 个终端组成, 每个小区包含一个基站、一个边缘服务器及若干待接入终端。其中, 基站为终端提供无线接入服务, 边缘服务器与基站位于相同位置并提供计算服务。基站集合表示为 $\mathcal{S} = \{s_1, s_2, \dots, s_J\}$, 终端集合表示为 $\mathcal{U} = \{u_1, u_2, \dots, u_K\}$ 。定义基站 s_j 所覆盖终端集合表示为 $\mathcal{N}_j \subseteq \mathcal{U}$, 基站 s_j 与基站 $s_{j'}$ 覆盖重叠区域内的终端集合表示为 $\mathcal{N}_{j \cap j'} = \mathcal{N}_j \cap \mathcal{N}_{j'}$ 。本文采用时隙模型, 总时间周期被划分为 T 个等长的时隙, 每个时隙的长度为 τ , 总时隙集合表示为 $\mathcal{T} = \{1, 2, \dots, T\}$ 。假设在一个时隙内基站与终端之间的信道状态信息是不变的, 而在时隙间发生变化。由于配用电力物联网终端的事件驱动激活或周期性激活等特点, 本文假设终端具有 2 种状态, 即活跃态与休眠态, 当终端具有数据传输需求时, 自动处于活跃态; 否则处于休眠态。在每个时隙初, 基

站预测处于活跃态的终端集合并选择终端分发上行链路许可, 获得授权的终端可与基站建立连接并进行数据传输。由于传统云计算技术在配用电海量终端并发接入时存在网络通信压力大、传输成本高等缺点, 本文考虑采用边缘计算技术^[16], 即终端将任务数据卸载到边缘服务器进行处理。系统参数如表 1 所示。

表 1 系统参数

参数	含义
J /个	基站/边缘服务器数量
K /个	终端数量
M /个	基站可授权的终端数量
T /个	时隙数量
τ /s	时隙长度
A_k /个	到达终端 u_k 的子任务数量
ρ /bit	子任务大小
q_k	终端 u_k 存储在本地缓存区的任务数据队列
Q_k /bit	终端 u_k 存储在本地缓存区的数据队列积压
$U_{k,j}$ /bit	终端 u_k 传输到基站 s_j 的任务数据量
U_k /bit	终端 u_k 的吞吐量
a_k	活跃指示变量, 表示终端 u_k 处于活跃态或休眠态
$x_{k,j}$	授权指示变量, 表示终端 u_k 是否获取基站 s_j 许可
γ_k /dB	数据传输信噪比
$P_{TX,k}$ /W	终端 u_k 的传输功率
$g_{k,j}$ /dB	终端 u_k 与基站 s_j 之间的信道增益
$B_{k,j}$ /kHz	终端 u_k 与基站 s_j 之间的传输带宽
N_0 (W·Hz ⁻¹)	噪声功率谱密度
$E_{k,j}$ /J	终端 u_k 接入基站 s_j 的传输能耗
$EE_{k,j}$ (bit·J ⁻¹)	终端 u_k 接入基站 s_j 的传输能量效率
η_k	终端 u_k 的接入服务质量约束
X_k τ /个	T 个时隙内, 终端 u_k 获得基站授权的总时隙数
T'_k /个	T 个时隙内, 终端 u_k 处于活跃态的总时隙数
F_k	终端 u_k 的接入服务质量需求赤字虚拟队列积压
$\theta_{k,j}$	终端 u_k 接入基站 s_j 时的能量效率和接入性能的加权和
$V_{k,j}$	基站 s_j 对终端 u_k 性能的估计值
P_k	预测终端 u_k 的活跃概率
$z_{k,j}(t)$	终端 u_k 被基站 s_j 授权的累积回报
$n_{k,j}(t)$ /个	终端 u_k 处于活跃态且被基站 s_j 授权的总时隙数
t'_k /个	终端 u_k 处于活跃态的总时隙数
ζ	算法探索权重
V_{EE}	能量效率权重
V_F	接入性能权重

2.2 任务传输模型

本文采用任务划分模型^[16], 将任务划分为有限个相同大小的子任务。假设每个时隙初到达终端 u_k

的子任务数量为 $A_k(t)$ 个，每个子任务的大小为 ρ ，所有的数据首先被存储在 u_k 的本地任务缓存区中。当 u_k 与基站建立连接时，将任务数据卸载到边缘服务器进行计算。存储在 u_k 本地缓存区中的任务数据被建模为队列 q_k ，队列积压为 $Q_k(t)$ ，其在第 $t+1$ 个时隙初的队列积压为

$$Q_k(t+1) = \max\{Q_k(t) - U_k(t) + \rho A_k(t), 0\} \quad (1)$$

其中， $U_k(t)$ 表示 u_k 在第 t 个时隙卸载到边缘服务器的任务数据量。定义活跃指示变量为 $a_k(t)$ ，当终端存在数据传输需求时，即 $Q_k(t) > 0$ ，终端处于活跃态，表示为 $a_k(t) = 1$ ；否则处于休眠态，表示为 $a_k(t) = 0$ 。定义授权指示变量为 $x_{k,j}(t)$ ， $x_{k,j}(t) = 1$ 表示在时隙 t 终端 u_k 获得基站 s_j 的许可，否则 $x_{k,j}(t) = 0$ 。

考虑上行数据传输，终端 u_k 与基站 s_j 之间数据传输的信噪比为

$$\gamma_{k,j}(t) = \frac{P_{\text{TX},k} g_{k,j}(t)}{B_{k,j} N_0} \quad (2)$$

其中， $P_{\text{TX},k}$ 是传输功率， $g_{k,j}(t)$ 是第 t 个时隙 u_k 与 s_j 之间的信道增益， $B_{k,j}$ 是传输带宽， N_0 是噪声功率谱密度。因此，传输速率^[17]为

$$R_{k,j}(t) = B_{k,j} \log\left(1 + \gamma_{k,j}(t)\right) = B_{k,j} \log\left(1 + \frac{P_{\text{TX},k} g_{k,j}(t)}{B_{k,j} N_0}\right) \quad (3)$$

终端 u_k 在第 t 个时隙传输到基站 s_j 的数据量为

$$U_{k,j}(t) = x_{k,j}(t) \min\{Q_k(t) + \rho A_k(t), \tau R_{k,j}(t)\} \quad (4)$$

终端 u_k 在第 t 个时隙的吞吐量为

$$U_k(t) = \sum_{s_j \in \mathcal{S}} x_{k,j}(t) \min\{Q_k(t) + \rho A_k(t), \tau R_{k,j}(t)\} \quad (5)$$

2.3 能量效率模型

在第 t 个时隙，终端 u_k 将任务数据卸载到基站 s_j 的能耗为传输功率与传输时延的乘积，即

$$E_{k,j}(t) = \frac{P_{\text{TX},k} U_{k,j}(t)}{R_{k,j}(t)} \quad (6)$$

u_k 将任务数据卸载到 s_j 的能量效率定义为吞吐量 $U_{k,j}(t)$ 与能耗 $E_{k,j}(t)$ 的比值，即单位能量可以传输的数据量大小，单位为 bit/J^[18]，表示为

$$EE_{k,j}(t) = \begin{cases} \frac{U_{k,j}(t)}{E_{k,j}(t)}, x_{k,j}(t) = 1 \\ 0, x_{k,j}(t) = 0 \end{cases} \quad (7)$$

2.4 接入服务质量需求模型

令 $X_{k,T}$ 和 T'_k 分别表示终端 u_k 在 T 个时隙获得基站授权的总时隙数和处于活跃态的总时隙数，则

$$X_{k,T} = \sum_{t \in T} \sum_{s_j \in \mathcal{S}} x_{k,j}(t) \quad (8)$$

$$T'_k = \sum_{t \in T} a_k(t) \quad (9)$$

接入服务质量需求模型可定义为

$$\frac{X_{k,T}}{T'_k} \geq \eta_k \quad (10)$$

其中， $\eta_k \in (0, 1]$ 表示 u_k 的接入服务质量约束。考虑到电力物联网终端差异化的接入服务质量约束，可以通过合理设置不同类型终端的 η_k 参数值，满足海量终端差异化的接入需求。

2.5 优化目标

本文的优化目标是在接入服务质量需求的长期约束下，最大化网络总能量效率。因此，优化目标表示为

$$\begin{aligned} \text{P1: } & \max_{\{x_{k,j}(t)\}} \sum_{s_j \in \mathcal{S}} \sum_{u_k \in \mathcal{U}_j} \sum_{t \in T} x_{k,j}(t) EE_{k,j}(t) \\ \text{s.t. } & \text{C}_1: \sum_{u_k \in \mathcal{U}_j} x_{k,j}(t) = M, \forall s_j \in \mathcal{S}, \forall t \in T \\ & \text{C}_2: \sum_{s_j \in \mathcal{S}} x_{k,j}(t) = 1, \forall u_k \in \mathcal{U}, \forall t \in T \\ & \text{C}_3: \frac{X_{k,T}}{T'_k} \geq \eta_k, \forall u_k \in \mathcal{U} \end{aligned} \quad (11)$$

其中， C_1 表示基站可授权的终端数量为 M ； C_2 表示在每个时隙，每个终端只能被一个基站授权进行数据传输； C_3 为接入服务质量需求长期约束。

3 算法设计

3.1 问题转化

由于短期的基站决策与长期的优化目标和约束相耦合，P1 难以直接求解。借助李雅普诺夫优化中虚拟队列^[19]的概念，可以将接入服务质量需求长期约束 C_3 转化为队列稳定性约束。定义一个接入服务质量需求赤字虚拟队列 $F_k(t)$ ，其更新计算式为

$$F_k(t+1) = \begin{cases} \max \left\{ F_k(t) - \sum_{s_j \in \mathcal{S}} x_{k,j}(t) + \eta_k, 0 \right\}, a_k(t) = 1 \\ F_k(t), a_k(t) = 0 \end{cases} \quad (12)$$

式(12)的意义为, 终端 u_k 在第 t 个时隙实际接入性能与规定要求之间的偏差。

根据李雅普诺夫理论, 定义向量 $\boldsymbol{\psi}(t) = [F_k(t)]$, 李雅普诺夫函数表示为

$$L(\boldsymbol{\psi}(t)) = \frac{1}{2} \left\{ \sum_{u_k \in \mathcal{U}} F_k(t)^2 \right\} \quad (13)$$

李雅普诺夫漂移定义为 $L(\boldsymbol{\psi}(t))$ 在连续 2 个时隙变化的期望值, 表示为

$$\Delta L(\boldsymbol{\psi}(t)) = \mathbb{E}[L(\boldsymbol{\psi}(t+1)) - L(\boldsymbol{\psi}(t)) | \boldsymbol{\psi}(t)] \quad (14)$$

在接入服务质量需求的长期约束下, 定义漂移减奖励来权衡最小化漂移或最大化奖励, 即最小化接入服务质量需求赤字或最大化能量效率, 表示为

$$\Delta_V L(\boldsymbol{\psi}(t)) = \Delta L(\boldsymbol{\psi}(t)) - V \mathbb{E}[\zeta(t) | \boldsymbol{\psi}(t)] \quad (15)$$

其中, $\zeta(t) = \sum_{s_j \in \mathcal{S}} \sum_{u_k \in \mathcal{N}_j} \text{EE}_{k,j}(t)$; V 是非负权重参数,

其值越大越倾向于最大化能量效率, 反之越倾向于最小化接入服务质量需求赤字。

将式(13)和式(14)代入式(15)并化简, 可得漂移减奖励的上界为

$$\Delta_V L(\boldsymbol{\psi}(t)) \leq C + \sum_{u_k \in \mathcal{N}_j} a_k(t) F_k(t) \left[\sum_{s_j \in \mathcal{N}_j} x_{k,j}(t) - \eta_k \right] - \sum_{s_j \in \mathcal{S}} \sum_{u_k \in \mathcal{N}_j} \text{EE}_{k,j}(t) \quad (16)$$

其中, C 为常数, 不影响李雅普诺夫优化。因此, 将 P1 转化为最小化漂移减奖励的上界(或最大化漂移减奖励上界的相反数), 表示为

$$\begin{aligned} \text{P2: } & \max_{\{x_{k,j}(t)\}} \sum_{s_j \in \mathcal{S}} \sum_{u_k \in \mathcal{N}_j} x_{k,j}(t) \boldsymbol{\theta}_{k,j}(t) \\ & \text{s.t. } C_1, C_2 \end{aligned} \quad (17)$$

其中, $\boldsymbol{\theta}_{k,j}(t)$ 是漂移减奖励上界的相反数, 即能量效率和接入性能的加权和, 表示为

$$\begin{aligned} \boldsymbol{\theta}_{k,j}(t) = & a_{k,j}(t) \left[V_{\text{EE}} x_{k,j}(t) \text{EE}_{k,j}(t) - \right. \\ & \left. V_F F_k(t) \left(\sum_{s_j \in \mathcal{S}} x_{k,j}(t) - \eta_k \right) \right] \end{aligned} \quad (18)$$

其中, V_{EE} 和 $V_F F_k(t)$ 分别为能量效率和接入性能的权重。相似的计算式推导可见文献[20-21]。

3.2 MAB 理论与 UCB 算法

转化后的问题 P2 可以建模为一个多臂老虎机 (MAB, multi-armed bandit) [22] 问题, 并利用上置信界 (UCB, upper confidence bound) [23] 算法求解。

传统 MAB 模型假设一个老虎机存在多个摇臂, 玩家反复从多个摇臂中选择一个执行, 每次执行后, 玩家都会从一个稳定的概率分布中得到一个数值奖励, 即回报。MAB 问题是指玩家如何在有限次的动作中最大化累积收益。

UCB 算法是强化学习中解决 MAB 问题的一个有效方法。玩家在每次决策中选择性能上界估计值最大的摇臂, 观察收益, 并更新其性能经验估计值作为下一次决策的依据。第 k 个摇臂的性能上界为

$$V_k(t) = \frac{z_k(t)}{n_k(t)} + \sqrt{\frac{\zeta \ln(t)}{n_k(t)}} \quad (19)$$

其中, $z_k(t)$ 表示到时隙 t 选择第 k 个摇臂获得的累积回报; $n_k(t)$ 表示到时隙 t 第 k 个摇臂被选择的总次数; $\zeta > 0$ 表示算法对探索的偏好, 其值越大代表越倾向于探索, 反之表示越倾向于利用。 $z_k(t) / n_k(t)$ 表示到时隙 t 第 k 个摇臂的性能经验估计值。置信区间 $\sqrt{\zeta \ln(t) / n_k(t)}$ 表示估计值的不确定度, 其值随着该摇臂被选择次数的增加而减小, 意味着性能经验估计值逐渐接近摇臂的实际期望值。

3.3 CLAC 算法

基于 MAB 理论, 本文将基站和终端分别建模为玩家与摇臂, 并利用基于终端状态感知的 UCB 算法求解基站累积收益最大化问题。

传统 MAB 问题假设所有终端在每个时隙都是可用的, 并不适用于终端具有活跃和休眠 2 种状态的场景。因此, 本文考虑一种改进的动态 MAB 问题, 即活跃终端的集合是随时间动态变化的。然而, 快速上行链路授权架构下, 基站无法感知所有终端的状态, 当基站在时隙 t 选择了一个处于休眠态的终端时, 由于该终端没有数据传输需求, 会造成资源浪费。因此, 本文考虑基站具有一个活跃终端预测算法 [24], 预测算法可根据网络流量模型在每个时隙预测终端 u_k 的活跃概率 $P_k(t)$, 并建立一个活跃终端集合。本文所提 CLAC 算法可与多种活跃终端预测算法组合使用, 具有较强的扩展性和兼容性。

考虑 2 种类型的网络流量, 即周期性流量和事

件驱动型流量。由于传统 UCB 算法在解决动态 MAB 问题时具有局限性，本文结合预测算法对其进行了改进，并在传统 UCB 算法的基础上加入了终端状态感知，其性能上界为

$$V_{k,j}(t) = P_k(t) \left(\frac{z_{k,j}(t)}{n_{k,j}(t)} + \sqrt{\frac{\zeta \ln t'_k}{n_{k,j}(t)}} \right) \quad (20)$$

其中， $z_{k,j}(t)$ 表示到时隙 t 基站 s_j 选择终端 u_k 获得的累积回报， $n_{k,j}(t)$ 表示到时隙 t 终端 u_k 处于活跃态且被授权的总时隙数， t'_k 表示到时隙 t 终端 u_k 处于活跃态的总时隙数。与传统 UCB 算法相比，本文算法只计算所选终端活跃的总时隙数，而不是算法迭代的总次数，能够保证 $V_{k,j}(t)$ 的计算更准确。 $z_{k,j}(t)$ 、 $n_{k,j}(t)$ 以及 t'_k 的更新式分别为

$$\begin{aligned} z_{k,j}(t+1) &= \begin{cases} z_{k,j}(t) + \theta_{k,j}(t), x_{k,j}(t) = 1 \cap a_k(t) = 1 \\ z_{k,j}(t), x_{k,j}(t) = 0 \cup a_k(t) = 0 \end{cases} \\ n_{k,j}(t+1) &= \begin{cases} n_{k,j}(t) + 1, x_{k,j}(t) = 1 \cap a_k(t) = 1 \\ n_{k,j}(t), x_{k,j}(t) = 0 \cup a_k(t) = 0 \end{cases} \\ t'_k &= \begin{cases} t'_k + 1, a_k(t) = 1 \\ t'_k, a_k(t) = 0 \end{cases} \end{aligned} \quad (21)$$

本文所提 CLAC 算法如算法 1 所示，包含 3 个阶段，分别为初始化阶段、决策阶段和学习阶段。

算法 1 CLAC 算法

输入 $\zeta, V_{EE}, V_F, N_0, T, M, \{\beta_k\}, \{\eta_k\}$

初始化

1) 设置 $Q_k(0)$ 为队列初始值，令 $F_k(t) = 0$ ， $x_{k,j}(t) = 0$ ， $a_k(t) = 0$ ， $\theta_{k,j}(t) = 0$ ， $\forall u_k \in \mathcal{U}$ ， $\forall s_j \in \mathcal{S}$ ， $\forall t \in \mathcal{T}$ 。

2) 循环 while $t \leq T$

决策

3) for $j = 1:1:J$

4) 基站 s_j 预测活跃终端集合

5) for $k = 1:1:K$

6) 基站 s_j 根据式(20)计算终端 u_k 的

$V_{k,j}(t)$

7) if $u_k \in \mathcal{N}_{j \cap j'}$

8) 比较 $V_{k,j}(t)$ 与 $V_{k,j'}(t)$ ，并将其归入函

数值较大的基站活跃终端集合中

9) end if

10) end for

11) 基站 s_j 对 $V_{k,j}(t)$ 进行排序，选择值最大的前 M 个终端分发上行链路许可

12) end for

学习

13) for $k = 1:1:K$

14) for $j = 1:1:J$

15) if $x_{k,j}(t) = 1$

16) 终端 u_k 观察 $U_{k,j}(t)$ 和 $E_{k,j}(t)$ ，计算能量效率 $EE_{k,j}(t)$

17) 终端 u_k 根据式(18)计算 $\theta_{k,j}(t)$ ，并将其值反馈给基站 s_j

18) end if

19) end for

20) 终端 u_k 根据式(1)和式(12)更新 $Q_k(t)$ 和 $F_k(t)$

21) end for

22) for $j = 1:1:J$

23) for $k = 1:1:K$

24) 基站依据式(21)更新 $z_{k,j}(t)$ 、 $n_{k,j}(t)$ 和 t'_k

25) end for

26) end for

27) $t = t + 1$

本文所提 CLAC 算法具有以下优点。

1) 终端状态感知。基站基于终端状态感知仅在活跃终端集合中选择终端授权，避免因将授权分发给休眠终端而导致资源浪费，降低网络性能。

2) 接入服务质量需求感知。基站基于接入服务质量需求感知动态优化终端授权策略。例如，当终端 u_k 实际接入性能与规定要求偏离严重时， $F_k(t)$ 会逐渐增大，使基站为其授权，保证其接入性能。

4 仿真分析

4.1 仿真参数设置

本文通过对比不同的仿真算法来验证所提 CLAC 算法的性能，对比算法设置如下。

1) 能量效率接入控制 (EEAC, energy-efficient access control) 算法。该算法基于终端状态预测算法最大化网络总能量效率，而未考虑接入服务质量需求长期约束。

2) 基于学习的接入控制 (LAC, learning-based access control) 算法。该算法在接入服务质量需求

长期约束下最大化网络总能量效率,但未考虑终端状态预测。

3) 快速上行链路授权(FULG, fast uplink grant)算法^[8]。该算法随机向终端分发上行链路授权,未考虑终端接入服务质量需求约束、能量效率以及状态预测。

仿真参数设置如表 2 所示。

4.2 仿真结果分析

不同算法下网络性能的对比如图 2 所示。图 2(a)~图 2(d)分别显示了 4 种算法下网络平均能量效率、平均队列积压、满足接入服务质量需求的终端比例和平均接入服务质量需求赤字积压随时隙的变化情况。

由图 2(a)可以看出,由于只考虑了能量效率优化,EEAC 算法的能量效率性能最优,但随着时隙的增加,所提 CLAC 算法逐渐逼近 EEAC 算法的性能。

由图 2(b)可以看出,所提 CLAC 算法性能最优,能够将队列积压维持在较低水平,相较于 EEAC 算法、LAC 算法和 FULG 算法,平均队列积压分别降低了 77.90%、97.68%和 83.83%。

参数	取值
J	3
K	3 000
M	300
T	500
τ / s	1
A_k	[4,6]
ρ / bit	10^4
$P_{\text{TX},k} / \text{W}$	0.1
$B_{k,j} / \text{kHz}$	200
$N_0 / (\text{W} \cdot \text{Hz}^{-1})$	10^{-12}
V_{EE}	5
V_F	5
ζ	0.2

由图 2(c)可以看出,所提 CLAC 算法性能明显优于其他 3 种算法,相较于 EEAC 算法、LAC 算法和 FULG 算法,满足接入服务质量需求的终端比例分别提高了 15.07%、77.46%和 54.95%。综合图 2(a)和图 2(c)可以看出,CLAC 算法实现了能量效率与

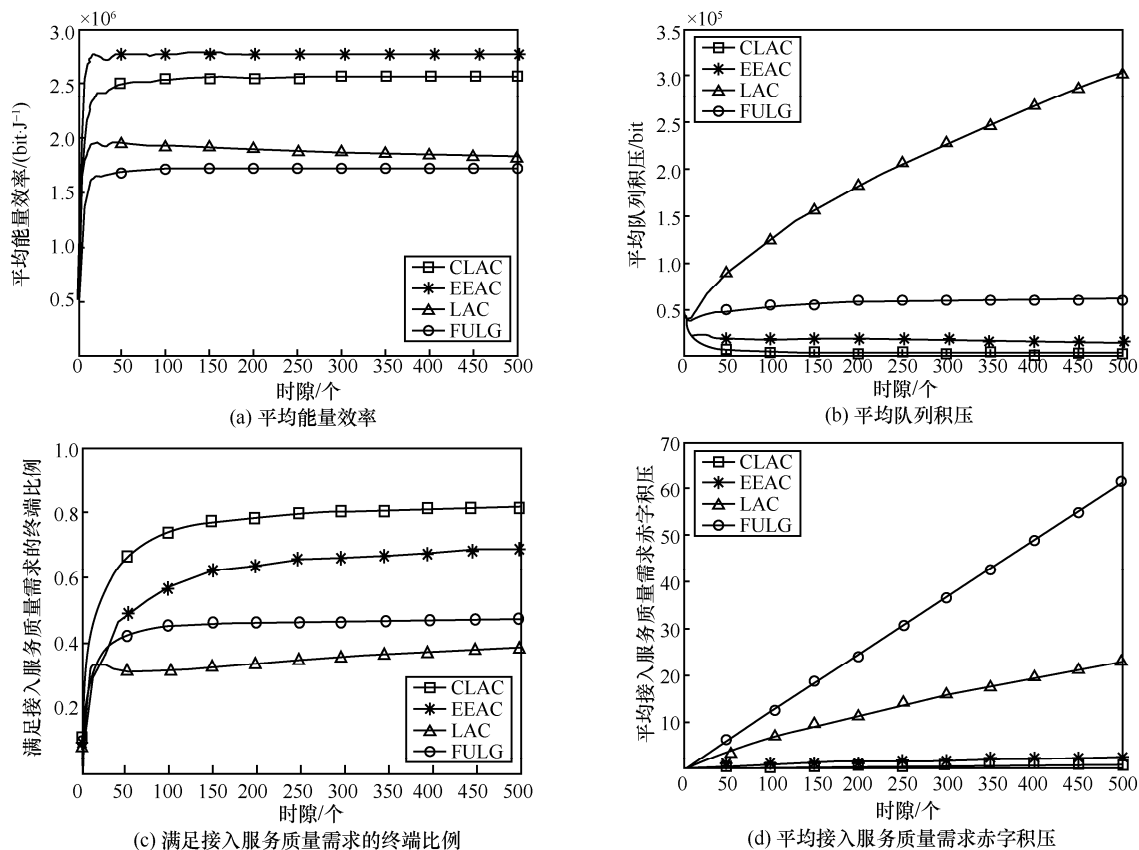


图 2 不同算法下网络性能对比

接入性能的平衡,在保证终端接入服务质量需求的前提下,尽可能使网络总能量效率最大化,而EEAC算法则通过牺牲终端接入性能换取更高的能量效率。

由图2(d)可以看出,由于具备终端状态感知和接入服务质量需求感知,CLAC算法可以使接入服务质量需求赤字积压最小。结合图2(c)和图2(d)可知,LAC算法满足接入服务质量需求的终端比例较低,但接入服务质量需求赤字积压较小,这是因为LAC算法考虑了接入服务质量需求约束,将授权频繁地分发给接入服务质量需求较高的终端,而忽略了大量接入服务质量需求较低的终端。综合图2可以看出,CLAC算法综合性能最优。

图3显示了CLAC算法终端平均能量效率和满足接入服务质量需求的终端比例随预测准确率的变化情况。可以看出,随着预测准确率的提高,终端平均能量效率与满足接入服务质量需求的终端比例均增大。原因在于,当预测准确率较低时,基站会频繁地为休眠终端分发授权,导致活跃终端无法接入,造成资源浪费,降低网络整体性能。当预测准确率达到0.6时,CLAC算法平均能量效率超过 2×10^6 bit/J,满足接入服务质量需求的终端比例可达0.6。结合图2可以看出,当预测准确率达到0.6及以上时,CLAC算法性能均优于LAC算法和FULG算法。文献[9, 24]研究了终端状态预测算法,且具有良好的预测准确率。因此,CLAC算法具有实用性。

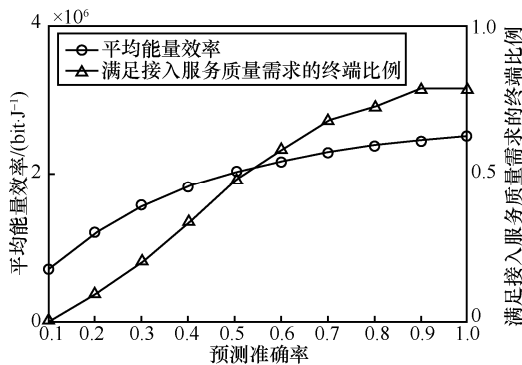


图3 CLAC算法性能随预测准确率的变化

图4显示了终端平均能量效率和满足接入服务质量需求的终端比例随 α 的变化情况,其中 α 为 V_{EE} 与 V_F 的比值,用于表征终端对能量效率和接入服务质量需求的关注度。仿真结果表明,随着 α 的增大,终端越来越关注能量效率而忽略接入性能,因此终端平均能量效率逐渐增加,满足接入服务质量需求的终端比例逐渐下降。此仿真结果为实际应

用中 V_{EE} 与 V_F 的设置提供了参考,在接入服务质量需求约束下,通过合理设置 V_{EE} 与 V_F 参数值可以提高网络性能。

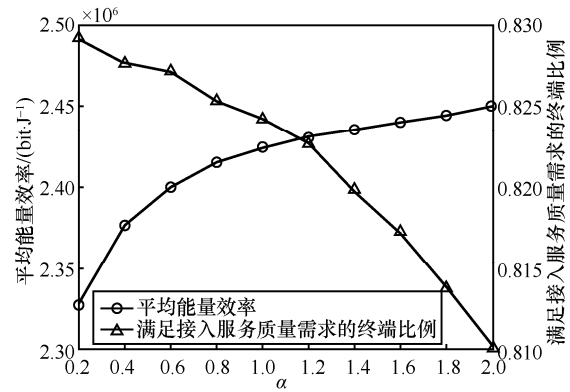


图4 CLAC算法性能随 α 的变化

定义终端活跃而未被授权的情况为接入失败。图5分别显示了4种算法下终端接入失败次数随接入服务质量需求的变化关系,以一个终端为例,图5中的点表示其接入失败。由图5(a)可以看出,在CLAC算法中,接入失败次数随着接入服务质量需求的增大而减小。从图5(b)可以看出,在EEAC算法中,接入服务质量需求为0.8、0.9的终端接入失败次数明显多于接入服务质量需求为0.6的终端,且与接入服务质量需求为0.2、0.3的终端基本相同,终端接入失败次数并不会随接入服务质量需求的增加而减少,因此EEAC算法不具有接入服务质量需求感知能力,无法满足终端差异化的接入服务质量需求。由于LAC算法与CLAC算法均具有接入服务质量需求感知能力,因此LAC算法接入失败情况的趋势与CLAC算法基本一致。但LAC算法不具有终端状态感知能力,导致活跃终端接入失败次数增加。由图5(a)和图5(c)可以看出,LAC算法较CLAC算法点更密集,即终端接入失败次数更多。分析数值结果表明,相较于LAC算法,CLAC算法的接入失败次数可降低45.97%。FULG算法下终端接入失败的总次数较多,且不同接入服务质量需求的终端接入失败情况大致相同,终端的差异化接入服务质量需求未得到满足。

5 结束语

本文针对电力物联网海量终端接入冲突严重、队列积压大、能量效率低等问题,提出了具有终端状态感知和接入服务质量需求感知能力的CLAC算

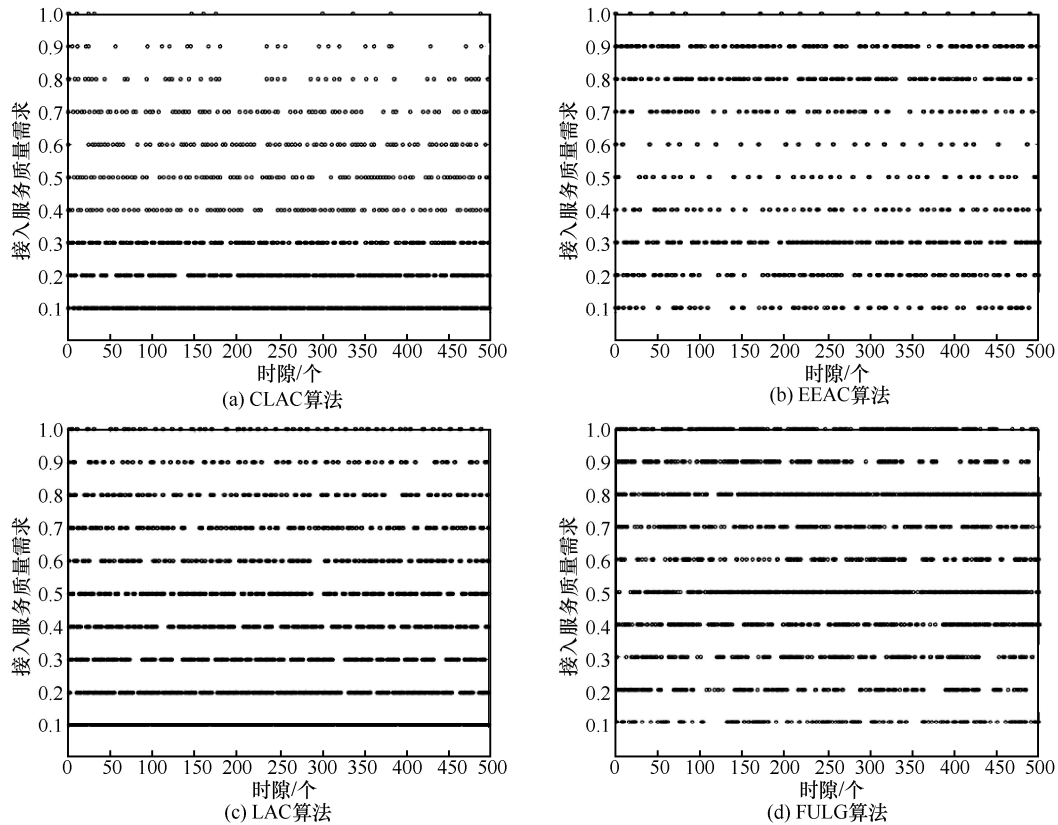


图 5 终端接入失败次数随接入服务质量需求的变化

法。该算法在满足终端接入服务质量需求的同时，能够提高网络总能量效率，降低队列积压。仿真结果表明，相较于 EEAC 算法、LAC 算法和 FULG 算法，CLAC 算法满足接入服务质量需求的终端比例可分别提高 15.07%、77.46%和 54.95%，平均队列积压可分别降低 77.90%、97.68%和 83.83%。未来将进一步研究基于流量建模和终端状态预测的 6G 电力物联网接入控制算法。

参考文献:

[1] 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): 5-16.
 FU Z X, LI X Y, YUAN Y. Research on key technologies of ubiquitous power Internet of things[J]. Electric Power Construction, 2019, 40(5): 5-16.

[2] 方敏, 段向阳, 胡留军. 6G 技术挑战、创新与展望[J]. 中兴通讯技术, 2020, 26(3): 61-70.
 FANG M, DUAN X Y, HU L J. Challenges, innovations and perspectives towards 6G[J]. ZTE Technology Journal, 2020, 26(3): 61-70.

[3] 王承祥, 黄杰, 王海明, 等. 面向 6G 的无线通信信道特性分析与建模[J]. 物联网学报, 2020, 4(1): 19-32.
 WANG C X, HUANG J, WANG H M, et al. 6G oriented wireless communication channel characteristics analysis and modeling[J]. Chi-

nese Journal on Internet of Things, 2020, 4(1): 19-32.

[4] HE H L, REN P Y, DU Q H, et al. Estimation based adaptive ACB scheme for M2M communications[C]//2015 International Conference on Wireless Algorithms, Systems and Applications. Berlin: Springer, 2015: 165-174.

[5] HARWAHYU R, WANG X, SARI R, et al. Analysis of group paging with pre-backoff[J]. EURASIP Journal on Wireless Communications and Networking, 2015, 1(2): 34-42.

[6] DHILLON H S, HUANG H C, VISWANATHAN H, et al. On resource allocation for machine-to-machine (M2M) communications in cellular networks[C]//2012 IEEE Globecom Workshops. Piscataway: IEEE Press, 2012: 1638-1643.

[7] BOCKELMANN C, PRATAS N, NIKOPOUR H, et al. Massive machine-type communications in 5G: physical and MAC-layer solutions[J]. IEEE Communications Magazine, 2016, 54(9): 59-65.

[8] 3GPP. Study on latency reduction techniques for LTE: 3GPP TR 36.881 version 14.0.0 Release 14[S]. 2016.

[9] ALI S, RAJATHEVA N, SAAD W. Fast uplink grant for machine type communications: challenges and opportunities[J]. IEEE Communications Magazine, 2019, 57(3): 97-103.

[10] KIM W, JI H, SHIM B. Fast uplink access in TDD systems for ultra reliable and low latency communications[C]//2018 IEEE Region 10 Conference. Piscataway: IEEE Press, 2018: 1616-1619.

[11] 严牧, 孙耀, 冯钢. 基于强化学习的无线网络智能接入控制技术[J]. 中兴通讯技术, 2018, 24(2): 10-14, 46.
 YAN M, SUN Y, FENG G. The intelligent access control mechanisms

in wireless network based on reinforcement learning[J]. ZTE Technology Journal, 2018, 24(2): 10-14, 46.

- [12] HASAN M, HOSSAIN E, NIYATO D. Random access for machine-to-machine communication in LTE-advanced networks: issues and approaches[J]. IEEE Communications Magazine, 2013, 51(6): 86-93.
- [13] SUN H, NALLANATHAN A, WANG C, et al. Wideband spectrum sensing for cognitive radio networks: a survey[J]. IEEE Wireless Communications, 2013, 20(2): 74-81.
- [14] BARVE S S, KULKARNI P. Dynamic channel selection and routing through reinforcement learning in cognitive radio networks [C]//IEEE International Conference on Computational Intelligence and Computing Research. Piscataway: IEEE Press, 2013: 1-7.
- [15] ALI S, FERDOWSI A, SAAD W, et al. Sleeping multi-armed bandit learning for fast uplink grant allocation in machine type communications[J]. IEEE Transactions on Communications, 2020, 68(8): 5072-5086.
- [16] 周振宇, 陈亚鹏, 潘超, 等. 面向智能电力巡检的高可靠低时延移动边缘计算技术[J]. 高电压技术, 2020, 46(6): 1895-1902.
ZHOU Z Y, CHEN Y P, PAN C, et al. Ultra-reliable and low-latency mobile edge computing technology for intelligent power inspection[J]. High Voltage Engineering, 2020, 46(6): 1895-1902.
- [17] LIAO H, ZHOU Z, ZHAO X, et al. Learning-based context-aware resource allocation for edge computing-empowered industrial IoT[J]. IEEE Internet of Things Journal, 2020, 7(5): 4260-4277.
- [18] VERDU S. Spectral efficiency in the wideband regime[J]. IEEE Transactions on Information Theory, 2002, 48(6): 1319-1343.
- [19] NEELY M J. Energy optimal control for time-varying wireless networks[J]. IEEE Transactions on Information Theory, 2006, 52(7): 2915-2934.
- [20] ZHOU Z Y, GUO Y F, HE Y H, et al. Access control and resource allocation for M2M communications in industrial automation[J]. IEEE Transactions on Industrial Informatics, 2019, 15(5): 3093-3103.
- [21] TANG L, WEI Y N, CHEN W, et al. Delay-aware dynamic resource allocation and ABS configuration algorithm in HetNets based on Lyapunov optimization[J]. IEEE Access, 2017, 5: 23764-23775.
- [22] MODI N, MARY P, MOY C. QoS driven channel selection algorithm for cognitive radio network: multi-user multi-armed bandit approach[J]. IEEE Transactions on Cognitive Communications and Networking, 2017, 3(1): 49-66.
- [23] SUTTON R S, BARTO A G. Reinforcement learning: an introduction[M]. Massachusetts: MIT Press, 1998.
- [24] ALI S, SAAD W, RAJATHEVA N. A directed information learning framework for event-driven M2M traffic prediction[J]. IEEE Communications Letters, 2018, 22(11): 2378-2381.

[作者简介]



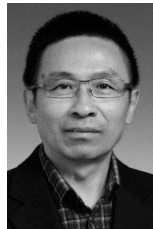
周振宇 (1983-), 男, 河北张家口人, 博士, 华北电力大学教授、博士生导师, 主要研究方向为无线通信网络与新技术、物联网与现代传感技术、能源互联网信息通信技术。



贾泽晗 (1998-), 男, 河北保定人, 华北电力大学硕士生, 主要研究方向为电力物联网、智能电网中无线通信技术等。



廖海君 (1997-), 女, 广西贵港人, 华北电力大学博士生, 主要研究方向为电力物联网、智能电网中无线通信技术等。



赵雄文 (1964-), 男, 陕西延安人, 博士, 华北电力大学教授、博士生导师, 主要研究方向为 5G 和后 5G 无线通信、电力系统通信等。



张磊 (1989-), 男, 河南新乡人, 国网山东省电力公司高级工程师, 主要研究方向为电力系统运行与控制等。