

## 面向数据中心绿色可靠运行的强化学习方法

贾庆山<sup>1</sup>, 唐静娴<sup>1</sup>, 吴俊杰<sup>1</sup>, 胡潇<sup>2</sup>, 林依挺<sup>3</sup>, 夏恒<sup>3</sup>

(1. 清华大学自动化系智能与网络化系统研究中心, 北京 100084;

2. 北京化工大学信息科学与技术学院, 北京 100029; 3. 腾讯 IDC 平台部, 广东 深圳 518052)

**摘要:** 数据中心的绿色可靠运行具有重大的社会经济价值。综述了面向数据中心绿色可靠运行的优化与控制方法, 提出一种事件驱动的强化学习方法, 用于提升运行能效; 提出一种电池寿命预测方法, 提升了预测精度。

**关键词:** 数据中心; 信息物理融合能源系统; 强化学习; 事件驱动的优化

**中图分类号:** TP29

**文献标识码:** A

**doi:** 10.11959/j.issn.2096-6652.202036

## Reinforcement learning for green and reliable data center

JIA Qing-Shan<sup>1</sup>, TANG Jingxian<sup>1</sup>, WU Junjie<sup>1</sup>, HU Xiao<sup>2</sup>, LIN Yiting<sup>3</sup>, XIA Heng<sup>3</sup>

1. Center for Intelligent and Networked Systems, Department of Automation, Tsinghua University, Beijing 100084, China

2. College of Information Science and Technology, Beijing University of Chemical Technology, Beijing 100029, China

3. IDC Platform Department, Tencent Corporation, Shenzhen 518052, China

**Abstract:** It is of significant social and economical impact to achieve green and reliable operation of data center. The optimization and control methods for green and reliable data center were reviewed briefly. An event-based reinforcement learning approach for improving the energy efficiency was developed. And a method to improve the accuracy of battery lifetime forecasting was developed.

**Key words:** data center, cyber physical energy system, reinforcement learning, event-based optimization

### 1 引言

数据中心已经成为现代社会的基础设施之一, 它不仅承载大量的互联网应用, 而且伴随着物联网技术的普及、5G 商用、边缘计算走进生活, 当今大型数据中心的计算能力、信息处理能力、网络通信能力等已经成为国家核心竞争力的标志之一。2020 年年初开始的新型冠状病毒肺炎疫情深刻改变了人们生活、学习、工作的方式, 线上购物、在线教学、远程办公成为重要趋势。数据中心作为承载线上教学、居家办公新模式的核心基础设施, 已

经获得世界各国的高度重视。数据中心已被列入我国 2020 年发布的“新基建”重点领域之一。

数据中心在过去 10 年间高速发展, 与此同时, 其运行能耗也与日俱增。据预测, 到 2030 年, 电信系统年用电量将占全球年用电总量的 20%以上, 其中数据中心的运行能耗将达到 3 000 TWh, 占全球年用电总量的 8%<sup>[1]</sup>。因而, 研究数据中心的运行节能具有重要的社会意义和经济价值。

世界上拥有大型数据中心的知名公司, 如国外的谷歌、亚马逊、脸书等, 国内的阿里、腾讯、华为、百度等, 均高度重视数据中心的运行能效。

收稿日期: 2020-11-18; 修回日期: 2020-12-02

通信作者: 贾庆山, jiaqs@tsinghua.edu.cn

基金项目: 国家自然科学基金资助项目(No.61673229, No.62073182); 国家重点研发计划基金资助项目(No.2016YFB0901900, No.2017YFC0704100); 111 创新引智基地计划(No.BP2018006)

**Foundation Items:** The National Natural Science Foundation of China (No.61673229, No.62073182), National Key Research and Development Program of China (No.2016YFB0901900, No.2017YFC0704100), 111 International Collaboration Project (No.BP2018006)

2016年,谷歌公司的DeepMind团队与数据中心部门合作,显著降低了数据中心的电源使用效率(power usage effectiveness, PUE)<sup>[2]</sup>。2020年,谷歌大型数据中心的平均PUE为1.11,个别数据中心的PUE甚至可以达到1.06<sup>[3]</sup>。然而,目前全球数据中心的平均PUE为1.8,我国为1.67。因而,数据中心的运行节能具有巨大潜力。

数据中心的运行节能面临多方面挑战。第一,高维信息。现代数据中心一般部署了大量传感器,传感器定期监测数据中心机柜的温度场分布、制冷系统的冷冻水水温与水流量等信息,并汇聚到大数据运行平台。高维海量数据中的信息方差小,为决策优化带来了挑战。第二,耦合动态过程。大型数据中心一般由多栋建筑物构成,建筑面积达到上万平方米,甚至更大。数据中心的发展呈现出模块化趋势,在每个模块化数据中心(modularized data center, MDC)内,由冷热通道将机柜群分隔,机柜整齐排列在通道两侧。冷热通道之间的空气流动相对独立,各机柜内的温度场分布受服务器所承载的计算任务影响,动态变化,关联耦合。这一复杂的动态过程虽有机理模型,但是精细的性能评价仍需基于仿真模拟甚至物理实验。制冷系统的循环过程也需仿真模拟方可评价控制策略的性能,这导致对策略性能的评价非常耗费时间。第三,计算业务的需求随机波动。数据中心支撑的计算任务种类丰富、差异明显,计算业务的需求量具有明显的随机波动性,难以精准预测。因而对制冷系统的控制策略和对计算作业的调度策略的性能评价,一般需要多次仿真、多次实验取均值才能获得较为精确的评价。这使得策略性能评价耗费了大量的时间与代价。如何在上述挑战下实现数据中心的运行节能是一项重大挑战。

作为基础设施,数据中心运行的稳定性与可靠性至关重要。为保障电源供应的稳定,数据中心一般配有备用发电机(柴油发电机、氢能源发电机等)及不间断电源(uninterruptible power supply, UPS)。UPS的核心构件是电池。数据中心一般建有针对电池工作状况的在线监测系统,将工作电压、温度等数据回传至大数据平台。虽然有许多针对电池寿命预测的机理模型与实验研究,但是现有方法一般需要通过充放电才可以获得较为详细的内部电化学反应数据。而数据中心UPS的电池一般处于浮充状态,很少处于放电状态。这导致电池监测数据中鲜有适用于经典分析方法的数据。如何将专家知识与

监测数据相结合,提升对电池潜在故障的预测精度,更准确地进行电池的预防性维护与更换,是一项重大的科学挑战,具有重要的实际意义。

## 2 数据中心的仿真优化方法综述

数据中心的绿色可靠运行有较多的相关研究工作。本节简要阐述与本文关系最密切的内容。

近年来,数据中心呈现出聚集化效应。大型数据中心的PUE低于中小型数据中心<sup>[4]</sup>。鉴于小型数据中心提升能效的潜力有限,越来越多的大型数据服务公司将其数据中心整合在一起。目前在全球数据中心的能耗中,大型数据中心的能耗占比已经超过1/4且在不断增加。为了提升数据中心的能效水平,可以在数据中心周边建设新能源发电厂(如太阳能、风能等),也可在电力市场购买新能源发电量,从而降低对常规电力的消耗。谷歌、脸书、亚马逊等公司先后宣布了可持续发展的路线图,计划逐步实现自有数据中心用新能源供电。2017年,谷歌实现了自有大型数据中心的100%新能源供电,更在2020年9月宣布实现了自有数据中心历史上所有碳排放的零中和(即自有数据中心历史耗电总碳排放已经通过植树、碳交易等方式中和)。可以预见,数据中心的碳中和的可持续发展模式是一种趋势。

数据中心的能耗与设计、选址、选型关系密切,因此利用海水、湖泊、极地、高寒等自然冷源是重要手段<sup>[5]</sup>。比起个案的特殊设计,推进数据中心的标准化是更重要的工作,更易被大量采用,有利于提升众多数据中心的平均能效。模块化数据中心(如图1所示)在这一发展中起到了重要作用,通过标准化接口设计、制冷、配电、UPS、封闭式机柜等模块,数据中心的布设成本与周期大大降低。腾讯的T-block系列产品是其中的代表之一。



图1 模块化数据中心结构示意图<sup>[6]</sup>

与设计优化相比,数据中心的运行节能是本文更加关注的内容。数据中心的内部温度场分布、制冷系统的构成包括众多原件,虽然有理论上的机理

模型，但是一般需要仿真模型进行更细致的刻画。常见的仿真软件包括 6SigmaDC、FLUENT、TileFlow、FloVENT、FloTHERM 和 EnergyPlus 等，这些计算流体动力学（computational fluid dynamics, CFD）商用软件常用于对数据中心气流分布的仿真与计算和热管理。

强化学习是解决策略优化问题的一种重要方法，特别适用于机理模型复杂或未知的系统，但是存在需要大量且丰富的观测数据的问题。利用上述 CFD 仿真软件，已经有学者基于强化学习方法对数据中心的运行节能进行研究。Li Y L 等人<sup>[7]</sup>基于深度确定性策略梯度（deep deterministic policy gradient, DDPG）和行动-评价（actor-critic）框架，提出了一种端到端的制冷控制算法，该算法包含 5 个动作（action）：直接蒸发冷却出口温度、间接蒸发冷却出口温度、冷却水回路出口温度、直接膨胀式冷却盘管出口温度、冷却器冷却空气回路出口温度；两个状态量（state）：环境空气温度、信息技术（information technology, IT）设备负荷系数；一个奖励函数：基于 PUE 和各温度数据计算得到的目标函数值。EnergyPlus 平台的仿真结果表明，该方法可达到较低的 PUE，与手动设定数据中心温度的基准控制方法相比，可节省约 11% 的冷却量。但这种方法是完全由数据驱动的，它的控制决策依赖于清洁的数据，在大型数据中心的清洁度是需要认真考虑的问题。Ran Y Y 等人<sup>[8]</sup>基于深度 Q 网络（deep Q-network, DQN）提出了一种参数化动作空间的方法，解决了离散（IT 设备）-连续（冷却装置）混合的动作空间问题，联合优化了 IT 系统的作业调度和冷却系统的风量调节。该方法使用了 6 个状态量，涉及 IT 系统与热管理，具体包括：可用 CPU 数、CPU 利用率、能量消耗、机架的两个入口温度和一个出口温度；两个动作：精密冷却单元的通风量、任务调度行为；奖励函数：基于 PUE 和惩罚项计算得到的目标函数值。6SigmaDC 软件的仿真结果表明，该算法可以节省 10%~15% 的能量，并且可以在节能和保证服务质量之间实现更好的权衡。Yi D L 等人<sup>[9]</sup>提出了基于深度 Q 网络的数据中心作业调度方法，使得数据中心在达到同样作业吞吐量的同时减小能源消耗。上述基于强化学习方法的数据中心均取得了一定的能源降耗成果，但在实际使用过程中还存在以下问题：第一，若数据中心规模较大、结构较复杂，则

原始数据可能含有较大噪声、缺失、异常值，数据清洗必须做到非常精细；第二，在训练阶段及强化学习算法收敛前，为了充分探索状态-动作空间，可能会尝试激进的动作，这些动作可能会导致严重偏离最佳状态和服务器过热等问题；第三，IT 系统往往在几秒甚至几微秒内做出响应，而冷却系统（机械设备）可能在几分钟内做出响应，这种反应时间的不匹配可能会导致不必要的波动等。

数据中心 UPS 的电池主要采用阀控式密封铅酸（valve-regulated lead acid, VRLA）蓄电池。在实际使用过程中，电池的寿命受板栅材料、运维状态、外部环境等多重因素的影响。常规的电池维护手段包括人工定期检测和故障维修。估计电池的剩余寿命一般需要通过深度充放电实验实现<sup>[10]</sup>，而数据中心 UPS 的电池长期处于浮充状态，完整的放电过程数据较少，这为电池维护带来了挑战。

### 3 节能运行优化

本节以 MDC 末端空调节能优化问题为例，讨论强化学习方法的应用问题。数据中心制冷系统结构如图 2 所示。冷水机组与数据中心机房之间：冷水机组产生冷冻水，经冷冻水循环泵输送到数据中心机房，经过与服务器风水换热后升温，携带热量回到冷水机组。冷水机组与冷却塔之间：高温冷却水经冷却水循环泵送至冷却塔，经风水换热后形成较低温度的冷却水，回到冷水机组，完成循环。

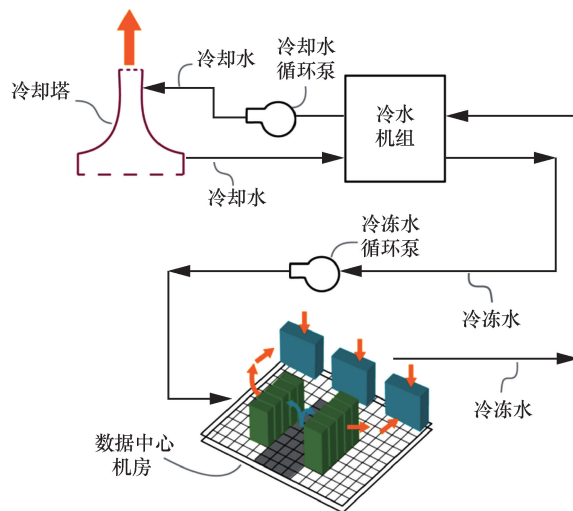


图2 数据中心制冷系统结构

其中，数据中心机房内的冷热交换是一个较为复杂的动态过程，一般需要基于 CFD 进行仿真模

拟。图3给出了6SigmaRoomDC软件中的仿真结果示例,其中,箭头方向代表气流流向,箭头颜色代表气流温度。其仿真时间随仿真实空间颗粒度呈立方增长。是否能用较为粗糙的模型进行较好的策略优化是一个重要的问题。

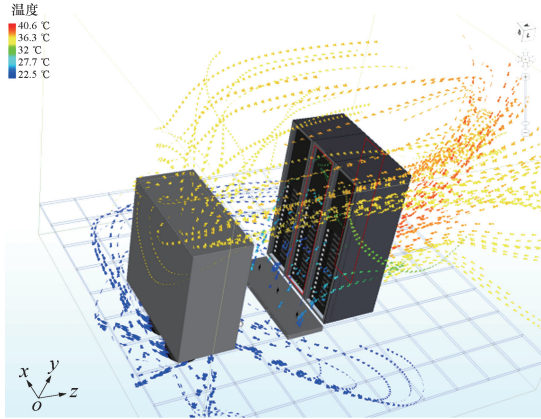


图3 数据中心温度场和气流组织仿真模拟效果

### 3.1 优化模型

将数据中心冷水系统的策略优化问题建模为马氏决策过程。考虑离散时间  $t=0,1,2,\dots,T$  的系统动态过程,每个阶段的时长为  $\Delta t$ 。可以观测的系统变量为  $O_t=(T_{CA}(t), T_{AR}(t), T_{CR}(t), F_C(t), H_{ITE}(t), T_{ITE,in}^i(t), T_{ITE,out}^i(t))$ ,其中  $T_{CA}(t)$  表示冷通道温度,  $T_{AR}(t)$  表示空调回风温度,  $T_{CR}(t)$  表示冷却水回水温度,  $F_C(t)$  表示冷却水流量,  $H_{ITE}(t)$  表示 IT 负载,  $T_{ITE,in}^i(t)$  表示第  $i$  台 IT 设备的进风口温度,  $T_{ITE,out}^i(t)$  表示第  $i$  台设备的出风口温度。工程中一般采用连续  $m$  个时段的系统观测值作为状态,即  $s_k=(O_{m(k-1)+1}, \dots, O_{mk})$ 。行为  $a_k=(S_F, SP)$ ,其中,  $S_F$  表示风机转速,  $SP$  表示送风温度。施加控制行为后,使用持续  $n$  个时段的系统性能计算收益,即:

$$r_k(s_k, a_k) = \lambda_1 r_k(\text{Power}) + \lambda_2 r_k(\text{Temperature}) + \lambda_3 r_k(T\_volatility) \quad (1)$$

其中,  $\lambda_1$ 、 $\lambda_2$ 、 $\lambda_3$  为正常数,表示权重;  $r_k(\text{Power})$  表示末端空调制冷功率对应的收益。

$$r_k(\text{Power}) = - \sum_{p=mk+1}^{mk+n} S_F(p)^3 \quad (2)$$

$r_k(\text{Temperature})$  表示机房绝对温度对应的收益,也是对机房温度越过安全边界的惩罚。比如可以设定冷通道温度的安全边界为  $27^\circ\text{C}$ ,热通道温度的安全边界为  $39^\circ\text{C}$ 。采用  $\log(1+\exp(\cdot)^2)$  表示对冷通道高温的惩罚,其中  $p$  表示时间点。

$$r_k(\text{Temperature}) = - \left\{ \sum_{p=mk+1}^{mk+n} \sum_{l=1}^L \log[1 + \exp(T_{ITE,in}^l(p) - 27)^2] + \log[1 + \exp(T_{ITE,out}^l(p) - 39)^2] \right\} \quad (3)$$

$r_k(T\_volatility)$  为 IT 设备温度波动性的收益,即对 IT 设备温度频繁大幅波动的惩罚。

$$r_k(T\_volatility) = - \sum_{l=1}^L \frac{\sigma[T_{ITE,in}^l(mk+1), \dots, T_{ITE,in}^l(mk+n)]}{\mu[T_{ITE,in}^l(mk+1), \dots, T_{ITE,in}^l(mk+n)]} \geq 0.04 \quad (4)$$

其中,  $\mu[\cdot]$  和  $\sigma[\cdot]$  分别表示序列的样本均值与标准差。考虑有限阶段总收益最大化的马氏决策过程,其目标函数为:

$$J(d) = \sum_{k=0}^{K-1} r_k(s_k, d(s_k)) \quad (5)$$

其中,  $K$  是在  $T$  时段内的决策次数,  $T=mk$ 。记上述马氏决策问题为 P1, 即:

$$\max_d J(d) \quad (6)$$

s.t. 由仿真软件刻画的动态过程及

$$\text{服务器环境温度等约束条件} \quad (7)$$

该模型考虑的是有限阶段的总收益,实际应用中一般会采用滚动优化的方式,前瞻式优化未来  $T$  时段的总收益,求得最优策略后,仅采用策略中针对当前时段的控制行为作用于环境。在下一个决策阶段  $k+1$  到来时,更新信息,重新对  $T$  时段进行策略优化。通过这种滚动方式融入动态信息,调整控制行为。

### 3.2 事件驱动的强化学习方法

问题 P1 对应的马氏决策过程面临高维状态向量的挑战。为了克服大状态空间的困难,可以定义事件并进行状态集结,采用事件驱动的强化学习方法。

事件驱动的强化学习方法适用于基于仿真的策略优化问题,特别是状态空间大、性能评价需要通过仿真进行以及随机的动态变化过程等难点。事件驱动的强化学习方法在能源互联网中的应用可参见参考文献[11]。下面仅简述其关键思想。

事件  $e \in E$  在数学上表示一组状态转移对构成的集合,即:

$$e = \{(s, s') | s, s' \in S\} \quad (8)$$

其中,  $s$  与  $s'$  表示任意两个状态,  $S$  表示状态空间,  $E$  表示事件构成的集合。事件驱动的优化关注如何由事件  $e$  驱动进行决策。可以定义策略  $d$  的事件-行为的  $Q$ -因子:

$$Q_K^d(e, a) = \sum_{i \in I(e)} \pi^d(i|e) \left( r(i, a) + \sum_{j \in O_i(e)} p^a(j|i, e) g_K^d(j) \right) \quad (9)$$

其中,  $I(e) = \{s | \exists s' \in S, \text{ s.t. } (s, s') \in e\}$  表示事件  $e$  的输入状态构成的集合,  $\pi^d(i|e)$  表示策略  $d$  下出现事件  $e$  时系统的真实状态为  $i$  的概率,  $r(i, a)$  表示在状态  $i$  和行为  $a$  下的收益,  $O_i(e) = \{s' | \exists s \in S, \text{ s.t. } (s, s') \in e\}$  表示事件  $e$  的输出状态构成的集合,  $p^a(j|i, e)$  表示系统在状态  $i$  下出现事件  $e$  时采取行为  $a$  且最终转移到状态  $j$  的概率,  $g_K^d(j)$  表示状态  $j$  在策略  $d$  下未来  $K$  阶段的势函数, 即:

$$g_K^d(j) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N r_k(s_k, d(e_k)) | e_0 = e, s_1 = j \quad (10)$$

然后决定  $k$  时刻的行为:

$$a_k = \max_a Q(e_k, a) \quad (11)$$

这便是事件驱动的 Q 学习的基本思想。该方法的性能受事件的定义影响较大。事件优选本身等价于状态转移对的着色问题<sup>[12]</sup>, 寻求最优事件定义一般而言并非易事, 应结合具体的问题结构, 选择恰当的状态转移对聚类方式。另外, 给定事件定义后, 可以与现有的其他强化学习算法相结合, 比如深度 Q 网络、深度确定性策略梯度、信任域策略优化 (trust region policy optimization, TRPO)、actor-critic 等。

### 3.3 仿真实验结果

以包含两个机柜的数据中心的空调末端控制问题为例 (如图 3 所示), 展示事件驱动的强化学习方法的性能, 并讨论需要注意的事项。本算例中的系统参数取值见表 1。采用事件驱动的 DQN 方法, 对该算例进行优化, 其收益曲线如图 4 所示。可以发现, 在第 1 600 步的迭代过程中, DQN 方法并没有收敛到较好的策略, 在第 1 000 步时出现了性能下降的现象。造成这一现象的主要原因是 CFD 仿真求解过程复杂、速度慢, 导致训练步数不足、模型缺少探索性。

为进一步提升算法性能, 将仿真优化的硬件平台从 2 核 Windows 服务器改为 64 核 Windows 服务器。改进后的 6SigmaRoom 仿真加速比从之前的 3:1 提升至 120:1 (即仿真数据中心 1 天的用能情况需耗时 0.2 h,  $24:0.2=120:1$ )。基于加速后的新仿真环境的 DQN 方法的性能如图 5 所示。可以明显看出, 改进后 DQN 输出的策略基本上收敛了。

表 1 算例 1 的系统参数

| 对比项 | 参数类型 | 参数名称                        | 参数值     |
|-----|------|-----------------------------|---------|
| 房间  | 设定值  | 尺寸/m <sup>2</sup>           | 36      |
|     | 设定值  | 天花板高度/m                     | 4.5     |
|     | 设定值  | 高架地板高度/m                    | 0.64    |
| 空调  | 设定值  | 最大显冷量/kW                    | 45      |
|     | 设定值  | 冷却水进水温度/℃                   | 14      |
|     | 设定值  | 冷却水最大流量/(m <sup>3</sup> /s) | 0.002 9 |
|     | 设定值  | 总标称制冷流量/(m <sup>3</sup> /s) | 3.6     |
|     | 控制值  | 送风温度设定点/℃                   | 18~27   |
|     | 控制值  | 风机转速                        | 0~100%  |
| 机柜  | 设定值  | 数量/个                        | 3       |
|     | 设定值  | 插槽数/个                       | 52      |
|     | 设定值  | IT 设备数/台                    | 1.5     |
|     | 设定值  | 单台 IT 设备最大功率/kW             | 0.6     |
|     | 设定值  | 单台 IT 设备负载率                 | 0~100%  |

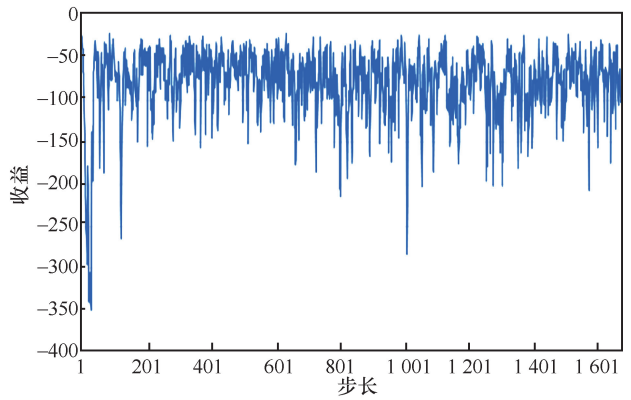


图 4 事件驱动的 DQN 方法的收益变化曲线

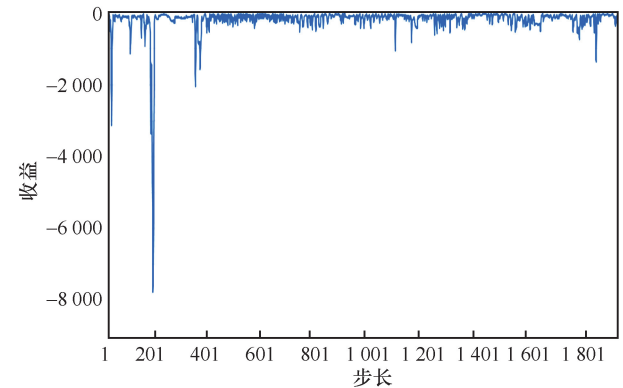


图 5 基于加速后的新仿真环境的 DQN 方法的收益变化曲线

## 4 电池预测性维护

本节介绍一种对数据中心 UPS 的 VRLA 蓄电池实施预测性维护的通用方法<sup>[12]</sup>。VRLA 电池失效模式主要有正极板栅腐蚀、正极活性物质软化、不可逆硫酸盐化等<sup>[13]</sup>, 这些过程通常持续数月, 为电池寿命预测提供了可能性。但是这一任务面临几方

面的困难：一是不同于传统蓄电池，UPS 电池罕有电池循环数据；二是数据中心的 UPS 电池故障鲜有发生，样本量稀缺；三是业界尚未形成统一的电池更换标准，强烈依赖专家经验。为了克服这些困难，需要综合监控数据和专家知识，借知识利用数据，以数据补充知识。

#### 4.1 问题建模

记 UPS 电池的历史数据为  $(\mathbf{x}_i, \mathbf{y}_i)$ ,  $i=1, \dots, N$ , 其中,  $\mathbf{x}_i$  为电池监控数据向量,  $\mathbf{y}_i \in \{0, 1\}$ , 标志着  $\mathbf{x}_i$  所表示的电池是否需要更换。希望找到一个预测函数  $f: \mathbf{x} \rightarrow \{0, 1\}$ , 使得该函数在历史和未来的数据上都能有相对准确的预测精度, 即:

$$\min \sum_{i=1}^N L(f(\mathbf{x}_i), \mathbf{y}_i) + E \left[ \sum_{i=N+1}^{\infty} L(f(\mathbf{x}_i), \mathbf{y}_i) \right] \quad (12)$$

其中,  $E(\cdot)$  表示数学期望,  $L(\cdot, \cdot)$  表示损失函数, 一般使用平均绝对误差 (mean squared error, MSE)。

#### 4.2 预测方法

电池的预测性维护主要分 4 个步骤进行。

第一步, 数据收集与清洗。收集丰富的电池监控数据, 剔除非浮充状态的样本点, 并根据专家经验处理异常样本。记清洗后的历史监控数据集为  $\{\mathbf{x}_i, i=1, \dots, N\}$ 。

第二步, 数据标注, 即对数据集  $\{\mathbf{x}_i, i=1, \dots, N\}$  添加标签  $\{\mathbf{y}_i, i=1, \dots, N\}$ 。由于  $N$  通常是上百万量级, 手动标注并不现实, 因此, 笔者提出了一种计算机辅助的标注方法<sup>[12]</sup>。该方法的核心思想是通过变点检测算法, 确定电池性能衰退事件发生的时间区间, 只需对该区间内收集的数据进行基于专家知识的手动标注, 其余数据可由计算机自动标注, 大幅减小了工作量。

第三步, 特征构造。寻找预测函数  $f$  的过程本质上是在  $\{\mathbf{x}_i, i=1, \dots, N\}$  所张成的向量空间中找到一个分类超平面。一般而言,  $\mathbf{x}_i$  只包含电压、电流、内阻三维, 相对于百万量级的样本量来说, 空间维度过低, 找到分类超平面的难度极大。因此需要首先设计映射函数  $H: R^3 \rightarrow R^D$  ( $D>3$ ), 将样本点从三维空间映射至  $D$  维空间。函数  $H$  的设计思路以专家经验为主, 综合考虑电池组内特征 (如组内电压和电阻的均衡性) 和离散度、时间序列特征 (如电压和电阻的变化率及组合特征)。

第四步, 样本平衡, 模型训练。电池故障在数据中心属于罕见事件, 因此收集的数据必然极为不平衡。以这样的数据集拟合出的映射函数  $f$  可能会

通过牺牲在少数故障样本上的精度来换取整体精度的上升, 因此对非故障样本的降采样很有必要。得到样本平衡的数据集后拟合分类函数, 常用的神经网络模型和决策树模型都可以用作分类模型。

#### 4.3 实验结果<sup>[14]</sup>

笔者使用了来自腾讯数据中心 292 节 UPS 电池两年多的监控记录, 数据清洗后样本量超 2 亿条。经过计算机辅助专家进行样本标注, 得到了 209 465 400 个健康样本点以及 447 215 个故障样本点。在特征构造步骤, 按照第 4.2 节描述的原则设计了映射函数  $H: R^3 \rightarrow R^{14}$ , 并通过  $k$  聚类对健康样本进行降采样, 通过训练梯度提升决策树 (gradient boosting decision trees, GBDT) 分类模型完成电池故障预测任务。

为了测试模型的性能, 笔者重新收集了 2 000 万条电池数据, 对比了基于本文模型和基于腾讯的专家经验的预测效果, 发现本文模型不仅比专家经验有 8% 的精度提升, 并且平均预测时间提前了 15 天。图 6 说明了特征构造对于模型性能的重要性。

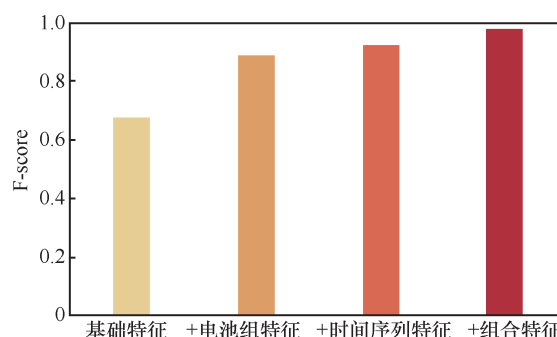


图6 逐渐增加特征对模型 F 得分 (F-score) 的影响

## 5 结束语

数据中心是典型的信息物理融合能源系统。随着当代社会对信息处理需求的迅猛增长, 数据中心的能耗将与日俱增, 因而实现数据中心的绿色可靠发展具有重要意义。现代数据中心具有丰富的监控数据和精细控制的潜力, 其运行决策仍面临高维信息、耦合动态过程、负荷的随机波动等挑战, 需要将机理模型与数据结合。本文围绕 MDC 末端空调控制问题, 展示了事件驱动的强化学习方法的性能; 围绕 UPS 电池故障预测问题, 展示了数据驱动的分析方法, 该方法可以进一步提升预测精度。

数据中心的绿色可靠运行仍有许多值得研究的问题, 比如, 用绿色能源给数据中心供电<sup>[15]</sup>、计算任务调度与新能源发电量的随机匹配、计算任务调

度等，在解决这些问题时将知识与数据融合可以获得比常规方法更好的效果。同时需要注意，以强化学习为例的数据驱动的方法可解释性不足，且具体性能受参数影响大，迁移能力有限。这些问题仍需进一步研究<sup>[16]</sup>。希望本文能为面向数据中心绿色节能运行的策略优化方法的研究抛砖引玉。

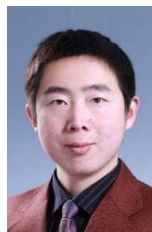
## 参考文献：

- [1] NICOLA J. How to stop data centres from gobbling up the world's electricity[J]. Nature, 2018, 561(7722): 163-166.
- [2] ISO/IEC. Information technology - data centers - key performance indicators - part 2: power usage effectiveness: 30134-2:2016[S]. 2016.
- [3] BULLARD N. Energy efficiency is a hot problem for big tech's data centers[Z]. 2019.
- [4] SHEHABI A, SMITH S, SARTOR D, et al. United states data center energy usage report[R]. 2016.
- [5] 张泉, 李震. 数据中心节能技术与应用[M]. 北京: 机械工业出版社, 2018.  
ZHANG Q, LI Z. Energy saving technology for data center and applications[M]. Beijing: China Machine Press, 2018.
- [6] 腾讯数据中心. 致敬经典——腾讯第三代数据中心[Z]. 2020. Tencent Data Centers. To classic - three generations of Tencent data centers[Z]. 2020.
- [7] LI Y L, WEN Y G, GUAN K L, et al. Transforming cooling optimization for green data center via deep reinforcement learning[J]. IEEE Transactions on Cybernetics, 2019, 50(5): 2002-2013.
- [8] RAN Y Y, HU H, ZHOU X, et al. DeepEE: joint optimization of job scheduling and cooling control for data center energy efficiency using deep reinforcement learning[C]//IEEE 39th International Conference on Distributed Computing Systems. Piscataway: IEEE Press, 2019: 645-655.
- [9] YI D L, ZHOU X, WEN Y G, et al. Toward efficient compute-intensive job allocation for green data centers: a deep reinforcement learning approach[C]//IEEE 39th International Conference on Distributed Computing Systems. Piscataway: IEEE Press, 2019: 634-644.
- [10] PE/ESSB-Energy Storage & Stationary Battery Committee. IEEE 1188-2005-IEEE recommended practice for maintenance, testing, and replacement of valve-regulated lead-acid (VRLA) batteries for stationary applications[S]. 2006.
- [11] 贾庆山, 杨玉, 夏俐, 等. 基于事件的优化方法简介及其在能源互联网中的应用[J]. 控制理论与应用, 2018, 35(1): 32-40.  
JIA Q S, YANG Y, XIA L, et al. A tutorial on event-based optimization with application in energy Internet[J]. Control Theory and Applications, 2018, 35(1): 32-40.
- [12] JIA Q S, WU J J, WU Z J, et al. Event-based HVAC control—a complexity-based approach[J]. IEEE Transactions on Automation Science and Engineering, 2018, 15(4): 1909-1919.
- [13] NEWMAN R H. Advantages and disadvantages of valve-regulated, lead/acid batteries[J]. Journal of Power Sources, 1994, 52(1): 149-153.
- [14] TANG J X, DU J H, LIN Y T, et al. Predictive maintenance of VRLA batteries in UPS towards reliable data centers[C]//IFAC World Congress. [S.l.:s.n.], 2020.
- [15] 夏元清, 高润泽, 林敏, 等. 绿色能源互补智能电厂云控制系统研究[J]. 自动化学报, 2020, 46(9): 1844-1868.  
XIA Y Q, GAO R Z, LIN M, et al. Green energy complementary based on intelligent power plant cloud control system[J]. Acta Automatica Sinica, 2020, 46(9): 1844-1868.
- [16] 王飞跃, 曹东璞, 魏庆来. 强化学习: 迈向知行合一的智能机制与

算法[J]. 智能科学与技术学报, 2020, 2(2): 101-106.

WANG F Y, CAO D P, WEI Q L. Reinforcement learning: toward action-knowledge merged intelligent mechanisms and algorithms[J]. Chinese Journal of Intelligent Science and Technology, 2020, 2(2): 101-106.

## [作者简介]



贾庆山（1980—），男，博士，清华大学自动化系智能与网络化系统研究中心副教授，主要研究方向为信息物理融合能源系统的优化理论与方法。



唐静娴（1996—），女，清华大学自动化系智能与网络化系统研究中心博士生，主要研究方向为绿色智能建筑的运行优化。



吴俊杰（1992—），男，清华大学自动化系智能与网络化系统研究中心博士生，主要研究方向为随机动态系统策略优化、强化学习。



胡潇（1999—），男，北京化工大学信息科学与技术学院本科生，主要研究方向为数据中心节能优化运行、强化学习。



林依挺（1995—），男，腾讯 IDC 平台部助理研究员，主要研究方向为数据中心 AI 能效分析、基础设施设备健康管理。



夏恒（1986—），男，腾讯 IDC 平台部产品管理组组长，主要研究方向为数据中心自动化运营管理平台。