受双脑世界模型启发的网络数字孪生：用于接入调度

Hrishikesh Dutta，*IEEE* 会员　电气与电子工程师，Roberto Minerva，*IEEE* 高级会员，以及 Noel Crespi，*IEEE* 高级会员

arXiv:2602.04566v1 [cs.NI] 4 Feb 2026

摘要——**工业物联网和实时信息物理基础设施等新兴网络系统，需要能够适应动态流量、截止时间和干扰约束的智能调度策略。本文提出一种受双心智世界模型（DMWM）架构启发的新型数字孪生赋能的调度框架，用于实现学习辅助、想象驱动的网络控制。与传统的基于规则或纯数据驱动的策略不同，所提出的 DMWM 将短时域预测规划与符号模型驱动的滚动推演相结合，使调度器能够预测未来网络状态，并据此调整传输决策。我们在一个可配置的仿真测试平台中实现了该框架，并在多种流量条件下，将其性能与传统启发式方法及强化学习基线进行对比评估。结果表明，在突发、干扰受限和截止时间敏感的环境中，DMWM 的性能更优，同时保持了可解释性和样本效率。该设计弥合了网络级推理与低开销学习之间的差距，标志着迈向可扩展、自适应的基于网络数字孪生的网络优化的重要一步。**

索引词——**网络数字孪生、双心智世界模型、物联网调度、强化学习、时限约束网络、无线传感器网络、网络优化。**

I. 引言

下一代无线通信系统，例如智慧城市基础设施、工业物联网和自主网络，需要前所未有的适应性、可靠性和智能水平。这些系统必须在突发流量、严格的服务质量（QoS）约束以及干扰受限环境下运行。因此，网络调度——即决定节点在任意给定时间可以访问信道的问题——成为一项至关重要但极其复杂的决策挑战。

无线网络中的传统调度算法，例如最大权重调度（MWS），致力于为没有严格时延约束的数据包最大化长期吞吐量。MWS 是吞吐量最优的，可实现任何可行的吞吐量向量；而最长队列优先（LQF）等更简单的策略 [1] 或分布式 CSMA（ [2], [3]）虽然降低了计算开销，但可能只能达到吞吐量区域的一部分。然而，新兴应用需要时限感知调度，这会使数据包效用出现不连续性，从而令调度比传统框架更具挑战性。现有的时限感知调度方法通常依赖严格的假设。基于帧的流量模型 [4],将时间划分为帧，并假设到达情况和时限均为先验已知。在这种情况下，每帧的最优解是最大权重调度方案，其权重基于亏欠计数器（用于跟踪尚未达到的交付率）。然而，除非流量具有周期性且同步，否则这种方法并非因果的。文献 [5] 中的部分泛化方法缺乏性能保证。贪心策略，如最大亏欠优先（LDF） [6]，将 LQF 扩展到实时调度；在独立同分布到达且不存在衰落的条件下，其效率比（即所保证的实时吞吐量区域比例）至少为 $\frac{1}{\beta+1}$，其中 β 是网络的干扰度 [7]。已有研究（ [8], [9]）提出了随机化算法（AMIX-ND、AMIX-MS），在放宽流量假设的同时提高效率比，但在大规模网络中仍面临实现障碍。

此外，强化学习（RL）等数据驱动方法（ [10]–[12]）展现出了应用前景，但往往无法保证满足领域特定约束条件下的可行性，例如干扰或分组截止期限。此外，纯无模型智能体通常缺乏可解释性，且需要大量训练数据，因此不适用于快速变化的网络条件。

为克服上述局限，并受到神经科学和人工智能中的认知架构启发，我们提出了一种新颖的混合调度框架，其构建基于双心智世界模型（*DMWM*）[13]的概念。该方法在网络数字孪生平台[14]中结合了反应式启发式方法与符号规划。受 DMWM 启发而开发的 NDT 包含一个快速反应模块 （“快思维”），用于基于启发式方法进行决策；以及一个审慎符号模块（“慢思维”），通过内部仿真（想象）预测有限时域内行动的后果。该框架通过知情约束导航（*ICN*）层确保可行性；该层充当符号约束检查器，用于检查干扰、截止期限和队列可用性。

DMWM 以双阶段循环运行：当约束简单或想象失败时，快思维 会触发轻量级调度；而慢思维 则在数字孪生环境中通过前向展开评估所有可行的调度候选方案，并选出能够优化长期吞吐量或时延的调度方案。与黑盒神经规划器不同，DMWM 提供可解释的规划、约束合规性和实时响应能力。其模块化设计可直接集成到真实或仿真的 NDT 中。

具体而言，本文的主要贡献如下：1）提出一种混合认知调度模型，将符号规划器与反应式回退机制相结合

作者来自法国巴黎理工学院 Telecom SudParis 的数据智能与通信工程实验室（电子邮件：hrishikesh.dutta@telecom-sudparis.eu；roberto.minerva@telecom-sudparis.eu；noel.crespi@telecom-sudparis.eu）

，从而兼具适应性与低时延执行能力。2）设计一种高效的符号状态转移函数，使智能体能够模拟未来的队列状态，在无需进行神经动力学建模的情况下提供长期前瞻能力。3）提出一种基于逻辑的约束筛选机制，在规划前确保满足逐节点截止期限、干扰规避和队列感知要求。4）构建一种仿真驱动的数字孪生环境，对突发流量、异构节点和调度约束进行建模，从而支持全面的性能基准测试。5）在多个真实场景下，将DMWM 与经典及基于学习的调度基线进行比较，证明其在时延管理、吞吐量和可行性遵循方面的优越性。

II. 相关工作

无线网络中在截止期限、干扰及时延约束下进行调度的挑战此前已有研究，但现有的许多方法在为动态异构网络开发解决方案方面仍有不足。

截止期限约束调度：传统调度策略（如最大权重调度（MWS））能够实现吞吐量最优性能，但不考虑实时约束 [4]。最大亏损优先（LDF） [6]等扩展方法及其实时局部池化因子（R-LPF） [9]解析推广形式，在特定流量假设下可提供有界效率比。近期研究探索了随机化调度 [7],，以提高交付率，尤其是在随机到达和衰落信道条件下。然而，这些方法往往假设流量是基于帧的，或依赖于流量统计的离线知识，因此在动态场景中的适应性有限。

基于学习的调度：强化学习（RL）已逐渐成为一种从与网络环境交互中学习调度策略的方法。例如，协作式强化学习方案 [12] 可在能量采集物联网网络中实现分布式调度；另有研究则在全双工通信环境中应用强化学习，以绕过信道状态信息估计 [11]。近期，研究者提出了离线强化学习方法 [8]，通过行为克隆和演员-评论家学习来训练截止期限感知策略。尽管前景可观，大多数基于强化学习的调度器仍是无模型的，需要大量训练样本，且可解释性有限，这给实时和安全关键型应用带来了挑战。

混合与基于图的方法：基于图卷积网络（GCN）的深度强化学习模型已用于时敏流量与尽力而为流量的联合优化 [10],，从而提升泛化能力和收敛速度。同样，快速混合CSMA变体 [4] 以及采用集中式协调的分布式CSMA协议 [3] 也展现出更强的可扩展性。然而，这些方法在调度循环中整合前瞻能力和约束推理方面仍存在局限。

符号与基于模型的方法：Deng等人 [5]的奠基性工作刻画了时延约束容-量区域，针对一般流量模式借助马尔可夫决策过程完成刻画，为基于模型的调度奠定了基础。尽管其方案达到了理论最优，无限时域马尔可夫决策过程的复杂度限制了实际部署。输入排队交换机调度策略如HRF和CHRF [1] 展现了基于等级的仲裁的潜力，但需要专门的队列结构，且难以推广到无线网络。

本文的定位：与上述技术相比，我们提出的双心智世界模型（*DMWM*）框架采用混合认知方法，在仿真驱动的网络数字孪生（*NDT*）环境中结合符号规划与反应式启发式方法。与纯无模型强化学习智能体不同，DMWM基于逻辑驱动世界模型进行前向滚动推演，从而以最少的训练实现约束感知调度决策。我们的设计融合了符号调度器的可解释性与效率，以及学习智能体的适应性，使其在突发、干扰受限和截止时间敏感的条件下表现更优。

III. 系统模型

我们考虑一个由N个传感器节点组成的时隙无线网络，节点集合记为 $\mathcal{N} = \{1, 2, \ldots, N\}$。每个节点都维护一个本地缓冲区（队列），并通过集中式调度竞争信道接入。该系统运行于资源受限物联网网络的典型约束下，包括干扰、缓冲区溢出和截止时间敏感流量。

A. 物理网络模型

在每个时隙 t，节点 $i \in \mathcal{N}$ 的队列长度记为 $q_i(t) \in \mathbb{N}_0$。数据包以速率 $\lambda_i(t)$到达节点 i ，到达过程遵循泊松过程；该速率可能随时间变化，以模拟突发或周期性行为。节点根据调度方案 $\mathcal{S}(t) \subseteq \mathcal{N}$进行传输，其中 $|\mathcal{S}(t)| \leq K$ 表示每个时隙同时传输的最大数量，该数量由时频资源数决定。

我们定义集合冲突对 $\mathcal{C} \subset \mathcal{N} \times \mathcal{N}$ 以表示干扰约束。有效调度必须满足：

$$\forall (i, j) \in \mathcal{C}, \quad i \in \mathcal{S}(t) \Rightarrow j \notin \mathcal{S}(t) \tag{1}$$

数据包会经历时延，部分节点还受到截止期限约束 D_i。在时间 t_a 到达的数据包必须在 $t_a + D_i$之前发送，否则将导致截止期限违规。目标是在最大化长期吞吐量的同时，尽可能减少队列积压、平均时延和截止时间违约。

B. 网络数字孪生环境

为实现规划与评估，我们构建了一个网络数字孪生，作为物理网络的虚拟镜像。它对以下动态特性进行建模：

- **流量**：具有可配置突发性的时变泊松到达。
- **排队**：具有最大缓冲区大小 B 的有界先进先出队列。

- **延迟与截止期限**：用于截止期限评估的逐分组时延跟踪。
- **约束**：干扰（通过 $\mathcal{C}$）和传输限制。

数字孪生支持符号想象——在给定状态下，针对假设动作进行模拟前向传递，而不与真实网络交互。这有助于在 DMWM 框架下实现前瞻性决策。

C. 调度动作与奖励

在每个时隙 t，调度器都会选择一个可行动作 $\mathcal{S}(t)$，该动作：

1) 满足所有逻辑约束（非空队列、冲突避免、截止期限）。
2) 根据队列状态和对未来拥塞的预测，最大化奖励信号。

每个成功传输的数据包都会带来单位奖励。在时间 t的系统总奖励为：

$$r(t) = \sum_{i \in \mathcal{S}(t)} \mathbb{I}[q_i(t) > 0] \tag{2}$$

为支持规划，使用数字孪生从当前状态进行H步滚动推演：

$$R(\mathcal{S}) = \sum_{\tau=0}^{H-1} \sum_{i \in \mathcal{N}} \mathbb{I}[q_i^{(\tau)} > 0] \tag{3}$$

其中，$q_i^{(\tau)}$ 表示未来第 τ步节点 i的模拟队列长度。

D. 双思维决策架构

调度器基于双心智世界模型（*DMWM*）运行：

- **慢思维（符号规划器）**：使用数字孪生对所有有效调度候选方案进行多步仿真。选择能够最大化预测奖励 $R(\mathcal{S})$ 的调度方案。
- **快思维（反应式启发式）**：当推演结果不明确、不可行，或在时间受限环境中计算成本过高时，退而采用简单规则（例如 LQF）。

约束检查通过知情约束导航（ICN）模块实现，该模块根据可行性筛选动作，以确保符合协议逻辑。

IV. 双重思维世界模型驱动的调度

所提出的调度架构基于一种受认知启发的框架，该框架模拟智能体中的双过程决策。如图1所示，这种双重思维架构使调度器能够在速度 与前瞻之间取得平衡，并根据当前的可行性、复杂度和紧迫性动态切换模式。

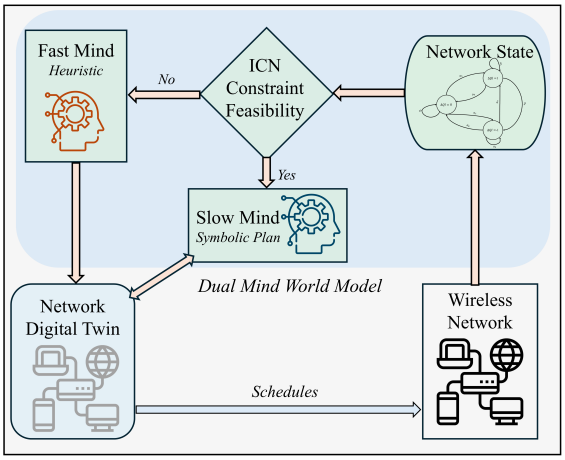

图 1：系统模型

A. 快思维：反应式调度

快思维使用基于当前队列长度的启发式方法选择调度 $\mathcal{S}_{\text{fast}}(t)$：

$$\mathcal{S}_{\text{fast}}(t) = \arg\max_{\mathcal{S} \subset \mathcal{N}} \sum_{i \in \mathcal{S}} q_i(t), \quad |\mathcal{S}| = K \tag{4}$$

该策略计算开销较低，在高负载或紧急情况下能够确保响应能力，但缺乏长期最优性。

B. 慢思维：符号展开规划

慢思维考虑一组可行的候选调度：

$$\mathcal{F}(t) = \{\mathcal{S}_c \subset \mathcal{N} : |\mathcal{S}_c| = K,\ \text{ICN}(\mathcal{S}_c, q(t), t) = \texttt{True}\} \tag{5}$$

其中，ICN 是一个约束检查函数，用于验证：

- 非空队列：$q_i(t) > 0$ 对所有 $i \in \mathcal{S}_c$
- 干扰约束：$(i,j) \notin \mathcal{C}$对所有 $i,j \in \mathcal{S}_c$
- 截止期限可行性：若 D_i 存在，则在允许时间内调度数据包

对于每个可行调度 $\mathcal{S}_c \in \mathcal{F}(t)$，符号世界模型模拟时域为 H 的滚动推演：

$$q_i^{(\tau+1)} = \begin{cases} \max(q_i^{(\tau)} - 1, 0), & \text{if } i \in \mathcal{S}_c \\ q_i^{(\tau)}, & \text{otherwise} \end{cases} \tag{6}$$

从 $q_i^{(0)} = q_i(t)$ 开始。时域内的累积符号奖励定义为：

$$R(\mathcal{S}_c) = \sum_{\tau=0}^{H-1} \sum_{i \in \mathcal{N}} \mathbb{I}[q_i^{(\tau)} > 0] \tag{7}$$

随后，慢思维选择滚动推演最优调度：

$$\mathcal{S}_{\text{slow}}(t) = \arg\max_{\mathcal{S}_c \in \mathcal{F}(t)} R(\mathcal{S}_c) \tag{8}$$

C. 组合决策规则

时刻 t 的最终调度决策取决于可行方案的可用性：

$$\mathcal{S}(t) = \begin{cases} \mathcal{S}_{\text{slow}}(t), & \text{if } \mathcal{F}(t) \neq \emptyset \\ \mathcal{S}_{\text{fast}}(t), & \text{otherwise} \end{cases} \tag{9}$$

该规则确保选择当前可用的最佳动作——可以依据预测的符号化结果（慢思维），也可以在符号规划失败或成本过高时采用快速启发式方法。

D. 复杂度与可行性

慢思维规划步骤的复杂度为 $O\left(\binom{N}{K} \cdot H \cdot N\right)$，这是因为需要为每个 $\binom{N}{K}$ 个候选调度方案模拟 H 步队列演化。实际上，ICN筛选器会通过及早排除不可行或不合逻辑的候选方案，大幅缩减搜索空间。

E. 与数字孪生的集成

所有符号模拟均在网络数字孪生中进行，该孪生体复现了物理系统的排队、干扰和流量动态。利用孪生体进行内部推演，慢思维得以预见未来，同时不影响实时运行。这种分离使统一决策框架能够实现预测性评估、安全检查和可解释性。所提出调度框架的详细工作流程见算法1，ICN模块见算法 2。

表I：基线仿真超参数

参数	值
节点数 (N)	5
最大调度节点数 (K)	3
队列容量 (B)	50
每次运行的时间步数	200
泊松到达基准范围（λ_i^{base}）	[0.5, 1.0]
想象视界（H）	3
Q学习 α （学习率）	0.1
Q学习 γ （折扣因子）	0.95
Q学习 ϵ (探索率)	0.2

V. 实验设置

为评估所提出的双心智世界模型（DMWM）调度器，我们实现了一个由仿真驱动的时隙无线网络数字孪生。除了转发由 DMWM 调度器学习到的调度策略外，该数字孪生还模拟了队列动态、随机到达、干扰约束和逐节点截止期限。所有实验均在多个具有不同约束的网络场景中进行，并与基线策略进行比较。

A. 网络仿真环境

我们模拟一个由 N 个节点组成的网络，每个节点都维护一个容量为 $B = 50$的有界先进先出队列。时间按时隙划分，每个时隙最多可调度 K 个节点。各节点的流量到达服从到达率动态变化的时变泊松过程

Algorithm 1 DMWM-Enabled Scheduling

Require: Current queue state $q(t)$, deadline vector $D = \{D_1, \ldots, D_N\}$, conflict graph $\mathcal{C}$, number of slots K, planning horizon H
Ensure: Selected schedule $S(t) \subseteq \{1, \ldots, N\}$ with $|S(t)| = K$

$\mathcal{F} \leftarrow \emptyset$ {Feasible schedule set}
for all $S \subseteq \{1, \ldots, N\}$ with $|S| = K$ **do**
  if ICN_CHECK(S, $q(t)$, D, $\mathcal{C}$, t) **then**
    $\mathcal{F} \leftarrow \mathcal{F} \cup \{S\}$
  end if
end for
if $\mathcal{F} \neq \emptyset$ **then**
  $R^\star \leftarrow -\infty$, $S^\star \leftarrow \emptyset$
  for all $S \in \mathcal{F}$ **do**
    $q^{(0)} \leftarrow q(t)$, $R \leftarrow 0$
    for $\tau = 0$ to $H - 1$ **do**
      for all $i \in S$ **do**
        $q_i^{(\tau+1)} \leftarrow \max(q_i^{(\tau)} - 1, 0)$
      end for
      $R \leftarrow R + \sum_{i=1}^{N} \mathbb{I}[q_i^{(\tau)} > 0]$
    end for
    if $R > R^\star$ **then**
      $R^\star \leftarrow R$, $S^\star \leftarrow S$
    end if
  end for
  return $S^\star$
else
  for $i = 1$ to N **do**
    $u_i \leftarrow q_i(t) \cdot (2 \text{ if } D_i < \infty \text{ else } 1)$
  end for
  $S_{\text{fallback}} \leftarrow$ top-K indices by u_i
  return S_{fallback}
end if

到达率 $\lambda_i(t)$。流量配置包含平滑和突发模式，以模拟现实世界中的变化：

$$\lambda_i(t) = \lambda_i^{\text{base}} \cdot \left(1 + 0.75 \cdot \sin\left(\frac{2\pi t}{50}\right)\right) + \delta_i(t)$$

其中，$\delta_i(t)$ 会在选定节点上引入偶发的突发尖峰。

调度动作受冲突图 $\mathcal{C}$(干扰模型)和可选的逐节点分组截止期限D_i约束。

B. 评估场景

我们在四种不同场景下评估这些策略：

- **默认**：中等程度的干扰（相互冲突的节点对），以及交替节点上的混合截止期限约束。
- **突发流量**：到达率的方差较高，且没有截止期限或冲突。
- **截止时间敏感**：截止期限交替较短（5）和较长（15）的节点，并存在成对干扰。
- **干扰受限**：无截止期限的环形冲突图。

算法 2 ICN CHECK$(S,\ q(t)\ D,\ \mathcal{C},\ t)$

对所有 执行
  如果 $q_i(t) = 0$，**则**
    返回 **假**
  结束如果
  如果 $D_i < \infty$ 且最旧数据包的时延为 $> D_i$ **则**
    返回 **假**
  结束如果
结束循环
对所有 $(i, j) \in S \times S$ **执行**
  如果 $(i, j) \in \mathcal{C}$ **则**
    返回 **假**
  结束如果
结束循环
返回 **True**

C. 基线调度策略

将所提出的 DMWM 调度器与以下基线方法进行比较：

- **随机策略**：均匀随机地选择 K 个节点。
- **最长队列优先（LQF）**：优先选择队列长度最长的节点。
- **截止时间优先**：根据分组截止期限分配紧迫性权重。
- **公平轮询**：通过基于记忆的轮转确保长期公平性。
- **强化学习智能体（Q学习）**：基于在线学习的调度器，采用 ϵ-贪心探索，并通过Q表进行更新。

DMWM 智能体使用符号前向模型（想象），在规划时域为 $H = 3$ 的情况下评估经 ICN 约束层筛选的候选调度方案。

D. 超参数与性能指标 s

所有策略均经过 30 次独立仿真运行测试，每次仿真持续 200 个时间步。仿真环境中使用的基准参数汇总于表 I。我们使用以下指标评估各项策略：平均吞吐量、队列长度、时延和截止时间违约。每项指标均报告多次仿真运行的均值和标准差。

VI. 结果与分析

为评估所提出的受 DMWM 启发的调度框架的有效性，我们在多种网络条件下进行了大量仿真，以模拟物联网和无线接入系统中的现实挑战。如第五节所述，我们在四种不同的流量和约束场景下，将 DMWM 的性能与传统基线方法以及一个强化学习（RL）智能体进行比较：默认、突发、截止时间敏感和干扰受限。首先，为深入了解 DMWM 的内部动态，我们分析了图2中展示的两种行为可视化结果。图2（a）所示的热力图展示了 DMWM 随时间选取的节点调度模式。该调度会根据队列状态和约束违规情况动态调整。

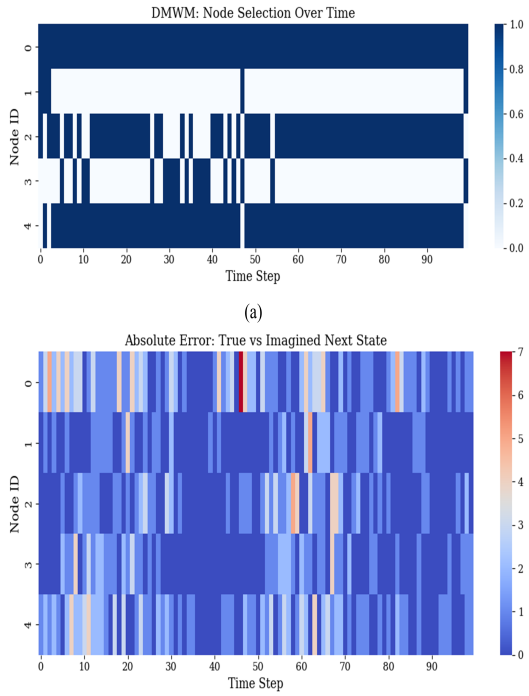

图 2：DMWM调度器行为与模型准确性：该图展示了 DMWM 策略的节点调度行为，其中每一行对应一个节点，每一列代表一个时间步。蓝色单元格（子图 (a) 中）表示相应节点在该时刻被选中进行传输。子图 (b) 展示了实际队列状态转移与慢思维内部仿真中使用的符号世界模型所预测转移之间的绝对差值。

激活模式的多样性和结构体现了反应式模块（快思维）与符号模块（慢思维）之间的相互作用。对比想象中的下一状态与实际的下一状态（图2（a））可见，各节点的绝对预测误差较低，证实了用于前向仿真的内部模型具有可靠性。这验证了数字孪生在规划时域内预测队列动态的能力。

图3展示了所有策略在四种场景下的对比结果，该结果为多次运行的平均值。观察到以下主要发现。首先，启用DMWM的调度在所有场景中始终取得最高或接近最高的平均吞吐量。值得注意的是，在突发和截止时间敏感负载下，它优于最长队列优先和强化学习智能体，这得益于其基于规划的想象模块能够预判拥塞并避免冲突。其次，就队列占用而言，与反应式策略相比，DMWM保持更低的平均队列大小，表明拥塞缓解更为有效。尤其在干扰严重的场景中，其符号约束模块使其能够比最长队列优先和随机策略更高效地应对竞争区域。最后，DMWM最显著的改进体现在时延敏感指标上。在截止时间敏感场景下，相对于强化学习智能体和启发式方法，DMWM大幅降低了平均时延和截止时间违约次数。这归因于其前向展开机制，该机制显式评估时延的影响

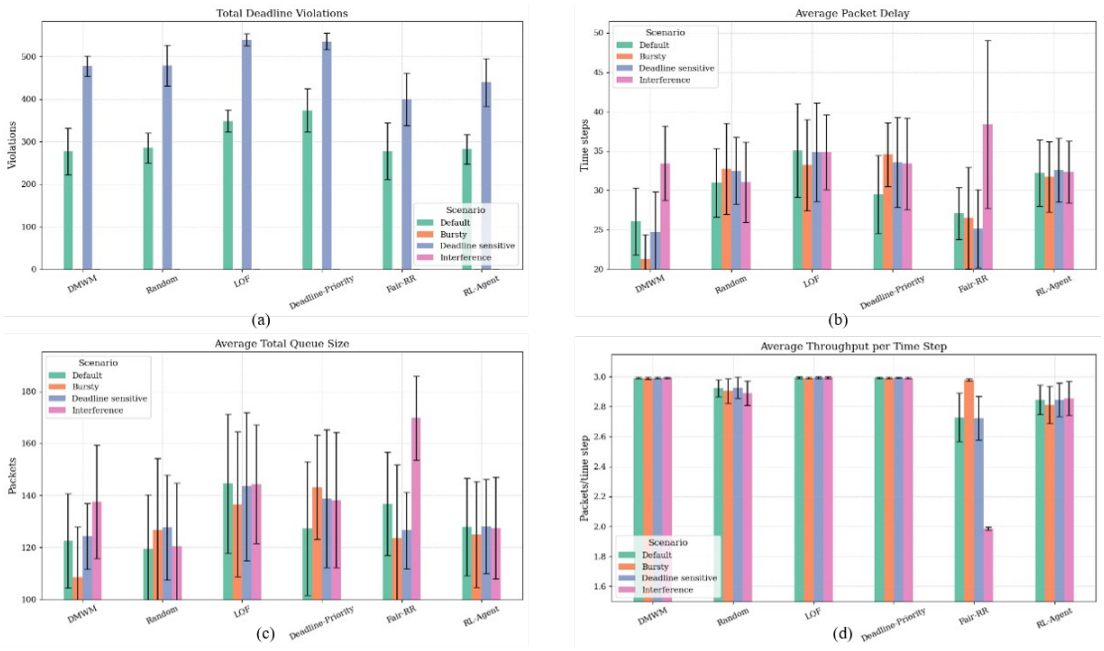

图 3：该图比较了六种调度策略——DMWM、随机策略、最长队列优先、截止时间优先、公平轮询和 Q-learning 智能体——在四种场景下的表现：默认、突发、截止时间敏感和干扰受限。结果展示了 30 次运行的平均性能及标准差误差棒。DMWM 在保持低时延和低队列占用的同时，实现了最佳或接近最佳的吞吐量，并在不违反干扰或截止期限约束的情况下实现了高效利用。

在调度过程中。需要指出的是，DMWM不仅平均表现良好，而且在各次试验中方差更低，表明其在动态网络条件下具有稳健性。

七、结论

本文提出了一种用于网络调度的新型双心智世界模型（DMWM）方法，旨在网络数字孪生（NDT）仿真环境中运行。DMWM 框架融合了符号规划与轻量级的基于模型的预测，在适应性和可解释性之间取得平衡。通过在较短的规划时域内推演想象中的状态转移，DMWM 使调度器能够预测并缓解网络拥塞、干扰和截止时间违约，而无需使用深度函数逼近器或进行大量重新训练。我们在多个模拟网络场景中评估了所提出的调度器，这些场景具有突发流量、截止期限约束和干扰等特点。与经典基线方法的对比结果表明，DMWM 具有较强的鲁棒性，尤其是在非平稳或受限环境中。除了在吞吐量、时延和截止时间遵循度等关键指标上表现优异外，DMWM 还通过可解释的滚动推演机制提供透明的决策过程。因此，它适用于需要问责性和洞察力的关键系统或资源受限系统。

参考文献

[1] B. Hu、F. Fan、K. L. Yeung 和 S. Jamin，“最高优先级优先：一类新型输入排队交换机单次迭代调度算法”，*IEEE Access*，第 6 卷，页 11 046–11062，2018 年。[2] Z. Zhang、S. Atapattu、Y. Wang 和 M. Di Renzo，“RIS 辅助多用户网络的分布式 MAC：CSMA/CA 协议设计与统计优化”，*IEEE Transactions on Mobile Computing*，2025 年。

[3] C. Huang 和 X. Wang，“面向可扩展无线网状网络的集中协调分布式调度”，*IEEE/ACMTransactions on Networking*，第 31 卷，第 1 期，页 436–451，2022 年。[4] B. Li 和 A. Eryilmaz，“时变条件下的最优分布式调度：一种快速 CSMA 算法及其应用”，*IEEE Transactions on Wireless Communications*，第 12 卷，第 7 期，2013 年。[5] L. Deng、C.-C. Wang、M. Chen 和 S. Zhao，“具有一般流量模式的及时无线流：容量区域与调度算法”，*IEEE/ACM Transactions on Networking*，第 25 卷，第 6 期，2017 年。[6] X. Kang、W. Wang、J. J. Jaramillo 和 L. Ying，“无线网络中最大亏损优先调度实时流量的性能”，载于*ACM* 移动自组织网络与计算国际研讨会，2013 年，页 99–108。[7] C. Tsanikidis 和 J. Ghaderi，“衰落信道上的无线网络实时流量随机化调度”，*IEEE/ACM Transactions on Networking*，第 31 卷，第 4 期，页 1688–1701，2022 年。[8] J. Wan、S. Lin、Z. Zhang、J. Zhang 和 T. Zhang，“实时无线流量调度：一种网络辅助离线强化学习方法”，*IEEE internet of things journal*，第 10 卷，第 24 期，2023 年。[9] C. Tsanikidis 和 J. Ghaderi，“无线网络实时流量调度中随机化的作用”，*IEEE/ACM Trans- actions on Networking*，第 29 卷，第 4 期，页 1703–1716，2021 年。[10] L. Yang、Y. Wei、F. R. Yu 和 Z. Han，“利用基于图卷积网络的深度强化学习优化时间敏感网络中的联合路由与调度”，*IEEE Internet of Things Journal*，第 9 卷，第 23 期，页 23 981–23 994，2022 年。[11] H. Fawaz、M. El Helou、S. Lahoud 和 K. Khawam，“全双工无线网络中队列感知调度的强化学习方法”，*Computer Networks*，第 189 卷，文章编号 107893，2021 年。[12] H. Dutta、A. K. Bhuyan 和 S. Biswas，“面向能量采集多跳网络能源管理的协作强化学习”，*IEEE Transactions onGreen Communications and Networking*，2025 年。[13] L. Wang、R. Shelim、W. Saad 和 N. Ramakrishnan，“DMWM：具有长期想象的双心智世界模型”，*arXiv preprint arXiv:2502.07591*，2025 年。[14] S. M. Raza、R. Minerva、N. Crespi、M. Alvi、M. Herath 和 H. Dutta，“利用网络数字孪生中的高效数据建模赋能边缘—云连续体”，载于*2024* 年第 *20* 届网络与服务管理国际会议（*CNSM*）。IEEE，2024 年，页 1–4。