

LIGHT: A Learned Join Order Selection Network Using Dynamic Graph Attention Mechanism

张秋鸿, 杨建文, 严瑾, 张硕, 丁治明

第42届CCF中国数据库学术会议 (NDBC 2025)

张秋鸿 qihong@iscas.ac.cn 18810931841

张硕 zhangshuo@iscas.ac.cn 15010525643

研究背景

- 连接顺序选择是数据库查询优化中的核心难题
- 随着表数量增加, 连接顺序组合呈指数增长, 导致查询优化计算开销巨大
- 传统优化器如PostgreSQL依赖静态启发规则和代价模型, 难以适应复杂查询场景

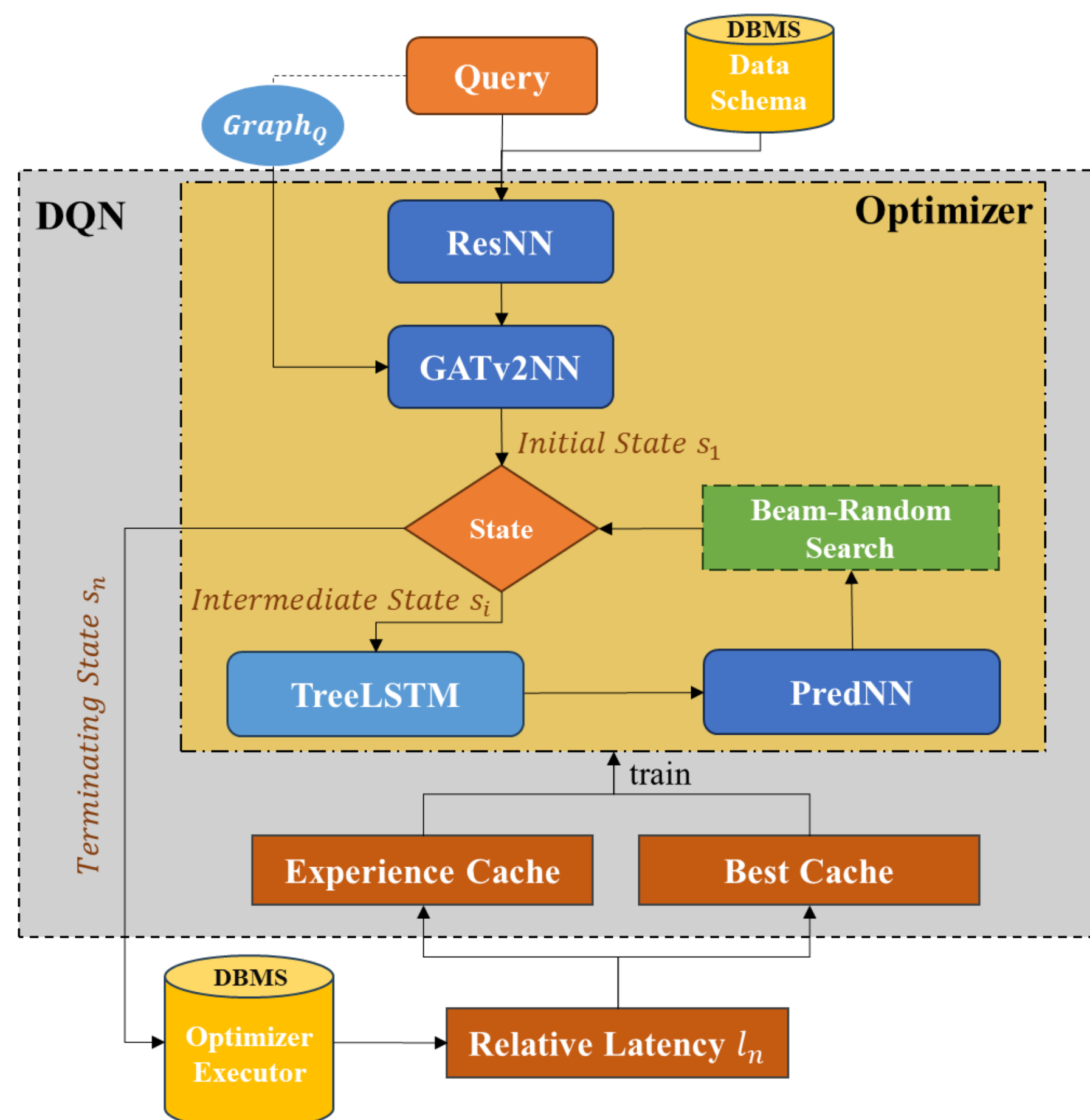
现有挑战问题:

- 查询图中不同节点重要性差异未被有效建模
- 探索策略过于保守, 难以找到最优连接顺序
- 奖励函数高方差使模型训练不稳定

所提出的方法 LIGHT

核心创新:

- ✓ 动态图注意力机制 (GATv2): 动态捕捉表间依赖关系
- ✓ Beam-Random Search探索策略: 融合beam search和加权 ϵ -greedy, 提升探索效率
- ✓ 对数奖励变换: 缓解奖励值剧烈波动, 提高模型收敛稳定性

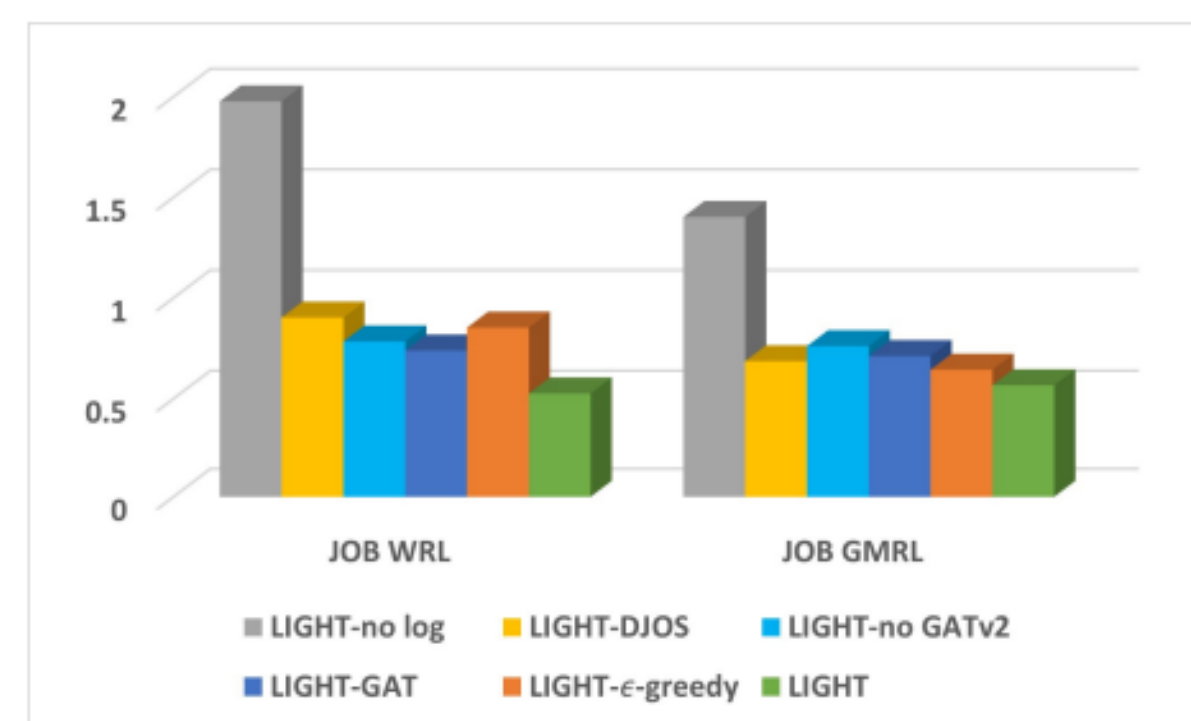


LIGHT总体架构图

查询解析 → 特征提取 (ResNN) → 图建模 (GATv2) → 动作选择 (DQN) → 奖励压缩训练

实验结果

Model	JOB				TPC-DS			
	WRL	GMRL	T_{inf}/ms	T_{train}/h	WRL	GMRL	T_{inf}/ms	T_{train}/h
DP	1	1	-	-	1	1	-	-
RTOS	0.885	0.911	65.4	7.1	0.846	0.852	25.6	1.2
Balsa	0.669	0.741	162.1	18.5	0.811	0.795	71.4	7.3
LIGHT	0.526	0.567	86.3	19.2	0.636	0.763	58.1	5.4



数据集: JOB、TPC-DS

对比方法: PostgreSQL、RTOS、Balsa

结果:

- JOB: WRL ↓47.4%, GMRL ↓43.3%
- TPC-DS: WRL ↓36.4%, GMRL ↓23.7%
- 显著优于传统与现有学习优化方法

