

# 基于深度强化学习的实时调度策略

## Developing Real-time Scheduling Policy by Deep Reinforcement Learning

薄紫彤, 乔颖, 冷昶, 王宏安, 郭超平, 张绍辉

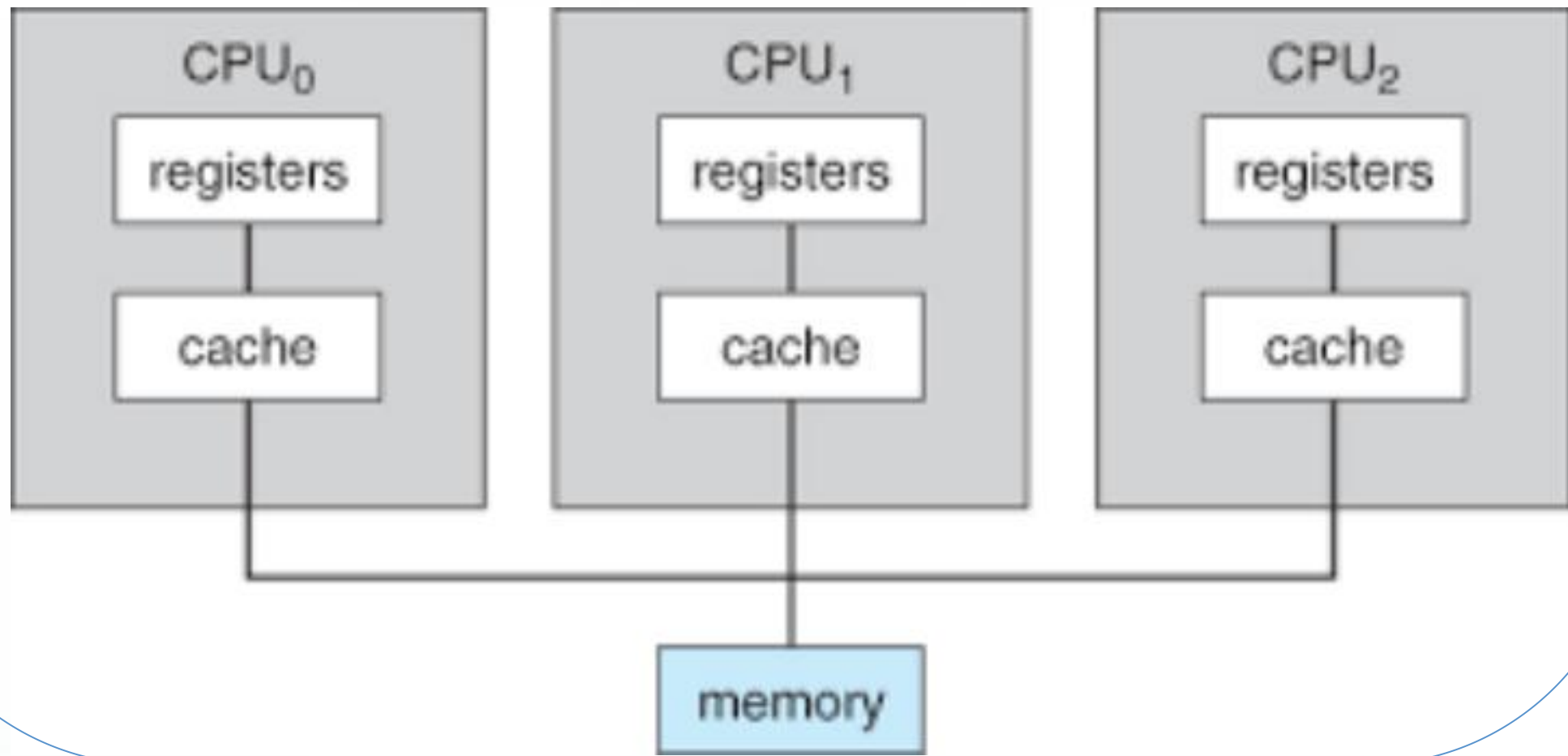
the 27th IEEE Real-Time and Embedded Technology and Applications Symposium (RTAS 2021)

Contact: 乔颖 qiaoying@iscas.ac.cn

### 背景

多处理器系统的实时任务调度是NPC问题。

经验驱动的启发式调度方法性能受专家水平影响, 定制成本高, 对不同负载的适用性差。

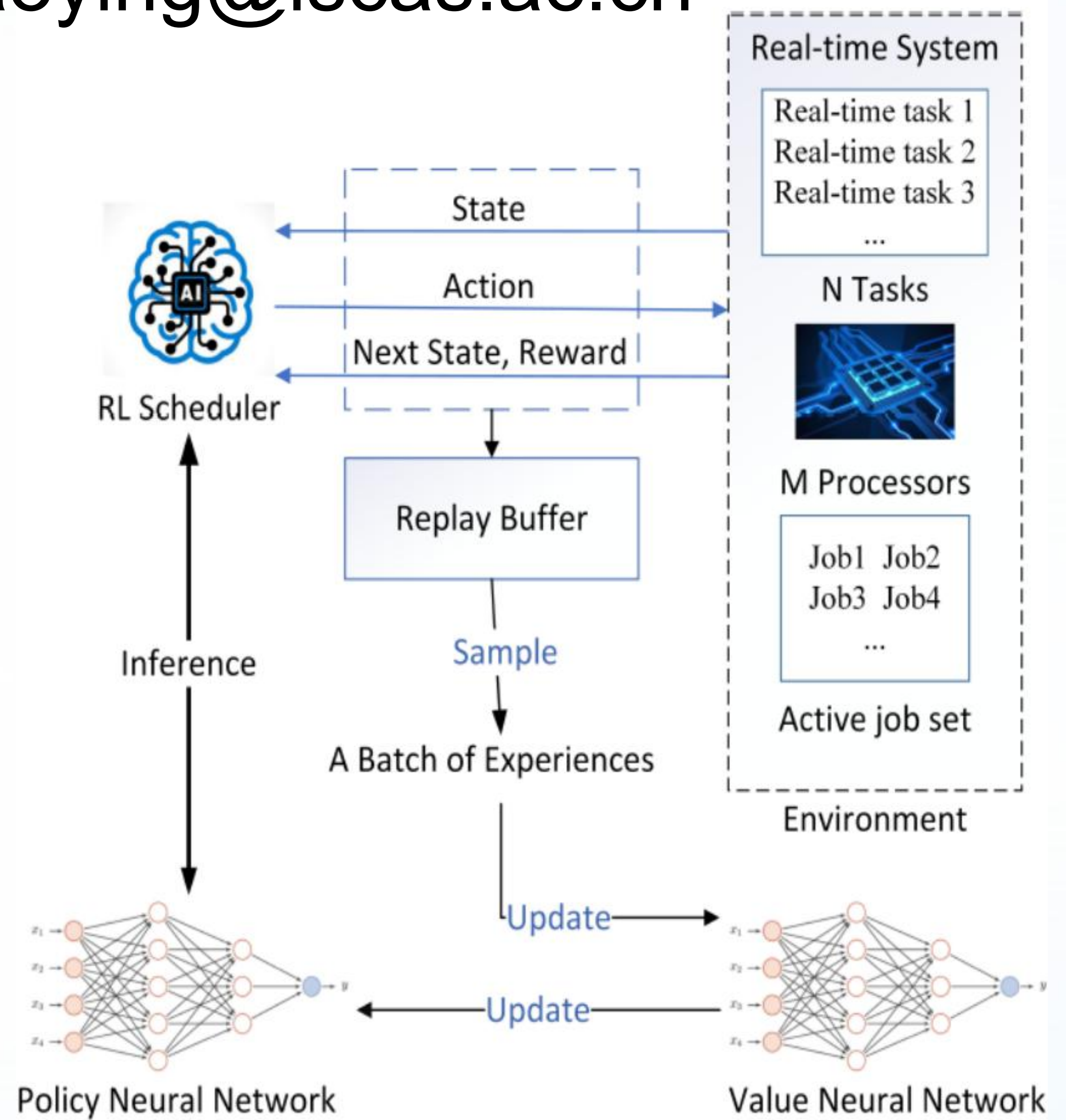


### 方法

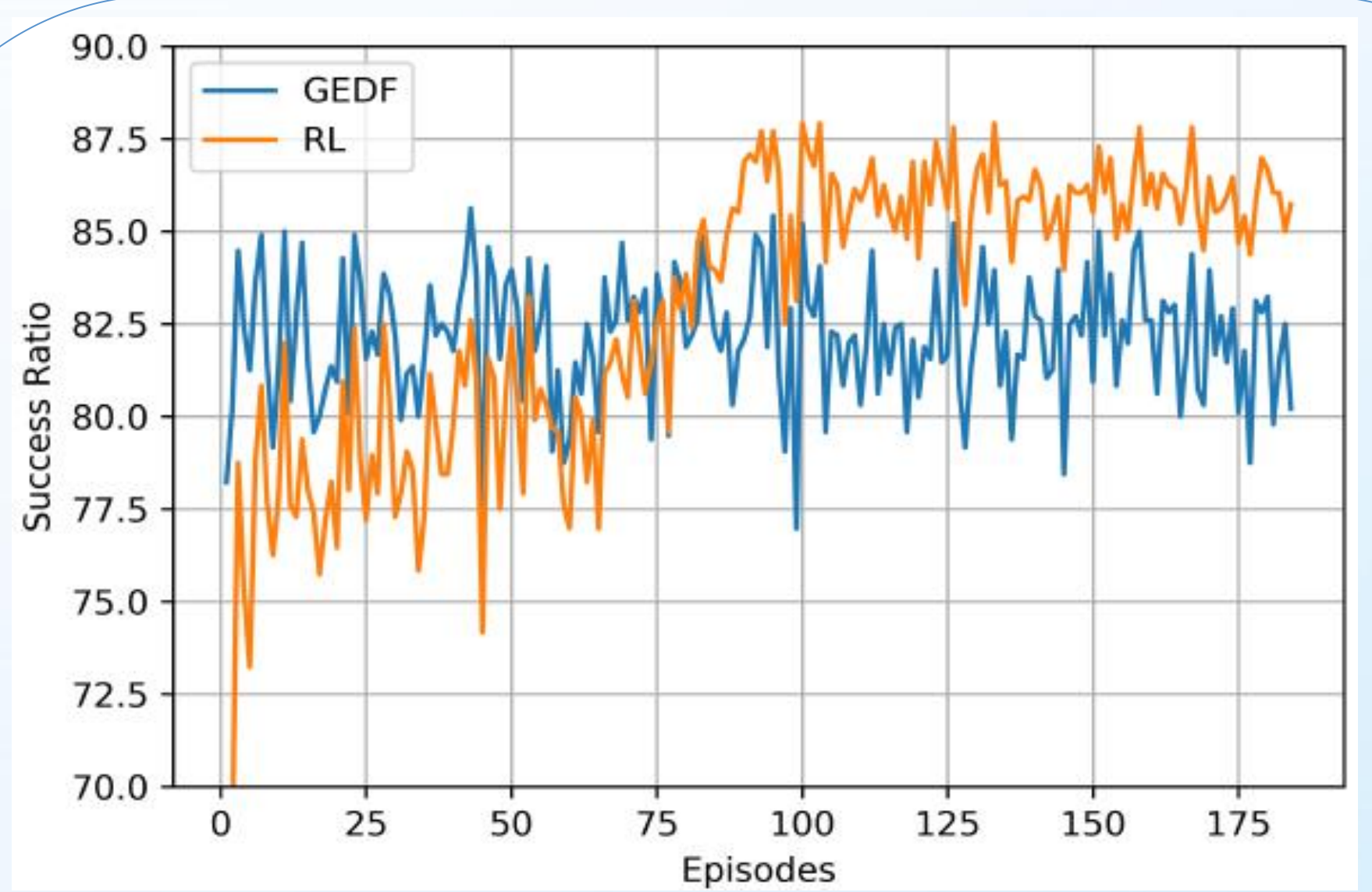
通过数据驱动的方法学习多处理器调度策略。

- 1) 由于实时调度过程具有马尔科夫性, 将实时调度过程建模为马尔可夫决策过程;
- 2) 针对调度过程中任务数量持续变化的特点, 将任务视作智能体。使用多智能体强化学习方法学习优先级策略。所有智能体互相合作, 最大化总收益。
- 3) 提出自合作多智能体学习方法, 使用对称函数对每个智能体的价值函数进行聚合, 全局价值函数 $Q(s, a)$ 的更新公式为:

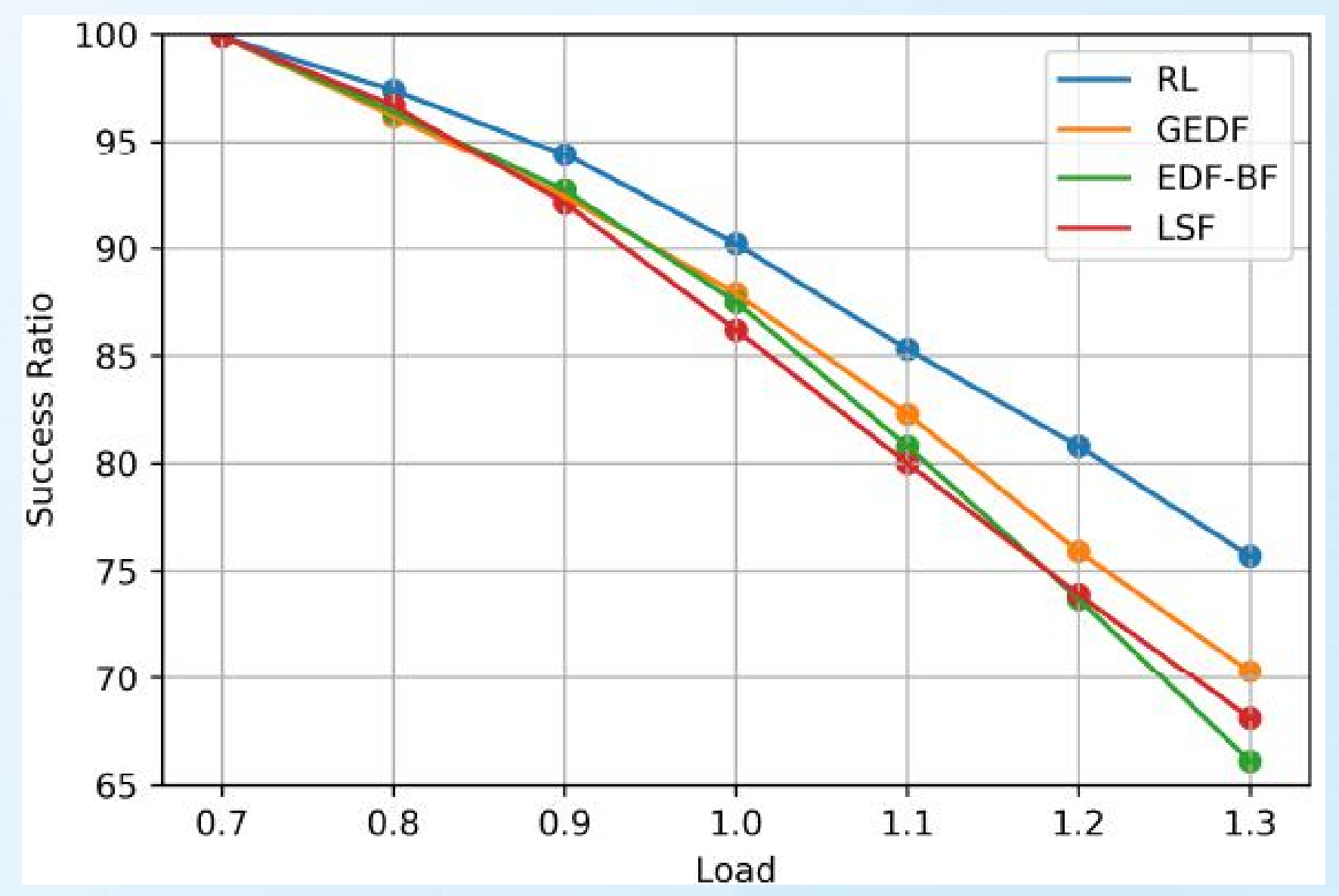
$$Q(s, a) = \frac{1}{N} \sum \dot{Q}(s_i, a_i) = r + \gamma \frac{1}{N} E_{s_i \sim p} [\sum \dot{Q}'(s'_i a'_i)]$$



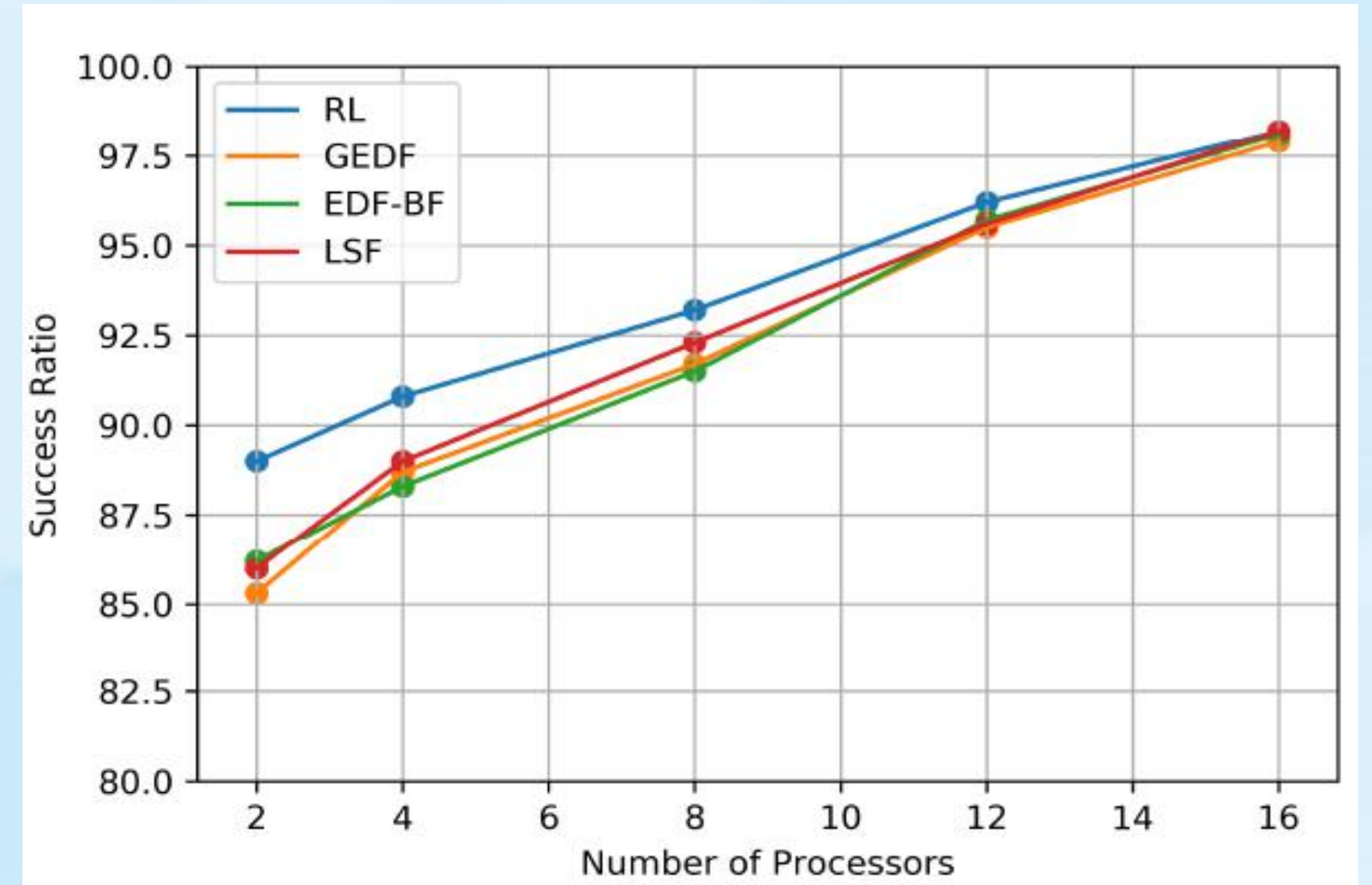
### 使用强化学习学习调度策略实验



学习过程中调度成功率逐渐上升



不同负载下的调度成功率优于传统方法



不同处理器数量下的调度成功率优于传统方法