UNIVERSITÄT PADERBORN

CN-UPB/PG : ARTIFICIAL INTELLIGENCE FOR COMPUTER NETWORKS

SEMINAR PAPER:

RESOURCE MANAGEMENT WITH DEEP REINFORCEMENT LEARNING

Presented by: Nayela Tasnim Labonno
Summer Term: 2020

UNIVERSITÄT PADERBORN

Contents

UNIVERSITÄT PADERBORN

Resource Management Challenges and Scopes of Reinforcement Learning

- Online decision making under massive workload
- Interdependency between tasks
- Dynamic and complex models

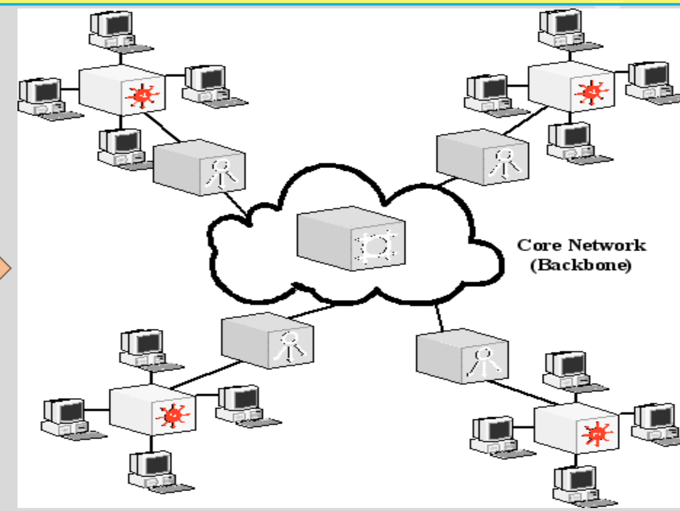

Figure 1 : Cluster computing [1]

Alternative to general heuristics

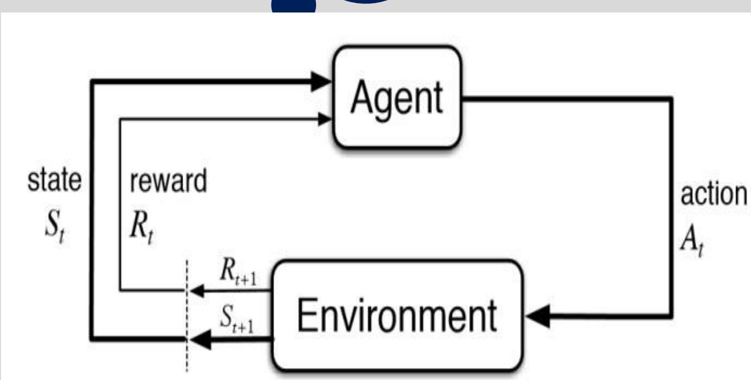

Figure 2 : Reinforcement Learning model [2]

- Repetitive decisions provide abundant training data
- Deep NN to model complex systems
- RL agent can be trained to optimize system objectives

Proposed Model: DeepRM

1. State Representation

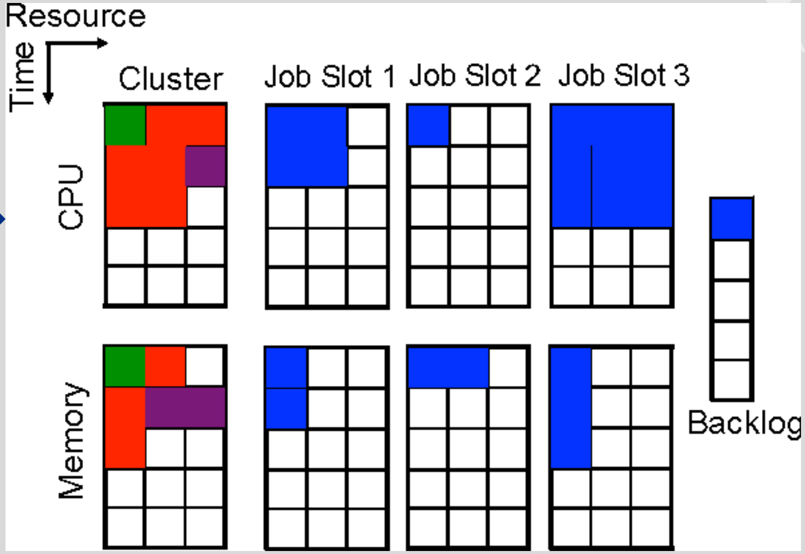

Figure 3 : State representation [3]

RL Formulation

2. Action Space,

A= { Ø, 1,… M}

3. System Objective and Reward

1

Minimize average job slowdown

Reward function,

$$\sum_{j \in J} \frac{-1}{T_j}$$

2

Minimize average job completion time

Reward function,

$- | J |$

UNIVERSITÄT PADERBORN

Training Algorithm

Notation Table	
Variable	Description
r_j	Resource requirements of a job, j
T_j	Duration (ideal) of the job
C_j	Completion time of job
J	Set of jobs (scheduled and waiting) currently in the system
γ	Discount factor
v_t	Discounted cumulative reward
b_t	Baseline value
t	timestep
M	subset of jobs
N	No. of episodes

for each iteration:
    $\Delta\theta \leftarrow 0$
    for each jobset:
        run episode $i = 1, \ldots, N$:
            $\{s_1^i, a_1^i, r_1^i, \ldots, s_{L_i}^i, a_{L_i}^i, r_{L_i}^i\} \sim \pi_\theta$
        compute returns: $v_t^i = \sum_{s=t}^{L_i} \gamma^{s-t} r_s^i$
        for t = 1 to L:
            compute baseline: $b_t = \frac{1}{N} \sum_{i=1}^{N} v_t^i$
            for i = 1 to N:
                $\Delta\theta \leftarrow \Delta\theta + \alpha \nabla_\theta \log \pi_\theta(s_t^i, a_t^i)(v_t^i - b_t^i)$
            end
        end
    end
    $\theta \leftarrow \theta + \Delta\theta$ % *batch parameter update*
end

Figure 4 : Policy training algorithm [3]

UNIVERSITÄT PADERBORN

Comparison

Comparables :

- a DeepRM agent
- a Shortest Job First agent
- a Packer agent
- A Tetris agent [4]

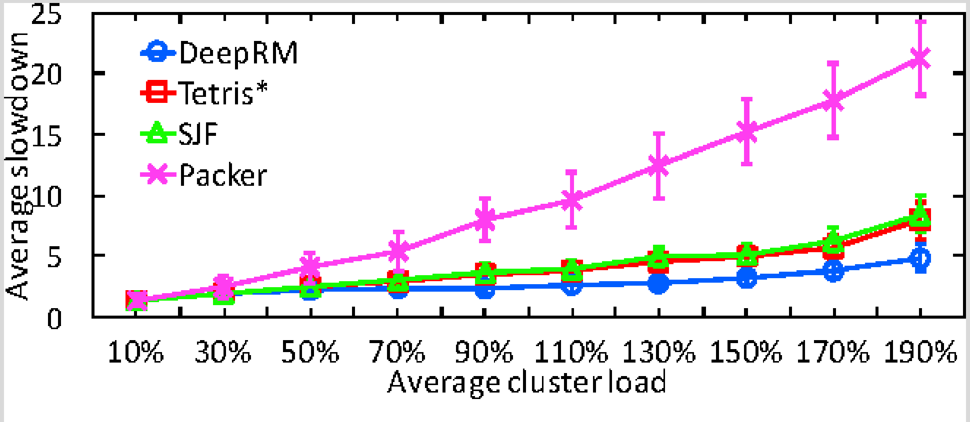

Figure 5 : Average job slowdown at different levels [3]

DeepRM is customizable for different objectives!

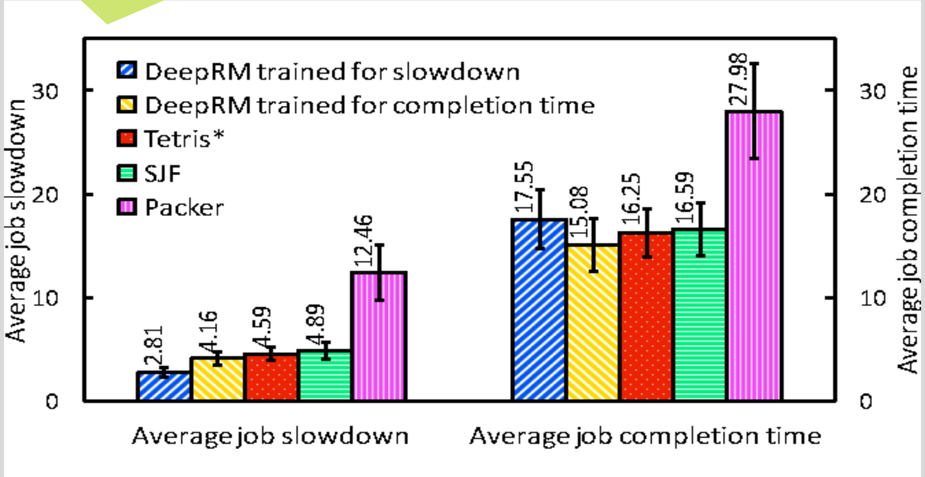

Figure 6 : Performance for different objectives [3]

Experimental parameters

M = 10

N = 20

Training iterations = 1000

Learning rate= .001

γ = 1

Convergence of DeepRM

After 200 iteration, DeepRM beats Tetris

Large gap between max and average value indicates the need for further learning

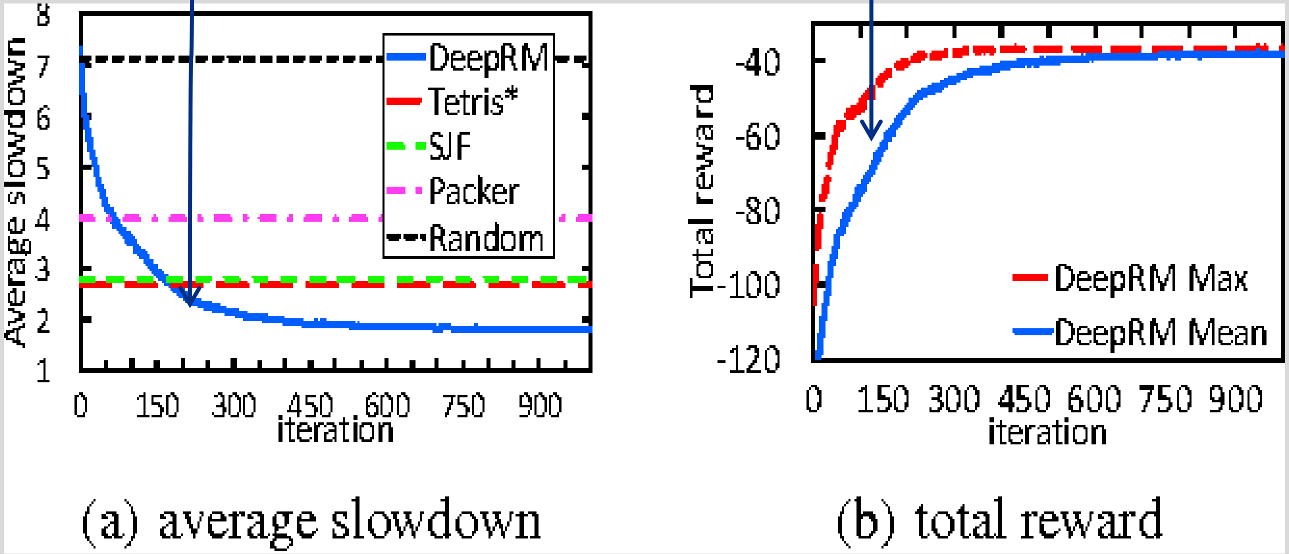

Figure 7 : Learning curve showing the slowdown and the total reward performance [3]

Performance Gain

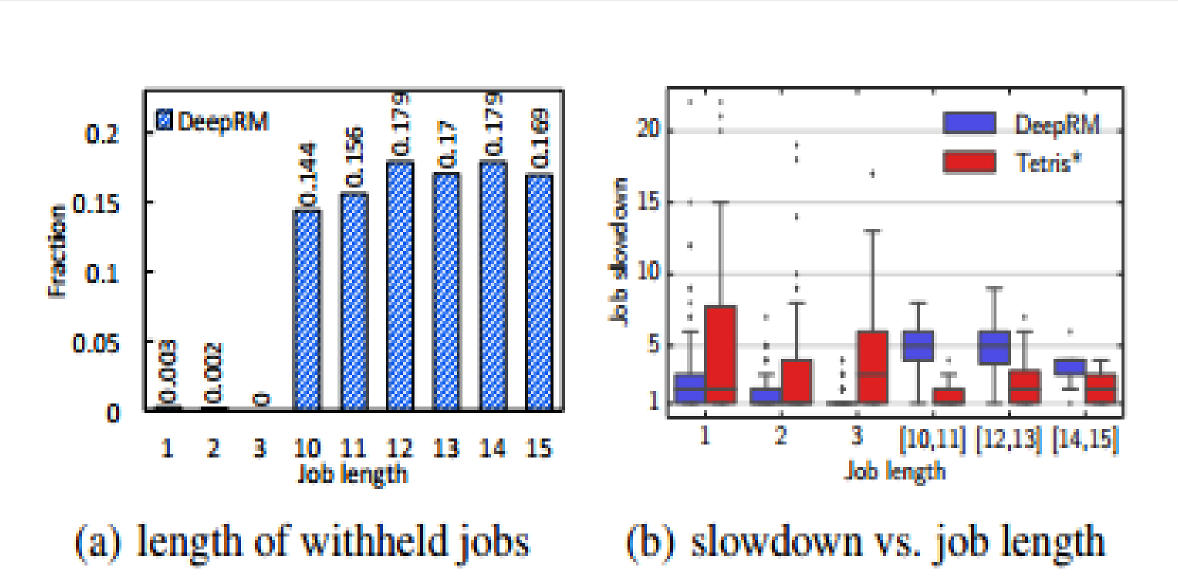

Figure 8 : Performance gain of DeepRM by withholding large jobs [3]

- ✓ DeepRM learns to withhold large jobs
- ✓ Slowdown for small jobs is significantly smaller with DeepRM than Tetris

UNIVERSITÄT PADERBORN

Limitation

- Overlooking machine boundaries and resource fragmentation

- Data local allocations

- Inter task dependency matters in real world

- Bounded time horizon to compute the baseline value

$$b_t = \frac{1}{N}\sum_{i=1}^{N} v_t$$

UNIVERSITÄT PADERBORN

Future Work

- Guiding the agent towards data local allocations with appropriate reward function
- Use of value network to compute the baseline

UNIVERSITÄT PADERBORN

References

[1] Yu, Lean & Wang, Shouyang & Lai, Kin Keung & Wu, Yue. (2005). A framework of Web-based text mining on the grid. 2005. 6 pp.-. 10.1109/NWESP.2005.3.

[2] https://www.kdnuggets.com/2018/03/5-things-reinforcement-learning.html

[3] H. Mao, M. Alizadeh, I. Menache, and S. Kandula, "Resource Management with Deep Reinforcement Learning," *Proceedings of the 15th ACM Workshop on Hot Topics in Networks*, pp. 50–56, 2016.

[4] R. Grandl, G. Ananthanarayanan, S. Kandula, S. Rao, and A. Akella. "Multi-resource Packing for cluster schedulers," *Proceedings of the 2014 ACM Conference on SIGCOMM*, pp. 455–466, 2014.

UNIVERSITÄT PADERBORN

Thank You !