Peter the Great St.Petersburg Polytechnic University

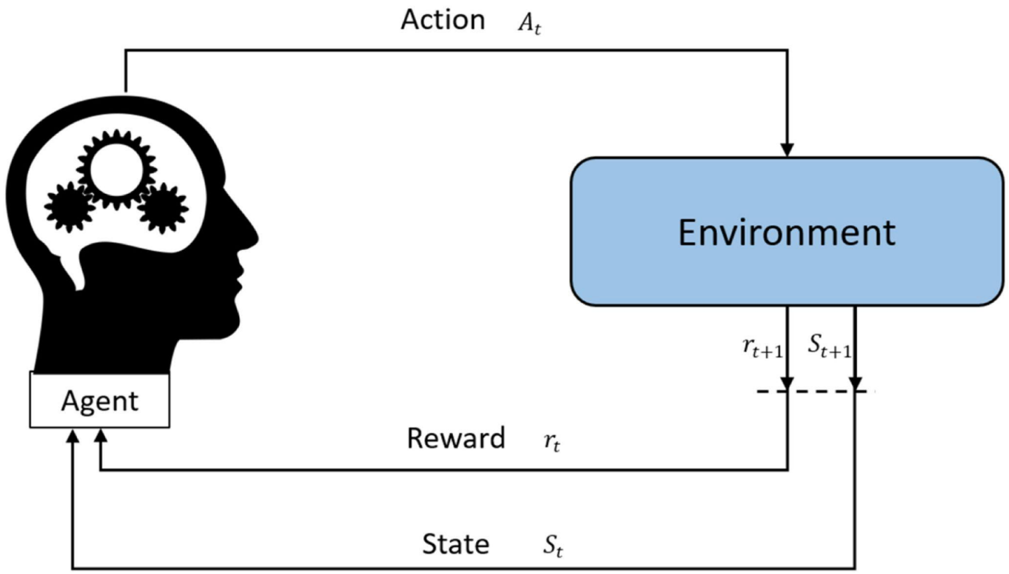
Resource scheduling optimization for industrial operating system using deep reinforcement learning and WOA algorithm

Daria Yashnova, 5120301/20102

2024

Problem statement

- Traditional approaches to resource instantiation scheduling rely on software-defined paradigms
- The large number of heterogeneous resources and their diverse QoS



Working of MDP

Whale optimization algorithm (WOA)

Encircling prey

$$\overrightarrow{D} = \left| \overrightarrow{C} . \overrightarrow{X^*} (t) - \overrightarrow{X} (t) \right|$$

t indicates the current iteration

X^* is the position vector of the best solution

$$\overrightarrow{X} (t+1) = \overrightarrow{X^*} (t) - \overrightarrow{A} \cdot \overrightarrow{D}$$

$$\overrightarrow{A} = 2\overrightarrow{a} \cdot \overrightarrow{r} - \overrightarrow{a}$$

$$\overrightarrow{C} = 2 \cdot \overrightarrow{r}$$

$\overrightarrow{a}$ is linearly decreased from 2 to 0 over the course of iterations

$\overrightarrow{r}$ is a random vector in [0,1].

Whale optimization algorithm (WOA)

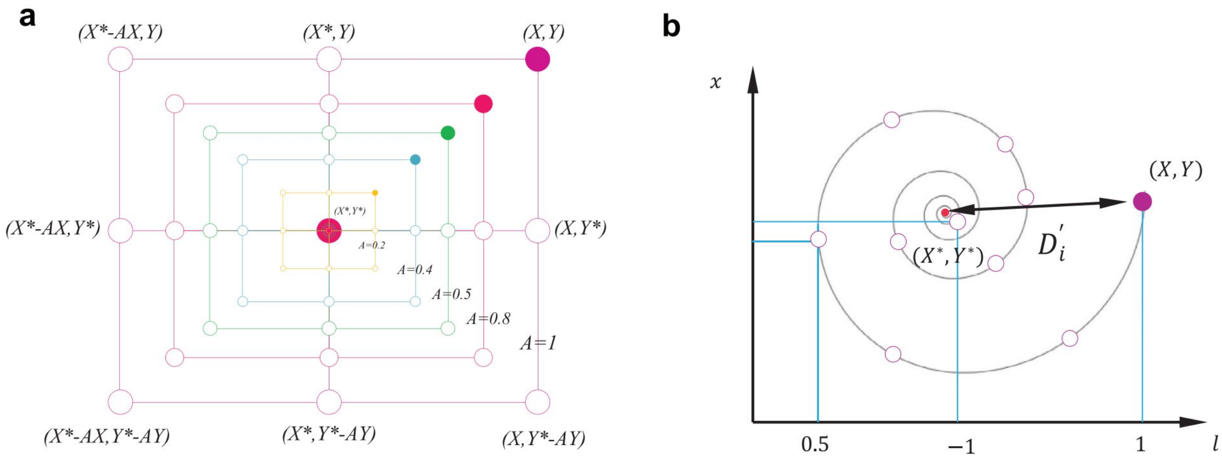
Bubble-net attacking the prey

$$\vec{X}(t+1) = \begin{cases} \vec{X^*}(t) - \vec{A} \cdot \vec{D} & if \quad p < 0.5 \\ \vec{D'} \cdot e^{bl} \cdot \cos(2\pi l) + \vec{X^*}(t) & if \quad p \geq 0.5 \end{cases}$$

$b = const$

$l \in [-1, 1] - \ random\ number$

$p \in [0, 1] - \ random\ number$



Bubble-net search mechanism implemented in WOA (X∗ is the best solution obtained so far):

(a) shrinking encircling mechanism and

(b) spiral updating position

Whale optimization algorithm (WOA)

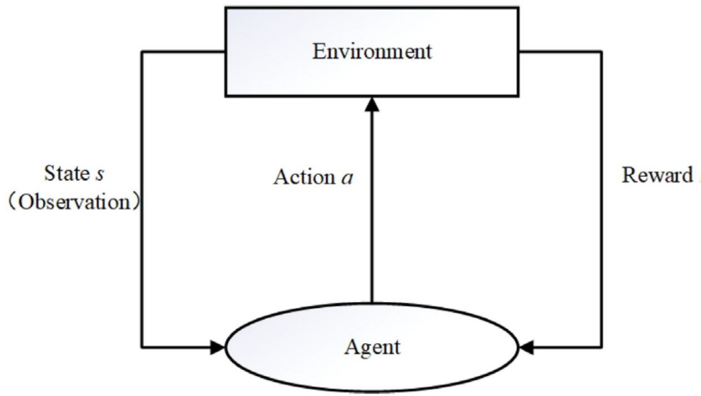
Searching prey

$$\vec{D} = |\vec{C} \cdot \vec{X}_{rand} - \vec{X}|$$

$$\vec{X}(t+1) = \vec{X}_{rand} - \vec{A}.\vec{D}$$

X_{rand} – *position vector of the randomly selected whale*

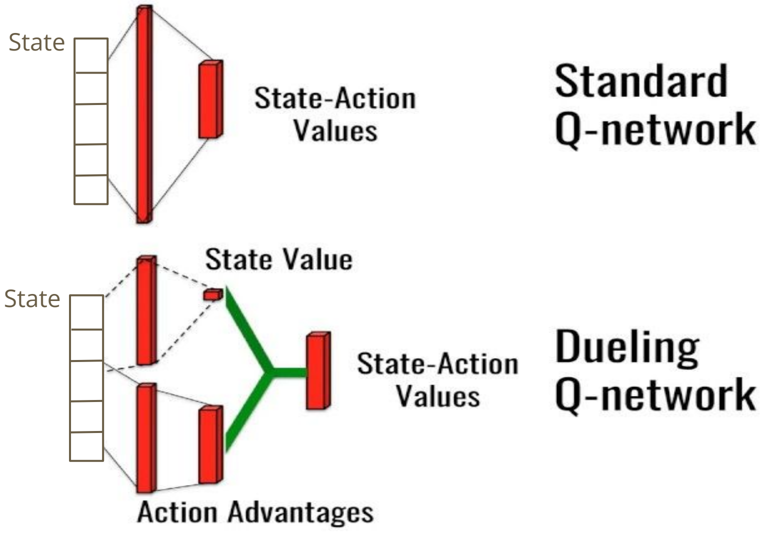
Deep reinforcement learning



Markov decision process of discrete events: $\langle S, A, P, r, \gamma \rangle$

- $S \subseteq R^k$ — *k-dimensional task state*
- $A \subseteq R^q$ — *q-dimensional action* $A = [a_1, a_2, \ldots, a_q]$
- $P_s(s_{t+1} | s_t, a_t)$ — *state transition probability*
- $r_s^a = E[r_{t+1} | s_t = s, a_t = a]$ — *reward function*
- $\gamma \in (0, 1)$ — *discount factor for future reward*

Deep Q-Network



$$Q_\pi(s,a) = \mathbb{E}\{\sum_{t=0}^{K} \gamma^t r_t | s_0 = s, a_0 = a, \pi\}$$

$$\pi^* = argmax\ Q_\pi(s,a)$$

$$Q(s,a) = V(s) + A(s,a)$$

V – *expected cumulative reward*

A – *relative advantage of action over other actions in a given state*

DWOA

State

$NDIV$ – *normalized population diversity*

$$NDIV = \frac{DIV}{|ub - lb|}$$

ub and lb – *upper and lower bounds on the parameters in the solution*

$$State = \{NDIV,\ |A|/2,\ |C|/2\}$$

$$A \in [-2, 2] \qquad C \in [0, 2]$$

Action

$$Action = \{\mathrm{SE},\ \mathrm{SU},\ \mathrm{RS}\}$$

SE – *shrinking encircling*
SU – *spiral updating*
RS – *random search*

Reward

$$Reward = \begin{cases} 1 & f(\vec{X}(t+1)) < f(\vec{X}(t)) \\ -1 & f(\vec{X}(t+1)) > f(\vec{X}(t)) \\ 0 & f(\vec{X}(t+1)) = f(\vec{X}(t)) \end{cases}$$

$f(X(t))$ – *fitness of current solution*

$f(X(t+1))$ – *fitness of the next solution*

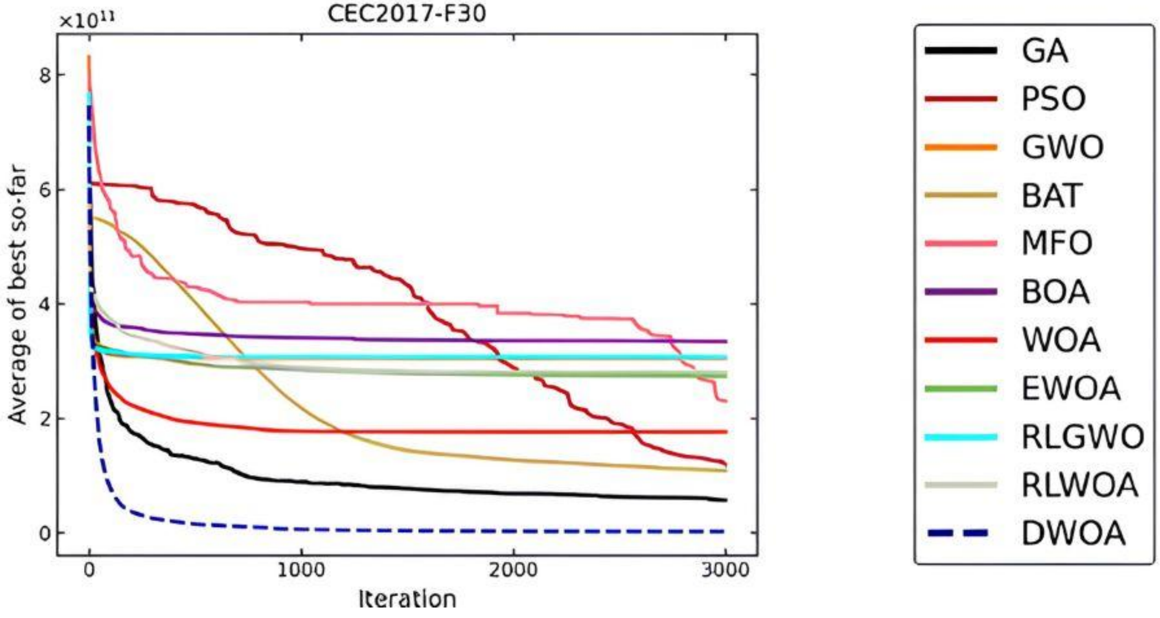
Friedman's test

F - rank – Friedman mean rank

Results of Friedman test on CEC2017 test suite for all algorithms.

Algorithms	F-rank				Final rank
	Dim=10	Dim=30	Dim=50	Dim=100	
GA	3.72	3.83	5.41	6.00	2/3/5/6
PSO	5.34	3.10	6.93	7.10	4/2/7/7
GWO	6.28	6.07	7.55	7.21	7/5/8/8
BAT	7.66	8.17	5.03	4.55	10/10/4/4
MFO	8.24	10.45	9.28	9.34	11/11/11/11
BOA	7.34	7.93	9.14	9.14	8/9/10/10
WOA	6.10	4.59	5.90	5.69	6/4/6/5
EWOA	7.52	6.52	8.03	7.45	9/6/9/9
RLGWO	5.34	7.14	3.31	3.93	4/8/2/3
RLWOA	5.24	6.83	3.55	3.86	3/7/3/2
DWOA	3.21	1.38	1.86	1.72	1/1/1/1

Convergence curves of the composition functions in CEC2017



Conclusion

A DWOA algorithm was introduced and analyzed with 11 well-known algorithms in the CEC2017 test functions and RIS tasks of different scales. Experimental results show that DWOA has significant advantages in terms of convergence rate, solution quality, and performance stability . The Wilcoxon rank test and Friedman test confirm the superiority of DWOA .