

Environment

Airborne-BSs (Agents)

Reinforcement Learning Algorithm

Observation



Control

SARSA or Deep Q Network

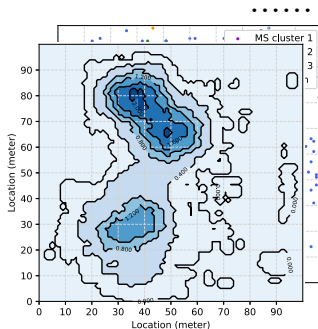
Algorithm 2: DQN implementation of two drones scenario

- 1: Initialization
- 2: **for** every episode j **do**
- 3: $s_1 \leftarrow$ random

Algorithm 1: SARSA implementation of two drones scenario

- 1: Initialization
- 2: **for** every episode j **do**
- 3: $s_1 \leftarrow$ random
- 4: **for** Every iteration t **do**
- 5: **for** Every drone δ **do**
- 6: $a_t \leftarrow$ Choose action based on current Q Table ($Q_{s_t}, \epsilon_i, \delta$)
- 7: $s_{t+1} \leftarrow$ Take the previous action (s_t, a_t, δ)
- 8: $a_{t+1} \leftarrow$ Choose action based on real scenario ($Q_{s_{t+1}}, \epsilon_i, \delta$)
- 9: $r_t \leftarrow$ Compute the reward (s_{t+1})
- 10: $Q(s_t, a_t, d) \leftarrow$ Update Q Table
- 11: $s_t \leftarrow s_{t+1}$
- 12: **end for**
- 13: **end for**
- 14: **end for**

Mobile Stations
Airborne-BSs
Building Position
Connection Condition



Environment information send to