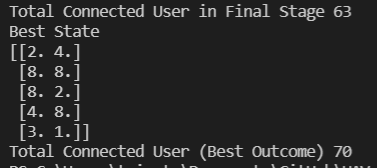
Based on the updated code the following different simulations are to done to obtain as much information as possible.

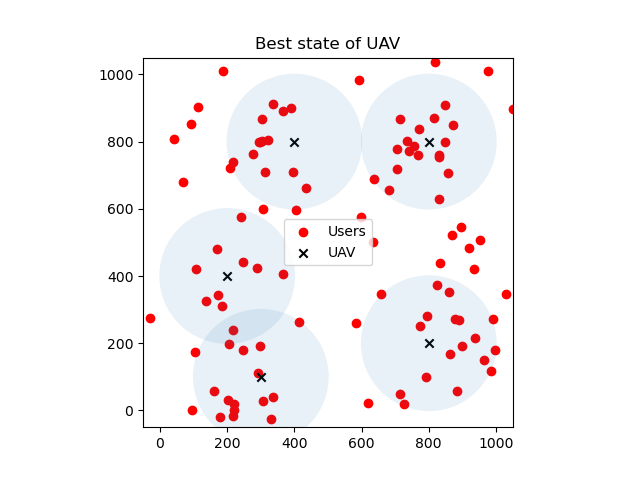
1. With collective number of connected users as the reward system

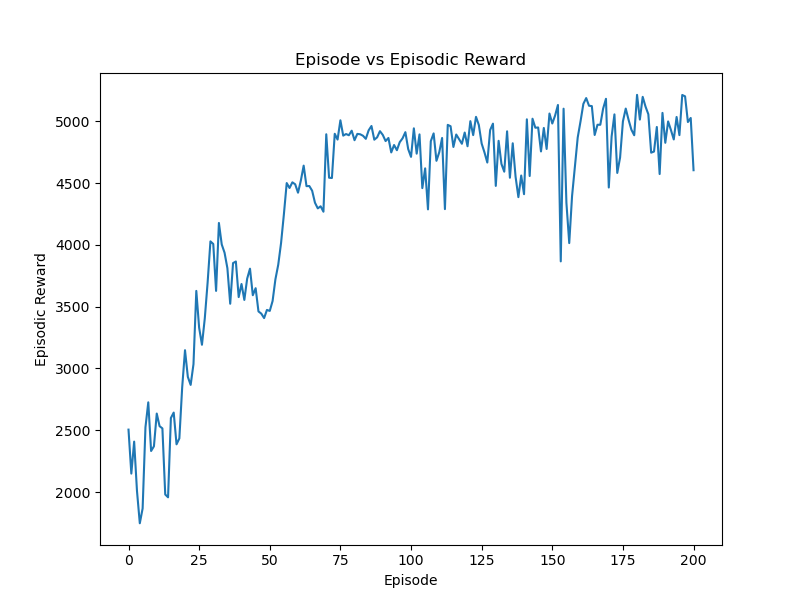
The reward of the individual UAV is flagged based on the boundary condition and is penalized accordingly by 2.

1. 1st User Distribution with hotspot points (Previously used value)
2. Starting position of the UAV at one corner of the target area (0, 0)

Since the start to the extreme corner (if optimal) requires 20 steps at minimum to complete so, having the knowledge of future 20 steps is necessary so a gamma value of 0.95 based on the previous formula from literatures.

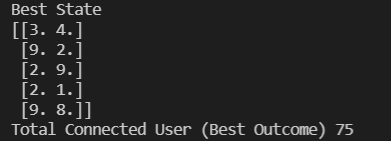


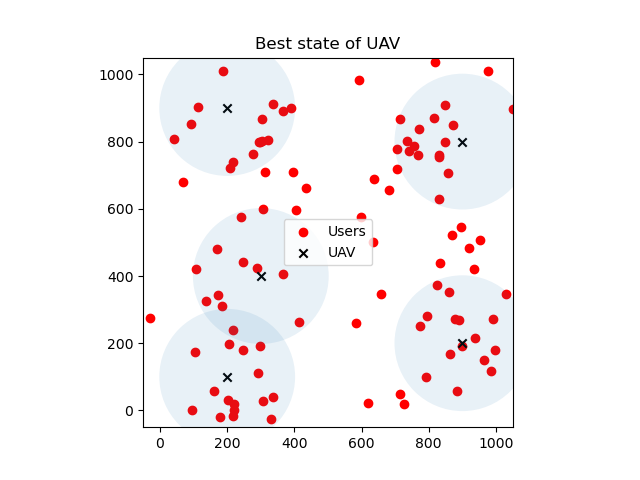


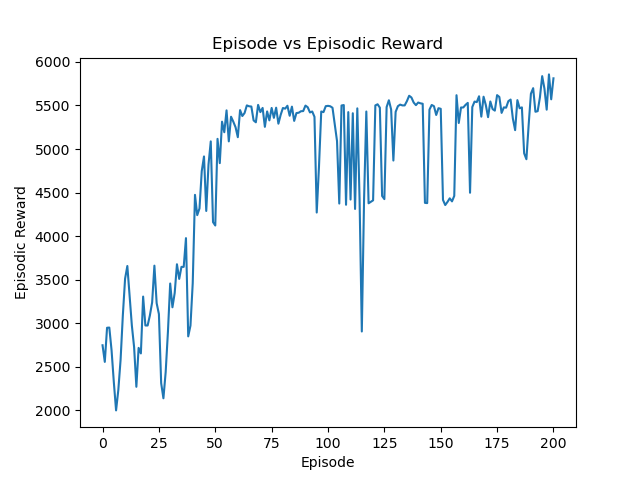


1. Starting at the center of the target area (5, 5)

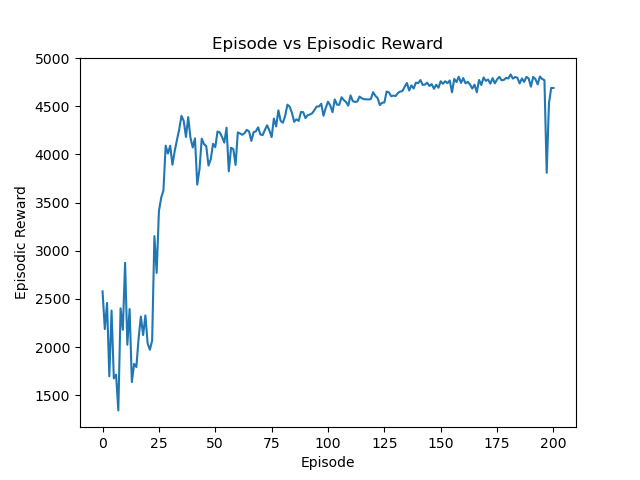
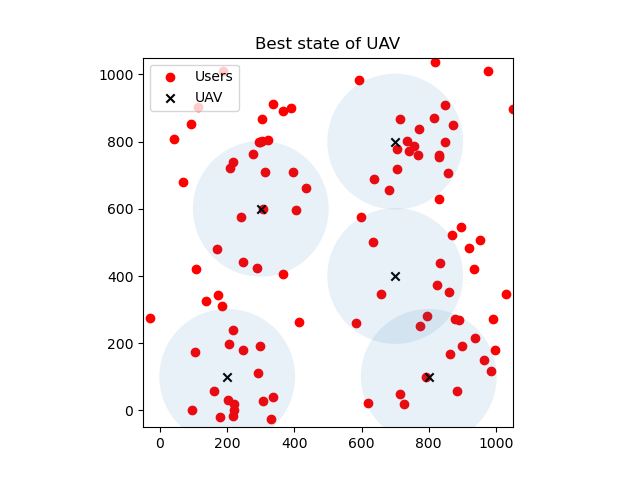
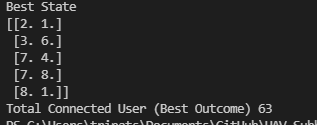
Since the start to the extreme corner (if optimal) requires 10 steps at minimum to complete so, having the knowledge of future 10 steps is necessary so a gamma value of 0.90 based on the previous formula from literatures.







1. 2nd User Distribution with uniform user along the target area
   1. Starting position of the UAV at one corner of the target area (0, 0)



* 1. Starting at the center of the target area (5, 5)

