

Human Swarm state space:

Separate state space for each group of civilians. We learn one policy for the interactions between a UAV and a civilian group and different instances of the same policy will run simultaneously with the different civilian groups. The only binding factor is the 'operator busy time' which will be a common variable across the different instances of the policy.

State variables:

1) Operator busy till...? (can be initialized with any of these values):

- 0 – Operator is free
- 1 – Operator will take a short time
- 2
- 3 - Operator will take a long time

2) Status of group (always start at status 0):

- 0 – Being monitored
- 1 – Being warned
- 2 – SPN
- 3 – ON
- 4 – Ready to move to safe zone
- 5 – Moving to safety/ saved

3) Dead Time i.e. time taken for the fire to reach group (can be initialized with any of these except 0)

- 0 – Fire is at the spot
- 1
- 2
- 3
- 4 – Fire is far away

4) Guidance information (initialized with any of these values)

0 - unknown

1 – self guide

2 – UAV guide

3 – vehicle guide

5) Negotiation status (initialized with 0)

0 – requires warn i.e. no warnings issued

1 – requires SPN i.e. warning issued

2 – requires ON i.e. SPN issued

Number of states = 1440

Actions:

- 1) Interrupt Operator
- 2) Warn
- 3) SPN
- 4) Query for guide details (only when group is ready to move?)
- 5) Guide
- 6) Call for vehicle guide
- 7) Wait

Rewards:

- 1) Group saved: +5000
- 2) Group dies: -5000