Scheduling Surveillance of Space Objects

Benedict Oakes, Supervised by Jason Ralph (UoL), Dominic Richards (STFC), Jordi Barr (Dstl)

EPSRC Centre for Doctoral Training in Distributed Algorithms, University of Liverpool, Liverpool, UK

Background & Aims

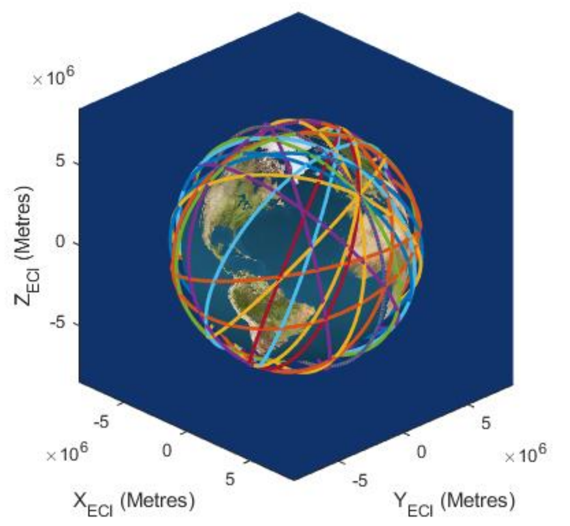
As the number of satellites in space continues to increase, it becomes harder to keep track of them with limited sensing capability. This project aims to use reinforcement learning to improve non-myopic (long-term) scheduling of limited sensor resources for better tracking performance.

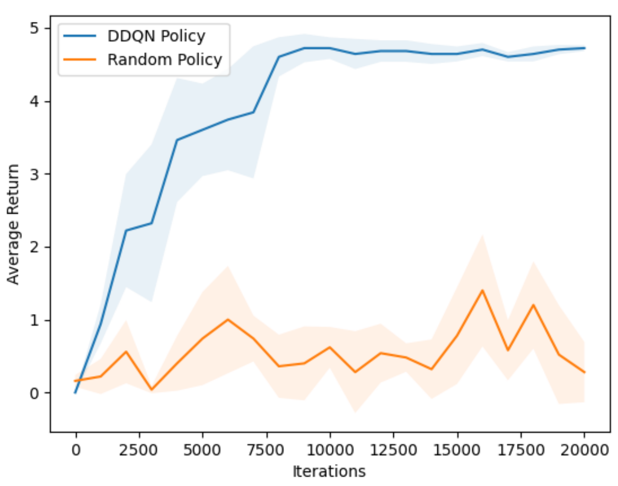
- The UK is interested in becoming more self-reliant in the space domain, including space situational awareness (SSA) [1].
- Between $^1/_3$ and $^1/_2$ of low Earth orbit space has already been filled [2], increasing the need for reliable tracking.
- The risk of collisions between satellites and producing debris becomes more likely, and needs to be understood to minimise damage.

- Reinforcement Learning (RL) offers an opportunity to model high-dimensional problems easily.
- RL: Intelligent agents making decisions to maximise a cumulative reward [3].
- Q-learning is a value iteration update on a Markov decision process. Q-values determine the reward gained by taking an action in a certain state.
- We use a Double Deep Q Network (DDQN) on simulated satellites.
- Paper: Double Deep Q Networks for Sensor Management in Space Situational Awareness
- We apply a DDQN to a controllable telescope attempting to observe multiple satellites, which learns the best actions to take.

Ongoing work

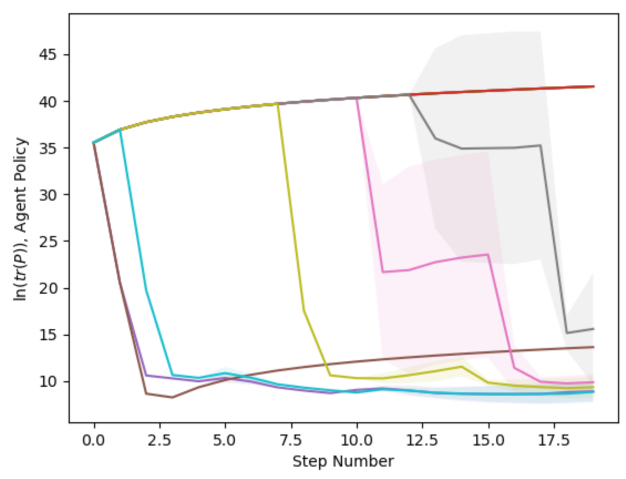
- The controllable telescope has a discretised action space that can point in the 4 cardinal directions.
- At teach time step, the agent can move the telescope in one direction, or do nothing.
- The agent is rewarded for capturing a satellite within the telescope's field of view.
- The agent will record measurements of captured satellites for use in an extended Kalman filter (EKF).

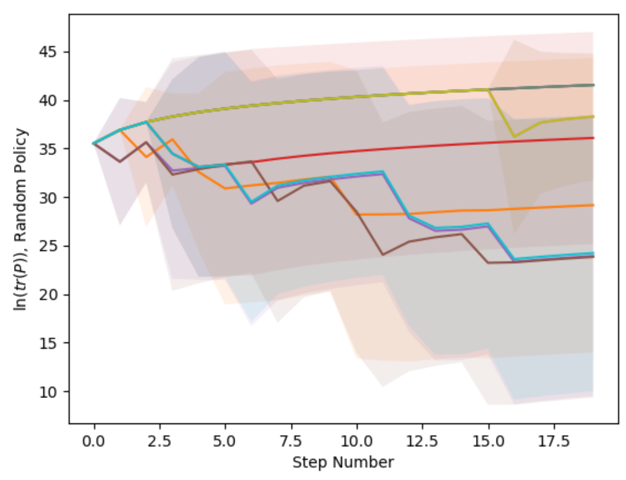




Preliminary Results

- As agent trains, we see smaller uncertainties of satellite position and velocity from the EKF as more successful detections and measurements of satellites are made.
- Compared to a random baseline policy, we see drastically improved tracking performance, using covariance trace as a performance metric.





Future Work

- Inclusion of more advanced satellite dynamics, including solar radiation pressure and atmospheric drag, and better filters.
- Use of more realistic/real world data, including angle only measurements.
- Continuous action agents to reflect real telescope properties.
- Tracking performance metrics, e.g. PCRB or GOSPA [4], [5].

Conclusion

This project will contribute to the UK's continued space capability and SSA using reinforcement learning and target tracking.

References

[1]: Ministry of Defence. National Security Strategy and Strategic Defence and Security Review. 2015. url: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/555607/2015_Strategic_Defence_and_ Security_Review.pdf (visited on 05/18/2021)

[2]: Carmen Pardini and Luciano Anselmo. "Evaluating the impact of space activities in low earth orbit". In: Acta Astronautica Volume 184 (2021), pp. 11–22.

[3]: R.S. Sutton, A.G. Barto, "Reinforcement learning: An introduction", Malaysia: MIT press, 2018

[4]: M. L. Hernandez, A. Farina, "Posterior Cramer-Rao Bound for Target Tracking in the Presence of Multipath", 2018 21st International Conference on Information Fusion (FUSION), pp. 151-158, 2018

[5]: A. S. Rahmathullah, A. F. Garcıa-Fernandez, L. Svensson, "Generalized optimal sub-pattern assignment metric", 2017 20th International Conference on Information Fusion (FUSION), pp. 1-8, 2017

Engineering and Physical Sciences Research Council

UNIVERSITY OF LIVERPOOL

[dstl] The Science Inside