IAF ASTRODYNAMICS SYMPOSIUM (C1)
Interactive Presentations - IAF ASTRODYNAMICS SYMPOSIUM (IP)

Author: Mr. Edward Tomanek-Volynets
University of Glasgow, United Kingdom


Dr. Matteo Ceriotti
University of Glasgow, United Kingdom
Prof. Colin R. McInnes
University of Glasgow, United Kingdom

# MULTI-TARGET SPACE MISSION SEQUENCE OPTIMIZATION WITH DEEP REINFORCEMENT LEARNING

## Abstract

Multi-target space missions have attracted interest in recent years due to their potential application in scenarios such as space debris removal or near-Earth asteroid rendezvous. Instead of visiting a single destination (target) per mission, a spacecraft visits several, saving money, emissions and time incurred by further launches. The trajectory design of such missions involves selecting optimal ordered subsets (sequences) from a large pool of targets.

This sequence optimization is addressed by the travelling salesman problem (TSP) and its variants. Obtaining good solutions to TSPs using conventional methods (such as ant colony optimization (ACO) or other bio-inspired algorithms) is extremely time-consuming for large numbers of targets. Applying the problem in an astrodynamic context involves another bottleneck because calculating the cost of transfers between two targets requires optimizing trajectories between pairs of orbits, which can be computationally expensive. Therefore, methods that minimize the number of these cost function evaluations result in significant time savings.

This paper will present a strategy to produce approximate solutions for the sequence optimization problem in very short amounts of time. The objective is to enable a spacecraft to visit a large number of selected near-Earth targets in a round trip, whilst minimizing the total $\Delta V$. Each target is on its own Keplerian orbit. The $\Delta V$ required to move from one target to another is given by a Lambert arc, where the time of flight, initial and final true anomalies are adjusted to minimize the $\Delta V$.

A sequence-to-sequence neural network is trained with reinforcement learning on problems with up to 40 targets. It is shown that the resulting model maintains a good solution quality for up to around 200 targets. It is found that the optimality surpasses the nearest neighbour algorithm and achieves within 10% of ACO. Importantly, it will be demonstrated that the method's solution time increases linearly with the number of targets. This makes the algorithm significantly more efficient than any conventional solution method, especially for medium to large problem sizes.

A related method is investigated, in which the trained network is post-trained directly on specific problems. It is found that the ACO solution can be approximately matched with around 70% fewer cost function evaluations.

Finally, a similar approach is applied to a more realistic problem where the total $\Delta V$ is fixed and the return from the mission must be maximized.