(Preprint) AAS 25-XXX

SPACE-TO-SPACE SURVEILLANCE USING AUTONOMOUS ON-BOARD SCHEDULING

Daniel Huterer Prats[*] and Hanspeter Schaub[†]

The increasing number of Resident Space Objects (RSOs) in Low Earth Orbit (LEO) presents new challenges for autonomous space-based Space Situational Awareness (SSA). Unlike traditional Earth observation, SSA requires agile imaging of dynamically moving targets under strict constraints such as limited power, line-of-sight, and lighting conditions. This work formulates the RSO inspection task as a Partially Observable Markov Decision Process (POMDP) and employs reinforcement learning (RL) to train an onboard agent for dynamic target selection and resource management. Using the BSK-RL environment and a high-fidelity spacecraft simulator, an actor-critic RL agent learns to autonomously image RSOs while maximizing coverage and adhering to subsystem limitations. The anticipated results demonstrate the agent's ability to generalize across orbital regimes and support future scalable SSA and Space Domain Awareness (SDA) missions.

INTRODUCTION

The exponential growth of resident space objects (RSOs), fueled by the deployment of numerous LEO constellations and the increasing sensitivity of modern sensors, has led to a surge in cataloged space assets. Estimates indicate over 9700 active satellites currently orbiting in LEO, with projections of more than 1700 launches annually by 2030.[1–3] This surge is creating significant strain on existing SSA capabilities, particularly for narrow FOV sensors that must efficiently track and revisit high-priority targets.

Unlike classical Earth observation satellites that monitor static ground locations, space-based SSA systems must address dynamically moving targets in diverse orbital planes. These scheduling tasks have been tackled using reinforcement learning (RL) approaches before.[4–8] The core of RL training involves agents receiving environment states, selecting actions based on policy networks, and receiving scalar rewards.

Building upon these foundations, this work focuses on RL-driven scheduling for a maneuverable space-based sensor in LEO. Such platforms must autonomously balance multiple mission constraints, including line-of-sight (LOS) access to the targets, power and momentum limits, eclipse periods, and finite onboard data storage. Actor-critic RL methods such as PPO are well-suited for this task, offering robust performance in large action spaces.[5] Our proposed agent autonomously selects imaging targets and manages resources across extended durations without ground intervention.

[*]PhD Student, Department of Aerospace Engineering Sciences, University of Colorado Boulder, 3775 Discovery Dr., Boulder, CO 80303.

[†]Professor, Department of Aerospace Engineering Sciences, University of Colorado Boulder.

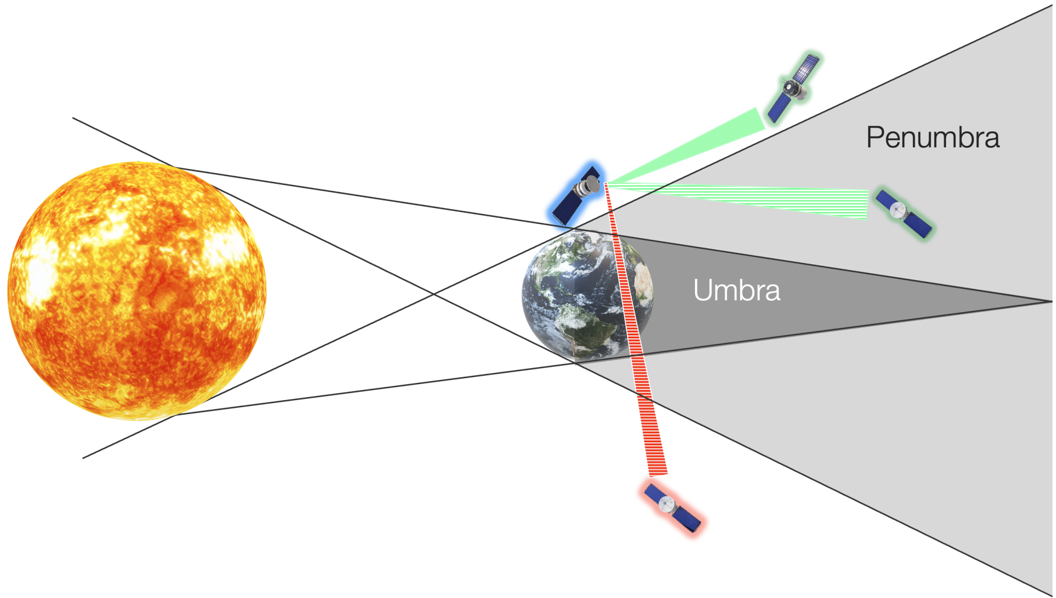

Figure 1. Space-based RSO imaging under eclipse and Line-of-Sight (LOS) restrictions. Green represent a succesful imaging whereas shaded would give partial reward to the agent as it is partially eclipsed. Red would violate the LOS constraint and hence no reward would be given for imaging that target.

PROBLEM FORMULATION

In the imaging task, an imaging spacecraft aims to image all RSOs in the environment considering also their illumination condition, namely eclipse status. To achieve this, the servicer must decide which action to take (i.e. which target to image based on their relative position and the relative motion with the RSO), while also being subject to battery, lighting and storage constraints.

This project addresses the problem of designing an autonomous spacecraft that can dynamically task itself to inspect multiple RSOs under resource constraints without human intervention.

The specific challenges addressed include:

- Continuous, real-time scheduling of multiple imaging tasks.
- Operating under limited energy, momentum, and data buffer resources.
- Respecting geometric constraints such as LOS and FOV.
- Adapting to orbital dynamics and Earth shadow (eclipse) conditions.
- Prioritizing targets based on historical data or policy goals.

Space-based SSA is very complex and cannot be accurately modeled with static targets, as is done with agile Earth observation satellites (AEOS), which focus on fixed ground targets. Instead, in the case of a space-to-space scanning satellite, the orbital dynamics and specifically the relative motion with respect to target RSOs are much more complicated and harder to predict as compared to the AEOS environment.[7,9] This is because space-based target are flying in and out of view from all possible directions since they have all types of possible inclinations at various different orbital regimes. Therefore it is a lot harder to predict which targets will be in the FOV and hence at every decision step of the agent, the list of targets that are within FOV needs to be updated (as will be seen later in this section)! Note that in this study only altitudes ranging from Low LEO to almost the edge of MEO are considered.

POMDP Formulation

The elements of the POMDP tuple for the inspection task are as follows:

- **State Space:** The underlying simulator state provides the generative model for the MDP. It includes the physical dynamics of the inspector spacecraft and Resident Space Objects (RSOs), internal subsystem states (e.g., battery level, data storage, momentum buildup), and external environmental states (e.g., lighting, eclipse status).
- **Observation Space:** The agent observes a partial view of the full simulator state. Observations are composed of selected, normalized quantities relevant to the imaging task. The components of the observation space are detailed in Table 1.

Element	Dim.	Description
s_{data}	1	Fraction of onboard data storage used
s_{batt}	1	Normalized battery charge level
$\boldsymbol{\epsilon}_i$	N	Elevation angles of visible targets, $i = 1, \ldots, N$
$\boldsymbol{r}^{\mathcal{N}}_{BR,i}$	$3 \times N$	Relative position vectors to targets in inertial frame
$\boldsymbol{\theta}_i$	N	Angle between boresight and target i
$\boldsymbol{d}_i$	N	Distances to target RSOs
$\boldsymbol{los}_i$	N	Binary flags for whether target i is observable (within Line-of-Sight)
$\boldsymbol{\mathcal{I}}_i$	N	Binary flags for whether target i has already been imaged

Table 1. Observation space elements provided to the agent at each timestep.

- **Action Space:** At each timestep, the agent selects one action $a \in \{0, 1, ..., N-1\}$ corresponding to a visible RSO in the sorted elevation list. The list includes up to $N = 32$ targets above $-14°$ elevation and is ordered in ascending elevation. The selected action indicates the spacecraft's intent to image the target at that index.
- **Reward Function:** The reward $r(t)$ at time t is determined by whether the selected target i was successfully imaged, which depends on satisfying multiple physical and operational constraints:

$$r(t) = \begin{cases} w_i \cdot \left(1 - \dfrac{e_i(t)}{e_{\text{thresh}}}\right), & \text{if } \mathcal{C}_i(t) = \texttt{true}, \\ 0, & \text{otherwise}, \end{cases} \tag{1}$$

  where:

  - w_i is the assigned priority weight of target i,
  - $e_i(t)$ is the eclipse value of the target i at time t,
  - $e_{\text{thresh}} = 0.6$ is the eclipse threshold for valid imaging,
  - $\mathcal{C}_i(t)$ is the set of constraints that must be satisfied for a successful image:

The term $\left(1 - \frac{e_i(t)}{e_{\text{thresh}}}\right)$ linearly scales the reward based on the lighting condition, assigning full reward under ideal lighting and reducing it to zero as eclipse conditions approach the threshold.

$$\mathcal{I}_i(t) = 0, \quad \text{(Target has not yet been imaged)} \tag{2}$$
$$e_i(t) < e_{\text{thresh}}, \quad \text{(Eclipse value below threshold)} \tag{3}$$
$$\text{LOS}_i(t) = 1, \quad \text{(Line-of-sight present at end of imaging)} \tag{4}$$
$$\theta_i(t) < \theta_{\max}, \quad \text{(Within pointing constraint)} \tag{5}$$

The agent's goal is to maximize the total accumulated reward over an episode, subject to spacecraft dynamics and subsystem constraints (e.g., battery, data storage, imaging cadence):

$$R = \sum_{t=0}^{T} r(t), \tag{6}$$

where T is the episode time horizon.

- **Transition Model:** Instead of a probabilistic transition function, the transition model is implemented as a deterministic generative model. When an action is taken, the environment propagates the simulation forward in time until the next action is to be taken. The Δt for the step is returned as the duration the simulator propagated that step.

  The environment is implemented in accordance with the Gymnasium API[10] using BSK-RL[*], a package for creating modular, high-fidelity spacecraft tasking RL environments.[11] The underlying spacecraft simulation is Basilisk, a high-performance spacecraft simulation package.[12] Rigid multibody dynamics in the perturbed orbital environment and flight-proven flight software algorithms are used to simulate the environment.

EXPECTED RESULTS

While the simulation has already been setup in a Basilisk environment, and the pointing-control of the scanning satellite has been verified, the training of the RL agent is currently in process. An example of the current pointing performance to 7 subsequently targeted RSOs is shown in Figure 2 through their 3D trajectories below. The precise pointing error over time is shown in Figure 3 along with the angular velocity produced by the reaction wheels on board the imaging satellite. Note that even in the worst case scenario of a very large initial pointing error and relative motion in the same direction as the rotation of the imaging satellite (target 6 in Figure 2 starting at timestep 1800), the alignment of the boresight axis to the target is achieved within 200 sec. Therefore, a duration of $a_{image} = 300$ sec was chosen for the imaging action. This will ensure that the imaging satellite has sufficient time to rearrange itself and align its boresight with the target, but at the same time it will not remain locked pointing to a target for much longer than necessary (for settling of the instruments and gimbals etc on board).

Based on similar work by Siew et al, it is plausible that the agent discovers an efficient routine where it essentially splits the imaging task into azimuthal and elevation components.[5] As seen in

[*] https://avslab.github.io/bsk_rl/

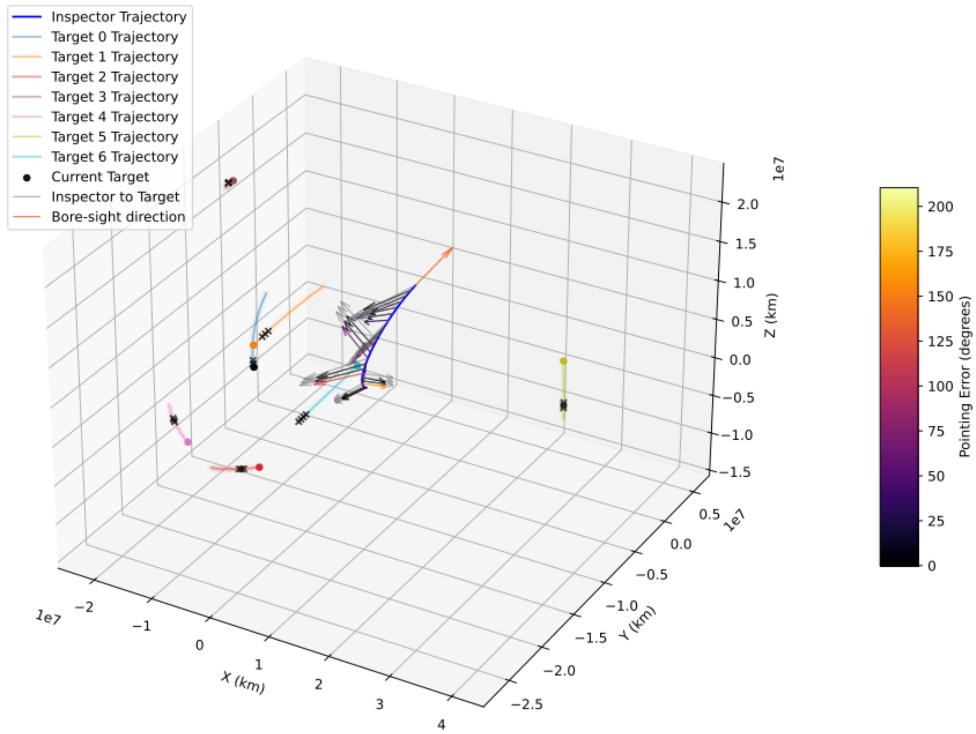

Figure 2. 3D trajectory of the scanning satellite and RSOs.

their paper, which focuses on imaging of GEO satellites from a LEO platform, the agent starts to continuously spin clockwise (or counter-clockwise) and then based on the targets tilts the boresight axis up/down in elevation. Note that their methodology is inherently different from the one proposed here as their action space chooses a grid within the 'Field or Regard' and is not looking at specifc targets. Moreove, they do not consider the eclipse conditions of the GEO targets and also are not taking into account the safety constraints such as power, buffer space etc. In this study, since all the targets are LEO RSOs it is possible that this 'strategy' is no longer feasible since the relative motion with respect to the scanning satellite is much faster and dynamic. Moreover, while the training will occur with the agent in a specific orbital inclination, the resulting Neural Network will also be tested by applying it on a platform with other initial orbital inclinations. Furthermore, the number of RSOs plays an important role in the optimization as in LEO only an average of about 6% of other satellites will be visible at any given time. This is in contrast with a much larger percentage of visible targets, when they are at a significantly higher altitude.

The anticipated outcome is an RL-based agent capable of continuously prioritizing RSOs, balancing energy and momentum resources, and autonomously adjusting imaging sequences in response to evolving mission conditions. The findings will support the development of intelligent, adaptive surveillance strategies for SSA/SDA, providing a scalable approach for future autonomous spacecraft tasked with monitoring space traffic and high-priority orbital regions.

SIGNIFICANCE TO ASTRODYNAMICS AND SPACE-FLIGHT MECHANICS

This work introduces an application of reinforcement learning (RL)-based scheduling, traditionally applied to Earth Observation platforms, to the domain of *space-to-space* surveillance. Contem-

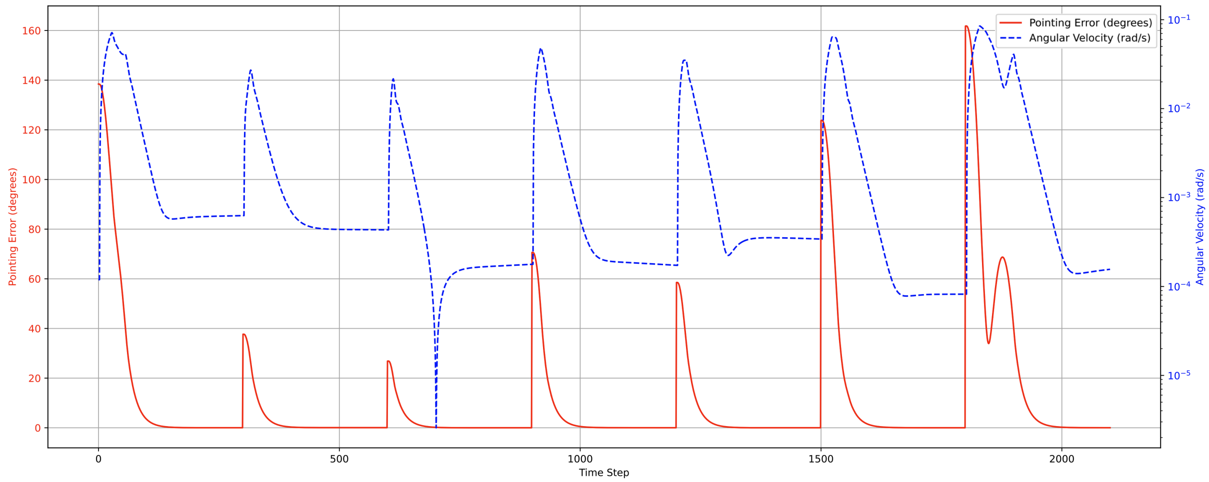

Figure 3. Camera pointing error over time for the same 7-target scenario.

porary SSA architectures heavily rely on ground-based telescopes and radars, which are limited by Earth-fixed fields of regard and atmospheric interferences. In contrast, the proposed space-based approach enables high-agility, dynamic observation of RSOs, unobstructed by terrestrial weather or lighting conditions.

From an astrodynamics perspective, this research addresses the complex control and planning challenges associated with maneuvering and pointing a spacecraft-mounted camera at targets in Low Earth Orbit. Unlike Earth-fixed imaging targets, RSOs are themselves in rapid orbital motion, necessitating predictive tracking under strict attitude constraints, variable lighting, and intermittent line-of-sight. The inclusion of eclipse-aware imaging, where reward diminishes linearly as eclipse conditions worsen, is an example of how orbital mechanics and sensor physics directly influence scheduling policies.

The approach also holds long-term implications for operational SDA capabilities. Future developments will explore *uncertainty-aware* imaging strategies, where the scanning satellite prioritizes targets with uncertain state estimates to enhance the accuracy of RSO catalogs. Additionally, this work can be extended to multi-agent systems, in which coordinated inspection constellations provide persistent surveillance over key orbital regions or specific geopolitical zones of interest.

Furthermore, the inclusion of multi-sensor fusion—e.g., radar, infrared, or hyperspectral payloads—may provide resilience under eclipse or obscured conditions, ensuring robust detection and classification of RSOs. This aligns with the broader goals of ensuring space safety and sustainability in the face of increasing orbital congestion and potential adversarial actions.

Ultimately, this research contributes a foundational capability toward intelligent, autonomous, and persistent space-based surveillance. By enabling adaptive space-to-space targeting, the proposed methodology strengthens the astrodynamics community's toolkit for real-time decision-making in dynamic orbital environments, fostering a secure and operationally effective SSA/SDA infrastructure.

REFERENCES

[1] “Over 480 Orbital Launches and 43,000 Active Satellites Expected by 2032,” 2024.

[2] F. ELMASRY, “New Satellite Market Forecast Anticipates 1,700 Satellites to Be Launched on Average per Year by 2030 as New Entrants and Incumbents Increase Their Investment in Space,”

[3] Union of Concerned Scientists, “UCS Satellite Database,” Online, 2023. `https://www.ucsusa.org/resources/satellite-database`.

[4] R. Linares, H. Khalilian, and C. Pieng, “Sensor Tasking for Space Situational Awareness via Reinforcement Learning,” *Advanced Maui Optical and Space Surveillance Technologies Conference (AMOS)*, 2021.

[5] P. M. Siew, D. Jang, T. G. Roberts, and R. Linares, “Space-Based Sensor Tasking Using Deep Reinforcement Learning,” *Journal of the Astronautical Sciences*, Vol. 69, No. 6, 2022, pp. 1855–1892, 10.1007/s40295-022-00354-8.

[6] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*. MIT Press, 2nd ed., 2018.

[7] L. Q. Mantovani, Y. Nagano, and H. Schaub, “Reinforcement Learning for Satellite Autonomy Under Different Cloud Coverage Probability Observations,” *AAS/AIAA Astrodynamics Conference*, 2023.

[8] M. A. Stephenson and H. Schaub, “Optimal Agile Satellite Target Scheduling with Learned Dynamics,” *Journal of Spacecraft and Rockets*, 2024, pp. 1–12, 10.2514/1.A36097.

[9] M. Stephenson and H. Schaub, “Reinforcement Learning for Earth-Observing Satellite Autonomy with Event-Based Task Intervals,” *AAS/AIAA Space Flight Mechanics Conference*, 2024.

[10] M. Towers, J. K. Terry, A. Kwiatkowski, J. U. Balis, G. de Cola, T. Deleu, M. Goulão, A. Kallinteris, A. KG, M. Krimmel, R. Perez-Vicente, A. Pierré, S. Schulhoff, J. J. Tai, A. J. S. Tan, and O. G. Younis, “Gymnasium,” Oct. 2023.

[11] M. A. Stephenson and H. Schaub, “BSK-RL: Modular, High-Fidelity Reinforcement Learning Environments for Spacecraft Tasking,” *75th International Astronautical Congress*, Milan, Italy, IAF, Oct. 2024.

[12] P. W. Kenneally, S. Piggott, and H. Schaub, “Basilisk: A Flexible, Scalable and Modular Astrodynamics Simulation Framework,” *Journal of Aerospace Information Systems*, Vol. 17, Sept. 2020, pp. 496–507, 10.2514/1.I010762.