Towards Realistic Earth-Observation Constellation Scheduling: Benchmark and Methodology

Luting Wang[1#] Yinghao Xiang[1#] Hongliang Huang[1]
Dongjun Li[1] Chen Gao[1*] Si Liu[1*]

[1]Beihang University
[#]Equal Contribution [*]Corresponding Authors

Friday, November 07

Table of Contents

1 Introduction

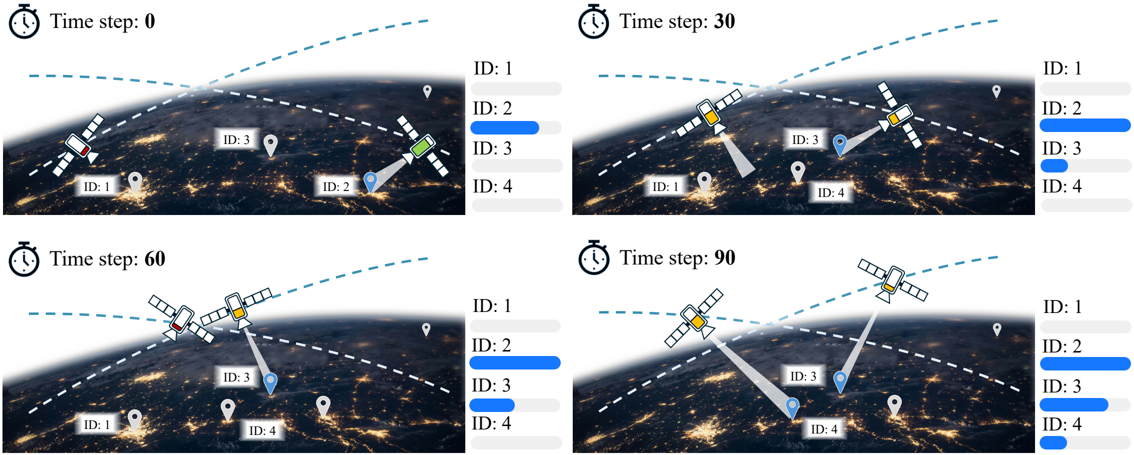
2 AEOS-Bench: Benchmark Suite

3 AEOS-Former: Proposed Method

4 Experiments

5 Conclusion

Problem Overview



Problem Overview

Agile Earth Observation Satellites (AEOSs)

- Rapid and flexible Earth surface monitoring
- Cooperative constellation operations
- Increased revisit frequency and coverage

Applications

- Disaster response
- Environmental monitoring
- Resource management

The Scheduling Problem

Goal

- Maximize task completion
- Minimize time and resource expenditure
- Satisfy all real-world constraints

Key Challenges

- Large Scale
- Dynamic Environment
- Strict Constraints

Key Challenges

Challenge 1: Large Scale

- Modern constellations: **large**
- Imaging requests: **hundreds of tasks**
- Exhaustive search is infeasible
- Strains heuristic and RL methods

Challenge 2: Dynamic Environment

- Tasks appear or expire at any moment
- Satellite status changing
- Satellites may join or leave

Challenge 3: Strict Constraints

- Available battery energy limits
- Satellite's attitude control
- Allowable time windows for each task
- Duration time differs from task to task
- Any violation $\Rightarrow$ task failure

Our Contributions

Contribution 1: AEOS-Bench

A standardized benchmark suite for realistic AEOS constellation scheduling

- High-fidelity simulation platform (Basilisk engine)
- Ground truth scheduling annotations including **16,410** diverse scenarios which contains **3,907** finely-tuned satellite assets
- First large-scale benchmark for realistic constellation scheduling

Contribution 2: AEOS-Former

A Transformer-based scheduling model with novel components

- Constraint-aware attention mechanism and Simulation-based iterative learning
- Internal constraint module for physical/operational limits
- Outperforms baselines in task completion and energy efficiency

Table of Contents

1 Introduction

2 AEOS-Bench: Benchmark Suite

3 AEOS-Former: Proposed Method

4 Experiments

5 Conclusion

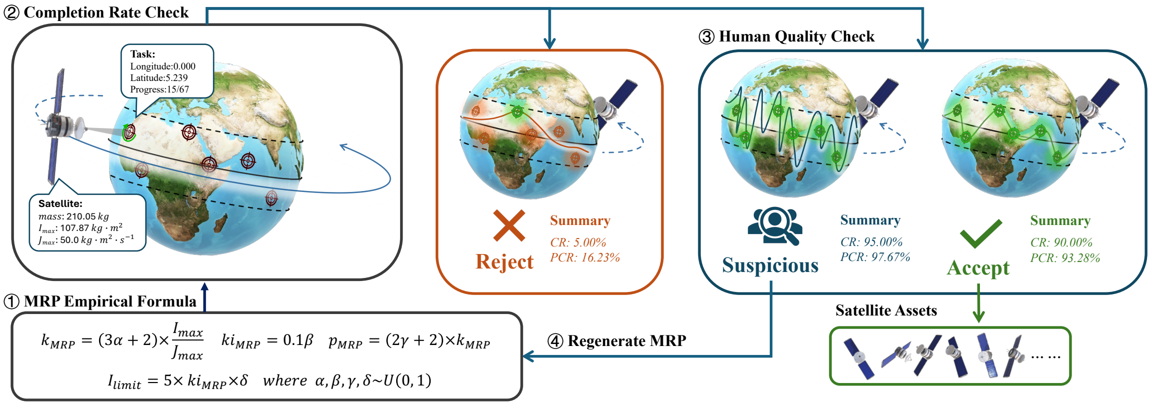
Motivation for a New Benchmark

AEOS-Bench Features

1. **Large-Scale:** 16,410 scenarios on various satellites
2. **Realism:** High-fidelity simulation based on Basilisk
3. **Comprehensiveness:** 6 evaluation metrics

Aspect	Previous Benchmarks	AEOS-Bench
Number of scenarios	< 10	**16,410**
Satellite assets	Limited	**3,907**
Physical simulation	Simplified	**High-fidelity**

Satellite Asset Generation



Satellite Asset Generation

Challenge: Attitude Control Stability

The MRP-based attitude control system requires careful parameter tuning:

- **Control gains:** govern adjustment speed
- **Actuator limits:** prevent destabilization
- Low gains or Overloaded actuators $\Rightarrow$ task failures

Asset Generation Process

1. Sample satellite parameters
2. Tune control parameters
3. Validate in simulation
4. Ensure reliable on-orbit performance

Annotation Pipeline

Three-Stage Pipeline

1 Distance-based Initialization

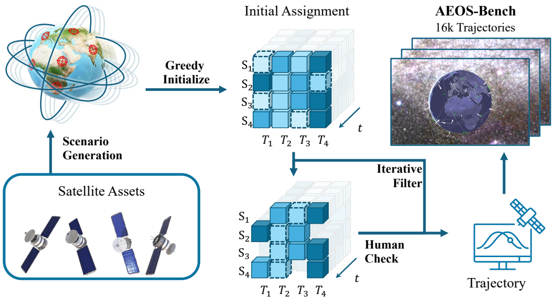
- Simple and intuitive baseline
- Problem: tasks too close ⇒ attitude control failures

2 Iterative Filter Stage

- Remove infeasible assignments
- Validate in simulation
- Refine scheduling strategies

3 Human Quality Review

- Final validation
- Ensure reliable annotations



Benchmark Statistics

Dataset Splits

Split	Scenarios	Satellites
Train	16,218	2,907
Val-Seen	64	Same as train
Val-Unseen	64	500 (new)
Test	64	500 (real data)

Scenario Composition

- Satellites per scenario: **1 to 50**
- Tasks per scenario: **50 to 300**
- Timesteps: **3,600**

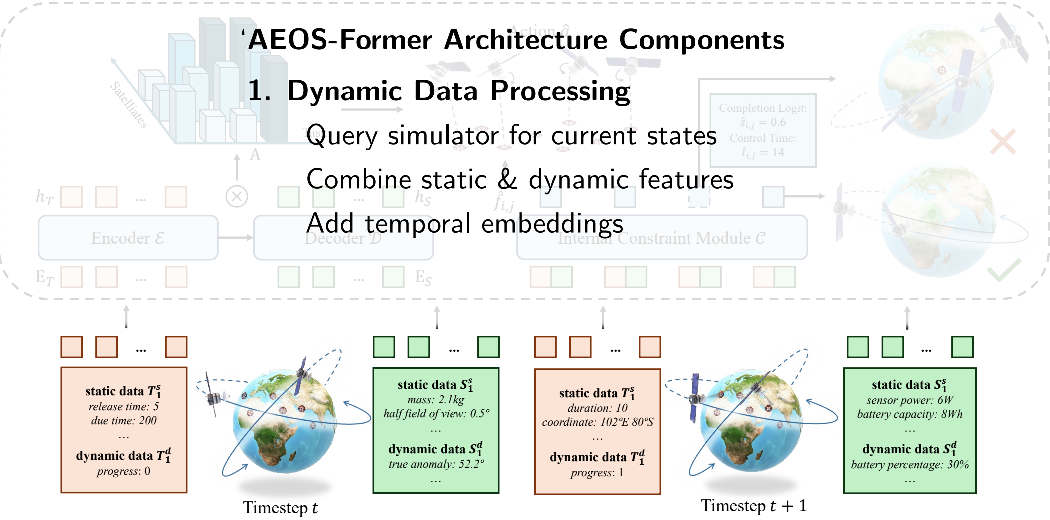
Satellite Asset Distribution

- Orbital parameters: approximately random distribution
- Smaller constellations: more frequent
- Larger task counts: more frequent
- Test split: real satellite data from N2YO and Gunter's Space Page

Table of Contents

1. Introduction
2. AEOS-Bench: Benchmark Suite
3. AEOS-Former: Proposed Method
4. Experiments
5. Conclusion

Architecture Overview



‘AEOS-Former Architecture Components

1. Dynamic Data Processing

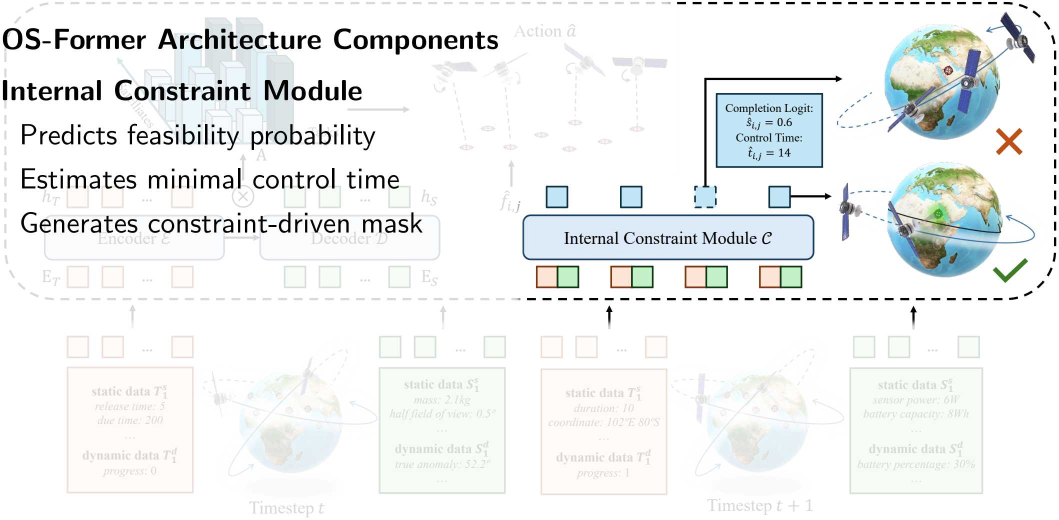
- Query simulator for current states
- Combine static & dynamic features
- Add temporal embeddings

Architecture Overview

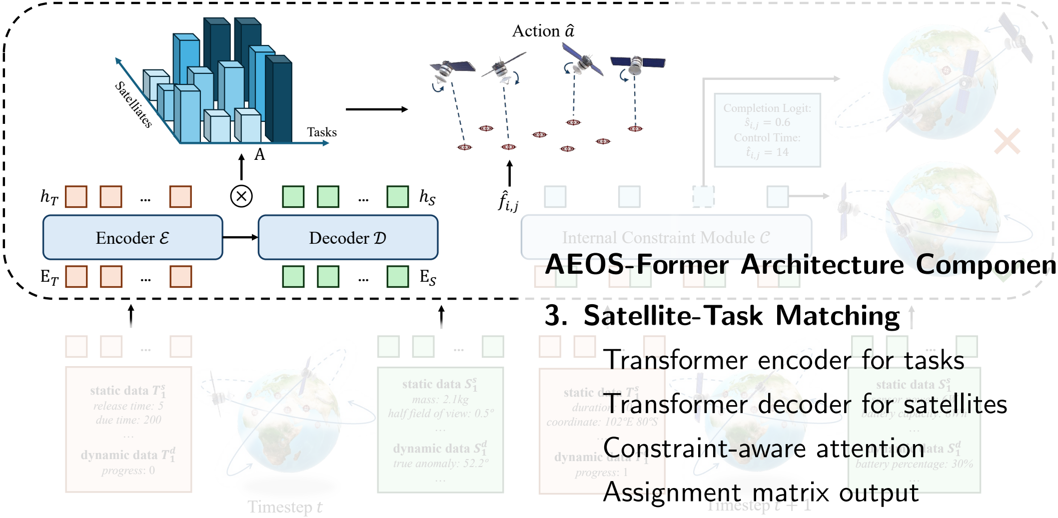
‘AEOS-Former Architecture Components

2. Internal Constraint Module

- Predicts feasibility probability
- Estimates minimal control time
- Generates constraint-driven mask



Architecture Overview

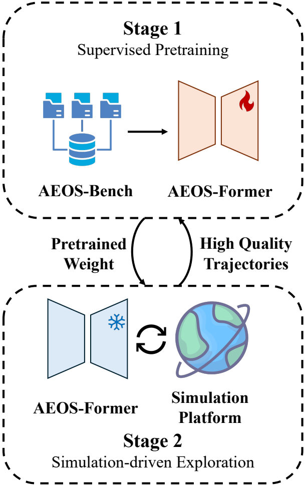


AEOS-Former Architecture Components

3. Satellite-Task Matching

- Transformer encoder for tasks
- Transformer decoder for satellites
- Constraint-aware attention
- Assignment matrix output

Architecture Overview



AEOS-Former Architecture Components

4. Iterative Learning

- Supervised pretraining
- Simulation-driven exploration
- High-quality trajectory collection
- Continuous refinement

Table of Contents

1 Introduction

2 AEOS-Bench: Benchmark Suite

3 AEOS-Former: Proposed Method

4 Experiments

5 Conclusion

Evaluation Metrics

Six Evaluation Metrics

Task Completion Metrics

1. **Completion Rate (CR)**
   Proportion of completed tasks
2. **Partial Completion Rate (PCR)**
   Ratio of maximum progress to required duration
3. **Weighted Completion Rate (WCR)**
   Weighted version considering task durations

Efficiency Metrics

4. **Turn-Around Time (TAT)**
   Average time to complete tasks
   Reflects scheduling efficiency
5. **Power Consumption (PC)**
   Total energy consumed by sensors
   Measured in Wh

Comprehensive Score (CS)

$$CS = (w_{CR} \cdot CR + w_{PCR} \cdot PCR + w_{WCR} \cdot WCR)^{-1} + w_{TAT} \cdot TAT + w_{PC} \cdot PC \tag{1}$$

where $w_{CR} = 0.6$, $w_{PCR} = 0.2$, $w_{WCR} = 0.2$, $w_{TAT} = 1/7$, $w_{PC} = 1/100$

Main Results

Split	Method	CS ↓	CR/%	PCR/%	WCR/%	TAT/h ↓	PC/Wh ↓
Val Unseen	Random	90.27	1.08	1.33	1.02	**0.17**	142.27
	HAAL	77.17	1.28	1.46	1.28	0.25	155.36
	REDA	21.54	4.83	5.75	4.85	0.71	153.95
	MSCPO-SHCS	5.21	35.35	**39.45**	34.85	7.27	140.83
	AEOS-Former (Ours)	**4.43**	**35.42**	38.93	**35.14**	6.78	**68.99**
Test	Random	113.53	0.85	1.02	0.88	**0.17**	150.54
	HAAL	94.83	1.05	1.17	1.03	0.25	155.56
	REDA	28.21	3.65	4.27	3.58	0.73	154.49
	MSCPO-SHCS	7.33	**19.44**	**24.00**	18.71	6.23	149.20
	AEOS-Former (Ours)	**6.28**	19.25	22.31	**18.73**	5.67	**40.91**

Performance Analysis

Performance Analysis: Scaling Behavior

Impact of Number of Satellites (N_S)

- As N_S increases ($1 \rightarrow 50$):
  - CS initially decreases
  - Stabilizes between 31–40
  - CR consistently increases
- Trade-off between completion and resource consumption

Impact of Number of Tasks (N_T)

- As N_T increases:
  - Lower completion rates
  - More resource consumption
  - CS slightly increases
  - CR slightly decreases
- More tasks $\Rightarrow$ harder scheduling

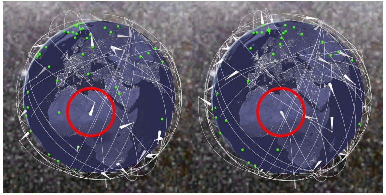


Table of Contents

1. Introduction
2. AEOS-Bench: Benchmark Suite
3. AEOS-Former: Proposed Method
4. Experiments
5. Conclusion

Summary and Future Work

Summary of Contributions

AEOS-Bench: Benchmark Suite

First large-scale benchmark for realistic AEOS constellation scheduling

3,907 satellite assets + 16,410 scenarios

High-fidelity simulation with realistic constraints

Comprehensive evaluation (6 metrics)

AEOS-Former: Scheduling Model

Novel **internal constraint module** for physical/operational limits

Transformer-based architecture with constraint-aware attention

Simulation-based iterative learning for continuous improvement