Spacecraft Data Wrangler

Austin Aden, India Hutson, Harshavardhini Manivannan, Olive Marangoni, Nicolas Morales, Ishaan Rao, Chinmaya Sahoo

PURDUE UNIVERSITY | The Data Mine

AEROSPACE

INTRODUCTION

The goal of this collaboration is to create a tool to automatically extract and normalize data from different spacecrafts for analysis, without requiring input from experts. Currently, there is no industry standard format for spacecraft telemetry data. The variation in formats between satellites makes it difficult to aggregate data and requires prohibitive effort from experts to fully utilize large data sets. The team experimented with telemetry data from a GPS receiver on FM-6. This is a low Earth orbit satellite, a part of the COSMIC-1 / FormoSat-3 constellation for gathering data on meteorology, the ionosphere, and climate. In preparation to create an automated data preprocessing tool, the team manually investigated temporal, relative, and spatial trends in GPS signal data reported by FM-6, NORAD ID: 29047.

Figure 1: GPSs in Contact

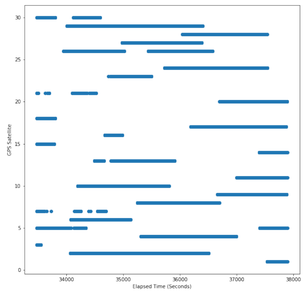


Figure 2: Number of GPS in Contact

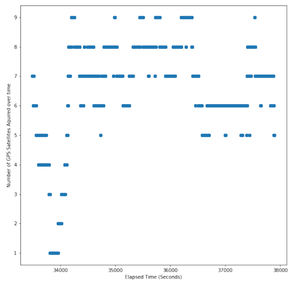


Figure 3: FormoSat-3

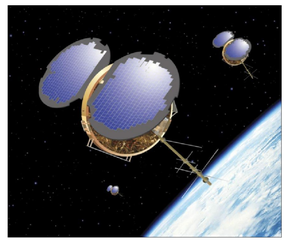
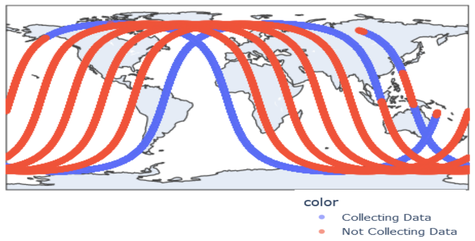


Figure 4: Data Collection Success over Time for Initial Data



METHODOLOGY

Objective 1: “Wrangle” extremely large, complex, unconditioned & unique spacecraft telemetry data sets.

- Parse given data set for short-duration Formosat-3 satellite operation and present sorted data using Python script.
- Plot time variation and relative variation of data parameters.
- Characterize behavior of recorded data and identify parameter types based on knowledge from basic radio frequency signal theory and orbital dynamics.

Objective 2: Detect and characterize interference impacting spacecraft in Low Earth Orbit.

- Expand data set to include signal information over two month, approximately 64 million observations, allowing for larger characterization of signal strength over Earth’s surface.
- Merge signal data set with publicly available latitude and longitude data acquired from GPS almanac YUMA archives.
- Map signal environment through binning of SA data in 2° “boxes” of latitude and longitude.
- Present global signal strength trends through a heatmap.

ANALYSIS AND RESULTS

The team preformed parametric analysis on GPS signal measurements. Without spatial information, it was difficult to interpret any meaningful information from the graphs. With spatial information, meaningful trends could be extracted from signal noise.

Figure 4 reveals gaps in data collection that could be indicative of maintenance activities or other equipment artifacts.

Figure 6 indicates that signal noise overwhelms data trends when displayed in aggregate, making feature extraction challenging.

Figure 8 bins the maximum signal strength over a 4° x 4° geographic area for two months. Although the data is very noisy, some patterns of higher signal strength regions are detected. An "L" shaped feature emerges as an out of family area. Regions such as these may hold value for further study.

Figure 5: 2° Binned Signal Strength of an Initial 24-hour Period

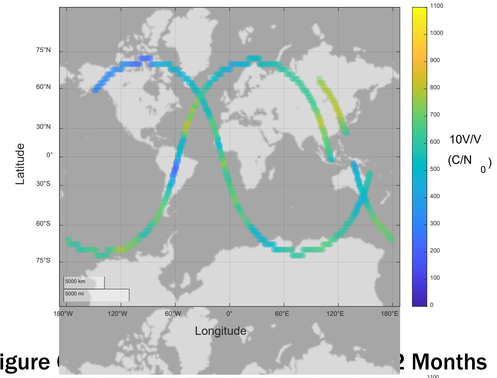


Figure 6: Signal Strength Below 20 Volts/Volt over 2 Months

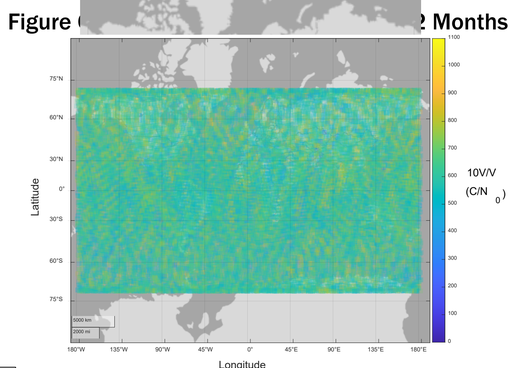


Figure 7: Signal Strength Below 20 Volts/Volt over 2 Months Period

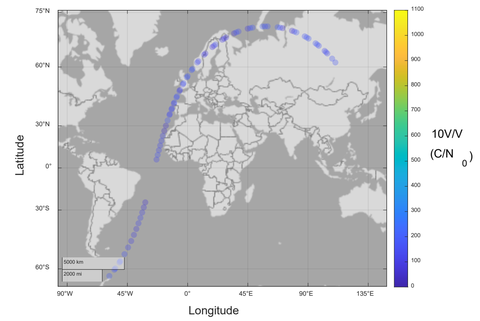
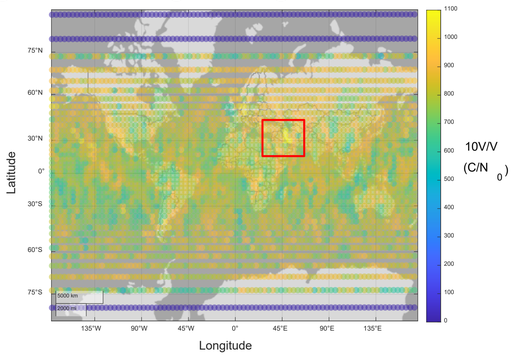


Figure 8: Binned Signal Strength over 2 Month Period



CONCLUSIONS

Over the course of this project the team has explored spacecraft telemetry data. By investigating and parsing an unfamiliar data set the team determined key correlations between spatiotemporal relations and telemetry data. Results were plotted across time and space to better visualize trends and outliers inherent in the data.

FUTURE WORK

FROM MANUAL TO AUTOMATED: Based on our initial exploration, our overarching goal for a longer project would be to automate the manual data preparation we performed by leveraging physical patterns in the data. This automation would minimize the expert time necessary for data preparation and optimize this complicated process. Eventually, the data wrangler will be expected to work on any given data set and interpret it in a normalized format.

Thank you, Dr. Keith LeGrand; Dr. Steven Lewis, Ms. Michelle Permann and Dr. George Pollock of The Aerospace Corporation; and Max Low and Maggie Betz of The Data Mine for your mentorship throughout this project.