JOHNS HOPKINS
APPLIED PHYSICS LABORATORY

Minimum Effort Telemetry Data Mining

An IRAD-funded effort to explore automated anomaly detection

Christopher O’Hare – Ground Software Engineer
Space Exploration Sector, Ground Applications Group
Johns Hopkins Applied Physics Laboratory

Agenda

- Overview
- Year 1 – Initial Investigation
- Year 2, 3 – Deep Dive into Characterization
- Proposed Mission work – plan overview

Problem Statement

Mission Telemetry is voluminous and growing. Much of this telemetry is provided “in case it’s needed” and never explicitly checked or trended. Advances in “Big data” data mining analytic techniques can help monitor and highlight interesting changes for subsystem analysis with minimal effort required to establish the checks.

Example: Van Allen Probes Mission*

- 50k TLM streams across 2 spacecraft
- Manual regular trending on ~300 pts
- Possible value in automated monitoring of other 99% of points

*Post-launch, the Radiation Belt Storm Probes (RBSP), RBSP-A and RBSP-B, were renamed the Van Allen Probes

Overview

Goal: **With little to no user input/supervision, identify novelty in telemetry using automated behavior characterization and analysis**

- No user input = no mission/system context, no tagged data, no feedback
- Novelty: “a localized or permanent change in the behavior of telemetry data, compared to a nominal characterization”
- Novelty includes:
  - Single-point outliers – ex: 3-sigma outlier values
  - Overall behavior changes – ex: clipping
  - Event-specific behavior changes – noise characteristic during a burn is 20% larger than what has previously been observed
- Identify novel events for further manual investigation, while minimizing the number of total novel events (false positives)

Effort:
- Year 1,2 – IRAD
- Year 3 – IRAD (Partial)
- Proposed – Mission Funding

Team:
- 3 Software Engineers
- Some Data Science/ML experience

Year 1 – Initial Investigation

Overview:

- Build up big data pipeline
- Focus on use of statistical techniques (min, max, avg) across constant time windows (3-hr, 1-day)
    - Input: nominal time range, test time range
    - All data treated as continuous numerical values
    - Calculation of an “outlier factor”
- Novelty determined at the spacecraft level (i.e.: across all telemetry)
- Broad search across early mission data
- Focus on known anomaly events, see if we can “find” them in Van Allen Probe spacecraft RBSP-B historical data
    - Get confirmation from SME

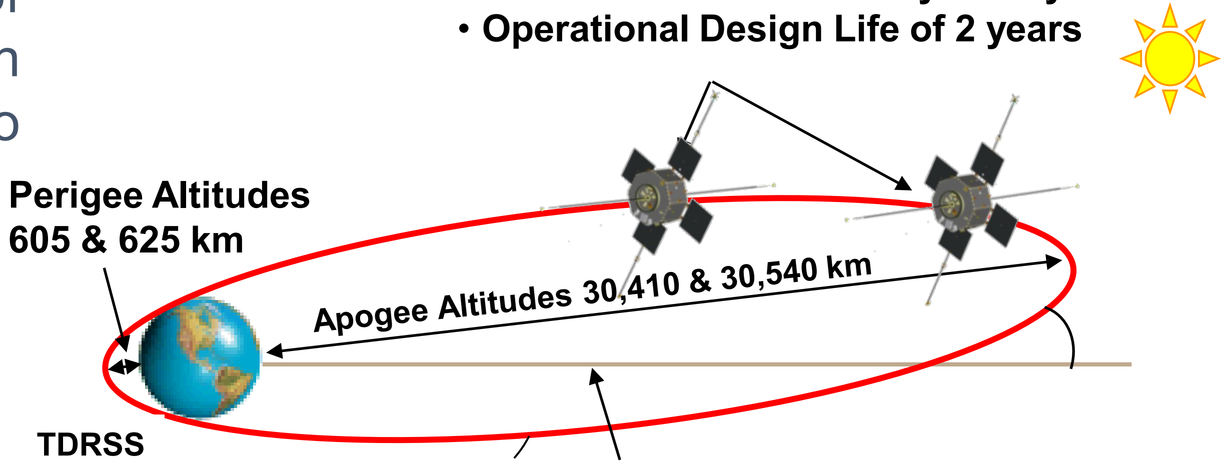
Van Allen Probes

- Objective:

  Provide understanding, ideally to the point of predictability, of how populations of relativistic electrons and penetrating ions in space form or change in response to variable inputs of energy from the Sun.

2 Observatories

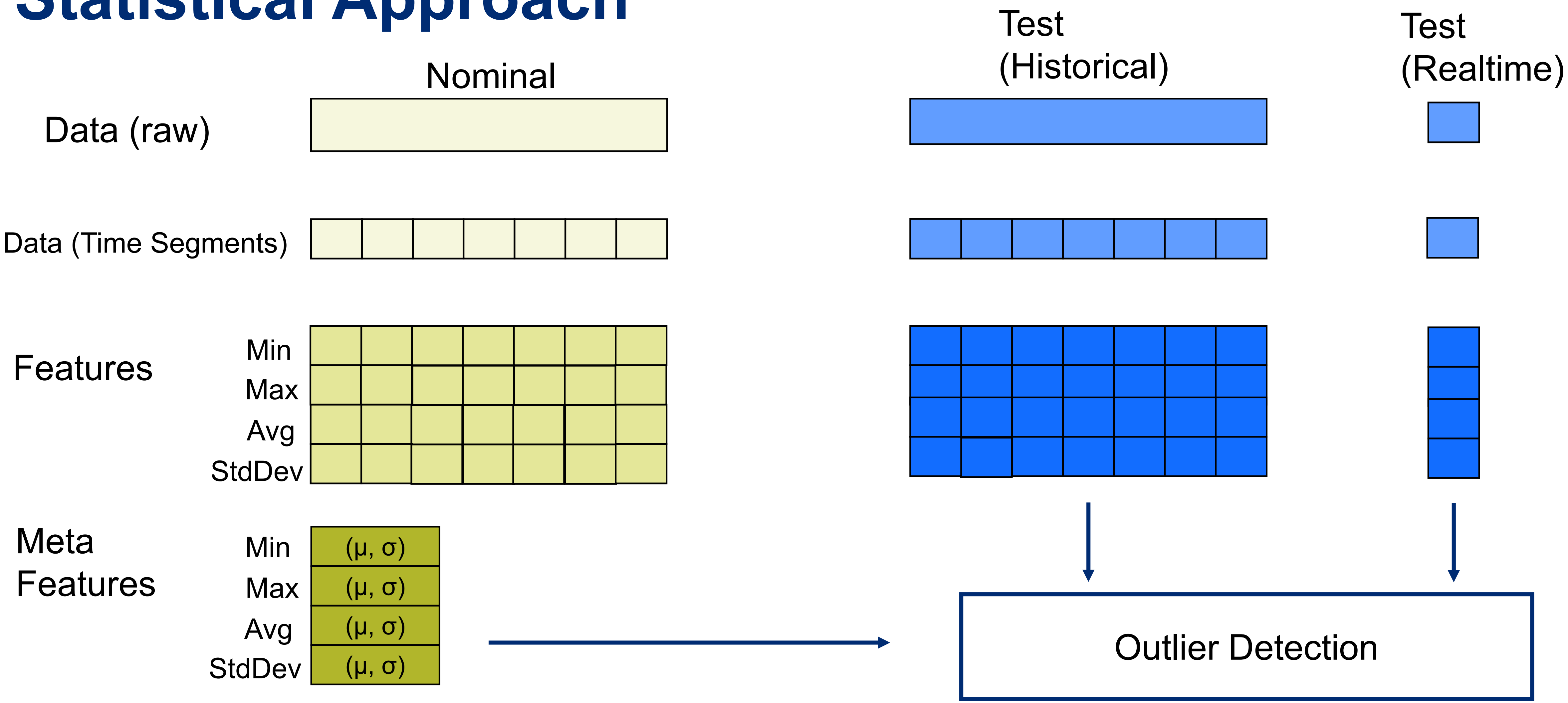
- **Spin Stabilized ~5 RPM**
- **Spin-Axis 15°-27° off Sun**
- **Attitude Maneuvers Every 21 days**
- **Operational Design Life of 2 years**



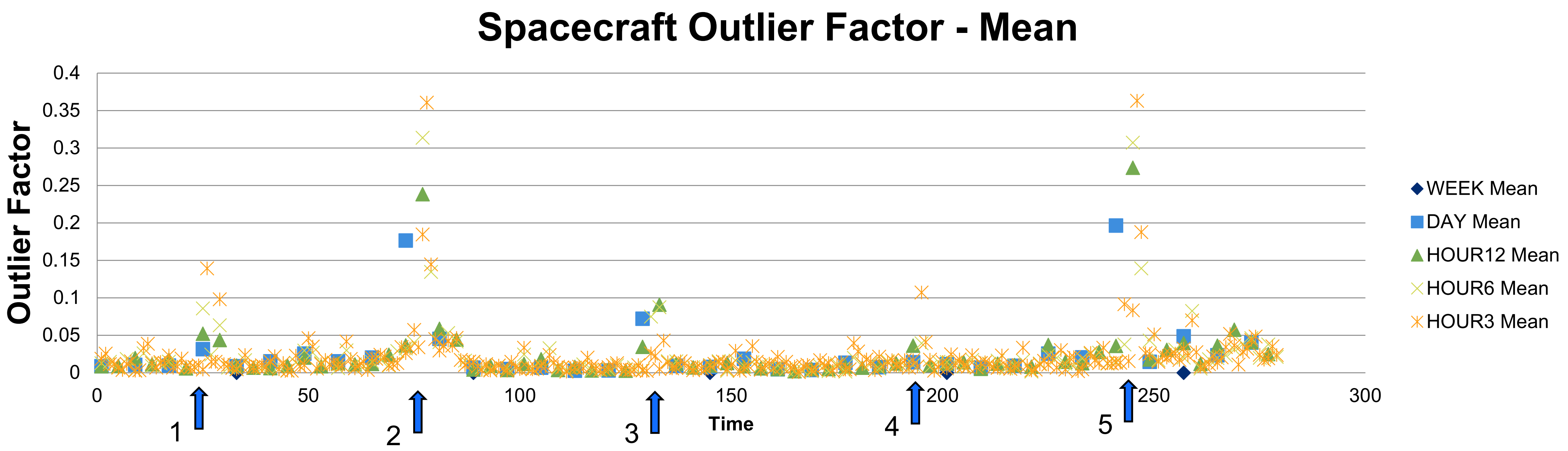
IRAD effort used a subset of historical data from Van Allen Probe spacecraft RBSP-B

- ~25,000 TLM points
- 622 mission days
- Ave:3.9 GB / day
- Total: 2.38 TB

Statistical Approach



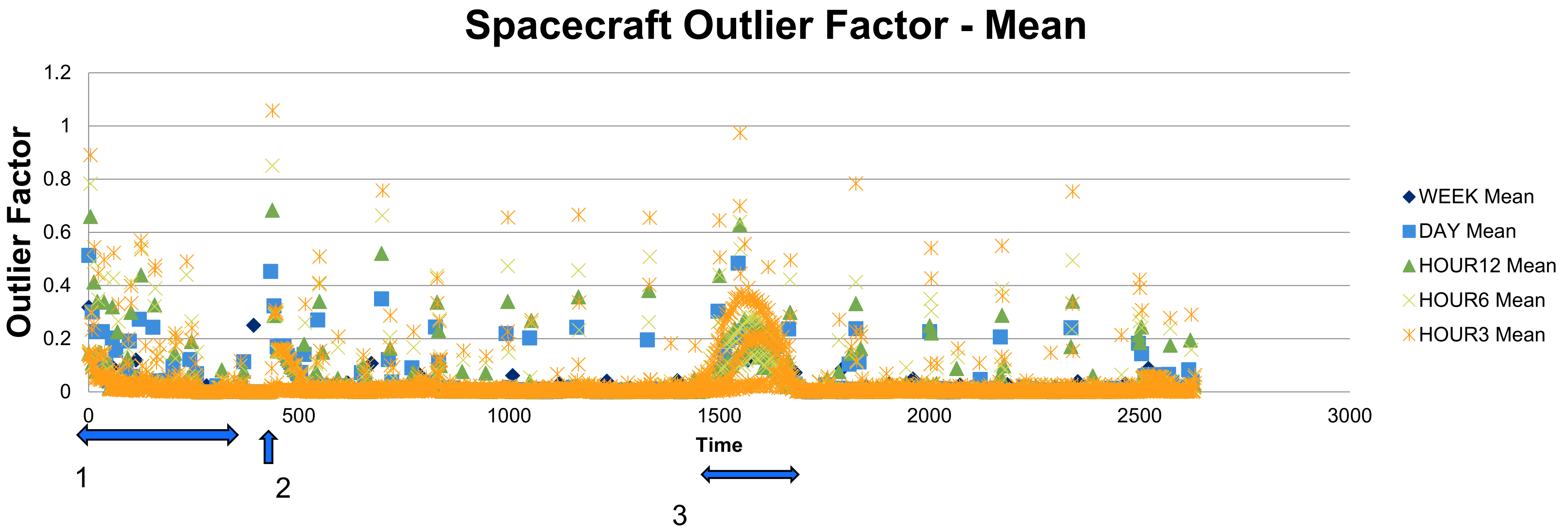
Sample Analysis 1: Jan 1, 2013 – Feb 4, 2013



(Note: Unless otherwise noted, dates are for the 3-hour features; all times represent window starts)

1. Isolated spike at Jan 3 2013, 22:00:00, secondary isolated spike at Jan 4 2013, 07:00:00
2. Extended spike starting at Jan 10 2013, 07:00:00, peak at Jan 10 2013, 10:00:00 extending through Jan 10 2013, 13:00:00 (9 hours)
3. Isolated spike at Jan 17 2013, 07:00:00 (12-hour feature)
4. Isolated spike at Jan 25 2013, 01:00:00
5. Extended spike starting at Jan 31 2013, 07:00:00, peak at Jan 31 2013, 10:00:00 extending through Jan 31 2013, 13:00:00 (9 hours)

Sample Analysis 2: Sept 1, 2012 – August 3, 2013 (48 weeks)



(Dates are general start-stop times of long-term trends, unless otherwise noted)

1. Launch and commissioning period (Aug 31, 2012 – Nov 1, 2012)
2. Isolated spike at Nov 1, 2012 11:00:00 (3-hour)
3. Extended period of outlier behavior (approx Mar 1, 2013 – Apr 4, 2013), showing a distinctive slope

Year 1 – Takeaways

- Success! - did find novelty that corresponded to identified events
- However, spacecraft-level novelty does not provide operational value
  - Most detected “events” were normal actions (ex: burns), or SC reactions to an anomaly (autonomy status changes)
- Determination: Current paradigm (fixed-window, statistical) does not provide sufficient granularity to characterize telemetry for target novelty

Year 2, 3 – Deep Dive into Characterization

Overview

- Refine novelty detection paradigm
- Deep dive into more detailed characterization of individual telemetry series
- Investigation into Machine Learning techniques

Challenges

Spacecraft data is subject to:

- Periodic and non-periodic environmental effects (known and unknown)
    - Orbit, solar events
- Periodic and non-periodic spacecraft events (known and unknown)
    - SC modes, commanded events (burns)
- Unknown correlation with other spacecraft subsystems (known and unknown)
    - Coupling

The problem is then the **generic time-series data problem**: can we characterize all the modes of a telemetry series well enough to understand when its behavior changes?

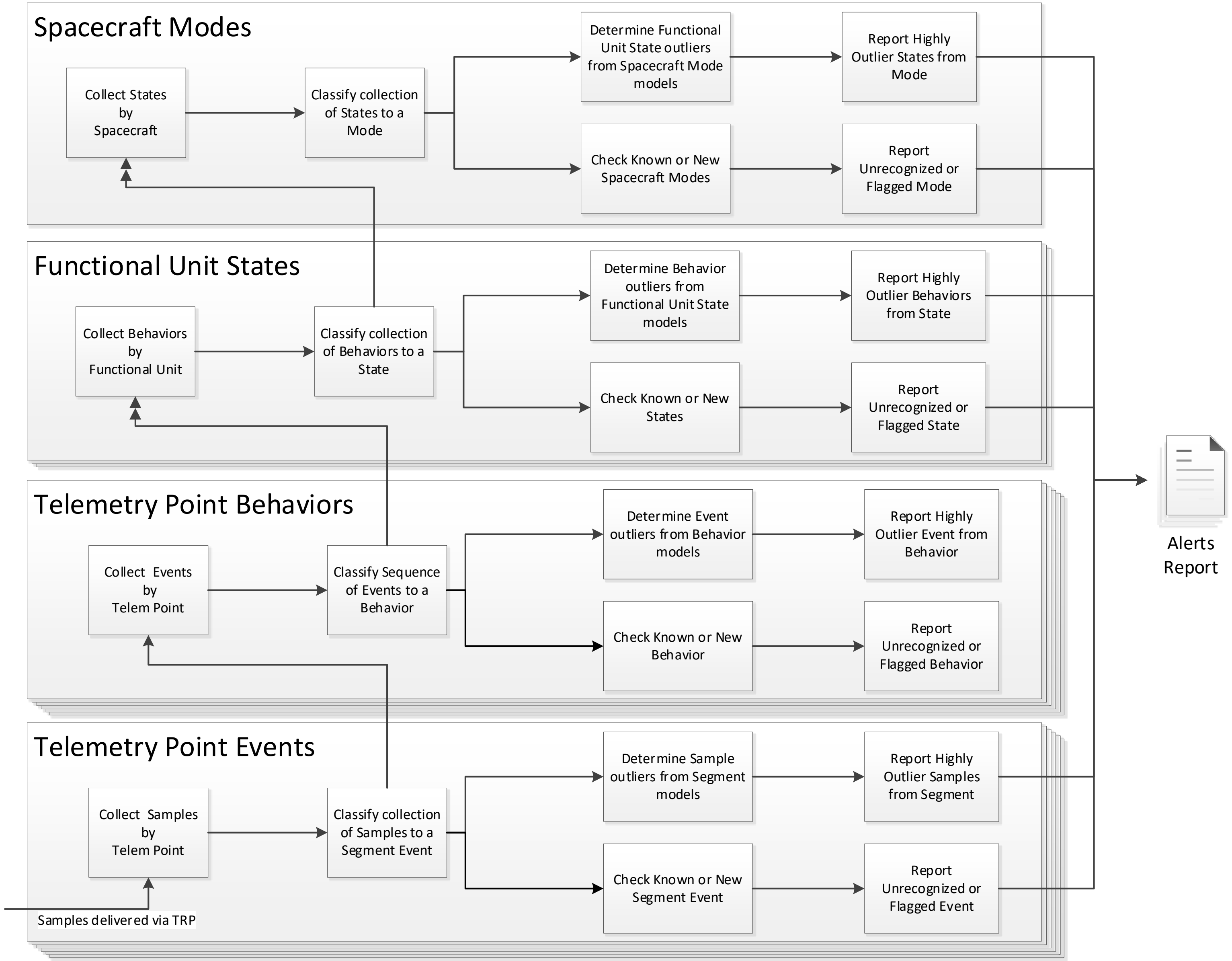
Spacecraft Behavior Classification Hierarchy

Subsystem outlier relative to other subsystems

TLM behavior outlier relative to other correlated TLM behavior

Behavior outlier relative to nominal behaviors

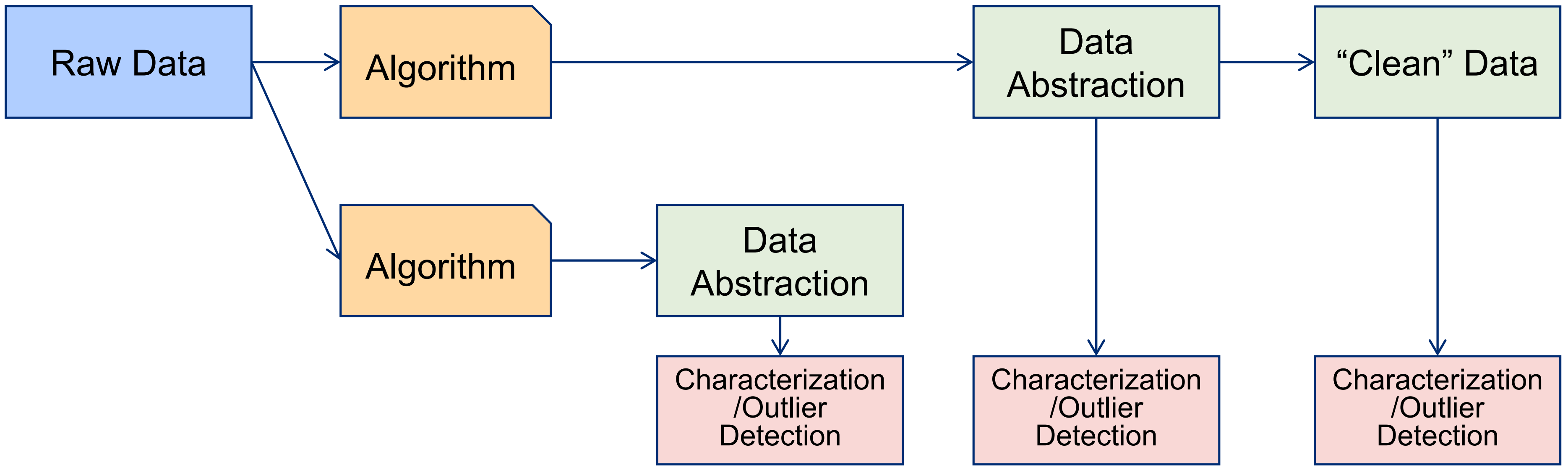
Single sample outlier relative to the local behavior



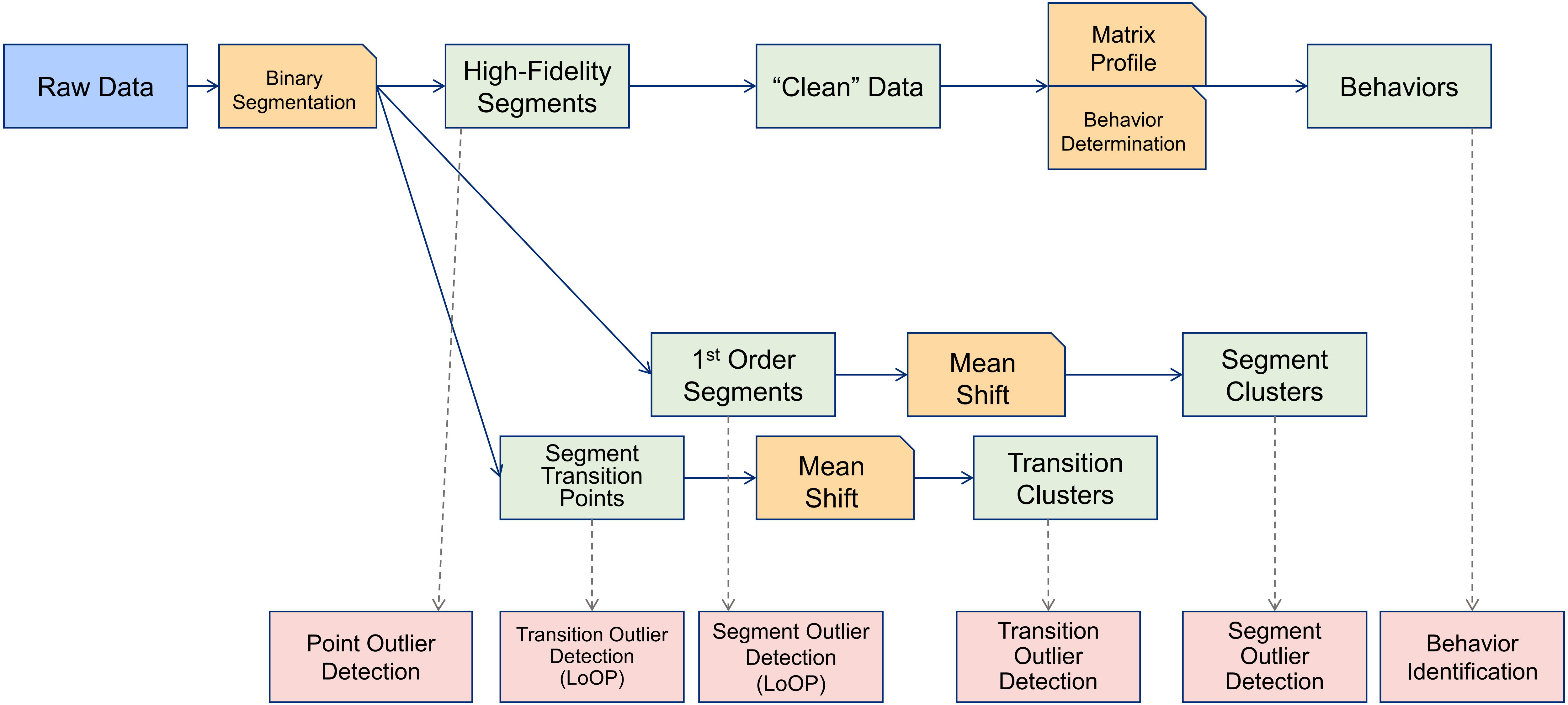
Data and Algorithms

- Generally, any algorithm will have certain requirements/restrictions on the input data
  - ex: Consistent frequency, consistent/no noise, relative scaling, volume
- Raw telemetry tends to not conform to these requirements/restrictions
- Higher levels of novelty detection generally require higher levels of abstraction

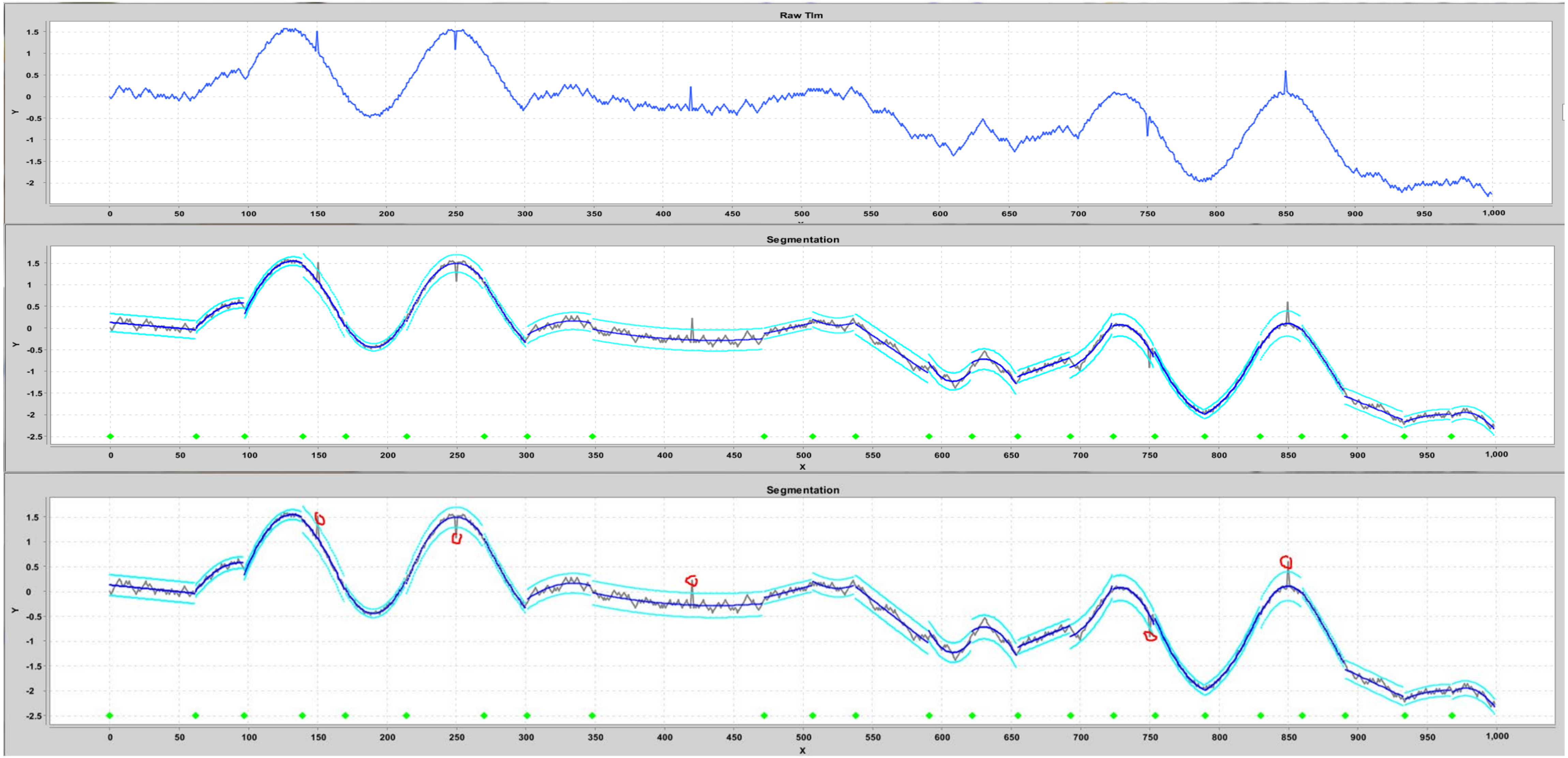
Thus, a significant portion of the effort was in data transformation to facilitate the use of more traditional ML algorithms.



Nominal Characterization – Single TLM Stream

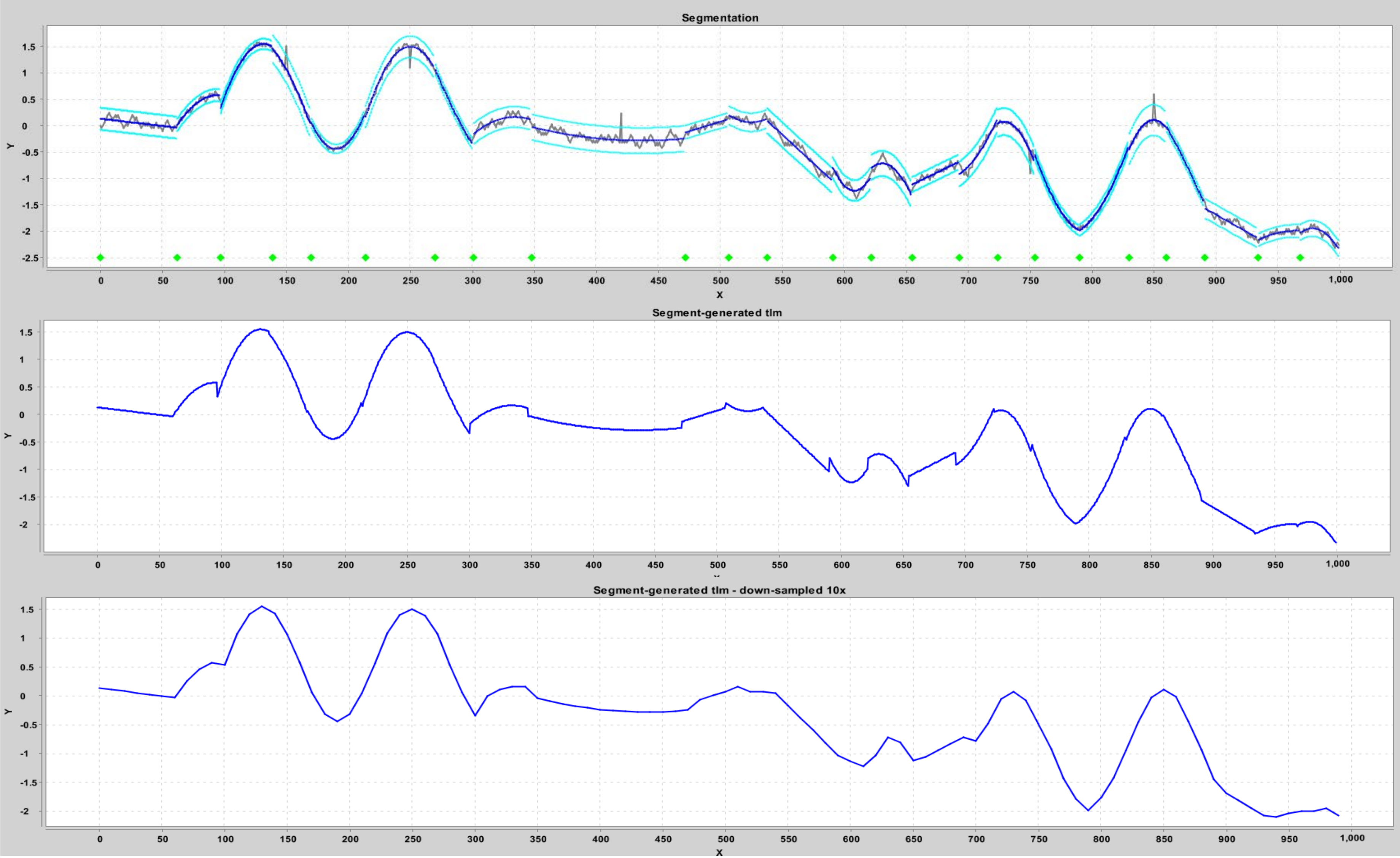


Change Point Detection, Segmentation, Point Outlier

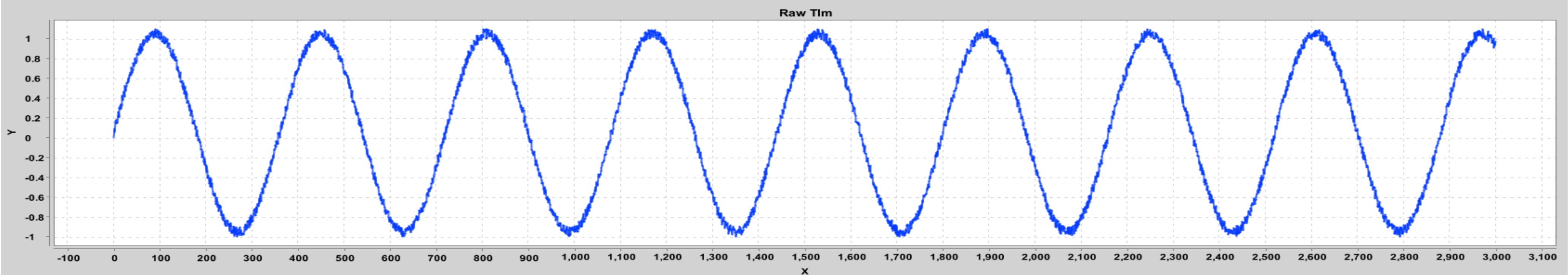


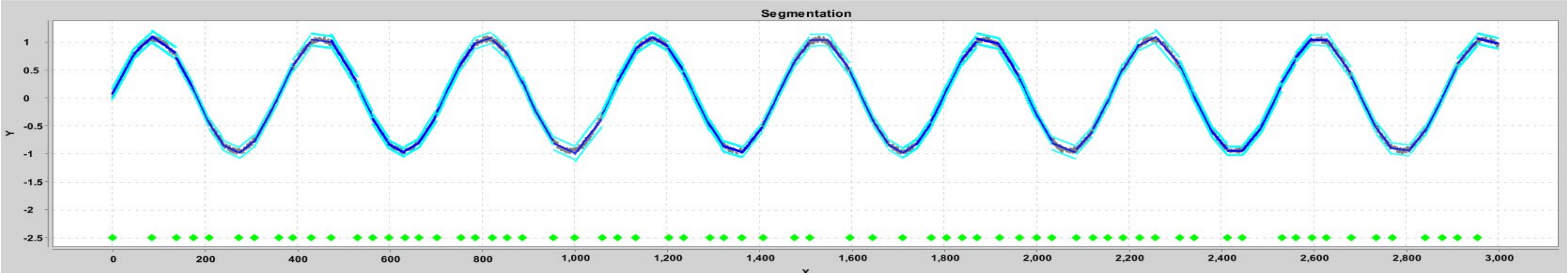
“Clean Data” Generation

- Generate a de-noised, constant frequency data set from segment representation
- Able to down-sample for performance



Segment-based clustering



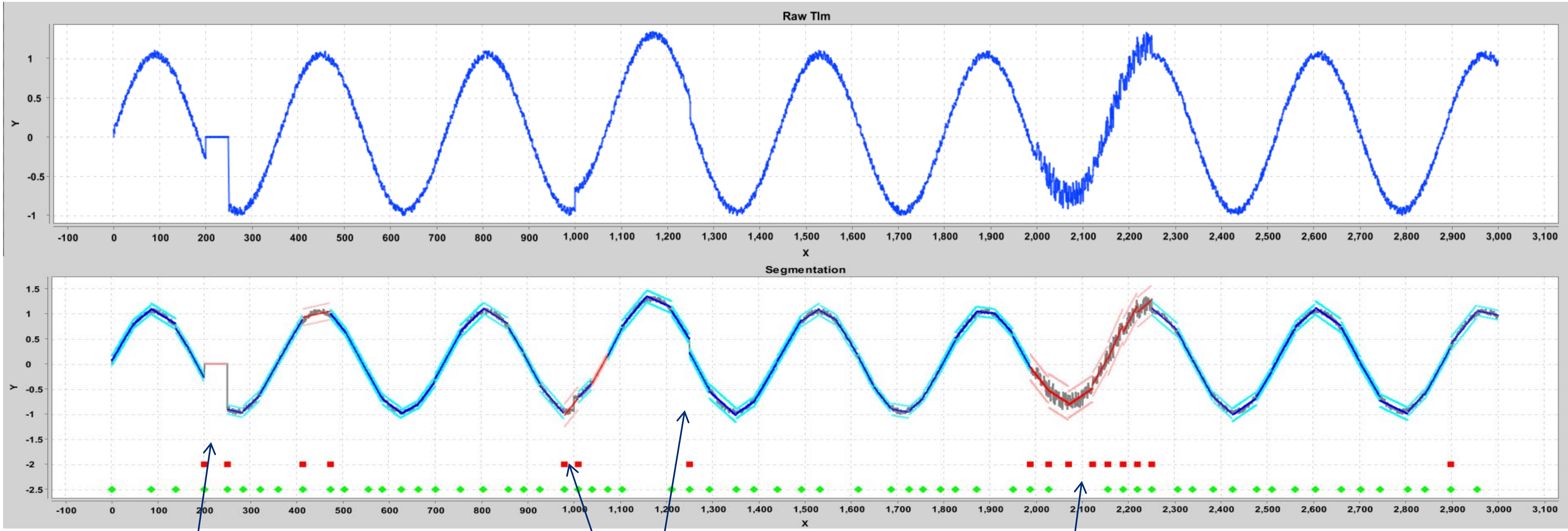


Convert raw data into segments with 1st order poly model: trend and noise
Represent segments and transitions as numerical arrays

Segment = [y-intercept, slope, stdDev, duration]
Transition = [transition value, slope1, slope2, stdDev1, stdDev2]

Segment-based novelty detection (LoOP)

Trained on the above slightly noisy sine wave



zero-signal

Abrupt jump in offset

High-noise section

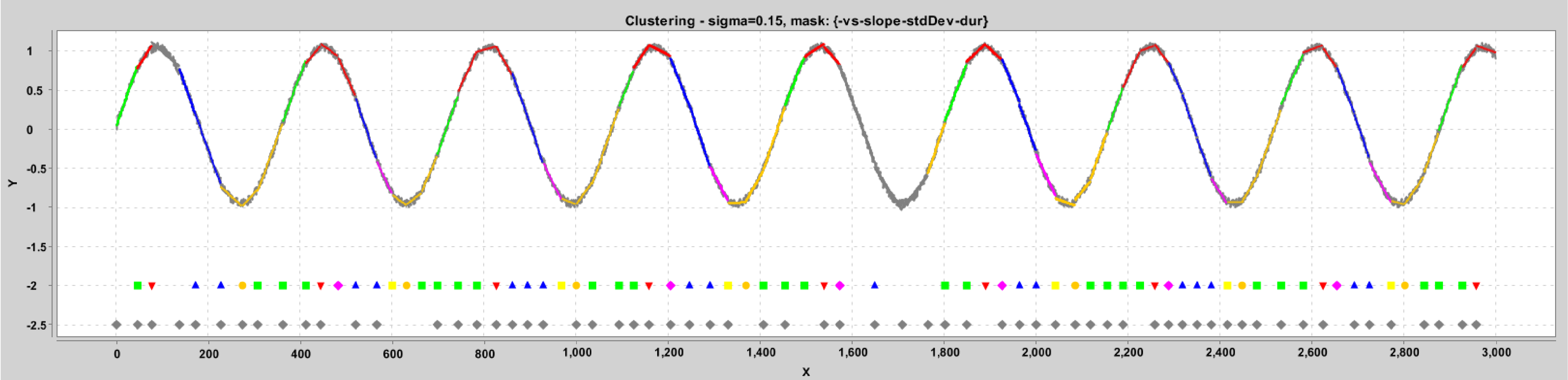
Segment-based clustering – Mean Shift

Goal: Cluster train segments and transitions:

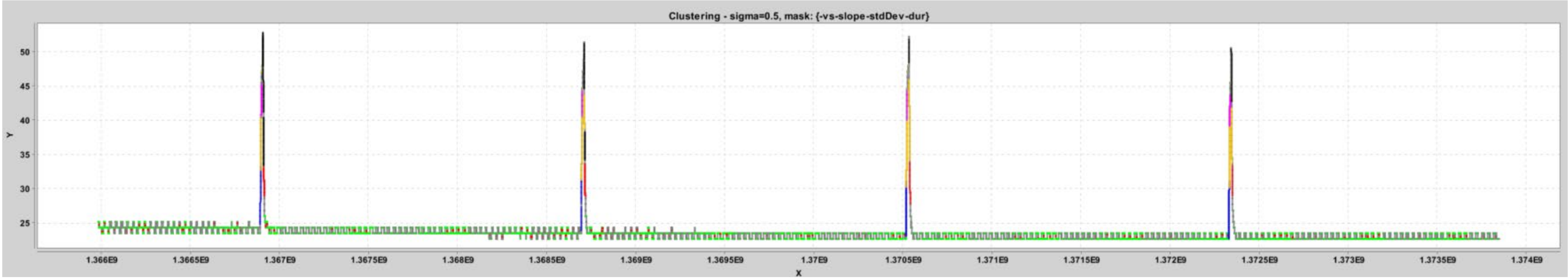
- To evaluate test outlier segments
- To provide layers of abstraction for further pattern characterization

Challenges

- Unknown number of clusters (k)



RBSPB data - TLM stream 101865
Segment clustering via Mean Shift

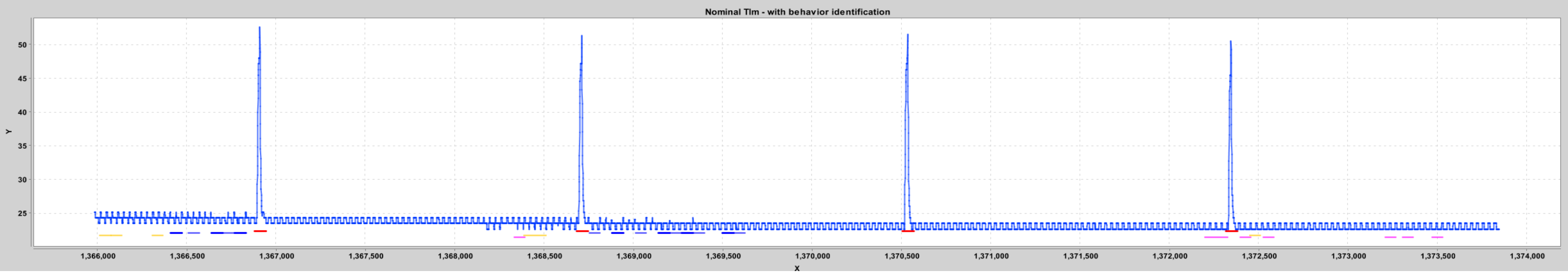


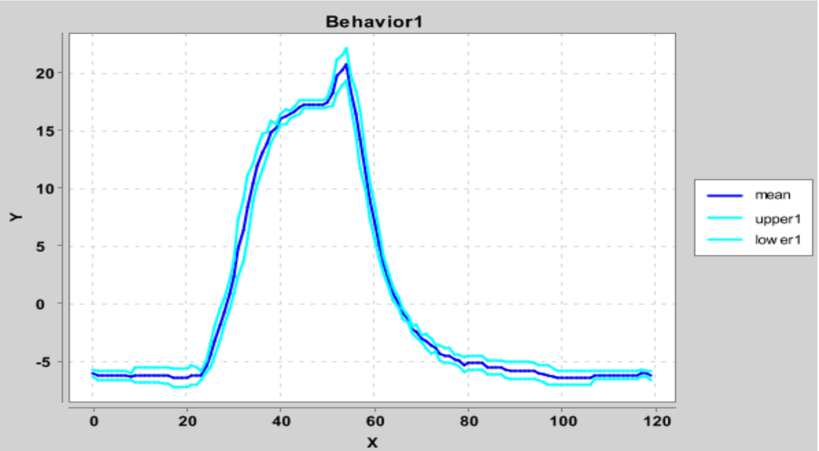
200k raw points -> 986 segments -> 8 clusters

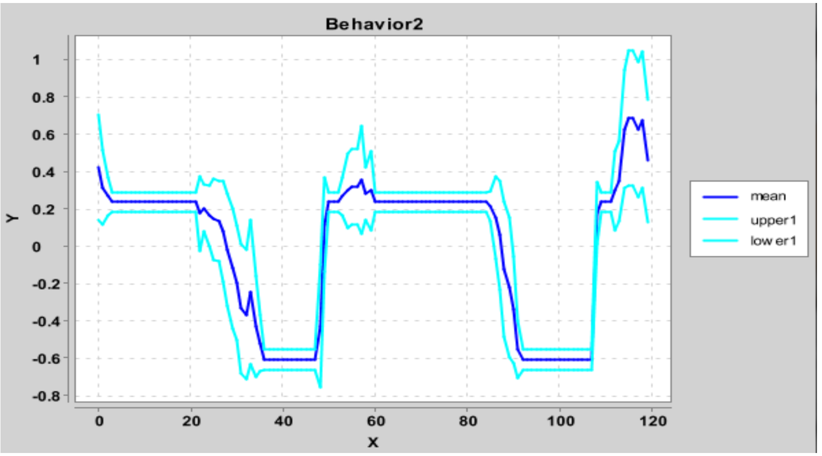
```
Cluster – Population - Color
0 - 611 - green
1 - 327 - red
2 - 5 - blue
3 - 4 - orange
4 - 4 - magenta
5 - 18 – yellow
…
```

RBSPB data - TLM stream 101865
Behavior Determination – Matrix Profile

Identify complex, repeated sequences to determine telemetry behaviors





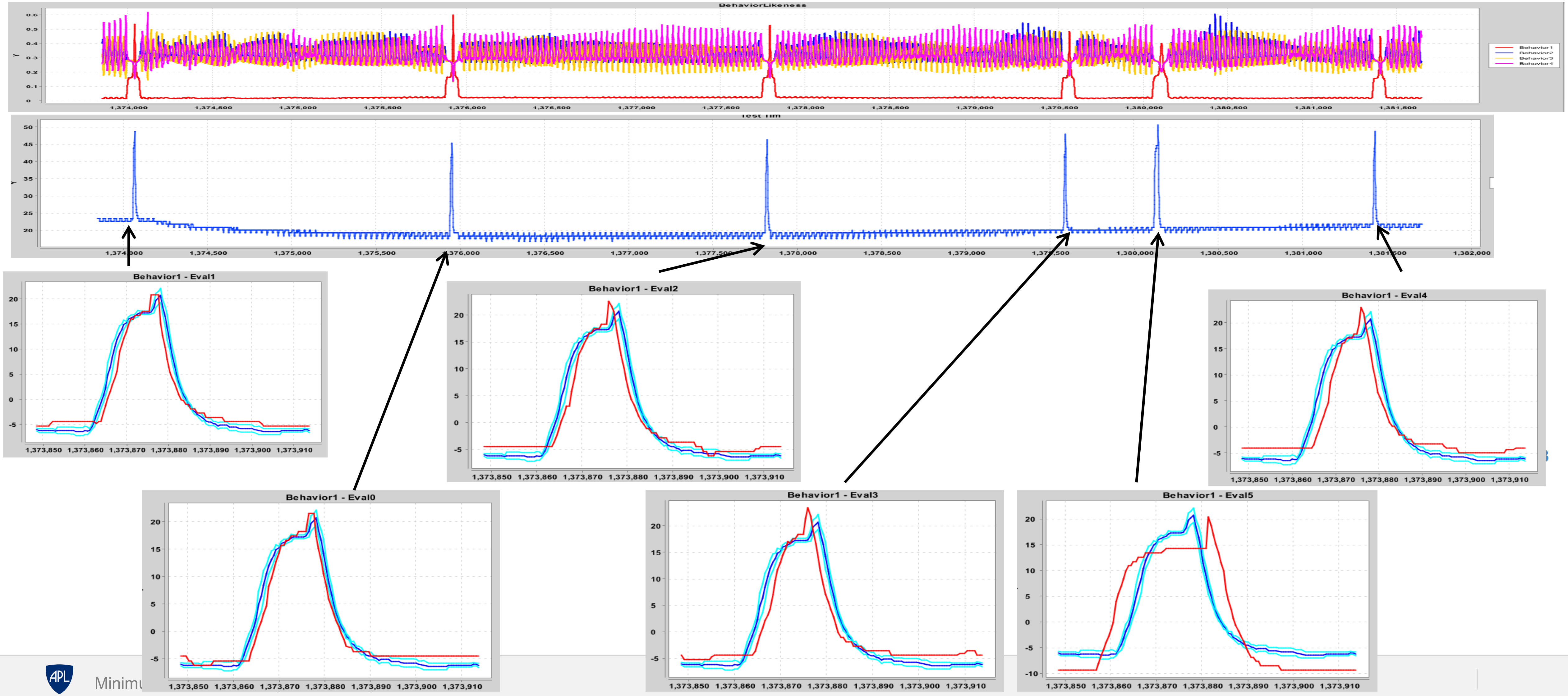


From red sequences

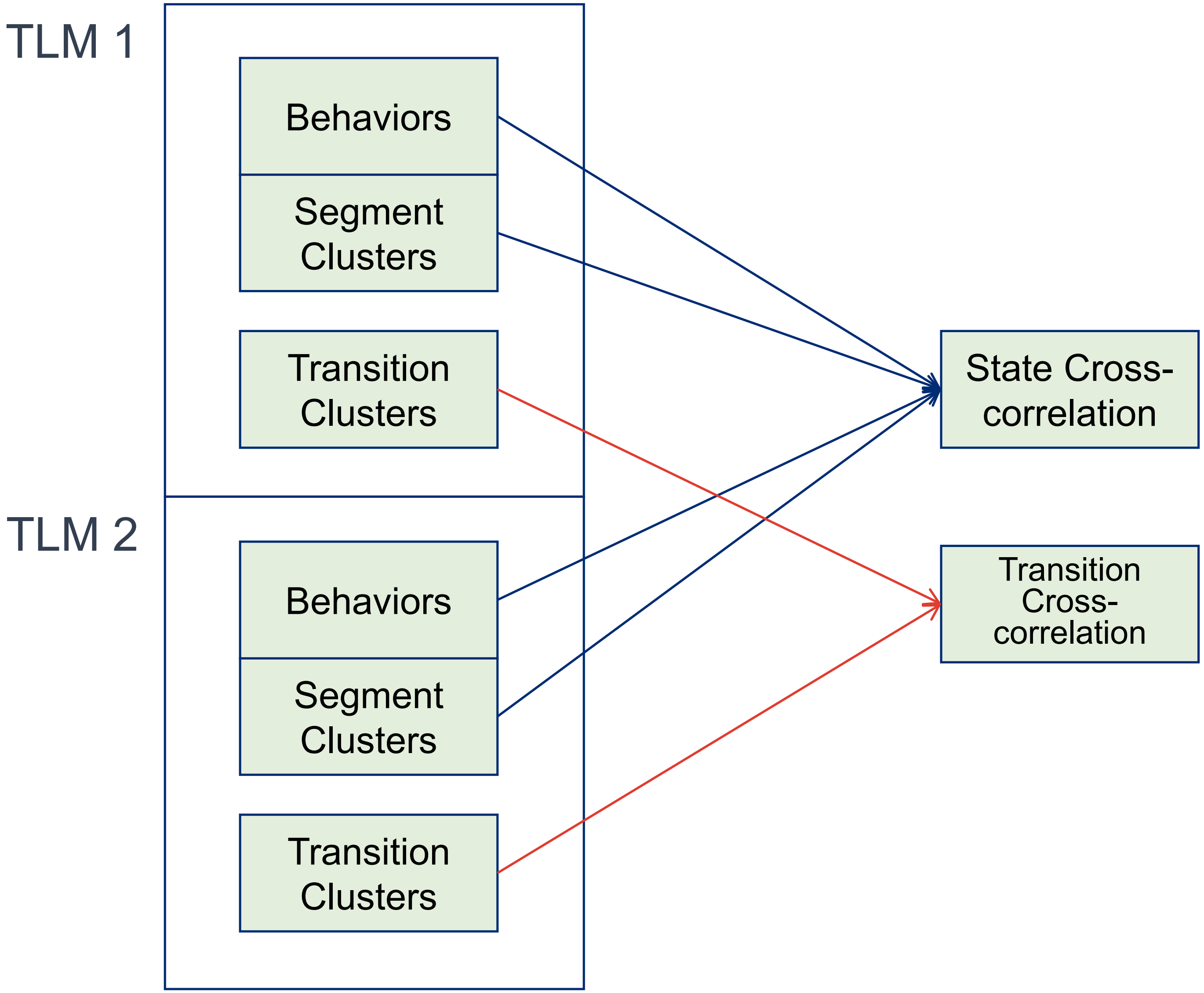
From blue sequences

Behavior Identification – Test Data

- Use identified behaviors to provide analysis context to test data
- Actual TLM in red, compared behavior in blue



Telemetry Cross-Correlation



Challenges:

- Downselection – what is a "meaningful correlation"
- Visualization of outliers in test data

Modifications to Algorithms

Algorithm	[Big O notation]	Modifications
Binary Segmentation	n^2	- Rolling Binary Segmentation - Heuristic Preprocessing - Model Demotion post-process - Noise-based segmentation post-process
Matrix Profile	n^2 – n^3	- Frequency scaling on clean data generation - “Input generic” matrix profile
Behavior Determination (using Matrix Profile)	n	- Match thresholding - z-shift vs z-norm options - “Null behavior” consideration

Lessons Learned – Part 1

- Algorithms can be very sensitive to hyper-parameters (tunables)
  - Our current paradigm has ~30 parameters
- A “universal parameter set” that works well for all possible telemetry streams is unlikely to exist
  - “Auto-tune” parameters themselves require parameters

- “No user input” makes accuracy and intended sensitivity difficult to achieve
  - Even small user insight can improve the characterization quality

- Data transformation is an important aspect of ML
  - Algorithms (even fancy ML ones) have some requirements/restrictions to the input data (content, quality, format, etc)

Lessons Learned – Part 2

- Becoming a “user” of an algorithm (i.e.: understanding limitations, tunables, req’s on input, interpretation of output) takes some bench time
  - The “art” is in understanding how to tune the algorithm, which comes with familiarity
  - You may not know how effective an algorithm will be with your data until you actually use it

- Visualization layer is a **<u>very important</u>** aspect of telemetry analysis
  - Ultimately a human needs to see and understand the results

- ML is not objectively better, statistical methods should be attempted, and are generally more transparent to understand/troubleshoot
  - Statistical methods better for determining “why” something is an outlier

Parker Solar Probe – Proposed Mission-Funded Effort

Goals

- Focus on specific subsystem - temperature sensor telemetry (~200 streams)
- Incorporate analyst feedback/context into characterization – nominal and realtime
- Develop visualization strategy
- Find "secondary value" in intermediate data transformations

References

Evans, D. J., Martinez, J., Korte-Stapff, Vandenbussche, B., Royer, P., & De Ridder, J. (2016). Data Mining to Drastically Improve Spacecraft Telemetry Checking: A Scientist's Approach. SpaceOps 2016 Conference.

Evans, D. J., Martinez, J., Korte-Stapff, M., Brighenti, A., Brighenti, C., & Biancat, J. (2016). Data Mining to Drastically Improve Spacecraft Telemetry Checking: An Engineer's Approach. SpaceOps 2016 Conference.

Keogh, Eamonn, and Jessica Lin. "Clustering of Time-Series Subsequences Is Meaningless: Implications for Previous and Future Research." Knowledge and Information Systems, vol. 8, no. 2, 2005, pp. 154–177., doi:10.1007/s10115-004-0172-7.

Killick, R., et al. "Optimal Detection of Changepoints With a Linear Computational Cost." Journal of the American Statistical Association, vol. 107, no. 500, 2012, pp. 1590–1598., doi:10.1080/01621459.2012.737745.

Kriegel, Hans-Peter, et al. "LoOP." Proceeding of the 18th ACM Conference on Information and Knowledge Management - CIKM 09, 2009, doi:10.1145/1645953.1646195.

Yeh, Chin-Chia Michael, et al. "Matrix Profile I: All Pairs Similarity Joins for Time Series: A Unifying View That Includes Motifs, Discords and Shapelets." 2016 IEEE 16th International Conference on Data Mining (ICDM), 2016, doi:10.1109/icdm.2016.0179.

Contributors

IRAD Team

- Jonathan Labin
- Shimon Gewirtz

Subject Matter Experts

- Kristin Fretz
- Daniel Ossing

JOHNS HOPKINS

APPLIED PHYSICS LABORATORY