Journal of Space Operations & Communicator (ISSN 2410-0005) Vol. 18 No. 1, Year 2022

A Machine Learning Solution for Satellite Health and Safety Monitoring

Zhenping Li

ASRC Federal, Beltsville, Maryland, 20705, USA

The machine learning (ML) solution consists of two ML processes running consecutively: the data training process to detect data pattern changes (DPC); and the clustering process to analyze DPCs generated from the data training process to detect anomalies. The data training process adopts a flexible approach by implementing different ML data models and data training algorithms for datasets with different data pattern types, since telemetry data have diverse types in data patterns with high complexity that makes data training in an operational environment more challenging. A ML framework for anomaly detections based on DPCs is presented, which consists of a ML representation to characterize both normal events and anomalies based on patterns of correlations among datasets with DPCs, and a hierarchical clustering algorithm to separate normal events and anomalies, since both normal events, such as an orbit maneuver, and anomalies cause DPCs in telemetry datasets. Patterns of correlations among telemetry datasets can be presented in an event plot in the ML framework, which provide engineers new insights into signatures of normal events and characteristics of anomalies. The results of the data training and the event clustering for Suomi National Polar-orbiting Partnership (NPP) satellite telemetry are presented and show that most events in NPP telemetry data are normal ones, while only a few events could be classified as potential anomalies. The ML solution provides effective and efficient event classifications and anomaly detections in satellite telemetry data.

Nomenclature

$d_j(t_i)$ = the value of raw telemetry dataset *j* at a time t_i, which is generally noisy

$S_j(t_i)$ = the value of telemetry dataset *j* at a time t_i.

$f_j(t)$ = the time dependent function for telemetry dataset *j*

σ_j = the standard deviation of telemetry dataset *j* obtained from a training session

$D_j(t_i)$ = the values of telemetry dataset *j* used in data training, which is derived from $d_j(t_i)$

W = a set of parameters in the time dependent function obtained from data training.

$F_j(t_i - t_0, \{S_k\}, W)$ = the time dependent function for telemetry dataset *j* in data training, and the function $f_j(t)$ is derived from $F_j(t_i - t_0, \{S_k\}, W)$.

$m_k(t_i)$ = the angular velocity of the reaction wheel *k* at the time t_i, which has the unit of RPM.

$M_a(t_i)$ = the reaction wheel momentum in x, y, z direction, which has the unit of NMS.

$O_i^N(d(t_i))$ = the value of outliers for telemetry dataset *j* at a time t_i, which is dimensionless.

ψ_j = the DPC metric for telemetry dataset *j*, which is dimensionless.

$\boldsymbol{E}$ = the event vector for characterizing the correlations among datasets.

$\boldsymbol{e}(\psi_S)$ = the unit event vector.

$\alpha_S^{i,j}$ = the scalar product of two event vectors

I. Introduction

The ML solution for monitoring satellite health and safety is a part of the application portfolio of the recently developed ML framework for space missions[1,2,3], which has been very successful in the onboard instrument radiometric monitoring and analysis[4]. The ML approach involves two ML processes running consecutively. The first process is a data training process to obtain data models, standard deviations, and outliers corresponding to DPCs in datasets. The second process is a clustering process to analyze the outputs from the data training process to evaluate data quality and detect potential anomalies. The focus of this paper is to provide a detailed presentation on how the ML framework is extended to satellite health and safety telemetry.

The extension of the ML framework for monitoring satellite health and safety needs to address the following challenges:

- Unlike instrument calibration data that are generally homogeneous, satellite health and safety telemetry data have more diverse data pattern types and higher complexities so that no single ML model can meet the efficiency and accuracy requirements in an operational environment for all telemetry datasets. The flexibility is needed to have different ML models for different data pattern types. Telemetry datasets in an operational environment generally contain outliers that could distort the data training outcomes. Data training algorithms are required to be robust not only to generate good training outcomes but also to accurately measure DPCs in telemetry data.
- Anomalies cause unexpected DPCs in telemetry datasets. However, the sources of DPCs in telemetry data could also come from the events that are part of normal operations, such as an orbit-maneuver. Changes in space environments may also cause disturbances in a satellite that lead to DPCs in telemetry datasets, which occur regularly so that they could be regarded as normal events since their effects are known and can be mitigated. The separation between normal events and anomalies is necessary for a ML solution. Anomaly detections without a separation of normal events from anomalies lead to false positives, which has been a major challenge for a ML application in satellite health and safety telemetry data.
- Interactions among subsystems in a satellite lead to correlations among telemetry datasets, and correlations among telemetry datasets generally lead to DPCs in multiple datasets during a normal event or an anomaly

so that signatures of normal events or anomalies cannot be determined by DPCs in a single dataset. Anomalies in satellite health and safety telemetry are determined by patterns of correlations among multiple datasets with DPCs.

There have been extensive studies in applying machine learning or data analytics techniques to satellite telemetry data for anomaly detections, which include the auto regression approach[5,6,7] and neural networks[8] to model telemetry datasets and to provide tighter limits of satellite health and safety telemetry and detects limit violations. The ML approach in anomaly detections includes clustering[9,10,11], kernel principle component analysis (PCA)[12], and dimensionality reduction and clustering techniques[13]. These studies generally follow the approach of learning the normal pattern and detecting deviations from normal patterns. Anomaly detections based on patterns of correlations among datasets with DPCs have not been studied in the literature. Because DPCs could be generated by both normal events, the separation of normal events from anomalies based on patterns of correlations is especially critical for large satellites with multiple payloads and subsystems: normal operational events, such as orbital maneuvers, happen frequently in satellite operations.

The contributions of the ML solution in this paper are the new ML framework for the data training of telemetry datasets with different pattern types and higher complexities and anomaly detections in satellite telemetry data based on patterns of correlations among datasets with DPCs. Different models and training algorithms for different data pattern types in telemetry data are implemented to meet the accuracy, efficiency, and robustness in an operational environment. The ML framework for anomaly detections in satellite telemetry data consists of the hierarchical event vectors as a ML representation to characterize patterns of correlations among datasets and the hierarchical clustering algorithm to separate normal events from anomalies. Both normal events and anomalies can be regarded as events in satellite operations that lead to DPCs in datasets and characterized by hierarchical event vectors that consist of metrics measuring DPCs occurred in the same period. The unit event vector is the ML representation to characterize patterns of correlations for datasets with DPCs, which provides signatures of normal events and characteristics of an anomalies. The separation of anomalies from normal events is achieved by the hierarchical clustering algorithm based on patterns of correlations. Normal events generally appear regularly in satellite telemetry so that events in an event type form its own cluster, while anomalous events occur rarely and don’t form clusters since each anomaly generally has its own characteristics. Thus, anomaly detections in this approach is to find events that don't belong to any cluster, and this can be achieved by the density based clustering algorithms[14], such as DBSCAN clustering. Because event vectors are hierarchical, the clustering is performed at subsystem level first to look for the correlation patterns among different subsystems, then clustering is further performed at mnemonic level for each cluster at subsystem level to look for correlation patterns among different datasets. The hierarchical clustering for datasets with DPCs reduces the algorithm complexity as the dimensions of event vectors for datasets with deviations from normal data patterns are significantly smaller. The high dimensionality in unsupervised learning approach to anomaly detections has been a significant challenge[13,15] in the literature.

Unit event vectors can be displayed in graphic form as event plots, which provide satellite engineers new insights into patterns of correlations among telemetry datasets and capture signatures of normal events and anomalies. This is a novel and unique feature from the ML approach that cannot be obtained from the traditional engineering analysis processes.

This paper is organized as following. Section II provides a brief discussion on the ML solution for monitoring satellite health and safety. Section III presents the data training results for typical data pattern types in satellite telemetry data. Different data types in satellite telemetry require different ML models, and the selection of specific model for a dataset is based on its noise level and pattern complexity. The data training process not only generates data models that can predict near future behavers in telemetry datasets but also detects DPCs in datasets based on data models that is crucial to establish patterns of correlations among different datasets. Section IV presents the quantitative metric to measure DPCs in telemetry datasets, hierarchical event vectors to characterize correlations among datasets with DPCs, the event plots based on the event vectors, and the hierarchical clustering on event-vector orientations for event classifications and anomaly detections. The results of event classifications and anomaly detections with NPP health and safety telemetry are presented in Section V. Section VI provides the summary and outlook.

II. The ML Solution for Monitoring Satellite Health and Safety

Fig. 1 shows the ML processes for satellite health and safety telemetry data in an operational environment. There are two ML processes and a data monitoring process in an operational environment. The two ML processes involve both supervised and unsupervised learning. The data training process learns normal data patterns in datasets and detects DPCs, which is a supervised learning. The data training process takes telemetry data from a telemetry data archive in a ground system. A satellite is a dynamical system characterized by its state variables A satellite is a dynamical system characterized by its state variables $\{S_j(t_i)\}$, which are generally time dependent. Datasets $\{d_j(t_i)\}$, such as satellite health and safety telemetry data, are measurements of state variables $\{S_j(t_i)\}$, which are generally noisy and follow the Gaussian probability distribution. State variables $\{S_j(t_i)\}$ are characterized by data models, which are obtained through the data training. A data model for a dataset with the Gaussian probability distribution consists of a time dependent function to predict near future values and a standard deviation representing its noise level. Data models for telemetry datasets generated from the data training process form tight data bounds determined by their noise level, which are highly sensitive to deviations from their expected behavior above the calculated noise level. This enables the dynamic monitoring that compares values of a dataset with predictions of its data model. Data points outside the data bounds determined by their data models are regarded as outliers. Persistent outliers in a dataset lead to DPCs, which can be quantitatively characterized. The data training process in this approach not only generates data models for telemetry datasets and but also detects the outliers corresponding to DPCs in telemetry datasets.

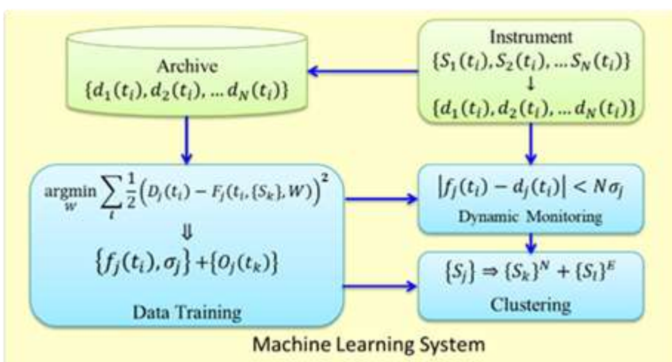

Fig. 1 The ML processes for satellite health and safety telemetry

Both data training and clustering processes in this approach are performed in sessions. Input training sets in a training session generally cover a period long enough to contain enough information for predictions of near-term behaviors. Because telemetry data used in data training are operational data and generally contain outliers that may distort the training outcomes, data training algorithms in an operational environment is required to be robust that can detect outliers to limit their impact on training outcomes. Training sessions are repeated periodically, which enable data models to adapt to long-term or seasonal pattern changes. Since daily changes in data patterns are small, the retraining of datasets is a minor adjustment of training outputs from previous training sessions. The data training in this approach uses the outputs from previous training sessions as inputs for the current training sessions. This is particularly important for the data training of nonlinear data models, such as neural networks: data retraining on existing data models makes the training algorithm much more efficient and enables data training in operational environments (near real-time processing).

The clustering process is performed after the data training to obtain actionable information from a large amount of data training outputs including data models and outliers, and this process is an unsupervised learning. The hierarchical event vectors are created from the outliers detected in the data training process, which consist of metrics measuring DPCs in multiple datasets at the same period. Both normal events and anomalies generally lead to DPCs in multiple datasets because of correlations among these datasets and are represented by event vectors. The nature of an event is determined by its composition and relative strengths of DPCs in datasets, which corresponds to the orientation of its event vector. The clustering is performed based on orientations of event vectors: the events with the same event orientation regardless of their magnitudes form a cluster. The event vectors are accumulated in multiple training sessions to form a large event set for clustering.

Dynamic monitoring process performs real time or near real time monitoring of incoming telemetry stream, which compares the incoming data values with the predictions of ML data models obtained from data training to detect deviations from normal data patterns that used as the input for clustering process for anomaly detections.

III. Data Training for LEO Satellite Telemetry

The data training in an ML system obtains data models for datasets $\{d_j(t)\}$ with the Gaussian probability distribution, which consist of a time dependent function $f_j(t - t_0)$ and a standard deviation σ_j,

$$\sigma_j = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(d_j(t_i) - f_j(t_i - t_0)\right)^2} \tag{1}$$

for its noise level. The quantity t_0 in Eq. 1 is the reference time for a given training session. The determination of the reference time t_0 depends on the orbital characteristics of a satellite. Data models $\{f_j(t-t_0), \sigma_j\}$ are obtained in an ML framework by training the time dependent function $f_j(t-t_0)$ with datasets $d_j(t)$. Data training for datasets with the Gaussian probability distribution follows the least square fitting routine with respect to a parameter set W:

$$argmin_W \sum_i \frac{1}{2}\left(D_j(t_i) - F_j(t_i - t_0, \{S_k\}, W)\right)^2 \quad (2)$$

where $D_j(t)$ and $F_j(t-t_0, \{S_k\}, W)$ are a training set for a dataset $d_j(t)$ and a data model for a time dependent function $f_j(t-t_0)$ respectively. A transformation between $D_j(t)/F_j(t-t_0, \{S_k\}, \boldsymbol{W})$ and $d_j(t)/f_j(t-t_0)$ is generally needed for data models within a ML framework and datasets with arbitrary scale. The data training finds a parameter set W so that the error function, Eq. 2, is minimized. This type of problems is generally referred as a regression problem in ML. The data patterns for satellite telemetry data are diverse and complex. There are many data models used in ML for time dependent trends, and each data model generally has its own data training algorithm. Different ML models are implemented for different pattern types in telemetry data to meet the accuracy, efficiency, and robustness requirements.

Because telemetry datasets generally contain outliers corresponding to DPCs, an iterative training approach is implemented, in which each data point in a training set is assigned a data quality flag (DFQ). The data quality flag is initialized to be 1 for each data point so that it is included in the initial round of data training. After the initial data training is performed, each data point in a training set is evaluated to determine if it is within the data bound predicted by its data model. Data points outside of their data bounds are outliers and their DQFs are assigned to be 0 so that they will be not included in the next round of data training. The iterative data training approach reduces the distortions of outliers on the data training outcomes.

Fig. 2 shows an example of data training outputs for the telemetry data in the power subsystem of the NPP satellite, which have periodic pattern. There are many datasets in the Power and Thermal subsystems with periodic patterns that are generally determined by satellite orbital characteristics. The pattern period for this pattern type follows the satellite orbital period, which is about 101 minutes for the NPP satellite. The neural network with two hidden layers is implemented for some of the datasets in the Power and Thermal subsystems. The Fourier expansion model[4] is also implemented for some datasets in the Power and Thermal subsystems, which generates highly accurate and efficient data training outputs.

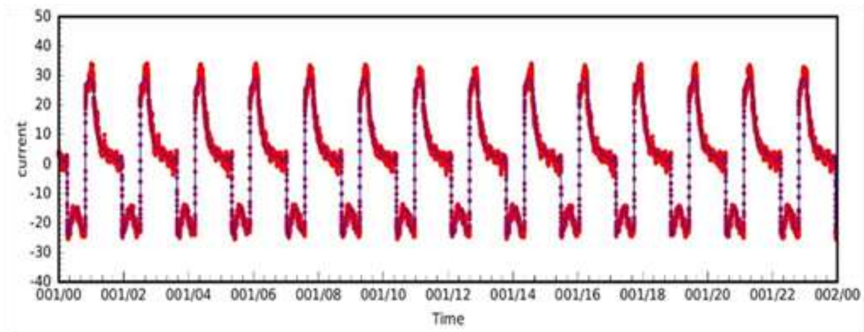

Fig 2 Battery current of the NPP power system. The red dots are the data values and blue lines are data training outputs. The neural network with two hidden layers is implemented for this dataset.

Fig. 3 shows the data training output of the second component of the quaternions, which is used to determine the attitude of a satellite. Accurate data training outputs for the quaternions is critical in creating event vectors for orbital maneuver events, as the DPCs in quaternions are highly sensitive to the data training accuracy. The data pattern for quaternions also follows the NPP satellite orbit characteristics, however, they are generally not continuous. The neural network with two hidden layers is implemented for quaternions. The data training outputs for other quaternion components are as accurate as the example shown in Fig 3.

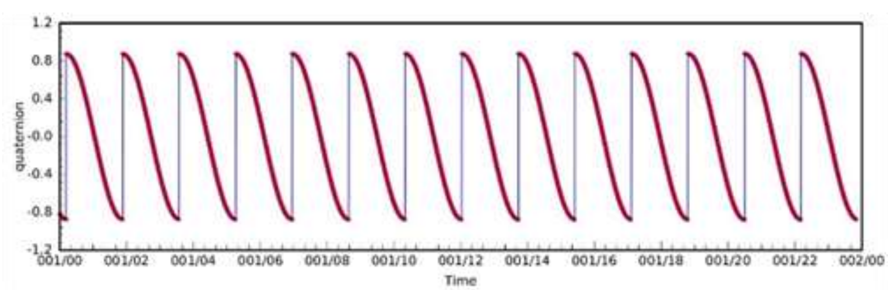

Fig. 3 The data training output for the second components of the quaternions, which has 4 components. The red dots are the data values, and the blue lines are data training outputs. The neural network with two hidden layers is implemented for the quaternions.

Fig. 4 shows an example of the data training outputs for the momentum profiles in the Reaction Wheel subsystem. The pattern complexity for the momentum profiles in the Reaction Wheel subsystem is much higher than the patterns shown in Figs. 3 and 4, and the pattern period for Reaction Wheel momentum profiles is about 14 times the NPP orbital period. The NPP satellite has a 4-Reaction Wheel configuration, and the data pattern complexities are different for the momentum profiles in each reaction wheel. For the datasets like the momentum profiles with very high complex patterns, the Fourier expansion model[4] is used since it is highly efficient in data training so that it can be performed in an operational environment.

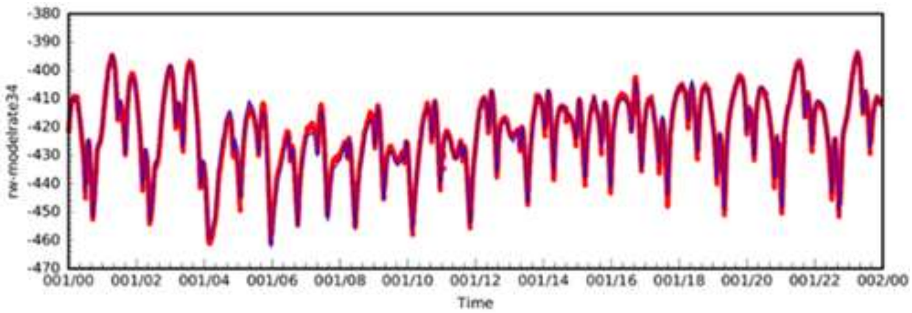

Fig. 4 The momentum profile of the NPP reaction wheel 3. The red dots are data values, and the blue lines are data training outputs. The NPP satellite has a 4-reaction wheel configuration.

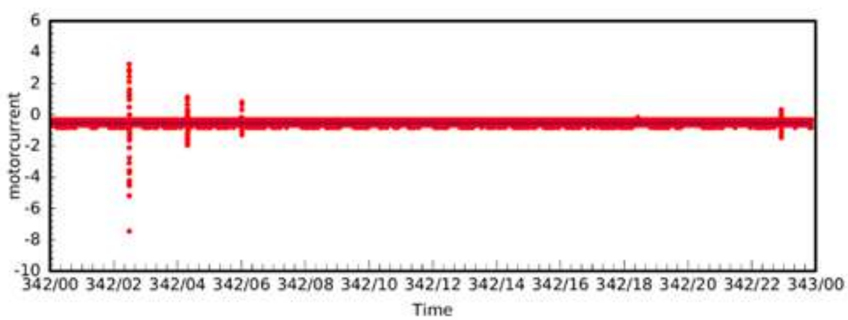

Fig. 5 The motor current for the Reaction Wheel subsystem. The red dots are the actual data values, and the blue line is the mean value of the dataset in data training periods. The data points along vertical lines are treated as outliers representing DPCs.

In addition to the Fourier expansion and neural network data models, more customized data models for satellite health and safety data are needed. Fig. 5 shows an example of the datasets that are constant and noisy, and many datasets in satellite telemetry belong to this category. The data model for this type of data is simply a statistical collection of input training sets, in which the time dependent function is the mean value of the input training set and the standard deviation in Eq. 2 is the same as the statistical standard deviation of the input training set. An important feature in Fig. 5 is the appearance of data points along vertical lines that cannot be fitted into a statistical model. These vertical lines are treated as outliers, and the presence of outliers in telemetry datasets, such as the ones in Fig. 5, does not necessarily correspond to an anomaly, and in fact, they are signals of normal events in most cases.

The ML algorithms for satellite telemetry also need to model relationships among datasets. Fig. 6 shows an example of the relationship modeling for the wheel momentum in x/y/z directions, which is a linear combination of the 4-reaction-wheel momentum. The relationship between the total wheel momentum and 4 reaction wheel momentum is determined by the reaction wheel geometry, which satisfies the general linear relationship

$$M_a(t_i) = c_a + \sum_k c_{a,k} m_k(t_i) \tag{3}$$

where the momentum M_a in $a = x, y, z$ is a linear function of momentum m_k in reaction wheels with $k = 1,2,3,4$. The parameter set $\{\boldsymbol{c_{a,k}}\}$ in Eq. 3 is obtained through data training. The data model in Eq. 3 does not have explicit time dependence. The data training algorithm for the data model in Eq. 3 is a linear multi-variate least square fitting algorithm, and it is highly efficient since only a subset of data sample in an input training set is needed for data training. Fig. 6 shows an example of the data training output of the wheel momentum in the y direction, which is a linear combination of the 4 reaction wheel momentums shown in Fig. 4.

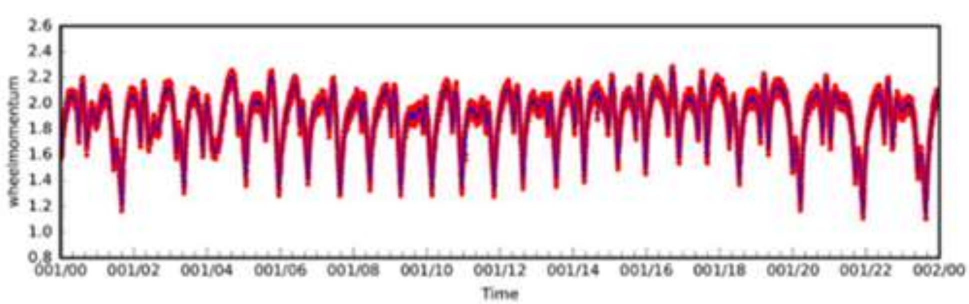

Fig. 6 The wheel momentum in y direction, which is a linear combination of 4 relation wheel momentum. The red dots are data values, and blue lines are data training outputs.

Fig. 7 shows an example of the data pattern type in telemetry data, which is a counter data type. The mnemonic **full search count** in the Star Tracker subsystem is an 8-bit counter, which increases monotonically until it saturates at 255 when it resets at 0 after the value of 255. The normal data pattern for the counter types is that the counter is increasing in its value at a nearly constant rate, such as the pattern during the period from 010/00:00 to 011/00 in Fig. 7. The higher rate of increases or the derivative during the periods from 011/00 to 012/00 indicates DPCs for the counter. Instead of performing data training on the original datasets shown in Fig. 7, the data training is performed on the rate of increases or the derivatives of the counter values. To highlight DPCs in the derivatives of the counter values, the accumulated derivatives are implemented for the processed data in Fig. 7 so that the values of the processed data are

$$p^a(t_i) = \begin{cases} 0 & \delta(t_{i-\Delta}, t_i) \leq 0 \\ p^a(t_{i-\Delta}) + \delta(t_{i-\Delta}, t_i) & \delta(t_{i-\Delta}, t_i) > 0 \end{cases} \tag{4}$$

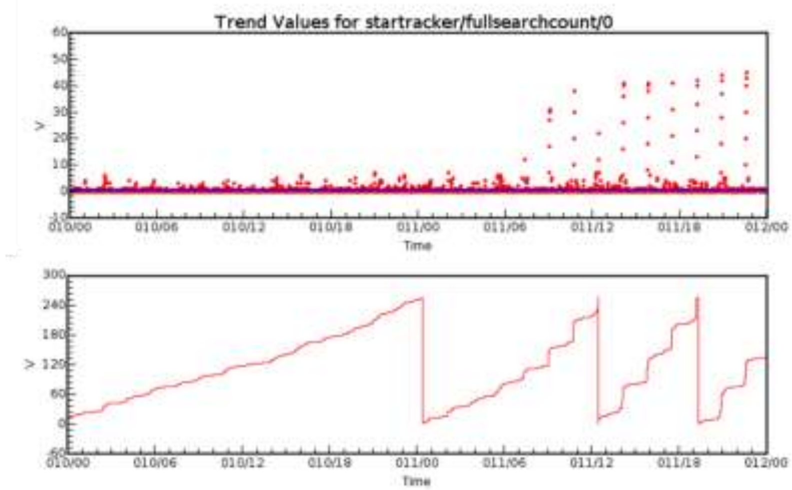

Figure 7 The full search count in the Star-Tracker subsystem. The plot at bottom is the original data. The plot at the top is the processed data used in the data training. The blue line in the top plot is the mean value of the processed data.

where $\delta(t_{i-\Delta}, t_i)$ is the derivative in the period Δ:

$$\delta(t_{i-\Delta}, t_i) = \frac{d(t_i) - d(t_i - \Delta)}{\Delta} \tag{5}$$

The period Δ is a parameter dependent on the specific datasets. The processed data shown in Fig. 7 are constants with small fluctuations in the periods with normal data patterns, and the higher derivatives in the dataset are converted to higher values in the processed data in the Fig. 7 that are treated as the outliers corresponding to DPCs. The data model and training algorithm for the processed data in Fig. 7 become a statistical collection of the processed dataset, which are simpler comparing those for the original data. This shows that the data training for the derivative of original data for the counter type data makes the ML data model simpler and highlights the DPCs that are difficult to evaluate in the original data.

IV. The Event Vector and Hierarchical Clustering Algorithm

The outputs from the data training process for each telemetry dataset include a data model and outliers, which are used as the inputs in the clustering process. The outliers from the data training process represent DPCs of individual data points. When a DPC happens in a dataset, it generally involves consecutive outliers for an extended period. This section provides the detailed discussions on how to define data pattern change matric based on the outliers generated in the data training process, and to create event vectors from data pattern change metrics, and to preform hierarchical clustering on event vectors.

A. Normalized Outliers and Data Pattern Change Metric

A data model for a dataset defines a data bound, and values of a dataset with a data bound should satisfy the following relationship

$$\left|f_j(t_i) - d_j(t_i)\right| < N\sigma_j \tag{6}$$

where *N* is a user defined threshold parameter. A data point with its value outside the data bound defined in Eq. 6 is defined as a normalized outlier, which can be quantitatively characterized by

$$O_i^N\big(d(t_i)\big) = \frac{\delta\big(|\Delta(t_i)| > N\sigma\big)}{\sigma}\left\{\Delta(t_i) - sign\big(\Delta(t_i)\big) \cdot N\sigma\right\} \tag{7}$$

where

$$\Delta(t_i) = f(t_i, \{S_k\}) - d(t_i). \tag{8}$$

Eq. 7 defines a normalized outlier that $O^N\big(d(t_i)\big) = 0$ when $|\Delta(t_i)| = N\sigma$, which is dimensionless so that one can compare the outliers among datasets with different scales.

The presence of an isolated outlier in a dataset generally does not constitute a data pattern change that impacts satellite operations. A data pattern change generally involves multiple consecutive outliers, which can be measured by the data pattern change metric

$$\psi_j = \frac{1}{T}\sum_i \frac{t_i - t_{i-1}}{2}\left(\left|O_j^N\big(d(t_i)\big)\right| + \left|O_j^N\big(d(t_{i-1})\big)\right|\right)\delta\left(t_i - t_{i-1} = {}^{1}/_{f_s}\right) \tag{9}$$

where $\delta\left(t_i - t_{i-1} = {}^{1}/_{f_s}\right)$ is a Boolean function, f_s is the data sampling frequency that is generally 1 Hz for most of telemetry data, and the quantity T is a time scale so that the metric ψ_j in Eq. 9 is dimensionless. Eq. 9 is in fact an integration of the quantity $O_j^N\big(d(t_i)\big)$ for consecutive outliers. The metric ψ_j represents the severity of DPCs in a telemery dataset, and it depends on the number of consecutive outliers and the quantity $O_j^N\big(d(t_i)\big)$.

B. Event Vectors and Event Plots

Both normal events and anomalies in telemetry datasets can be regarded as events, which generally involve DPCs in multiple datasets in the same period due to correlations among datasets in multiple subsystems. An event can be mathematically characterized by an event vector $\boldsymbol{E}$ with non-zero DPCs of correlated datasets aligned in the same period:

$$\boldsymbol{E} = \{\psi_1, \psi_2, \dots \psi_n\} \tag{10}$$

The event vector $\boldsymbol{E}$ in Eq. 10 is defined in a space with the data pattern change metric ψ_i in Eq. 9 as its coordinate, which could be regarded as an outlier space. A dataset $d_j(t_i)$ generally in our ML approach has the hierarchical structure *Subsystem/Mnemonic/Index* so that it can be uniquely mapped to its physical state in a satellite. For example, the name *reaction-wheel/motor-current/2* represents the telemetry dataset motor-current for reaction wheel 2 in the Reaction-Wheel subsystem, as there are generally multiple reaction wheels in a satellite. An event vector $\boldsymbol{E}$ is written in the following hierarchical form:

$$\boldsymbol{E} = \begin{cases} \boldsymbol{\psi}_{S_1} \begin{cases} \boldsymbol{\psi}_{S_1,M_1} \begin{cases} \boldsymbol{\psi}_{S_1,M_1,I_1} \\ \cdots \\ \boldsymbol{\psi}_{S_1,M_1,I_l} \end{cases} \\ \cdots \\ \boldsymbol{\psi}_{S_1,M_k} \begin{cases} \boldsymbol{\psi}_{S_1,M_k,I_1} \\ \cdots \\ \boldsymbol{\psi}_{S_1,M_k,I_l} \end{cases} \end{cases} \\ \cdots \\ \boldsymbol{\psi}_{S_n} \begin{cases} \boldsymbol{\psi}_{S_n,M_1} \begin{cases} \boldsymbol{\psi}_{S_n,M_1,I_1} \\ \cdots \\ \boldsymbol{\psi}_{S_n,M_1,I_l} \end{cases} \\ \cdots \\ \boldsymbol{\psi}_{S_n,M_k} \begin{cases} \boldsymbol{\psi}_{S_n,M_k,I_1} \\ \cdots \\ \boldsymbol{\psi}_{S_n,M_k,I_l} \end{cases} \end{cases} \end{cases} \tag{11}$$

where S stands for subsystems, M for mnemonic, and I for index, as each subsystem in a satellite consists multiple mnemonics, and a mnemonic generally has multiple elements. The indexes (S, M, I) can uniquely determine a dataset in a system with hierarchical structure. The strength of an event $\boldsymbol{E}$ can be written as

$$\psi = \sqrt{\sum_S \psi_S^2}, \tag{12}$$

which determines the overall severity of an event. The outlier value at the subsystem level ψ_S in Eq. 12 is an aggregated value of its components at the mnemonic level:

$$\psi_S = \sqrt{\sum_M \psi_{S,M}^2} \tag{13}$$

similarly, the metric $\psi_{S,M}$ at the mnemonic level is an aggregated value of its elements at the index level:

$$\psi_{S,M} = \sqrt{\sum_I \psi_{S,M,I}^2}. \tag{14}$$

Fig. 8 shows an outlier plot at the subsystem level, which consists plots of outliers for multiple subsystems with a common time axis, which highlights correlations among different subsystems. The outlier plots in Fig. 8 involves outliers for the Reaction-Wheel and the Ephemeris subsystems, and outliers in the two subsystems aligned in the same period are associated with a single event involving correlations between Reaction-Wheel and Ephemeris systems. An ephemeris subsystem contains datasets that determine attitude, position, velocity, and angular momentums of a satellite. Reaction-Wheel subsystems in a satellite are used to maintain the attitudes in a satellite, changes in momentum profiles in reaction wheels are generally responding to changes of satellite momentum profiles due to external forces to keep overall attitude constant or responding to command performing an attitude maneuver. Thus, changes in the ephemeris subsystem lead to changes in the reaction-wheel subsystem or vice versa.

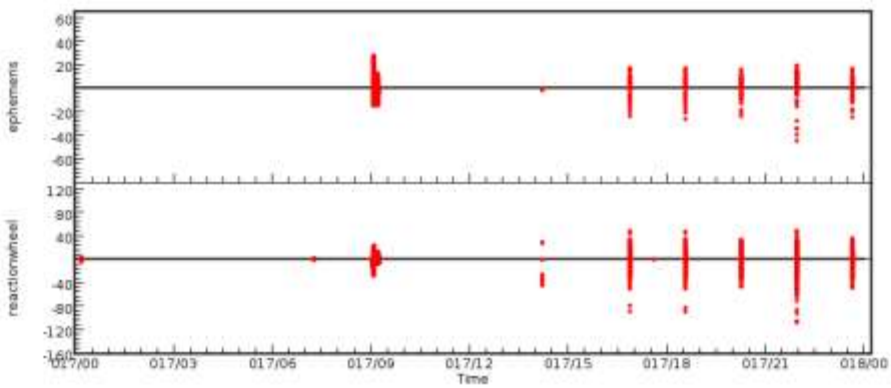

Fig. 8 The outlier plot at the subsystem level for the reaction-wheel (bottom) and ephemeris (top) subsystems of NPP satellite. The outliers at the subsystem level include the outliers of all datasets in corresponding subsystem. Values of outliers are determined by Eq. 8.

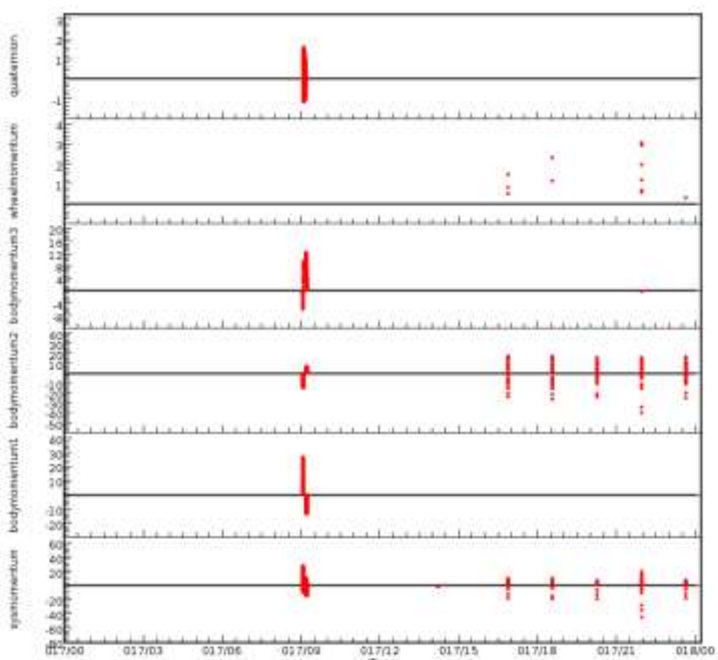

Figure 9 The distribution at the mnemonic level for the outliers in the ephemeris subsystem in Fig. 8

Outliers at the subsystem level involve all outliers present in its datasets, and one can examine how outliers are distributed in each subsystem at the mnemonic level to show correlations among different mnemonics in the same subsystem. Fig. 9 shows the distribution of outliers in the Ephemeris subsystem in Fig. 8 at the mnemonic level. The changes in the Reaction-Wheel subsystems generally leads to changes in angular momenta in a satellite, such as the system and body momenta. A significant feature for the event at 09 hours in Figs. 8 and 9 is the DPCs in the quaternions, which indicates the attitude of the corresponding satellite has been changed. This event involves the changes in Reaction-Wheel subsystem and the attitudes of a satellite, which suggests an attitude maneuver operation that is part of normal operations. The other events in Figs. 8 and 9 don't involve attitude changes, which indicates a small change in system momentums leads to changes in the Reaction-Wheel subsystem to keep a satellite in balance. Since this type of events happens frequently in a satellite, they are regarded as normal events that are part of satellite operations.

An event vector $\boldsymbol{E}$ in Eq. 10 can be further characterized by the amplitude ψ and the orientation of event vectors in an outlier space. The orientation of an event vector is expressed as a unit event vector

$$\boldsymbol{e}(\psi_S) = \left\{\frac{\psi_{S_1}}{\psi}, \frac{\psi_{S_2}}{\psi}, \dots \frac{\psi_{S_n}}{\psi}\right\} \tag{15}$$

at the subsystem level, where ψ is given in Eq. 12 representing the overall strength of an event vector so that

$$\boldsymbol{e}(\psi_S) \cdot \boldsymbol{e}(\psi_S) = \frac{1}{\psi^2} \sum_{S_i} \psi_{S_i} \psi_{S_i} = 1 \tag{16}$$

The unit vector in Eq. 15 consists of the composition and relative strengths of datasets with DPCs corresponding to patterns of correlations in normal events or anomalies. The same type events can have the same orientations in an outlier space but may have different amplitudes ψ, since the periods of events could be different that leads to different amplitudes. For example, two thruster firing events have the same composition and relative strengths in their event vectors, although they may not have the same amplitudes because one thruster firing event may take longer time than the other.

The unit vector in Eq. 15 can be shown in an event plot for a given period, and Fig. 10 shows the event plot for the events in Fig. 8. An event plot consists histograms of subsystems with a common time axis. Each event in an event plot is characterized by a time, a relative strength, and an event period. The events in Fig. 10 show the DPCs in the Reaction-Wheel and Ephemeris subsystems, while there is no DPC present in other subsystems. The vertical bars in different subsystems aligned in the same period represent the relative strengths of DPCs in an event in each subsystem. The values of relative strengths in an event from different subsystems or mnemonics have a range from 0 to 1 and correspond to the values of the vertical bars in Fig. 10. The width of vertical bars aligned in the same period represents the event period, and the longer period of an event generally leads to the larger event amplitude. The event plot in Fig. 10 shows two types of events. The relative strengths of the attitude maneuver event around the 09 hours in Fig. 10 are different from those in other events, in which the DPCs in the Ephemeris subsystem are stronger than those in the Reaction-Wheel subsystem. The relative strengths for other events in Fig. 10 are dominated by the DPCs in the Reaction-Wheel subsystem. This shows that different types of events in Fig. 10 have different relative strengths in DPCs, and relative strengths in DPCs among datasets are the signatures of different types of events. Fig. 10 shows event plots in the subsystem level, and event plots can also be shown at the mnemonic level as well for events with dominant strength in DPCs from a single subsystem to study how relative strengths are distributed at the mnemonic level.

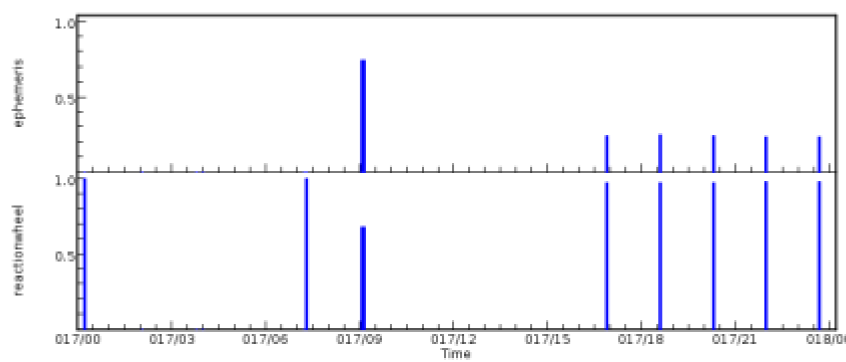

Fig. 10 The event plot at subsystem level for the events displayed in Fig. 7, which shows the time, period, relative strength from 0 to 1 for each subsystem in an event.

C. Hierarchical Event Clustering

The unit event vectors in Eq. 15 can be used as the ML representation in clustering algorithms to separate normal events from anomalies. One can define an event similarity metric α as the scalar product of two event vectors at a given hierarchical level

$$\alpha_S^{i,j} = \boldsymbol{e}^{\boldsymbol{i}}(\psi_S) \cdot \boldsymbol{e}^{\boldsymbol{j}}(\psi_S) = \frac{1}{\psi^i \psi^j} \sum_S \psi_S^i \psi_S^j \tag{17}$$

where the indices *i* and *j* are the event index. The metric α_S has a range from 0 to 1. The value 1 corresponds to two parallel vectors corresponding to the same type of events, while the value 0 for the metric α_S represents two very different events. Therefore, the metric $\alpha_S^{i,j}$ in Eq. 17 can be used as the clustering metric for event classifications and anomaly detections: the same type events with $\alpha_S^{i,j} \approx 1.0$ belong to the same cluster. Normal events in satellite operations are generally repeatable and occur regularly to form their own clusters, while anomalous events happen rarely generally and don't form their own clusters since each anomaly has its own characteristics. Fig. 11 provides pseudo code of a hierarchical clustering algorithm used in separating normal events from anomalies. The DBSCAN clustering algorithm[14] is implemented as the clustering algorithm at each hierarchical level.

```
Hierarchical Clustering Algorithm
    Perform clustering for all events at subsystem level
        Output noise events
        For each cluster at subsystem level
        Perform clustering for events mnemonic level
          Output noise events
```

Fig. 11 Hierarchical event clustering for event classification and anomaly detections.

V. Clustering Results for NPP Health and Safety Telemetry

The data training and clustering analysis for the NPP satellite telemetry includes telemetry datasets in the Reaction Wheel, Star Tracker, Ephemeris, Gyro, Power, Propulsion, and Thermal subsystems. Events are generated in daily training sessions by converting outliers into the data pattern change metrics, and data pattern change metrics aligned in the same period form an event vector. The number of events being generated in each training session ranges from 0 to 30 events so that events must be accumulated from multiple training sessions to have large number of events for clustering purpose. The scalar product $\alpha_S^{i,j}$ in Eq. 17 is the metric in DBSCAN clustering, and the threshold $\alpha_S^{i,j}$ value in DBSCAN clustering is set to be 0.98 for two events in the same cluster.

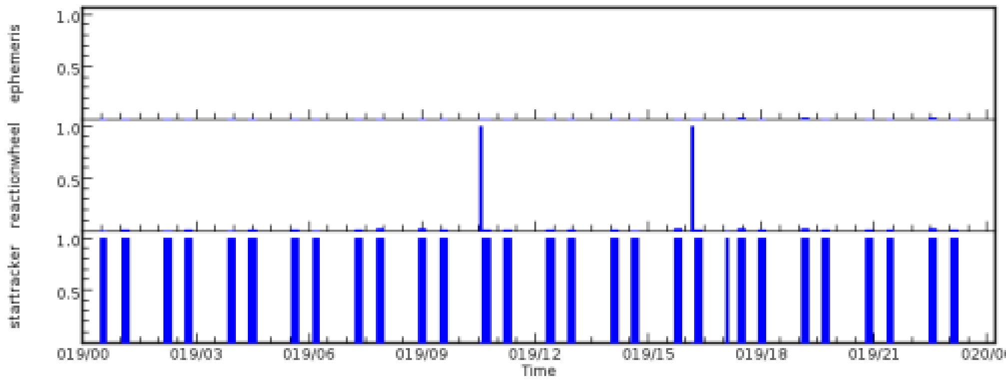

Fig. 12 The event plot for NPP telemetry data on 2019/019. The events are mostly dominated by the DPCs in the Reaction-Wheel or the Star-Tracker subsystems. There are small contributions from DPCs in the Ephemeris subsystem.

The clustering is performed for events accumulated over 75 training sessions from 2019/005 to 2019/079. There are more than 300 significant events accumulated over 75 training sessions. Most of events show strong correlations among datasets in Reaction-Wheel, Ephemeris, and Star-Tracker subsystems, which are used to determine satellite attitudes. The clustering outputs at the subsystem level consist of two main clusters, and the noise events that don't belong to any cluster account for less than 1 percent of total number of events. The two clusters show very distinct characteristics: one cluster is dominated by DPCs in the Reaction-Wheel subsystem with small contributions from the Ephemeris subsystem; while the events in other cluster are dominated by the DPCs in the Star-Tracker subsystem with small contributions from both Reaction-Wheel and Ephemeris subsystems. The events dominated by the Reaction-Wheel subsystem in Fig. 10 are the examples of the cluster 1 events. The events in the cluster 2 account for 80 percent of total events generated in 75 training sessions and have their own temporal patterns. The daily event distribution for the cluster 2 shows a burst of events (>10 events) in some training sessions, while the number of events is generally small (<10) in most of training sessions. The large number of events in some daily sessions are the events in two clusters. Fig. 12 shows an event plot at the subsystem level on 2019/019 with around 31 events, which is typical for outliers in these subsystems aligned in the same time periods. The relative strengths for most of 31 events in Fig. 12 are dominated by the Star Tracker subsystem, while other 2 events in Fig. 12 show the dominance of the Reaction-Wheel subsystems in DPCs. The event distribution in Fig. 12 shows a periodic pattern that points to specific positions on the satellite orbit. The overall strength ψ for a typical event in the cluster dominated by DPCs in the Star Tracker subsystem have much larger amplitudes than that in the cluster dominated by the DPCs in the Reaction-Wheel subsystem. The events in Fig. 12 are most likely due to changes in the NPP space environments, and a study[16] suggested that this might be linked to the low energy electrons and geomagnetic perturbations in a LEO environment. Although there are a lot of data patteren changes in this training session, the clustering analysis shows that most events belong to either cluster so that they are normal events.

Fig. 13 shows an example of a noise event on 2019/055. The event at 21:16z shows that the relative strengths of DPCs are not dominated by either Star-Tracker or the Reaction-Wheel subsystem so that it does not belong to the two clusters, which indicates a potential anomaly that requires further investigation.

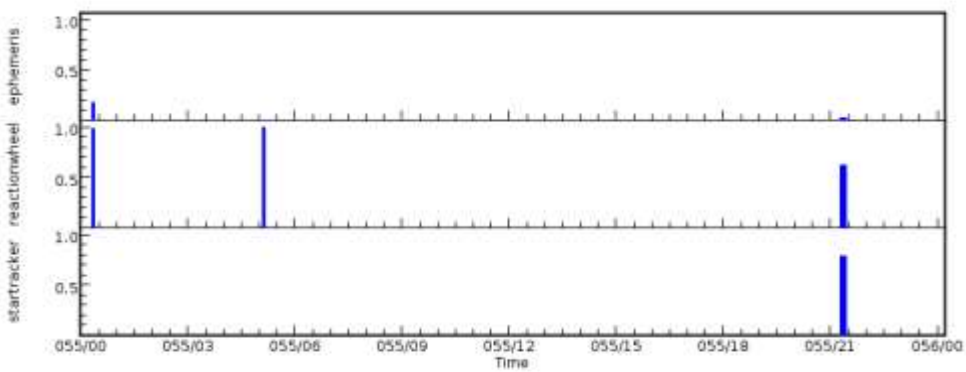

Fig. 13 The event plot for NPP data on 2019/055. The around 21:16z shows that the relative strengths of DPCs are not dominated by the Star-Tracker or Reaction-Wheel subsystem, and there is also a small data pattern change in the Ephemeris Subsystem in the event.

VI. Summary and Outlook

Satellite health and safety telemetry data are highly complex, which bring considerable challenges in developing ML data models for different data pattern types with high complexity. The data training outputs presented in this paper implement different data models and data training approach for different data pattern types, which enables data training to be performed in an operational environment with sufficient accuracy and efficiency. Our investigation shows that correlations among datasets lead to DPCs in multiple datasets during the same period, and patterns of correlations among datasets determine the nature of events. The hierarchical event vectors provide a ML representation to characterize patterns of correlations for normal events as well as anomalies. The event plots based on unit event vectors capture signatures of normal operation events and anomalies and provide satellite engineers a new ML tool for engineering analysis in satellite operations. NPP telemetry data also show the importance in separating normal events from anomalies, as only few of the events (less than 1 percent) in NPP telemetry data are potential anomalies. Our investigation shows that the hierarchical clustering algorithm is very effective and efficient in separating normal events from anomalies, and this is especially critical for a large and complex satellite like NPP spacecraft.

This study focuses on the telemetry data for LEO satellites. The application of the ML approach to anomaly detection for satellite with different orbit characteristics, such as the geosynchronous satellite, should address the same issues or challenges discussed in this paper. The data patterns for telemetry data in different orbital characteristics have different complexities, and the operation concepts for satellite operations might be different, which may lead to different ML algorithms in data training.

The algorithms discussed in this paper has been implemented in Advanced Intelligent Monitoring System (AIMS). AIMS is a flexible, extensible, and scalable ML platform for space missions. ML algorithms in AIMS are treated as plugin and play components to enable low cost, low risk, and rapid deployment in a mission. AIMS has a wide range application portfolio including satellite onboard instrument monitoring, satellite health and safety telemetry data monitoring, and satellite launch vehicle, such as Space Launch System(SLS), monitoring. AIMS is currently deployed in NOAA GOES-R ground system for monitoring satellite health and safety telemetry and instrument calibration process, which has shown to provide considerable benefits to satellite operations.

Acknowledgments

This work is supported by ASRC Federal, and the author are grateful for the support from the satellite engineers in ASRC Space Operating Group.

References

[1] Zhenping Li, "A Machine Learning Framework for Space Missions", 35th Space Symposium, April. 8-11, Colorado Springs, Colorado. https://www.spacefoundation.org/what-we-do/research-and-analysis/tech-track-paper-archive/developing-machine-learning-solutions.

[2] Zhenping Li, “Developing Machine Learning Solutions for Spacecraft Ground Systems”, Proceedings of 6th International Conference on Space Mission Challenges for Information Technology (SMC-IT), Alcala de Henares, Spain, September 2017.

doi: 10.1109/SMC-IT.2017.20.

[3] Zhenping Li, “Creating Situational Awareness in Spacecraft Operations with the Machine Learning Approach”, 2016 AMOS Conference, September, 20-23, Maui, Hawaii. http://www.amostech.com/TechnicalPapers/2016/Poster/Li.pdf.

[4] Zhenping Li, Ken Mitchell, Biruh Tesfaye, and Dave Pogorzala, “Monitoring GOES-R Advanced Baseline Imager (ABI) Radiometric Performance with a Machine Learning System”, CALCON 2018 Conference Proceedings, June, 2018. Logan Utah. https://digitalcommons.usu.edu/calcon/CALCON2018/all2018content/15/.

[5] D. DeCoste “Automated learning and monitoring of limit functions”, In Proc. Int. Symp. Artif. Intell., Robot. Autom. Space., 1997, pp. 287–292.

[6] T. Yairi, M. Nakatsugawa, K. Hori, S. Nakasuka, and K. Machida “Adaptive limit checking for spacecraft telemetry datausing regression tree learning” In Proc. Int. Conf. Syst., Man Cybern., 2004, pp. 5130–5135.

[7] R. Fujimaki, T. Yairi, and K. Machida “An anomaly detection method for spacecraft using relevance vector learning”, In Proc. 9th Pacific-Asia Conf. Knowl. Discovery Data Mining., 2005, pp. 785–790.

[8] S. K. Ibrahim, A. Ahmed, M. A. E. Zeidan and I. E. Ziedan, "Machine Learning Methods for Spacecraft Telemetry Mining," in IEEE Transactions on Aerospace and Electronic Systems, vol. 55, no. 4, pp. 1816-1827, Aug. 2019.

doi: 10.1109/TAES.2018.2876586

[9] Y. Gao, T. Yang, M. Xu and N. Xing, "An Unsupervised Anomaly Detection Approach for Spacecraft Based on Normal Behavior Clustering," 2012 Fifth International Conference on Intelligent Computation Technology and Automation, Zhangjiajie, Hunan, 2012, pp. 478-481,

doi: 10.1109/ICICTA.2012.1

[10] Schwabacher, N. Oza, and B. Matthews Unsupervised anomaly detection for liquid-fueled rocket propulsion health monitoring J. Aerosp. Comput., Inf. Commun., vol. 6, no. 7, pp. 464–482, 2009.

[11] L. Iverson et al. “General purpose data-driven monitoring for space operations” J. Aerosp. Comput., Inf. Commun., vol. 9, no. 2, pp. 26–44, 2012.

[12] R. Fujimaki, T. Yairi, and K. Machida, “An approach to spacecraft anomaly detection problem using kernel feature space”, In Proc. 11th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining, 2005, pp. 401–410.

[13] Yairi, Takehisa et al. “A Data-Driven Health Monitoring Method for Satellite Housekeeping Data Based on Probabilistic Clustering and Dimensionality Reduction”, IEEE Trnsactions on Aerospace and Electronic Systems, vol. 53, no. 3, 2017.

[14] Martin Ester , Hans-Peter Kriegel , Jörg Sander , Xiaowei Xu, “A density-based algorithm for discovering clusters in large spatial databases with noise”, Proceedings of the Second International Conference on Knowledge Discovery and Data Mining (KDD-96). AAAI Press. pp. 226–231, 1996.

[15] A. Zimek, E. Schubert, and H.-P. Kriegel A Survey on Unsupervised Outlier Detection in High-Dimensional Numerical Data Statist. Anal. Data Mining, vol. 5, no. 5, pp. 363–387, 2012.

[16] Ahmad, N., Herdiwijaya, D., Djamaluddin, T., Usui, H., Miyake, Y., “Diagnosing low earth orbit satellite anomalies using NOAA-15 electron data associated with geomagnetic perturbations”. Earth Planet Space 70, 91 (2018).
doi: 10.1186/s40623-018-085