Evaluation Testing of Learning-based Telemetry Monitoring and Anomaly Detection System in SDS-4 Operation

T.Yairi*, T. Oda*, Y. Nakajima**, N. Miura**, N. Takata**

*The University of Tokyo, Japan
e-mail: {yairi, oda}@space.rcast.u-tokyo.ac.jp

**Japan Aerospace Exploration Agency (JAXA), Japan
e-mail: {nakajima.yuta, miura.yoshiyuki, takata.noboru}@jaxa.jp

Abstract

Health monitoring and anomaly detection techniques for artificial satellites are very significant, as it is very hard to repair those space systems on orbit. Authors have proposed the framework of learning-based anomaly detection that applies statistical machine learning and data mining techniques to the satellite telemetry data to automatically obtain normal behavior models which can be used for monitoring the health status of the system. In this study, we evaluated the learning-based anomaly detection method in the on-going operation of JAXA's satellite, SDS-4 (Small Demonstration Satellite 4), and obtained miscellaneous insights for putting this technology into practical use.

1 Introduction

Health monitoring and anomaly detection are very significant issues for the artificial satellite operation. If they miss a single symptom of anomalous behavior, the whole mission may become a failure. Although the telemetry data sent from the spacecraft potentially contains abundant information on the system health status, it is very difficult for the operators to monitor the huge amount of data in detail thoroughly. It is needless to say that the classical limit checking technique is insufficient for this purpose.

Based on these backgrounds, the authors have proposed and studied the framework of learning-based (or data-driven) anomaly detection for artificial satellites. In this framework, ones apply some machine learning or datamining algorithms to the past telemetry data and acquire normal behavior models. Once the normal models are learned, they check the subsequent data using them and judge whether any anomalous patterns are contained in the data. Especially, in our previous works [1], we proposed a learning-based anomaly detection method based on the clustering and dimensionality reduction techniques. We applied the method to the data of several past artificial satellites and confirmed its effectiveness. Through these experiences, the authors are now developing a learning-based anomaly detection system named ADAMS.

In this paper, we report the intermediate results of evaluation testing of ADAMS conducted in the post-mission phase of the Small Demonstration Satellite-4 (SDS-4) of JAXA.

2 Learning-based Anomaly Detection for Artificial Satellites

2.1 Anomaly Detection with Machine Learning

Roughly speaking, anomaly detection or health monitoring techniques for artificial systems (including spacecraft) are divided into two categories – knowledge-driven and data-driven methods. The former category includes the rule-based methods [2][3] and model-based methods [4]. They perform anomaly detection and diagnosis using rule or model bases obtained from human experts. While these knowledge-driven approaches have succeeded in some areas, it is generally expensive to construct and maintain high-quality rule and model bases.

On the other hand, the latter approaches obtain statistical models representing the normal behavior of systems by applying some techniques of statistics, multivariate analysis and machine learning to the past normal data of the systems. In this paper, we call this approach “learning-based” anomaly detection. Fig.1 illustrates the basic procedure of the learning-based (or data-driven) anomaly detection. Following the terminology in machine learning, we call the data from which the normal model is learned “training data”, and the data which is monitored by the learned model "testing data", respectively. A comprehensive overview of learning-based anomaly detection methods in general can be found in [5].

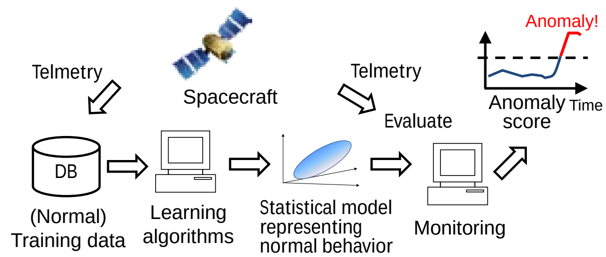

Figure 1. Learning-based Anomaly Detection

2.2 ADAMS: Anomaly Detection and Monitoring System

The authors also have developed several anomaly detection methods using a variety of machine learning techniques[6][7][8][9], and applied them to the telemetry data of past artificial satellites. Based on these results, we are now developing a health monitoring system named ADAMS (Anomaly Detection and Monitoring System), which is meant to assist the satellite operators in monitoring and understanding the health status of systems. Fig.2 shows an example of snapshots of ADAMS during analyzing satellite telemetry data.

So far, several kinds of anomaly detection methods using machine learning and probabilistic inference algorithms such as MPPCA (mixture of probabilistic PCA), SLDS (switching linear dynamical systems) and so on have been implemented in ADAMS, in addition to the functions of computing basic statistics and visualization of the telemetry data.

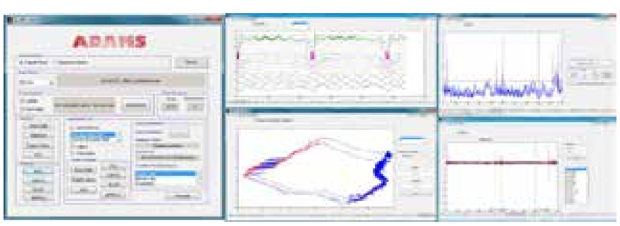

Figure 2. A Screenshot of ADAMS

2.3 Anomaly Detection using Mixture of Probabilistic PCA

In this study, we focus on MPPCA among the machine learning algorithms implemented in ADAMS. MPPCA, originally developed by Tipping and Bishop [10], is a probabilistic model which has features of both clustering and dimensionality reduction.

In our previous work [11], it was shown experimentally that MPPCA is well-balanced in performance and efficiency compared with other clustering and linear / non-linear dimensionality reduction methods, when modeling the high-dimensional and multi-modal satellite telemetry data.

In the rest of this section, we describe the procedure of the learning-based anomaly detection method using MPPCA.

2.3.1 Training Phase: Modelling the normal behavior of data by MPPCA

In the training phase, a MPPCA model is learned from "normal" telemetry data $\boldsymbol{Y}_{\text{train}} = [\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_N]^T$. The probability density of an observation vector is represented as,

$$p(\boldsymbol{y}_i \mid \Theta) = \sum_{k=1}^{K} \pi_k \cdot N\left(\boldsymbol{y}_i ; \boldsymbol{\mu}_k, \boldsymbol{W}_k \boldsymbol{W}_k^T + \sigma_k^2 \boldsymbol{I}\right) \tag{1}$$

Here, $\boldsymbol{y}_i \in R^D$ denotes the i-th observation vector, and D denotes the original dimensionality of the telemetry variables. $\Theta = \{\boldsymbol{\mu}_1, \boldsymbol{W}_1, \pi_1, \sigma_1^2, \ldots, \boldsymbol{\mu}_K, \boldsymbol{W}_K, \pi_K, \sigma_K^2\}$ is the set of model parameters of MPPCA. $N(\cdot; \boldsymbol{\mu}, \boldsymbol{\Sigma})$ stands for the probability density of a multivariate Gaussian distribution with the mean μ and covariance matrix Σ.

As can be seen from Eq. (1), MPPCA is a special case of the Gaussian mixture model (GMM) whose covariance matrix is constrained. Moreover, Eq. (1) can be represented in another form, using a discrete latent variable z_i and a continuous latent variable vector $\boldsymbol{x}_i$ as follows,

$$p(\boldsymbol{y}_i \mid \Theta) = \sum_{k=1}^{K} p(\boldsymbol{y}_i \mid \boldsymbol{x}_i, z_i, \Theta) p(\boldsymbol{x}_i \mid \Theta) p(z_i \mid \Theta) \tag{2}$$

$$p(\boldsymbol{y}_i \mid \boldsymbol{x}_i, z_i = k, \Theta) = N\left(\boldsymbol{y}_i ; \boldsymbol{W}_k \boldsymbol{x}_i + \boldsymbol{\mu}_k, \sigma_k^2 \boldsymbol{I}\right) \tag{3}$$

$$p(\boldsymbol{x}_i \mid \Theta) = N(\boldsymbol{x}_i ; \boldsymbol{0}, \boldsymbol{I}) \tag{4}$$

$$p(z_i \mid \Theta) = \mathrm{Cat}(z_i ; \boldsymbol{\pi}) \tag{5}$$

Here, $\mathrm{Cat}(\cdot; \boldsymbol{\pi})$ stands for the categorical distribution. A maximum likelihood estimation of the MPPCA parameters $\hat{\Theta}$ can be obtained by the EM algorithm [10] from the training data $\boldsymbol{Y}_{\text{train}}$.

2.3.2 Testing Phase: Monitoring of new data

In the testing phase, we need to determine whether each observation vector $\boldsymbol{y}_i$ in the testing data $\boldsymbol{Y}_{\text{test}} = [\boldsymbol{y}_{N+1}, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_{N+N'}]^T$ is normal or anomalous. Once the MPPCA model representing the normal data is obtained, we can compute the likelihood of $\boldsymbol{y}_i$ using Eq. (1). Therefore, we define the anomaly score a($\boldsymbol{y}_i$) as the negative log likelihood as follows,

$$a(\boldsymbol{y}_i) \equiv -\log p\left(\boldsymbol{y}_i \mid \hat{\Theta}\right) \tag{6}$$

If the value of $a(\boldsymbol{y}_i)$ is larger than some threshold, it means that $\boldsymbol{y}_i$ is an "unseen" pattern and is likely to be anomalous.

2.3.3 Diagnostic Analysis

An advantage of using MPPCA is that not only the anomaly score $a(\boldsymbol{y}_i)$ but also several other kinds of useful information are available.

Firstly, we can tell the global behavior of the system by looking into the distribution as to the discrete latent variable $p\left(z_i \mid \boldsymbol{y}_i, \hat{\Theta}\right)$, as z_i indicates which cluster (or local model) the observation vector $\boldsymbol{y}_i$ belongs to.

Besides, a "reconstructed" observation vector $\hat{\boldsymbol{y}}_i$ can be easily computed in MPPCA. When an anomalous pattern is detected, we can guess which variables are more related to it by examining the contribution of each variable to the reconstruction error, i.e., $\|\boldsymbol{y}_i - \hat{\boldsymbol{y}}_i\|^2$.

2.3.4 Extension and Limitation

From the view point of modeling the satellite telemetry data, an important extension of MPPCA is the treatment of telemetry variables that take discrete (or qualitative) values. This is because the telemetry data contains a number of discrete variables such as "on-off" status of system instruments, as well as continuous variables. While the original MPPCA model assumes only continuous variables, it is not very difficult to enable it to cope with discrete variables. In ADAMS, we combine the mixture of categorical distributions with the original MPPCA model to deal with both discrete and continuous telemetry variables.

Another important enhancement of MPPCA is the treatment of dynamics. In our previous work [1], we showed that MPPCA can be combined with (semi-)hidden Markov model (HMM) to capture the temporal transition of the latent discrete variables from z_i to z_{i+1} . In ADAMS, we have also implemented the learning and inference algorithms of SLDS (switching linear dynamical systems) [12], which also takes into account the continuous state transition of the system. In this study, however, we do not consider these dynamics modeling techniques, because the sampling periods of SDS-4 telemetry often change due to the operation restrictions. In other words, we treat each observation vector $\boldsymbol{y}_i$ as an i.i.d. sample.

3 Preliminary Testing of ADAMS on SDS-4 Telemetry Data

3.1 Overview

In July 2013, we conducted a preliminary testing of ADAMS in the developer's side, by applying the anomaly detection method to the telemetry data of SDS-4 satellite. The main purpose of this preliminary testing is to evaluate the functions of ADAMS qualitatively and obtain miscellaneous hints for improving it.

In this experiment, we first obtain the MPPCA model from the training data $\boldsymbol{Y}_{\text{train}}$ that consists of the SDS-4 telemetry of a week (7 days). Then we monitor the test data $\boldsymbol{Y}_{\text{test}}$ that consists of the telemetry of one subsequent day using the learned MPPCA model. We repeated this training-test cycle for 2 weeks (Fig. 3). Tab. 1 summarizes the training / testing periods of this preliminary experiment. Subsequently, we presented the results such as anomaly scores to the SDS-4 operators, and asked them to evaluate the results.

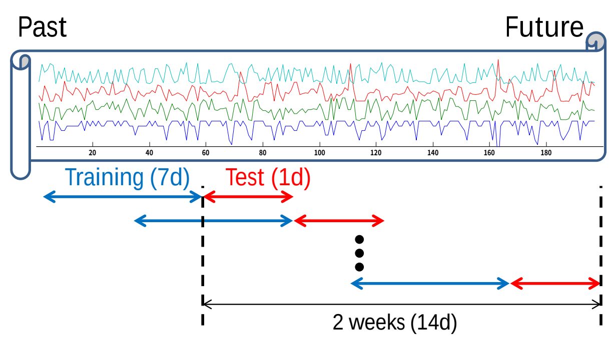

Figure 3. Training-Testing Cycles of Preliminary Experiment

While 1458 variables are originally defined in the SDS-4 telemetry, the number of valid variables which can be used for health monitoring is limited, because most of them are missing or remain constant in the most of time. Therefore, we manually chose the set of variables. Tab.3 summarizes the number of monitored variables grouped by subsystems. Due to the limitation of communication capacity, the number of valid continuous telemetry variables varies, depending on whether the satellite is visible or not visible from the ground station.

As the original telemetry variables have a large variety of physical units and ranges (e.g., electrical current, voltage, temperature, rotational speed, etc.), we have to "standardize" or "nondimensionalize" them in order to deal with them in a uniform manner. In this experiment, each variable is scaled and centered so that the 1st and 99th percentiles correspond to -1 and +1 respectively, in the period of two months from May to June of 2013.

In the remaining of this section, we show three examples of results in this preliminary experiment.

Table 1. List of Training / Testing Periods of Preliminary Experiment (Results of cycle #2,3,5 are described in sections 3.2, 3.3, and 3.4, respectively.)

Cycle No.	Training		Testing	
	Period (UT)	# samples	Period (UT)	# samples
1	6/24 4:52 - 6/30 23:58	10298	7/1 0:00 - 7/1 12:09	1744
2	6/24 4:52 - 7/1 12:09	12000	7/1 12:09 - 7/2 0:23	230
3	6/24 4:56 - 7/2 0:23	12000	7/2 0:23 - 7/2 12:26	2762
4	6/25 7:05 - 7/2 12:26	12000	7/2 13:45 - 7/2 13:53	465
5	6/25 15:01 - 7/2 13:53	12000	7/3 3:20 - 7/4 3:20	2423
6	6/27 3:28 - 7/4 3:20	12000	7/4 3:24 - 7/5 3:17	1792
7	6/28 5:00 - 7/5 3:17	12000	7/5 3:22 - 7/6 3:13	1887
8	6/29 1:07 - 7/6 3:13	12000	7/6 3:17 - 7/7 3:12	405
9	6/30 1:46 - 7/7 3:12	12000	7/7 3:16 - 7/8 3:11	441
10	7/1 2:05 - 7/8 3:13	12000	7/8 3:13 - 7/9 3:07	1967
11	7/2 4:19 - 7/9 3:07	12000	7/9 3:12 - 7/10 3:07	2135
12	7/2 12:18 - 7/10 3:07	12000	7/10 3:07 - 7/11 3:03	2841
13	7/4 2:03 - 7/11 3:03	12000	7/11 3:03 - 7/12 3:02	2841
14	7/4 11:30 - 7/12 3:02	12000	7/12 3:03 - 7/13 3:01	2075
15	7/5 7:29 - 7/13 3:01	12000	7/13 3:02 - 7/14 2:58	780
16	7/6 3:03 - 7/14 2:58	12000	7/14 3:02 - 7/15 2:57	598

Table 2. Numbers of Monitored Telemetry Variables in Preliminary Experiment (Grouped by Subsystems)

	# of Continuous Variables		# of Discrete
Subsys.	Visible	Not visible	Variables
ACS	76	34	21
AIS	0	0	14
CDH	0	0	5
DER	0	0	37
DSS	2	1	0
EPS	8	8	14
MAG	3	0	0
MCR	0	0	5
MIU	0	0	3
MTQ	0	0	6
STS	0	0	54
STT	4	4	5
TCS	24	24	0
TRX	1	1	11
VSG	12	0	3
WHL	15	0	0
Total	145	72	178

3.2 Result 1: Successful Detection of an Anomalous Event

In the early testing period of cycle No.5, a sharp increase in the anomaly score was detected (Fig. 4). By identifying the variables that contribute to the anomaly score in this period, we found that some of the temperature measurements of reaction wheels and other instruments decreased significantly in the corresponding time period (Fig.5). According to SDS-4 operators, this was likely to be due to the change in the battery heater threshold, which had been performed by them in reaction to the preceding error that occurred in the on-board computer (OBC). This result implies that ADAMS successfully detected the unexpected change in the data.

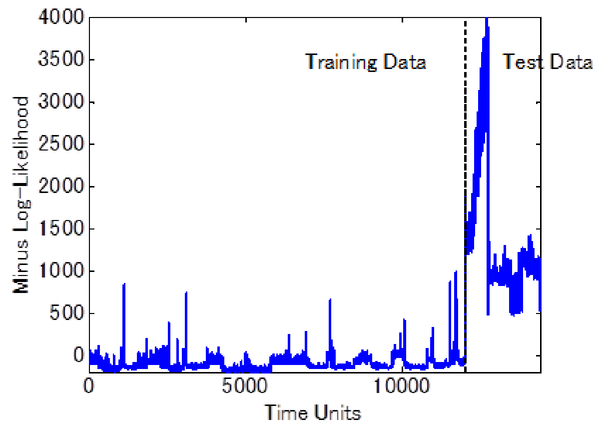

Figure 4. Anomaly Score in Preliminary Testing Cycle No. 5 (Result 1)

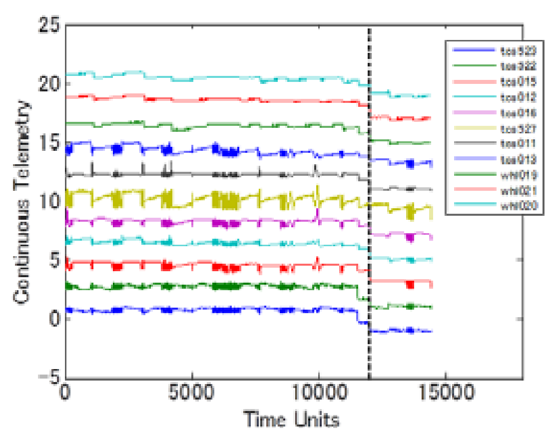

Figure 5. Identified Variables Contributing the Anomaly Score in the testing period of No.5 (Result 1)

3.3 Result 2: Detection of Unseen Modes

In the middle of testing period of cycle No.3, the anomaly score became considerably large over a certain duration (Fig.6). We found that two telemetry variables ACS044 (target attitude angle Y) and ACS046 (VSGA measurement w_y) especially contributed to the anomaly score during the period in question. According to the operators, this is considered to be due to the attitude maneuvering for an instrument experiment (IST experiment). Although the attitude maneuvering itself is not an anomalous event as it was planned, the anomaly detection worked reasonably in a sense that it detected a data pattern that had not been contained in the training data.

behavior of the variables during the period in question is obviously different from that in the training data. The SDS-4 operators, however, pointed out that this detection was a false alarm, because the detected pattern was not so rare in a long-term view. After all, we concluded that this was a typical example of false detections caused by inappropriate setting of training data and standardization of variables.

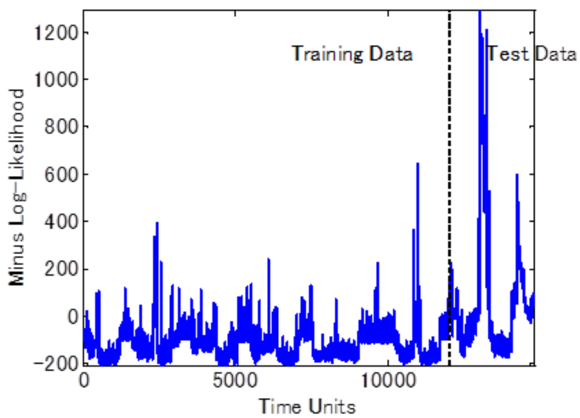

Figure 6. Anomaly Score in Preliminary Testing Cycle No. 3 (Result 2)

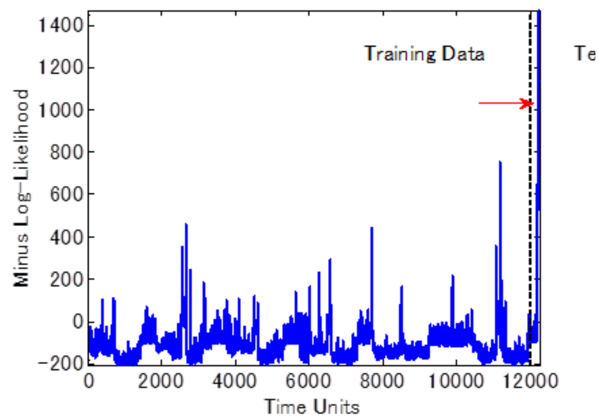

Figure 8. Anomaly Score in Preliminary Testing Cycle No. 2 (Result 3)

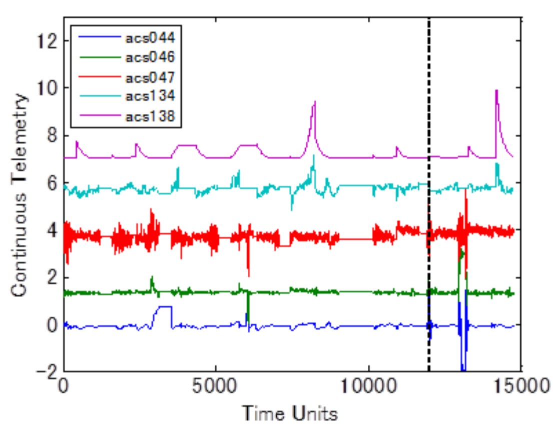

Figure 7. Identified Variables Contributing the Anomaly Score in the testing period of No.3 (Result 2)

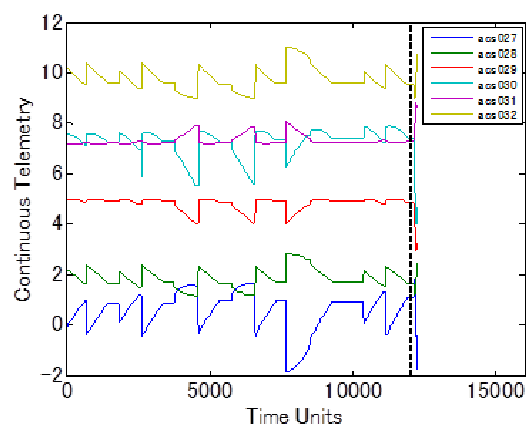

Figure 9. Identified Variables Contributing the Anomaly Score in the testing period of No.2 (Result 3)

3.4 Result 3: False Alarms

In the last part of testing period of cycle No. 2, the anomaly score increased sharply (Fig. 8). We found that some of the variables related to the satellite position information contributed to the anomaly score. Judging from the original data of those variables (Fig. 9), it seems that this result is also a reasonable one, because the

4 Improvements of ADAMS

Based on the the preliminary testing results described above, several refinements were made on ADAMS. In this section, we explain them.

4.1 Selection of Variables

The framework of learning-based anomaly detection implicitly assumes that the characteristics of the data are stationary, if the system remains normal. The SDS-4 telemetry, however, contains variables whose characteristics slowly change over time due to some reasons such as seasonality, regardless of whether the system is normal or not. These variables unnecessarily increase the anomaly score of test data. Therefore, we first compare the distributions of each variable between training and testing periods, and exclude the variables whose distributions are not stationary from the set of variables used in the MPPCA model. In other words, advanced multivariate analysis methods such as MPPCA should be utilized only when the anomaly cannot be found by examining each variable individually.

4.2 Variable Standardization

While we standardized each telemetry variable based on the data of the past two months as explained in section 3.1, there were some "inactive" variables that stayed almost constant during the period. Those variables were not appropriately standardized, which lead to performance degradation. To overcome this issue, we simply used the data of the past one year to standardize each variable. Of course, this strategy is not applicable if there are not sufficient amount of data. In that case, the variables should be standardized based on other information sources such as the design specifications and the results of the ground tests.

4.3 Selection of Training Data

In the preliminary experiment, we used the data of the past one week as the training data. However, we found that this amount of data was insufficient, as the basic operation cycle period of SDS-4 was four weeks. As a result, possible behavior patterns of the satellite were not fully covered by the training data, which leads to a lot of false alarms. Therefore, we changed the basic period of the training data to one month.

Moreover, we enabled ADAMS to compose the training data more flexibly and dynamically so that it contains the same operation patterns as the testing data does, taking advantage of the prior information on the operation schedule

4.4 Model Library

By extending the dynamical selection of training data explained above, we introduced a new idea of "model library". In this framework, we save all trained models in past into the "library", then pick up a suitable model from it that contains similar operation patterns as the testing data. This enables us to "reuse" the learning results (i.e., MPPCA models), whereas the former framework treated the trained models as "disposable".

4.5 Multiresolution Monitoring

In the preliminary testing, we basically used all telemetry variables that meet a certain condition. Although it has an advantage that we can monitor all the variables at the same time, it also has a disadvantage that we tend to miss subtle anomalies that are related to a small set of the variables. Furthermore, it is hard to identify the causal variables even if anomalies are detected. Therefore, in addition to the conventional monitoring mode using all variables, we adopted a "distributed" monitoring mode in which we partition the variables into several groups, train multiple MPPCA models corresponding to the groups, and use them to monitor the test data in parallel. There are several ways of grouping based on (1) subsystems, (2) properties of variables such as physical units, and (3) statistics such as correlation and mutual information.

5 Evaluation Testing of Improved ADAMS

5.1 Overview

We provided the revised version of ADAMS to the SDS-4 operation team in November of 2013, and they started an evaluation testing on the user's side. While this evaluation testing is still on the way, we show an example in which they found ADAMS was helpful for their understanding the satellite health status.

5.2 A Case Study

As SDS-4 is a demonstration satellite, a variety of experiments testing mission instruments have been performed. It is known that the telemetry data behaves differently for each instrument testing. Exploiting the characteristics, the operation team conducted a "virtual" anomaly detection experiment as described below.

They first composed the training data by gathering up the telemetry data during the period between May 7 and Jul 28, 2013, corresponding either to (1) routine operation or (2) FOX instrument testing operation. Then they had ADAMS learn a MPPCA model from the training data. After that, they specified the telemetry data in the period between Jul 29 and Sep 1, 2013 as the testing data. This testing data contains not only the two operational patterns contained in the training data, but also another operation mode - (3) MCMR (monitoring camera) experiment, in which the attitude of the satellite is controlled so that MCMR points to the earth. If ADAMS works correctly, the data during this MCMR experiment is expected to be detected as anomalous.

Fig. 10 shows the anomaly score in the training and testing data. In this figure, the period of MCMR experiment is indicated by a red rectangle. Fig. 11 shows the same score during the period and the differences between the original and reconstructed data of the top 10 contributing telemetry variables. Interestingly, most of the contributing variables belong to the TCS (thermal control system), while the MCMR experiment is mainly related to the attitude control. Fig. 12 shows the original data of these variables. Through this analysis, the SDS-4 operators came to the conclusion that the data during MCMR experiment is characterized by a unique pattern in some temperature-related variables, because during the experiment the sunlight shines on the satellite differently from other operation modes.

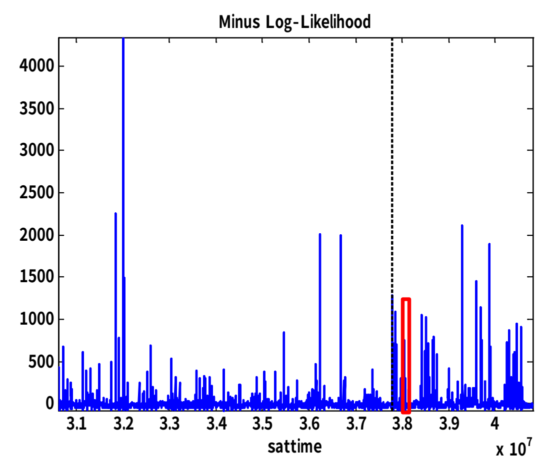

Figure 10. Anomaly Score in the Training and Testing Data

5.3 Discussion

So far, the SDS-4 operators have evaluated ADAMS favorably. They especially admit that ADAMS is helpful for them to understand the health status of the satellite. It was observed, however, that ADAMS was still sensitive to the selection of the training data and the set of telemetry variables, and standardization of the variables, which lead to many false positives. This made us realize ADAMS needs more improvements to be accepted as semi-automatic and reliable health monitoring system for artificial satellites. Especially, we think the way of making full use of the operators' "feedback" to improve its performance or reduce false detections is important.

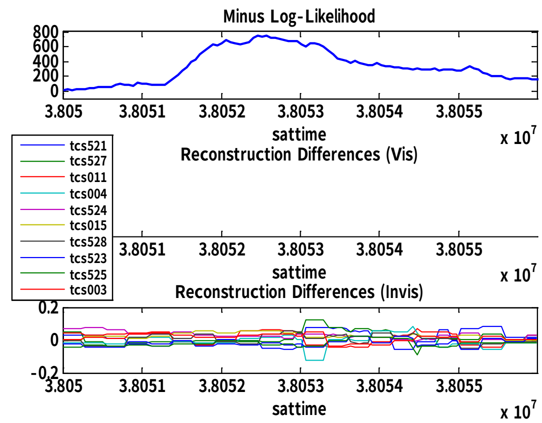

Figure 11. Anomaly Score in MCMR Experiment Period (Top) and Differences between Original and Reconstructed Data of Top 10 Contributing Variables (Bottom)

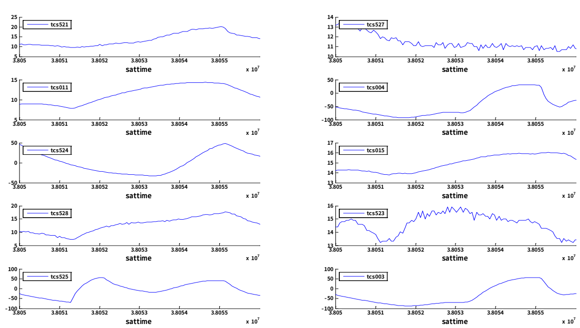

Figure12. Original Data of Top 10 Contributing Variables during MCMR Experiment

6 Conclusions

In this paper, we reported intermediate results of the evaluation testing of the learning-based health monitoring and anomaly detection system named ADAMS that have been performed in the post-mission phase of SDS-4. While some problems remain to be solved, we confirmed that ADAMS provides the satellite operators with valuable information on the system health status. We are going to continue this evaluation testing along with a lot of improvements on ADAMS.

References

[1] T. Yairi, T. Tagawa, and N. Takata, "Telemetry Monitoring By Dimensionality Reduction and

Learning Hidden Markov Model", i-SAIRAS 2012, Turin, Italy, 4 - 6 September 2012.

[2] F. Ciceri, and L. Marradi, “Event Diagnosis and Recovery in Real-time On-board Autonomous Mission Control”, Proceedings of the 1st Ada-Europe Symposium, 1994.

[3] D. P. Tallo, J. Durkin and E. J. Petrik, “Intelligent Fault Isolation and Diagnosis for Communication Satellite Systems”, Telematics and Informatics, Vol. 9, pp.173-190, 1992,.

[4] B. C. Williams and P. P. Nayak, “A Model-Based Approach to Reactive Self-Configuring Systems”, Proceedings of the Thirteenth National Conference on Artificial Intelligence and Eighth Innovative Applications of Artificial Intelligence Conference, pp.971-978 , 1996..

[5] V. Chandola, A. Banerjee and V. Kumar, "Anomaly Detection : A Survey", ACM Computing Surveys, Vol. 41(3), 2009.

[6] R. Fujimaki, T. Yairi, K. Machida, "An Approach to Spacecraft Anomaly Detection Problem Using Kernel Feature Space", The Eleventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining , pp.401-410, 2005.

[7] R. Fujimaki, T. Yairi and K. Machida, "Adaptive Limit-checking for Spacecraft using Relevance Vector Autoregressive Model", i-SAIRAS 2005, Munich, 2005.

[8] Y. Kawahara, T. Yairi and K. Machida, "Spacecraft fault diagnosis based on switching estimation of parameters using sequential Monte Carlo methods", i-SAIRAS 08, Los Angeles, 2008.

[9] T. Tagawa, T. Yairi, N. Takada and Y. Yamaguchi, "Data Monitoring of Spacecraft Using Mixture Probabilistic Principal Component Analysis and Hidden Semi-Markov Models", The 3rd International Conference on Data Mining and Intelligent Information Technology Applications, Macau, 2011.

[10] M.E. Tipping and C.M. Bishop, “Mixtures of Probabilistic Principal Component Analyzers”, Neural Computation, 11, pp.443-482, 1999.

[11] T. Yairi, M. Inui, A. Yoshiki, Y. Kawahara and N. Takata, "Spacecraft Telemetry Data Monitoring by Dimensionality Reduction Techniques", The SICE Annual Conference, pp.1230 - 1234, Taipei, 2010.

[12] K. Murphy, “Learning Switching Kalman Filter Models”, Compaq Cambridge Research Lab Tech Report 98-10, 1998.