Applying Data Mining for Detecting Anomalies in Satellites

Denise Rotondi Azevedo, Ana Maria Ambrósio

Space Engineering and Technology
National Institute for Space Research, INPE
São José dos Campos, Brasil
denise@dss.inpe.br, ana@dss.inpe.br

Marco Vieira

CISUC, Department of Informatics Engineering
University of Coimbra
Coimbra, Portugal
mvieira@dei.uc.pt

Abstract—Telemetry data is the only source for identifying/predicting anomalies in artificial satellites. Human specialists analyze these data in real time, but its large volume, makes this analysis extremely difficult. In this experience paper we study the hypothesis of using clustering algorithms to help operators and analysts to perform telemetry analysis. Two real cases of satellite anomalies in Brazilian space missions are considered, allowing assessing and comparing the effectiveness of two clustering algorithms (K-means and Expectation Maximization), which showed to be effective in the case study where several telemetry channels tended to deliver outlier values and, in these cases, could support the satellite operators by allowing the anticipation of anomalies. However for silent problems, where there was just a small variation in a single telemetry, the algorithms were not as efficient.

Keywords - space systems; anomaly detection; clustering.

I. INTRODUCTION

Artificial satellites normally render important services in communication, remote sensing, scientific experiments, etc. Satellite damage implies not only a financial loss, but also loss of essential, and sometimes strategic, services. In this scenario, the early detection, diagnosis, and prevention of anomalies and faults promotes reliability and availability of space systems, extending their lifetime and enduring continuity of service.

A satellite may be logically seen as a set of integrated subsystems (orbit and attitude control, thermal, power supply, structure, payload, on-board computer, etc.). Aimed at monitoring the satellite, each subsystem has a set of sensors (thermistors, switches, battery depletion, etc.) for measuring the satellite state and conditions. These measurements are transmitted to the ground stations on every satellite passage and are commonly called housekeeping telemetry data.

The telemetry data is analyzed and monitored by human operators and analysts, mainly to assess if the values are not out of a pre-defined range or pattern. Values out of limits normally indicate a potential anomaly in a given satellite subsystem, but the large amount of telemetry data makes it almost impossible to perform a careful and detailed analysis in real time.

This practical experience paper studies the hypothesis of using data mining techniques to help operators and analysts to perform telemetry analysis. In practice, the goal is to assess the possibility of detecting satellite anomalies through the analysis of telemetry data using clustering algorithms.

Anomaly detection refers to the problem of finding patterns in data that do not conform to the expected behavior [1]. There are many different techniques that can be used to find a non-conforming pattern in data, such as classification-based techniques, statistical based techniques, clustering, among others. Clustering-based techniques are quite adequate for satellite anomaly detection as they can operate in a semi-supervised or unsupervised way. These techniques also have an efficient test phase based on comparing a test instance with a small number of clusters defined in the previous training phase [1]. This efficiency is an essential characteristic for real-time monitoring.

Clustering [2][3] is used to group similar data instances into clusters. Even though clustering and anomaly detection appear to be different from each other, several clustering-based anomaly detection techniques have been developed. These techniques can be divided into three groups, which are related to the way they classify anomalies: 1) normal data instances belong to a cluster, while anomalies do not belong to any cluster; 2) normal data instances belong to large and dense clusters, while anomalies belong to small or sparse clusters; and 3) normal data instances are closer to their closest cluster centroid, while anomalies are far from their closest cluster centroid [1].

The key reason behind using clustering algorithms for detecting anomalies in satellites is that the telemetry data from normal satellite operation periods form clusters of characteristics and any telemetry instances that do not belong to these clusters are a potential indication of anomaly (normally, this is the type of analysis that human operators and analysts do manually). This way, it seems possible to train a cluster algorithm with data from periods of normal operation and then evaluate new telemetry data by assigning an index that indicates how much of the new instances belong (or not) to the learned normal clusters. This index is normally a measure of similarity, dissimilarity or deviation.

In this paper we present the results of the application of clustering techniques in the context of two real cases of satellite anomalies in Brazilian space missions. In practice, the experimental study evaluates and compares the effectiveness of two algorithms (K-means and Expectation Maximization) in detecting real anomalies in telemetry data. Although there are many alternative algorithms, we selected K-means [4][5] because it is stable and quite well known, and Expectation Maximization [6] as it uses statistical mixture models and is less sensitive to the lack of data standardization.

The outline of the paper is as follows. Section 2 presents an overview of related work. Section 3 introduces the case studies and Section 4 presents the experimental evaluation.

Section 5 summarizes the lessons learned. Finally, Section 6 concludes the paper and proposes lines for future work.

II. Background and Related Work

Anomaly detection refers to finding patterns in data that are not in accordance with a well-defined notion of normal behavior [1]. The automatic anomaly detection research area presents a number of challenges, including the availability of historical data, the definition of what are normal and abnormal patterns, and even the consideration that in many fields the concept of normality can evolve. Along these lines, we can say that one of the main issues associated with automatic anomaly detection task is the characterization of the problem to be addressed, including identifying the application domain characteristics, the available data and the appropriate techniques.

To identify and apply the most appropriated techniques for automatic detection of anomalies, the problem domain should be characterized in regard to some key aspects [1]: (i) *data type*: data can be uni or multivariate according to the number of attributes, and the attributes may be binary, categorical, discrete or continuous; (ii) *anomaly type*: it could be a point anomaly when a variation in a value indicates a problem; a context anomaly when a value is normal in a context but abnormal in another; or a collective anomaly when the anomaly indication is the occurrence of some values in an abnormal sequence and not the values themselves; (iii) *training data*: according to the available label of the input dataset (i.e., supervised, semi-supervised or unsupervised); (iv) *expected results*: this aspect defines what kind of result would be interesting or feasible in a given domain. The results can be classificatory (instances anomalous or normal) or an index (the probability of an instance being anomalous or not).

In [7] is proposed a method that summarizes telemetry data automatically by preserving the important information while discarding the rest. It considers as important information two kinds of events: immediate events where values are different from their neighbors and permanent events where there are permanent changes or trends in the telemetry behavior. In the method applied, first, each time-series in the telemetry data is divided into a set of subsequences with a fixed length. Then all the subsequences are grouped into clusters based on the DTW (Dynamic Time Warping) distance measure. Finally, each cluster is assigned a unique symbol, and the subsequences contained in "small" clusters are detected as event patterns.

In [8] the authors used the Kernel PCA statistical technique to expand conventional limit checking to adaptive limit checking by considering the relationship of telemetry data series. The proposed method creates new limits for the new-composed attributes, it may be interesting in two ways: for the reduction itself, and for the attributes combination that can reveal some abnormal interaction between attributes.

In [9] the same authors have compared the Kernel PCA techniques with data mining regression algorithms. Both have been employed to adapt telemetry measurement ranges in an automatic way.

NASA projects use ORCA [10] and IMS (Inductive Monitoring System) [11]. ORCA is a data mining tool that searches for unusual data points, or outliers, in multivariate data sets by calculating the distance of each data point from neighboring points, while the IMS tool uses clustering techniques to generate a knowledge base about the normal operation for health monitoring [11]. In particular, the IMS is a tool that applies clustering to extract models of normal system operation from archived data. It then characterizes how the parameters relate to one another during normal operation by finding areas in the vector space where nominal data tends to fall. These areas are called nominal operating regions and correspond to clusters of similar points found by the IMS clustering algorithm. The nominal operating regions are stored in a knowledge base used by IMS for real-time telemetry monitoring or archived for posterior data analysis [11].

In this work we follow an approach similar to the IMS tool, as we apply clustering techniques to detect anomalies in two different real-world problems. In practice, we evaluate and compare two clustering algorithms for anomaly detection in telemetry data: K-means and Expectation Maximization (EM).

K-means is a stable and well-known algorithm that aims at partitioning the data into k-clusters so that the within-group sum of squared errors is minimized. The simplest form of the algorithm is based on alternating two procedures. The first is to assign an object to the group whose mean is the closest in terms of the similarity measure. The second is the calculation of new group means based on the assignments. The process terminates when no movement of an object to another group reduces the within-group sum of squares [4][5]. The algorithm is efficient but sensitive to the first assignments.

The EM is a statistical clustering algorithm used to approximate a probability function. Its foundation is a statistical model called finite mixture, where a mixture is a set of k probability distributions, representing k clusters. EM assigns a probability distribution to each instance, which indicates the probability of it belonging to each of the clusters [5][6]. This algorithm also uses a heuristic approach and may be affected by the correlation between attributes.

Both algorithms make Gaussian assumptions about the underlying distribution of the data. However, they use a different index to determine if an instance belongs or not to a cluster. While K-means uses dissimilarity measures (Euclidean, Manhattan, etc.), the EM checks the probability of a given instance belonging to a Gaussian distribution. The algorithms are implemented via Weka API [12].

III. Case Studies: Two Real Satellite Failures

Thanks to satellite specialists, two distinct problems, in two different satellites (named here as S1 and S2), were identified for our study. Although not related, both problems concern the power supply subsystem which, according to statistics presented in different papers addressing failures in satellites [13][14], this is one of the most critical subsystems in a satellite. In both studies, the power subsystem was responsible for a considerable percentage of failures occurrences: 27% [13] and 22% [14].

A. Power Supply Subsystem

The power subsystem is responsible for generating and conditioning the satellite primary energy (solar panels), for storing energy on secondary sources (batteries), and for supplying energy to other subsystems at different voltages.

The power supply subsystem consists of two energy sources: a primary source that converts the light striking on the solar panel (SAG) during sunlight periods into electrical energy, and a secondary source (batteries) that can store energy through chemical processes. The control of the battery charge status is made by the End of Charge (EOC) circuit, which limits the maximum battery charge turning the charging current on and off. The conditioning of the main bus is done by SHUNT and Battery Discharge Regulator (BDR), which both control the energy supplied by the SAG and by the battery, providing the satellite and the DC/DC converters with a stable main bus.

During the eclipse phase, the batteries discharge as in these periods they are the only source of energy for the satellite. In periods of sunlight, the batteries are charged up to the charge limit indicated by the end-of-charge (EOC) curve. This process takes a certain time t0, which is often less than the total period of satellite illumination (t0 + t1). Fig. 1 (a) illustrates this.

The charging circuit checks the temperature and charges the battery to the limit specified by the charging curve (curve EOC). The value of this charge limit is calculated according to the curve chosen and is a function of the observed temperature. Verification pulses are launched to obtain the level of battery charge and charging current is applied as a function of the observed temperature. These pulses are short and have a small current. The application of this small charge current is required to measure the present level of battery discharge (Deep-of-discharge). When the circuit detects that the battery is not fully charged, the charge is made, always checking the maximum battery charge (EOC), i.e., a current (pulse) is applied to the battery until it reaches the maximum charge value indicated by the EOC curve (always observing the battery temperature). Once charged, during the entire sunlight period, the verification pulses continue to be applied to the battery at regular intervals to verify whether it is still completely charged. Whenever it is partially discharged, the charging process is repeated to ensure that the battery enters the eclipse period at full capacity. Fig. 1 (b) presents this process, showing the pulses as green lines.
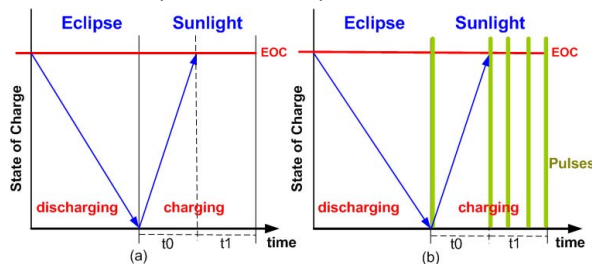
B. Power Subsystem Telemetry Channels

Figure 1.   Battery Charge – Normal State.

Telemetry indicates satellite status and health. In the case of the satellites studied in this work, there are more than 2000 different types of telemetry channels to be analyzed by ground operations. These represent values measured by sensors and equipment, so they are classified as: **Analog**: are related to analog values such as temperatures, currents, etc. **Binary**: indicates equipment states, such as on or off, main or redundant, etc.

In this work we used only the analog telemetry channels that indicate the measurements of interest, namely the telemetry channels related to the battery, solar panel, Shunt and BDR equipment, since they are the most important for detecting anomalies in the power subsystem. Table I introduces the list of telemetry channels that were analyzed.

C. Case Study 1 - Satellite S1

During the entire operation of the S1 satellite, a silent problem was detected on the solar panel (SG1-B). During some periods, the SG1-B solar panel telemetry channel (TM023) presented values lower than the expected ones for sunlight time, showing a weakness in the power supply of battery 2. Fig. 3 shows the behavior of TM023 during the entire life of the satellite S1. In the figure the y-axis represents the months since the satellite launch (October, 1999), while the x-axis represents the values taken by telemetry. The frequency of these values can be observed through colors, darker colors represent a higher occurrence of a given value in time.

As seen in Table I, the normal range values for telemetry channel TM023 is between 0 and 7.2A. The telemetry channel values are normally close to zero during the eclipse period, when the solar panel does not receive any energy, and near 6A in the sunlight period. By analyzing the behavior of this telemetry channel throughout the satellite lifetime (Fig. 2), we observe the following: (i) in the region of high values there is a seasonal effect that occurs due to the normal solar panel degradation; (ii) there are intermediate values between the measurements expected during eclipse and sunlight periods; (iii) over time, the number of measurements in the intermediate region increases; (iv) in the periods of January to March each year, the phenomenon of intermediate measurements did not occur.

TABLE I.   POWER SUBSYSTEM TELEMETRY CHANNEL

| Telemetry	Description	Normal Values
TM001	Main Bus voltage	27 to 29V
TM002	Main Bus current. This indicates if the current is being supplied by the BDR (batteries) or by Shunt (main solar panel)	0 to 36A
TM003	Mean Error Amplifier output (MEAS)	8 to 23.5V
TM013/017	BDR input current (battery discharge current)	0 to 13A
TM014/018	Battery voltage	43.2 to 56.5V
TM015/019	Battery temperature	0 to 10°C
TM016/020	3 Cell minimum voltage	3.6 to 4.65V
TM021	BDR output current (same as MEAS at eclipse time)	0 to 36A
TM022/023	Solar panel current (SG1 and SG2)	0 to 7.2A
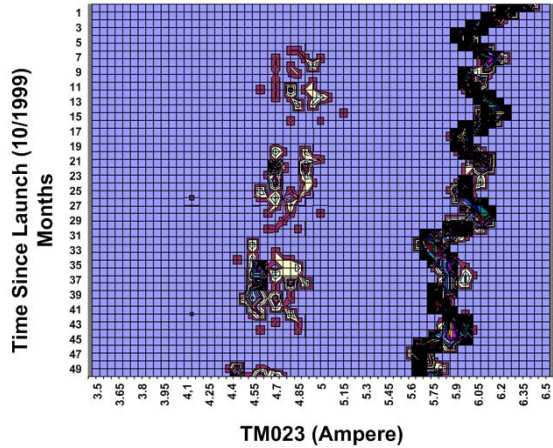
Figure 2. TM023 Behavior.

By analyzing the telemetry channel values, experts concluded that the most probable cause of the reduced value of TM023 was a real reduction of the SAG current due to an intermittent contact problem caused by fatigue crack in the welded junction of 4 string return lines (due to thermal stress and improper use of the stress relief technique). The failed welded junction is part of the wing power harness and can also be stressed during manipulation of the harness before flight.

D. Case Study 2 – Satellite S2

In this case, a failure was identified in two different operation periods of satellite S2. From the power subsystem point of view, this is not a primary error, since it first appeared in the Attitude and Orbit Control subsystem. In practice, a problem in this subsystem led the satellite to an emergency state. In this state, the priority is to keep the satellite's power guaranteed. To ensure that, the solar panel is better pointed towards the sun, and, to save energy, all equipment is turned off.

As mentioned before, the satellite's batteries are charged during the sunlight periods to be utilized during eclipse periods. Charging a battery takes place through the BDR and the EOC circuit. This circuit applies a current pulse to check the battery deep-of-discharge (DOD). If necessary, the battery is charged. However, in the emergency state the batteries charge faster than normal, since the solar panel is well pointed and all the equipment is turned off. Due to the
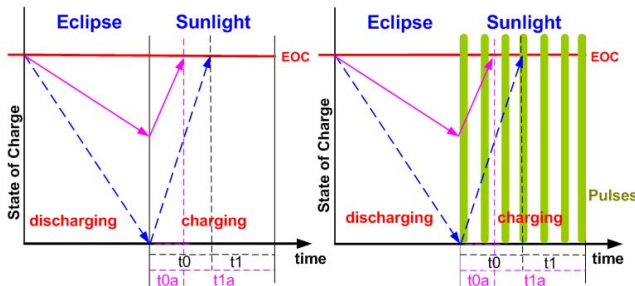

Figure 3. Battery Charge – Emergency State.

combination of two factors – the application of pulses in the battery for longer than expected and a problem in the pulses duration – the battery received more charge than normal, which started being transformed in heat. This heat was not completely dissipated through the thermal subsystem and the battery warmed a little more every pass (Fig. 3).

As shown, in every charging process, the EOC circuit indicates the level of battery charge, which is calculated through the EOC curve and is a temperature function. During some satellite revolutions, the temperature rose but the charge was being corrected. As the time went by, the temperature reached a maximum value for which the curve fails to compensate the voltage to be applied to the battery. At this point, the problem got worse and the satellite could have been lost.

This problem, classified by the satellite analysts as being catastrophic, happened during a few satellite passages. This failure could be seen through many different telemetry channels, especially those connected to battery 2. In reality, from a certain point in time, the affected telemetry channels became out of range and started being monitored by operators.

IV. EXPERIMENTAL STUDY

This section discusses the results of the experimental study, showing the application of the anomaly detection algorithms to real stored telemetry data, for normal and anomalous operation periods. This experimental campaign included two steps, as explained further.

In the **first step**, the algorithms were trained and tested with data from normal operation periods. We conducted three different experiments: (i) trained and tested the algorithms with normal data from the early periods of satellite life; (ii) trained and tested the algorithms with normal data from some years later; (iii) trained the algorithms with normal data from the early periods of the satellite life and tested the algorithm with normal data from some years later (in order to assess the aging effect in the power subsystem equipment).

The data used in the K-means algorithm training and testing phase were standardized relative to their mean and standard deviation. For this algorithm, the dissimilarity indexes were calculated using both Euclidean and Manhattan distances and the results showed that clustering and detection capacity were similar (only Euclidean Distance results are presented).

The clustering algorithms behaved as expected for case studies 1 and 2 in normal operation periods. For both studies the clustering algorithms indexes were in a normal range, i.e., the result indexes are low indicating that test data are close to the trained clusters. Fig. 4 shows the results for case study 2 (similar results were observed for case study 1).

Another observation regarding the values of the clustering index is that, when trained with data from the satellite early life and checked with data from later periods, the algorithms pointed correct data as anomalies, i.e. raised false positives (Fig. 5). In this case, the indexes are greater than they were in the previous test (Fig. 4) and exceeded the threshold. This is especially true in Case Study 2, where the characteristic of
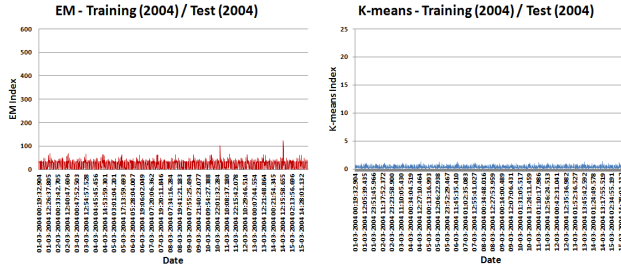
Figure 4.   Case Study 2 – Training (2004)/Test(2004).

the battery changed considerably over the years.

Since there are normal changes in the values of telemetry due to the satellite aging or due to on-earth operations, in order to incorporate this evolutionary aspect of the data, the index threshold used was calculated as been the average of dissimilarity indexes with two standard deviation variation. The use of this threshold proved to be satisfactory for both normal and anomalous phase. In the normal phase, it presents just spurious false positives (they are in the range of at most 5% and did not show constancy as it normally shows in anomaly periods).
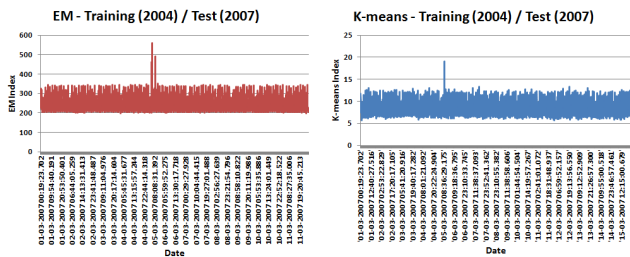

Figure 5.   Case Study 2 – Training (2004)/Test(2007).

In the **second step** of the experimental evaluation, the algorithms were trained with data from normal operation periods and tested with data from anomalous periods. This step was aimed at evaluating the anomaly detection capability of the clustering algorithms.

In case study 1, the behavior of both algorithms was very similar. In fact, the algorithms did not point the low value regions for TM023 as observed by the specialists, but showed the real unexpected unbalance values in telemetry channels TM023 and TM022. This showed that both algorithms failed in detecting values slightly outside the standard in a single telemetry channel and confirmed what was observed before in the simulated tests: a single telemetry channel with erroneous values between valid cluster regions
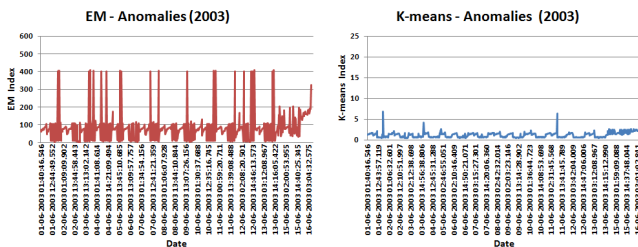

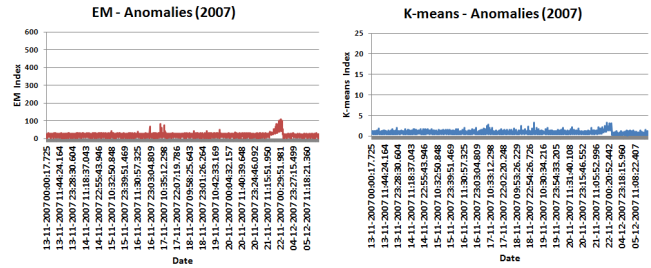Figure 6.   Case Study 1 – Anomaly Detection (2003).

Figure 7.   Case Study 2 – Anomalous Detection (2007).

is not well identified. Fig. 6 shows the results for both algorithms in Case Study 1.

In Fig. 7 (related to case study 2), we can see that both EM and K-means algorithms have detected the anomalous values. It can be observed that the indexes that measure the instances' distance to the cluster tend to grow in these anomalous periods, allowing the detection of the anomalies.

Fig. 8 compares the telemetry channel values with the K-means algorithm index showing the algorithm anomaly detection capability. The index grows in the regions where there was unbalance values between TM022 and TM023 that indicates a problem since these telemetry channels may have similar values. However, as shown in the figure, the algorithm does not identify the low unexpected values for TM023.

Fig. 9 shows a comparison between the main telemetry channel values and the index pointed by the algorithms. For the EM algorithm, the index started growing at least 6 hours before the alarm provided by telemetry channel TM015. So, the algorithms could anticipate the anomaly. Although short, this anticipation could allow taking measures to prevent propagation.

V. LESSONS LEARNED

The following points summarize the most important lessons from the experimental study with the clustering algorithms EM and K-means, both in a controlled environment (using simulated data) and in a real environment:

- The use of clustering algorithms for detecting anomalies allows operators to monitor just one index (measurement of distance or deviation from the known cluster) to observe an anomalous trend in data. This can
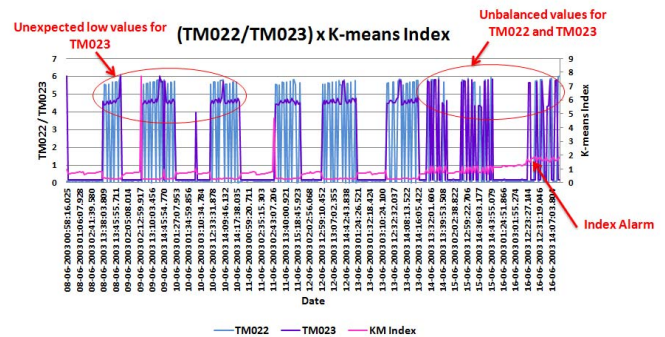

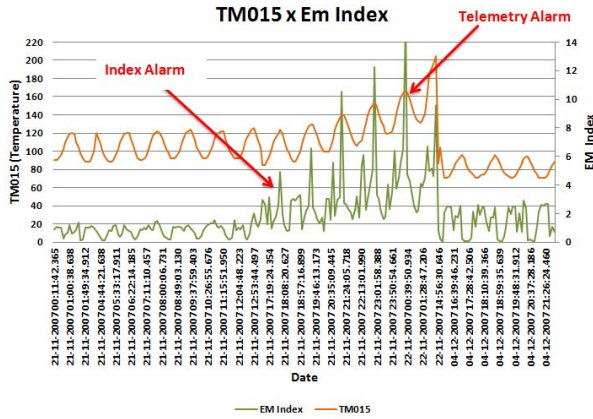Figure 8.   Case Study 2 – Anomalous Detection (2007).

Figure 9.   Case Study 2 – Anomaly Detection.

be a contribution to the routine of a satellite operation, since currently an operator has to monitor trends in many telemetry channels at the same time, which, under certain circumstances, can be extremely difficult (if not impossible);

- In Case Study 1, the algorithms were not able to detect the anomaly described by the specialists. Although we have used only 16 power telemetry channels, the slight value variation of a single telemetry channel could not be detected, since the weight of a variation in just one attribute, combined with the differences of the other ones in relation to the cluster centroids, is very low. If we had used a greater number of telemetry channels this problem would be worse. So, the use of clustering techniques for anomaly detection in satellites telemetries has to be combined with feature selection techniques, for example, to address this kind of problem. The feature selection algorithms reduce the feature space by selecting or combining the most relevant features in a new set of attributes;

- When using real anomalous data in Case Study 2, where telemetries exceeded the expected limits, both algorithms were able to detect the anomaly and there was a considerable anticipation (in relation to the time these telemetry channels would be alarming);

- Clustering algorithms are sensitive to telemetry channel changing behavior resulting from the satellite aging, so a continuous learning is required to help correct operation. However, continuous learning may lead the algorithms to be trained with anomalous behavior, since it may be difficult to distinguish, in advance, normal operation periods from anomalous operation periods;

- Both algorithms cluster data using Gaussian models and, as observed, worked similarly in both case studies EM algorithm seemed to work better when we simulated data with alternating trends and inter-group regions, although this kind of simulated data were not in the case studies.

VI.   CONCLUSIONS

The experimental study showed that the clustering algorithms are good to anticipate anomalies if applied in situations where several telemetry channels behave abnormally and at least one of them tends to go out of limit. Two advantages were observed in this approach: the anticipation itself and the possibility of observing the anomaly through a single index. In this case, both algorithms behaved in a similar way. However, we also observed that a small abnormal change in a single telemetry channel could not be easily perceived by the clustering algorithms. Thus, the issue of feature selection has a great importance, as many times a serious problem manifests just through small variations in one or two telemetry channels.

Another weakness of using clustering algorithms for satellite anomaly detection is that they are sensitive to data changes related to satellite aging. This way, permanent learning is required, which could lead to the learning of anomalies as normal behavior, causing false negative results.

REFERENCES

[1]   V. Chandola, A. Banerjee & V. Kumar, "Anomaly Detection: A Survey," ACM Computing Surveys, 2009.

[2]   A. K. Jain, and R. C. Dubes, Algorithms for Clustering Data, Prentice-Hall Inc, 1988.

[3]   P. N. Tan, M. Steinbach, and V. Kumar, Introduction to Data Mining. Addison-Wesley, 2005.

[4]   A. Webb, Statistical Pattern Recognition, 2nd ed, Jonh Willey & Sons, LTD, 2002, pp.377.

[5]   I. H. Witten and E. Frank, Data Mining: Practical Machine Learning Tools and Techniques, 2nd ed. Elsevier,2005.

[6]   A. P. Dempster, N. M. Laird, and D. B. Rubin. "Maximum likelihood from in-complete data via the EM algorithm," Journal of the Royal Statistical Society:Series B, 39(1):1–38, November 1977.

[7]   K. Hori, N. Ishihama, S. Nakasuka, S. Ogasawara and T. Yairi, "Summarization of Spacecraft Telemetry Data by Extracting Significant Temporal Patterns," The Eighth Pacific-Asia Conference on Knowledge Discovery and Data Mining, 2004.

[8]   K. Machida, R. Fujimak and T.Yairi, "Adaptive Limit-Checking for Spacecraft Using Sequential Prediction Based on Regression Techniques," Journal of the Japan Society for Aeronautical and Space Sciences, Volume 54, Issue 630, 2006, pp. 312-318.

[9]   K. Machida, R. Fujimak, T. Yairi, Y. Kawahara and Y. Sato, "Telemetry-mining: A machine Learning Approach to Anomaly Detection and Fault Diagnosis for Space Systems," 2nd IEEE International Conference on Space Mission Challenges for Information Technology, IEEE, 2006.

[10] S. D. Bay and M. Schwabacher, "Mining Distance-Based Outliers in Near Linear Time with Randomization and a Simple Pruning Rule," Proceedings of The Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2003.

[11] D. L. Iverson, "Data Mining Applications for Space Mission Operations System Health Monitoring," in Proc. of the SpaceOps Conference, 2008.

[12] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I.H. Witten, "The WEKA data mining software: an update," SIGKDD Explor. Newsl, vol. 11, 2009, pp. 10-18.

[13] M. Tafazoli, "A study of on-orbit spacecraft failures," Acta Astronautica 64, 2009. p. 195–205.

[14] B. Robertson and E. Stokening, "Satellite GN&C Anomaly Trends," in AAS Guidance and Control Conference, 2003.