New Telemetry Monitoring Paradigm with Novelty Detection

José-Antonio Martínez-Heras[1]
Black Hat S.L., Córdoba, 14002, Spain

Alessandro Donati[2]
European Space Agency, Darmstadt, 64293, Germany

Marcus G. F. Kirsch[3]
European Space Agency, Darmstadt, 64293, Germany

Frederic Schmidt[4]
Vega Space GmbH, Darmstadt, 64293, Germany

Currently, automatic telemetry monitoring is performed by Out-of-Limits (OOL) alarms. This approach consists of defining an upper and lower threshold so that when a measurement goes above the upper limit or below the lower one, an alarm is triggered. We discuss the limitations of the Out-Of-Limits approach and propose a new monitoring paradigm based on Novelty Detection. The proposed monitoring approach can detect novel behaviors, which are often signatures of anomalies, very early – allowing engineers in some cases to react before the anomaly develops. A prototype implementing this monitoring approach has been implemented and applied to several ESA missions. The operational assessment from the XMM-Newton operations team is presented.

I. Introduction

The most widely extended approach for automatically detecting anomalous behaviour in Space Operations is the use of Out-Of-Limits (OOL). The OOL approach consists of defining an upper and lower threshold so that when a measurement goes above the upper limit or below the lower one, an alarm is triggered. Then engineers will inspect the parameter that is out of limits and determine if it is an anomaly or not and decide which action to take (e.g. run a procedure). This is the original Out-Of-Limits concept. The current OOL concept has evolved to cope with more situations:

- Distinguish between soft and hard limits: a soft OOL triggers a warning to pay attention, a hard OOL triggers an error that demand attention. Soft limits are contained within hard limits.
- OOL thresholds (soft and hard) can be configured so that different thresholds are applicable in different situations (e.g. depending on the working mode of a given instrument).

While Out-Of-Limits are useful and they successfully trigger alarms when parameter readings go out the defined thresholds, they suffer from the following limitations:

1. Some behaviours are anomalous even if they are within the defined limits.
    - A typical example is shown in Figure 1. The parameter values reach the upper limit but do not hit it. Since engineers don't know when this will happen they have to closely monitor key telemetry parameters even if in most cases everything would be nominal.
    - Paradoxically, sometimes the anomalous behaviour is more in limits than the nominal one. Figure 5 shows an example of this situation. More information about this anomaly can be found in section VI.

[1] Software Engineer / Researcher, HSO-OSA, Robert Bosch Strasse 5, Darmstadt Germany.
[2] Advanced Mission Concept & Technology Manager, HSO-OSA, Robert Bosch Strasse 5, Darmstadt Germany.
[3] Head XMM-Newton Spacecraft Operations Unit, HSO-OAX, Robert Bosch Strasse 5, Darmstadt Germany.
[4] Spacecraft Operations Engineer, HSO-OAX, Robert Bosch Strasse 5, Darmstadt Germany

2. OOL are not defined for every parameter. Engineers only define OOL for a subset of parameters for which they want to receive alarms if they exceed the limits. Therefore, OOL is not systematic in the sense that it does not cover every parameter.
3. Quite often engineers receive OOL alarms that are completely expected. A typical example is the OOL defined for the Automatic Control Gain (AGC) during a pass. At Acquisition of Signal (AOS) and Loss of Signal (LOS) the AGC goes outside limits. However, it is expected to happen and in every pass these two OOL alarms will be raised. Ideally, engineers should only have to investigate real potential anomalies.
4. It requires effort to adapt OOL to useful values as the mission goes through different phases or simply degrades with time.

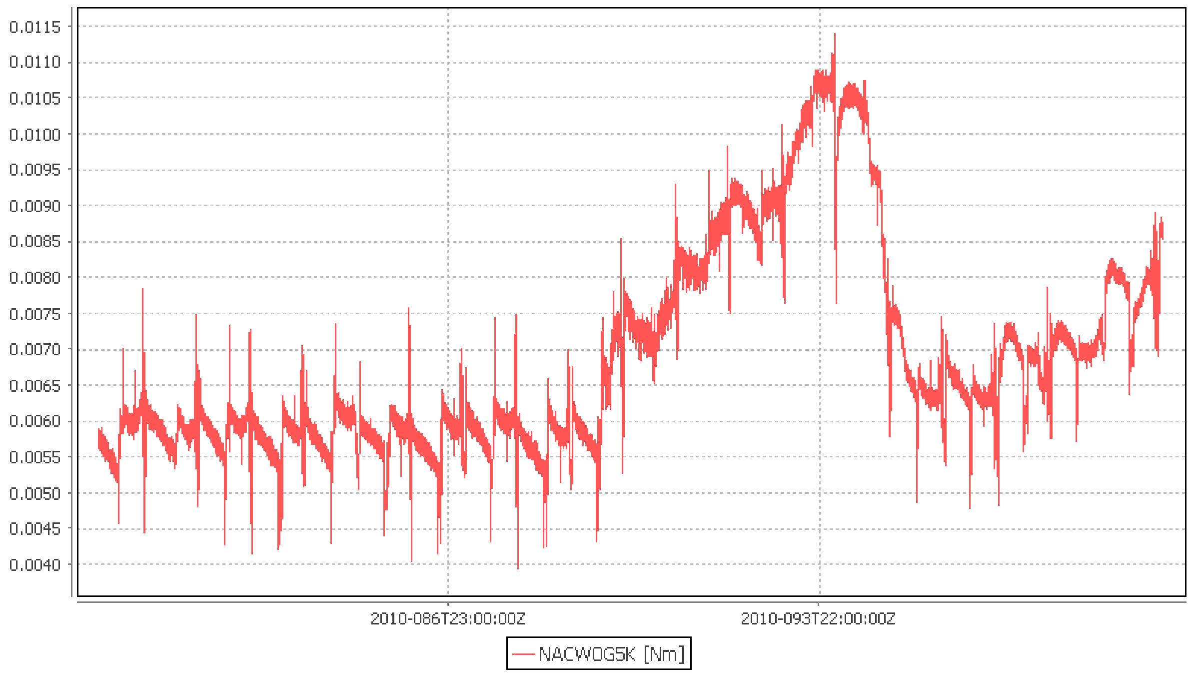

Figure 1. Venus Express Reaction Wheel 4 Friction increases starting in Day of Year 2010.89. *This behavior was not detected by Out-Of-Limits alerts as the upper limit was set to 0.02 at that time. In this case, this behavior was recognized by Venus Express engineers as they closely monitor the reaction wheels even if they do not trigger any Out-Of-Limits.*

The Novelty Detector project has been developed to cope with the current OOL limitations. The Novelty Detector main goal is to automatically detect anomalies and report them to engineers for further investigation. The ideal Novelty Detector should have these features:

- It should take into account that parameters can behave nominally in several different ways.
- It should not make any assumption on what kind of behaviour or how many different behaviours a parameter will have. This have two advantages:
    - It will work with any parameter.
    - No prior knowledge is required.

II. Approaches to automatic anomaly identification

There are mainly three approaches to identify novel behaviour (possibly anomalies):

- *Supervised*: this approach consists of having labelled examples of both nominal and anomalous behaviour. This approach works quite well if we need to recognize previously labelled nominal and anomalous behaviour. Its major drawback is that it can only identify anomalies occurrences for the anomaly types that it knows already. This is a major limitation since the anomalies that generally will impact operations the most are the ones that are happening for the first time. Being able to recognize first time ever anomalies quickly allow Flight Control Teams to take action immediately.

- *Unsupervised*: this approach consists of having unlabelled examples of data - no prior knowledge is provided. The implicit assumption made by the systems using a non-supervised approach is that anomalies happen far less often than nominal behaviours. So they attempt to automatically distinguish what is nominal and what it anomalous. The major drawback is the risk of missing anomalies: if an anomaly happened several times in the past, a non-supervised system may consider it a normal behaviour and do not report it in the future.
- *Semi-supervised*: this approach is a combination of the supervised and unsupervised approaches. It consists of providing only with nominal behaviours examples. This approach has several advantages:
    - Engineers are in full control to specify what should be consider as nominal.
    - Repeated anomalies can be detected since they are not in the nominal set.
    - Since no assumptions are made about the possible behaviour of the anomalies, any anomalous behaviour can be potentially detected.

The proposed monitoring paradigm follows a semi-supervised approach to perform Anomaly Detection. We will use the term Novelty Detection instead of Anomaly Detection since the only thing that can be said is that a behaviour is novel when compared to a set of behaviours known to be nominal. The new behaviour might well be also nominal but so far not present in the nominal set. The decision of classifying a new behaviour as nominal or anomalous is left to the flight control engineers.

III. Novel Behavior Detection

A. Period Representation

In order to characterize behaviours we compute 4 statistical features (average, standard deviation, maximum and minimum) of fixed length periods. The duration of the time period is chosen so that it represents a natural time span (e.g. orbit period or time covered by the short-term planning). The exact duration is not critical; however, it should be long enough to allow behaviours to develop and not so long that many different behaviours happen in it.

While there are other ways to characterize behaviour in a given time period (e.g. Fourier transformations, wavelets, etc.) we used statistical features for a number of reasons:

- They are robust to sampling rate changes, behaviour order and work even if very few samples are available.
- They are compatible with the future ESOC Infrastructure Data Archive (DARC). DARC pre-computes and make available these statistical features.

Figure 2 shows representation of how these fixed time periods look like in a side by side 2 dimensions comparison. We are showing this representation in this document only as an example. In reality, 4 dimensions (average, standard deviation, maximum and minimum) are used simultaneously.

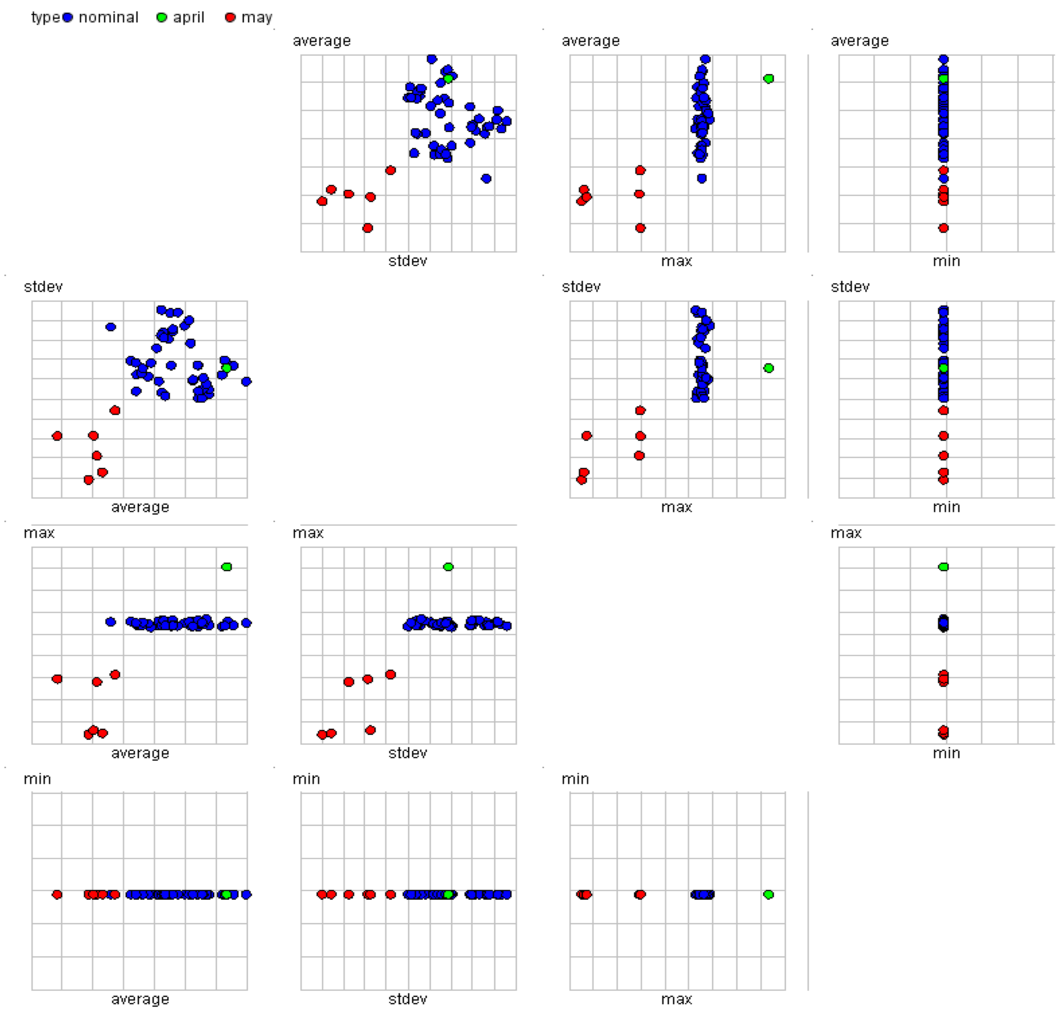

Figure 2. Feature extraction and side by side comparison of the XMM anomaly shown in Error! Reference source not found. *Every point represents a time period; its position in the chart is given by its statistical features. Legend: blue consists of nominal periods [1st January 2009 – 31st March 2009], green represents a spike in April and red represents the thermostat dithering anomaly*[3]*. This 2-dimensions example is only an easy to visualize this representation. In reality, 4 dimensions (average, standard deviation, maximum and minimum)*

B. Periods Distance

Once we have defined the representation of a time period for a given parameter we need to be able to compare time periods. We need a distance measurement so that we can say that for a given parameter A, the period X is closer to the period Y than to the period Z. Mathematically: d(X, Y) < d(X, Z). We use the Euclidean Distance as distance measurement. Mathematically:

$$d(X,Y) = \sqrt{(avg_x - avg_y)^2 + (std_x - std_y)^2 + (\max_x - \max_y)^2 + (\min_x - \min_y)^2}$$

Equation 1. Euclidean Distance between two periods representations.

C. Outlier Detection

We make use of Outlier Detection techniques to find which periods have anomalous behaviour. The general assumption is that anomalous behaviors will have greater distances to known nominal behaviors than known nominal behaviors among them. The question is how big the distance should be so that it can be considered a novel behavior. If the distance is too small many false anomalies will be reported. If the distance is too big then some anomalies will be missed.

The solution to overcome the problem of having to define an outlier distance is to use local density outlier detection techniques. The most widely used is called LOF: Local Outlier Factor[1]. LOF computes a factor that gives an indication of the degree of outlierness (novel behaviour). It takes into account the density of the k closest points. If they are very dense, little distance is required to consider a new behaviour a novelty. If the neighbours are sparse a bigger distance is required to consider a new behavior a novelty.

The major disadvantage of LOF is that the resulting factor values are quotient-values and hard to interpret. A value of 1 or even less indicates a clear inlier, but there is no clear rule for when a point is an outlier. In one data set, a value of 1.1 may already be an outlier; in another dataset and parameterization (with strong local fluctuations) a value of 2 could still be an inlier. These differences can also occur within a dataset due to the locality of the method[4].

To overcome the LOF limitations we will use LoOP: Local Outlier Probabilities[2]. It is a relatively new method derived from LOF. It uses inexpensive local statistics to become less sensitive to the choice of the parameter *k*. In addition, the resulting values are scaled to a value range of [0:1][2] that can be directly interpreted as the probability of being a new behavior. Figure 3 shows an example using 2 dimensions (e.g. average and standard deviation). LoOP has the advantage of being more robust and providing a much more intuitive output. By using LoOP we can rank novel behaviors by novelty probability showing first the behaviors with higher chances to be truly novel.

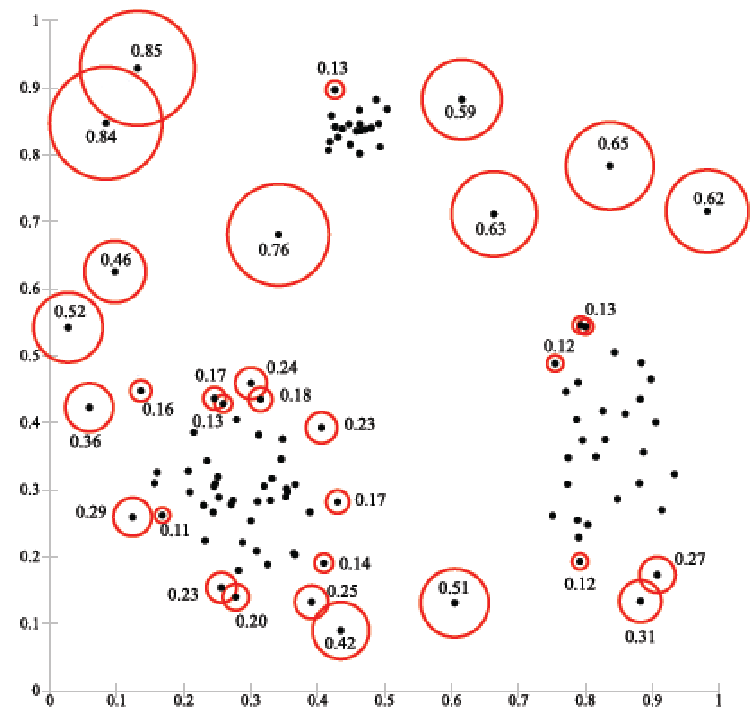

Figure 3. Local Outlier Probabilities outlier detection technique graphical example with 2 dimensions. *Points with high outlier probability indicate a potential novel behavior.*

IV. Novel Behavior Detection - Summary

We use LoOP[2] (Local Outlier Probabilities) to find if a new behavior is novel with respect to a set of given behaviors. If the set of given behaviour consists of nominal behaviours only and LoOP finds that the new behaviour is an outlier with high probability it can only mean two things:

- The new behaviour is an anomaly
- The new behaviour is nominal in a new way (e.g. it behaves differently than other nominal behaviours in the nominal set).

The more nominal time periods the nominal set has, the more the chances that the reported novelties are really anomalies.

Behaviours are characterized by simple statistical features (average, standard deviation, maximum and minimum) in a fixed size time period (e.g. duration of an orbit). This means performing LoOP in 4 dimensions. Euclidean distance is used to measure how different two time periods are. The same procedure is applied to all parameters and those with high probability of being novelties are notified to engineers.

In order to be mode independent of the choice of the parameter k (number of closest points to compute density) we try several k = {10, 20, 30} and use as outlier probability the minimum value of these different runs. The assumption is that if new behaviour is really an outlier it should be an outlier independently of the number of closest points (k) used to compute the local density. This procedure minimizes the chances of getting false alarms.

Limitations: this approach does not consider novelties in the combination of two or more parameters; it works in parameter by parameter basis only. It is, however, systematic in the sense that it can be applied to every parameter.

V. Prototype

If flight control engineers would be able to look every day at every parameter they would be able to identify all novelties. Unfortunately, they cannot. There are way too many parameters (in the order of several thousands) and the trend is that this number will increase in future missions. The objective of this prototype is to automate the process of noticing novel behaviour at parameter behavior level.

New behaviors are often signatures of anomalies either happening now or in the way to develop. Noticing them early is of uttermost importance for planning corrective measurements and keeping the spacecraft healthy. We should take into account that not every new behavior correspond to an anomaly: it could be related to a new environmental condition (e.g. extremely high radiation) or be totally expected as result of planned new operations (e.g. Venus orbit insertion).

The functionality of being able to automatically detect anomalies has been the driver for this project. However, we understood that we could not build such system. The closest we can get is identifying a new behaviour as novel when compared to a set of known behaviours. Hence the name Novelty Detection.

In order to get closer to our goal of being able to automatically detect anomalies, we choose the known behaviour set so that it only contains nominal behaviours. With this configuration, when a new behaviour is classified as novel it can only mean two things:

- It is an anomaly
- It is a new nominal behaviour

As the time goes, the set of known nominal behaviours will grow. This have a positive impact in reducing the number of Novelty alerts as many behaviours will be classified as nominal.

The Novelty Detection prototype makes uses of MUST[5,6] (Mission Utilities & Support Tools) as housekeeping telemetry & ancillary data provider. The MUST's performance allows to perform advanced monitoring with novelty detection efficiently.

A. Functionalities

This section discusses the two major functionalities of the Novelty Detection prototype. The underlying principle is the same but one can achieve one functionality or the other depending on which configuration is used.

i. Identification of potential anomalies

The main purpose of the Novelty Detection prototype is to detect potential anomalies. For fulfilling this objective we will use as known periods set the collection of all known nominal behaviours. This way, when a new behaviour is classified as novel with a high probability, it is very likely that it would be an anomaly. It could be still a new kind of nominal behavior but, as the time goes, this should happen less and less frequently.

ii. Verification of expected new behavior

In addition of potential anomalies identification, the same Novelty Detection technique can be used to verify the occurrence of an expected new behavior. For instance, let say that certain behaviour is expected as consequence of a manoeuvre and we would like to verify it. A way of doing it with the Novelty Detection prototype is by using the recent past as the known behaviour set. The output of the Novelty Detection will be the novel behaviours compared with the recent past. These novelties should contain the expected new behaviours and possibly other parameters. These parameters that were not initially foreseen can be considered side effects of the expected behaviours.

B. Input

Two inputs are required to run the Novelty Detection prototype:

1. *Periodicity*: the statistical features needed to characterize a fixed length time period has to be computed over a large enough time period. A typical example is to use the periodicity of the orbit or the amount of time that the short term planning covers.
2. *Set of known behaviours*: the Novelty Detection will detect if a new behaviour is novel as compared with the set on known behaviours specified as input. Two options are recommended:
   a. Use all nominal behaviours: this is ideal to perform Anomaly Detection.
   b. Use the recent past: this should be used to verify expected new behaviour.

C. Output

The output consists of text file that contains the list of parameters that are believed to be novel. They are grouped by period and by novelty probability within periods. Figure 4 shows and example of such output file. The same information is stored in a database for further analysis by client applications.

```
0.935057, no-Gaps, 2010-10-14T00:00:00Z [2010.287.00.00.00], NTTX1000, Loop Error Term. 1
0.817176, no-Gaps, 2010-10-14T00:00:00Z [2010.287.00.00.00], NACW0G15, RW4_SWR estimat friction
0.746310, no-Gaps, 2010-10-14T00:00:00Z [2010.287.00.00.00], NACW0G5K, RW4_SWR estim abs frict
0.738150, no-Gaps, 2010-10-14T00:00:00Z [2010.287.00.00.00], NACW0R08, AUT_GUID Cmd Quater Vy
0.728106, no-Gaps, 2010-10-14T00:00:00Z [2010.287.00.00.00], ses_pred, predicted SES-Angle
0.712610, no-Gaps, 2010-10-14T00:00:00Z [2010.287.00.00.00], NPWG0060, Incoming Solar Flux
0.712455, no-Gaps, 2010-10-14T00:00:00Z [2010.287.00.00.00], NACW0S01, EPH_MGR Sun-SC dist
0.701722, no-Gaps, 2010-10-14T00:00:00Z [2010.287.00.00.00], NAAD0331, ACM1 PT1 press (He-High)
0.935141, no-Gaps, 2010-10-15T00:00:00Z [2010.288.00.00.00], NTTX1000, Loop Error Term. 1
0.804666, no-Gaps, 2010-10-15T00:00:00Z [2010.288.00.00.00], NACW0R08, AUT_GUID Cmd Quater Vy
0.782812, no-Gaps, 2010-10-15T00:00:00Z [2010.288.00.00.00], ses_pred, predicted SES-Angle
0.771778, no-Gaps, 2010-10-15T00:00:00Z [2010.288.00.00.00], NACW0R09, AUT_GUID Cmd Quater Vz
0.766928, no-Gaps, 2010-10-15T00:00:00Z [2010.288.00.00.00], NPWG0060, Incoming Solar Flux
0.766873, no-Gaps, 2010-10-15T00:00:00Z [2010.288.00.00.00], NACW0S01, EPH_MGR Sun-SC dist
0.753964, hasGaps, 2010-10-15T00:00:00Z [2010.288.00.00.00], NVVD103E, Spare
0.742360, no-Gaps, 2010-10-15T00:00:00Z [2010.288.00.00.00], sse_pred, predicted SSE-Angle
0.742312, no-Gaps, 2010-10-15T00:00:00Z [2010.288.00.00.00], NAWG0052, Sun-S/C-Earth Angle
0.742307, no-Gaps, 2010-10-15T00:00:00Z [2010.288.00.00.00], NAWG0053, Sun-Earth-S/C Angle
```

Figure 4. Novelties found for Venus Express using the Novelty Detection prototype for verifying expected new behaviour. *The format of the file output is: probability of being an outlier, did this parameter have data gaps during this period?, start time of the period (in two time formats), parameter mnemonic, parameter description.*

VI. Operational Validation & Current Usage

In order to proof the feasibility of the monitoring paradigm with novelty detection we applied it to an already documented anomaly that the ESA satellite XMM-Newton experienced in 2009 and checked when the Novelty Detection prototype would have been able to detect the anomaly.

Here is an excerpt of the paper that describes the anomaly and measures taken by the FCT (flight control team) to cope with it[3]: “We noticed that the thermostat T6073 started to have a strange behavior since mid-May 2009, already 2 months before the failure was spotted. The thermostat range reduced, the temperature where it opened started to decrease, sign of a deterioration of the high threshold, even if the bottom limit was respected quite well, until mid of July, when the upper limit and the lower limit went very close to each other. The thermostat started to oscillate, in a narrow temperature range, until it did not close anymore at the correct temperature, and it let the temperature go down to almost 22 deg. This first temperature drop was not spotted because it did not generate any OOL. After that the thermostat had some cycles with a nominal behavior, but on the 13 July 2009 the temperature went deeper down to 21.25, triggering an OOL and allowing the FCT to spot the problem."

We configured the novelty detection prototype to consider data in the range [January 2009, March 2009] as nominal. We used as time period 48 hours since it is the duration of an XMM-Newton orbit. Then we run the novelty detection prototype for the period [April 2009, July 2009]. The results is that the novelty detection prototype managed to find unusual behavior 2 months before the Out-of-Limit triggered. This is remarkable not only because it allows to react to anomalies early, but also because it matches flight control engineers diagnosis results and mimics the effect of having somebody looking every day at every parameter and noticing if something new is happening. Figure 5 shows where the OOL triggered and where the novel behavior was found.

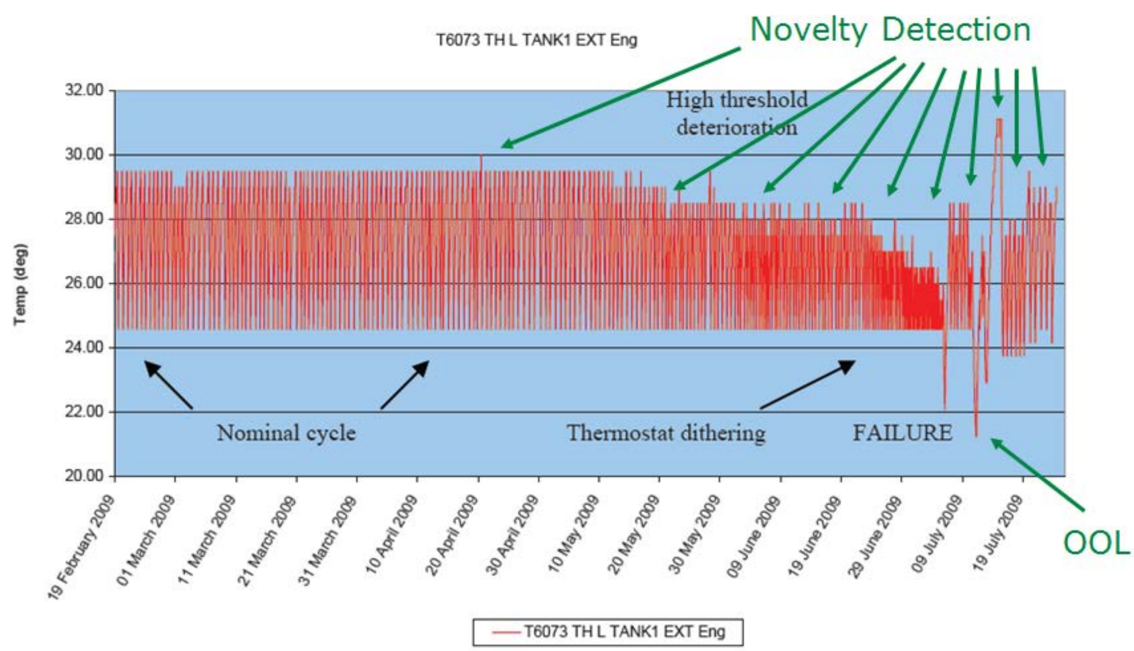

Figure 5. Monitoring with Novelty Detection a known XMM-Newton anomaly. *Nominal behavior: February – April, OOL on 13th July. This thermostat has been properly working showing the same behavior for 10 years. However, it started to have a strange behavior since mid-May 2009 and it was only noticed 2 months after (July 2009) when it crossed the lower limit. For this type of anomaly, the Out-Of-Limits checks are not effective because, paradoxically, the behavior of the anomaly was "more in limits" than before. The proposed novelty detection monitoring technique could find this anomaly 2 months before the Out-of-Limit alarm triggered.*

Currently, the Novelty Detection prototype checks every day around 2000 XMM-Newton housekeeping telemetry parameters and reports which of them, if any, has a new behavior. The results are sorted by probability of certainty of been a new behavior. The novelties reported so far [March 2012 – April 2012] correspond to expected new behaviors as confirmed by the XMM flight control team.

The Novelty Detection for XMM is integrated in a wider scope project, XEWS[7] (XMM Early Warning System). XEWS is developed to perform near-real-time trend analysis of spacecraft parameters in order to detect early degradation of components. XEWS will enable the mission to perform early counter measures in case degradation is detected.

VII. Conclusion

We have introduced a new monitoring paradigm based on novelty detection. In this approach, every day every telemetry parameter is scanned and a list of the parameters that exhibit a novel behavior is reported to flight control engineers. New behaviors are often signatures of anomalies either happening now or in the way to develop. Noticing them early is of uttermost importance for planning corrective measurements and keeping the spacecraft healthy.

The major benefits of the proposed monitoring paradigm are the following:

- *It requires little engineering effort*: the only inputs required consists of the period duration (e.g. orbital period) and ranges of times to be used as examples of nominal behavior.
- *It generates very few false alarms*: the fact that it uses a local density outlier detection technique avoids the need of using a distance threshold to detect new behaviors. Therefore, this approach does not suffer from the problem of having to define a threshold where having it too small would lead to many false alarms and having the threshold too big will lead to missing new behaviors.
- *Systematic*: it is systematic in the sense that it detects new behaviors in all parameters.
- *Automatic*: it runs unattended; user intervention is only required to investigate the new behaviors found (if any).
- *Generic*: this monitoring approach is not mission specific. We have successfully applied it to XMM and Venus Express. In fact, we can easily adapt it to any ESOC controlled mission since we use MUST[5,6] as data provider.

The proposed Novelty Detection monitoring approach has been successfully validated with XMM anomalies, finding them before they were triggered by the Out-of-Limits alarms, sometimes as early as 2 months in advanced. Currently, the XMM mission uses the novelty detection prototype to detect new behaviors on about 2000 parameters on daily basis.

We believe that every mission will benefit from the adoption of the novelty detection monitoring paradigm as complement to the classic Out-Of-Limits mechanism. Being able to know which few parameters (out of several thousands) exhibit a new behavior helps flight control engineers to efficiently direct their monitoring efforts.

The ESA's patents group has decided to protect the proposed monitoring paradigm by filing an international patent.

References

[1]Breunig, M. M.; Kriegel, H. -P.; Ng, R. T.; Sander, J. (2000). "LOF: Identifying Density-based Local Outliers". *ACM SIGMOD Record* 29: 93.

[2]Kriegel, H. -P.; Kröger, P.; Schubert, E.; Zimek, A. (2009). "LoOP: Local Outlier Probabilities". *Proc. 18th ACM Conference on Information and Knowledge Management (CIKM)*: 1649.

[3]Pantaleoni, M.; Kirsch, M.; Martin, J.; Weissmann, U.; Krusenstiern, N. "The new operational strategy for XMM-Newton's tank thermal control, after Low Temperature Protection thermostat failure". In proceedings of the SpaceOps 2010 Conference, Huntsville, Alabama, April 25-30.

[4]Wikipedia contributors. "Local Outlier Factor." *Wikipedia, The Free Encyclopedia*. Wikipedia, The Free Encyclopedia, 5 Oct. 2010. Web. 26 Oct. 2010.

[5]Martínez-Heras, J.; Baumgartner, A.; Donati, A.; "MUST: Mission Utility and Support Tools", Proceedings DASIA 2005 conference, Edinburgh, UK.

[6]Baumgartner, A.; Martínez-Heras, J.; Donati, A.; Quintana, M.; "MUST – A Platform for Introducing Innovative Technologies in Operations", Proceedings of the ISAIRAS 2005 conference, in Munich, Germany.

[7]Kirsch, M.; Martin, J.; Pantaleoni, M.; Southworth, R.; Schmidt, F.; Webert, D.; Weissmann, U.; "Cage instability of XMM-Newton's reaction wheels discovered during the development of an Early Degradation Warning System". Proceedings SpaceOps 2012 Conference, Stockholm, Sweden, June 11-15 (to be published).