SMART TELEMETRY DEMONSTRATOR

Marc Bernot [(1)], Bernard Alison [(2)]

[(1)] *Thales Alenia Space, 100Bvd du Midi, BP99, 06156 Cannes La Bocca Cedex, France, Email: Marc.Bernot@thalesaleniaspace.com*

[(2)] *Thales Alenia Space, 100Bvd du Midi, BP99, 06156 Cannes La Bocca Cedex, France, Email: Bernard.Alison@thalesaleniaspace.com*

ABSTRACT

A R&T study has been performed in 2011-2012 by Thales Alenia Space for ESOC (ESA contract n° 4000102889/10/NL/AF) .

The main objective of this study was to investigate methods to better understand the structure of the information contained in a satellite telemetry. Indeed, with the understanding of the data comes the ability to give a sense to what is a significant/unexpected event and to identify the parameters giving rise to such events.

This self-detection of unexpected events together with the focus on implicated parameters is what makes this way to tackle telemetry smart.

To avoid a theoretical study, it has been decided to evaluate these methods on the complete telemetry volume provided during 24 hours by two satellites currently in orbit.

Several algorithms have been evaluated and a method has been devised to structure a significant part of the parameters into a limited number of clusters in which all the parameters are correlated together on a pair basis. This allows for a synthetic description of what is a nominal state of the telemetry and fulfils the main objective.

TELEMETRY BASED ON PACKETS

Up to now, the downloaded telemetry is based on packets. Monitoring parameters are gathered together on a fixed manner to facilitate the de-commutation of the packet by the ground. The layout of the packets on the download stream is fixed and generally dependant of the current mode: it is defined as the TM plan.

The operators have a preference for raw data associated to the sampling date. This approach allows better investigation in case of trouble, the provided values are close from the value delivered by the on-board source with a limited risk of degradation due to the on-board treatment. Few parameters do not correspond directly to the acquired parameters: they are called calculated parameters. For example the On Board Computer calculates the instantaneous power by multiplying the measured current by the measured voltage. This new parameter facilitates the understanding of the spacecraft behaviour by the operators. Even in this case, raw parameters (current and voltage) are also downloaded in parallel.

The sampling frequency of parameter is generally fixed and defined in relation with on-board need. The download frequency could be identical to the sampling frequency but usually is lower (sampling frequency divided by a power of 2) in order to save the TM bandwidth.

On ground demand, it is possible for a predefined sub-set of parameters to increase the downloaded frequency by increasing the sampling frequency if needed. This is called the dwell. It is used during a limited period of time and increases the volume of available data for expertise and investigation by the operators.

The PUS services, especially the service 12, gives to the operators the capability to modify the TM plan in orbit. This capability is facilitated thanks to the standardization of PUS services. The real limitation for the ground is the obligation to store in a Data Base the definition of different TM plans and the date of the change. This is required to allow the de-commutation of old telemetry after retrieval.

OBJECTIVE OF THE STUDY

The study aims at devising what could be called a smart telemetry. This would be a telemetry that either knows or learns what is a nominal state and a non nominal one. Thus, it could self-detect unexpected events and communicate only the few parameters that go wrong so that the operator can quickly narrow down the scope of the problem analysis.

In order to tackle this ambitious goal, we considered with ESOC that the main objective of the study is to better understand the structure of a nominal TM plan through the analysis of downloaded telemetry of two ESA missions. Indeed, this allows to make the information content of the telemetry apparent so that it results in a synthetic description of what is a nominal state vector of the telemetry.

This naturally leads to the other objectives of smart telemetry:

- Detection of significant/unexpected events.
- Identification of the parameters that are the cause of the deviation from nominal.

A FIRST ANALYSIS OF THE TELEMETRY

ESOC has provided all the telemetry data downloaded during a consecutive 24 hour window for two ESA missions: Mars Express (MEX) and GOCE. The current mode was the nominal one and the selected period was without anomaly.

Mars Express

The main objectives of the Mars Express mission are :

- provide high-resolution, 3D images for the study of the Martian surface and geology
- reveal the structure of the subsurface with the aid of a radar able to penetrate the crust
- precisely determine the composition of the Martian atmosphere to give an accurate picture of its climate and meteorology
- study the interactions between the atmosphere and outer space. Collecting all this information will provide a better understanding of our own ecosystem, for example in reference to the spread of deserts which happened on a global scale on Mars and is now taking place in some parts of the Earth.

GOCE

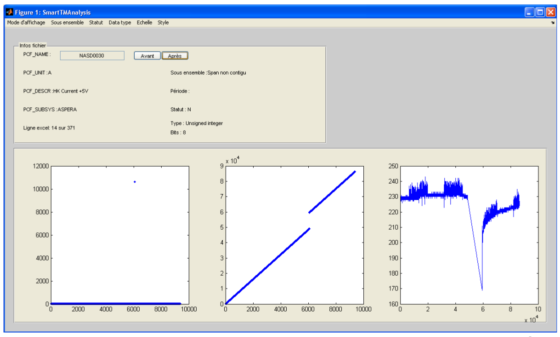
The Gravity Field and Steady-State Ocean Circulation Explorer (GOCE) mission will measure high-accuracy gravity gradients and provide global models of the Earth's gravity field and of the geoid.

The 24 h excerpt of Mars Express telemetry started at T0=Sun, 10 Apr 2011 00:00:01 UTC. The telemetry contains 8101 parameters. More than 6000 of them present a unique value. The analysis was concentrated on the remaining 1972 parameters. Most of the parameters are periodically sampled with 8 second multiple periods (from 8 s to 2048 s)

The 24 h excerpt of GOCE telemetry started at T0= Mon, 14 Mar 2011 00:00:00 GMT. It contains 4620 parameters, including 3230 parameters with a unique value. The parameter frequencies mainly present different values: 0.5, 1, 2, 5, 8,16 and 32 seconds.

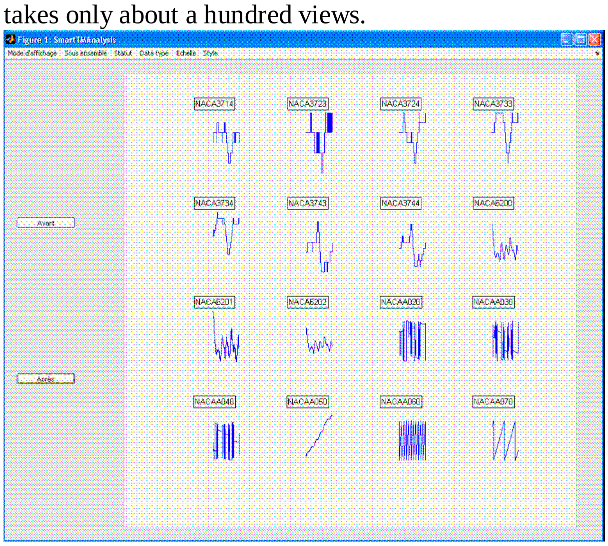
For both missions some parameters present irregular sampling periods, or non contiguous span. A pre-treatment of the parameters have been performed to re-synchronise the different parameters allowing further correlation analysis.

A first classification has been performed on a visual basis, thanks to a tool based on Matlab, specifically developed for this purpose. The tool, called smartTMviewer permits to have a look at each parameter and its properties like unit, subsystem, number of bits, type, etc.



The smartTMviewer permits to view all the useful information for each parameter

It is also possible to look at 16 parameters simultaneously so that going through all parameters takes only about a hundred views.



The smartTMviewer allows to compare several parameters

This first step has shown a large heterogeneity of the telemetry parameters:

- Time sampling could be simple and contiguous, a mix of several intricate sampling periods, non-contiguous span…
- Dynamic behaviour could be stable with/ without spikes/noise or variable into the full dynamic range.

Due to this complexity of the dynamical behaviour of the parameters, it has been decided to focus on the understanding of the static aspect of state vectors of the satellite. In other words, within this study, we analysed the information present in the set of state vectors and not the information given by the way a state vector evolves with time.

MAKING THE STRUCTURE IN THE DATA VISIBLE : ENUMERATED PARAMETERS

The case of enumerated parameters is a good starting point for the analysis since parameters only take a few (most often two) number of values. In the case of GOCE data, they are 67 non constant enumerated parameters and they are all booleans.

Clearly, if a parameter takes a value it has never taken before, this can be considered as a significant event. However, only monitoring each parameter separately is not sufficient. For example there is in the GOCE data some redundant parameters that code for an ON/OFF state. Thus this pair only takes the values (0,1) and (1,0). Surely for this pair to take the value (0,0) or (1,1) would be an anomaly.

This means that it is quite important to group correlated parameters together into subsystems so that a significant event can be defined as a new state of this subsystem. A new state of a subsystem cannot be detected by looking at each parameter in isolation.

In the case of GOCE data, since non constant parameters are all booleans, we defined a pair to be strongly correlated if it takes 2 values, weakly correlated if it takes 3 values and uncorrelated if it takes 4 values.

We then clustered the parameters into subsystems such that all parameters are correlated one another. This led to a decomposition of the parameters into 4 groups of correlated parameters with respective sizes 47,4,4 and 2. An additional group of 8 parameters are uncorrelated and do not need monitoring.

We tested this way of grouping the parameters during the 24h span and these four groups only take 17,3,5 and 3 values. This means that this way to present the information is extremely synthetic while it well approximates the true content of the information at the same time.

Thus, only monitoring the new values of these 4 subsystems (and of all constant parameters) is a good compromise between not missing important events and not raising too much irrelevant events.

MAKING THE STRUCTURE IN THE DATA VISIBLE : NUMERICAL PARAMETERS

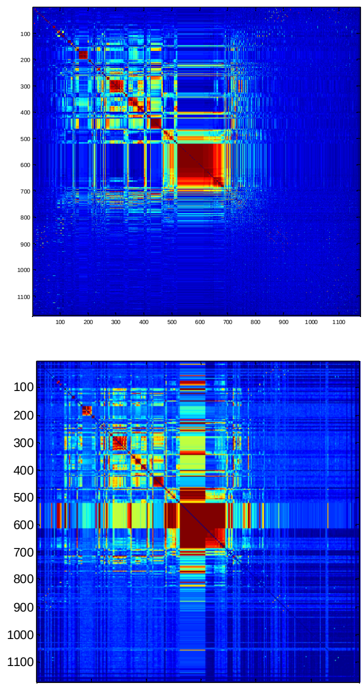
In the case of numerical parameters, the main strategy is based on the same pragmatic approach as for the enumerated by first analyzing pairs of variables by looking at their scatterplot. We want to consider two parameters as being correlated if the scatterplot is concentrated along some structure rather than being spread out quite uniformly.

We considered two recent methods for carrying out such correlation analysis :

- MIC (maximum Information Coefficient) method (2011)
- Dcorr method (2009)

These methods have been applied to find correlation on the 1174 numerical parameters of GOCE mission. Both methods were found to be complementary. They don't always give the same results but the use of both methods allow to identify pertinent correlations. In particular, MIC gives rise to a lot of false positives and is not as good as dcorr for linear correlation. However MIC can spot very structured scatterplots (e.g. absolute value relation between two parameters) for which dcorr does not give a high score.

The first step is to calculate correlation of all the pairs of parameters by both methods. This analysis shows that many parameters are correlated but a sorting of these parameters is needed to ease the reading of the results. The second step consists to apply a sorting / clustering algorithm.

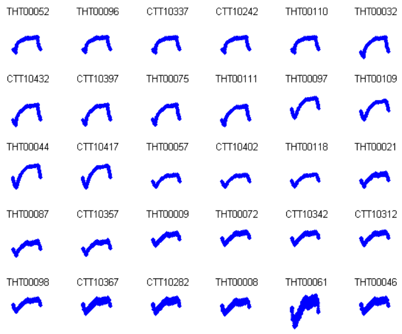


DCOR and MIC correlation matrices after reordering (red/blue=correlated/uncorrelated)

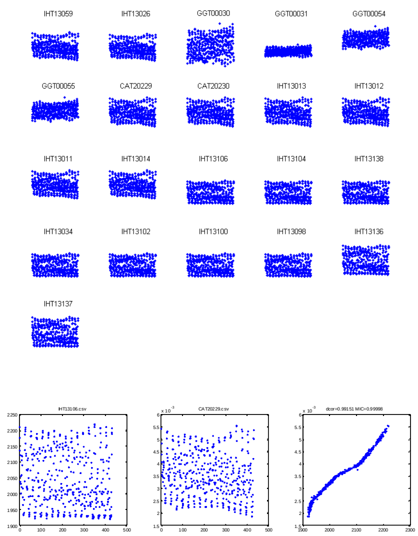
The above figures show that roughly 400 parameters are completely uncorrelated to the others whereas the remaining 700 parameters clearly show some structure. These 700 parameters can be grouped into different clusters such that all the parameters are correlated together on a pair basis (or more indirectly through a chain of correlation within the cluster) .

To have a clearer view of the correlation matrix, we applied some threshold that results in a binary matrix. With a 0.9 threshold, we were able to partition 444 parameters that show some correlation into 62 clusters (from 2 to 120 parameters per cluster).
Within these 62 clusters, there are 27 clusters of more than 3 inter-correlated parameters (the other 35 clusters are only clustering of pairs or triplets of parameters).

A visual inspection of the parameters contained in the clusters shows that the clustering algorithm was able to group together obviously (that you can spot by looking at the two parameters plots) or less obviously (that you can only spot by looking at the scatterplot) correlated parameters.



A cluster of 30 parameters obviously correlated



An example of cluster of non obviously correlated parameters and the scatterplot of a non linear dependence between two parameters

Finally, it remains to define a notion of significant event. Clearly, if two parameters show a very structure scatterplot such as on the above figure, it would be a significant event to depart from this scatterplot.
Such a structured scatterplot can be encoded by the way of numerical constraints that two parameters of a nominal state have to verify. More generally, it is the set of nominal values of a whole state vector of a clustered subsystem that can be efficiently described with such numerical constraints.

CONCLUSION AND PERSPECTIVE

We presented in this document an approach that gives a first answer to the three aims of the study :

- make the information apparent
- detection of the significant events
- identification of the culprit parameters that give rise to a significant event.

In this study we proposed a practical and successful approach that can be developed in many different directions (use more statistics in the enumerated case, use higher dimensional scatterplots in the case of numerical parameters to name a few).
The fact that this approach is promising also comes from the fact that it was obtained through a deep inspection of the provided data and not from a preconception of what could be done on a theoretical basis.
The principal limitation that made this study feasible is that we focused only on the static analysis. Thus the only information we used and made apparent does not depend on the time parameter. It is certain that many phenomena are of a dynamical nature in the behavior of these parameters and that a dynamical analysis would yield other types of information.

REFERENCE

MIC (maximum information coefficient) :
Detecting Novel Associations in Large Data, SetsReshef et al., Science (16 December 2011)
Dcorr measurement :
Brownian distance covariance,
Székely et al. Ann. Appl. Stat. (2009).