International Journal of Applied Engineering Research ISSN 0973-4562 Volume 13, Number 11 (2018) pp. 10124-10128

A Novel Application of Rule Induction Algorithm for Parametric interpretation of Satellite On-orbit performance

S.A.Kannan[#], T.Devi[*]

[#]*UR Rao Satellite Centre, Bangalore, India.*
[*]*Head, Dept of Computer Applications, Bharathiar University, Coimbatore, India.*

Abstract

This paper presents a novel method of applying Rule Induction mining algorithm for analysing the Telemetry Data received from a satellite in orbit. Health of a satellite in orbit is very important for it to deliver its intended space-based services. It is monitored by a set of house-keeping parameters which is received as telemetry by Satellite Ground Control Centre. This data is received continuously throughout the mission life of the satellite. Hence a huge repository of data is accumulated and archived. Analysis of this data by conventional means is tedious and time consuming. In addition, manual analysis can give information about related parameters. Data Mining techniques can be applied to explore the performance relationships by looking for unknown relationships between apparently unrelated telemetry parameters. The parameters of interest are collected in a multidimensional database. From the inputs provided, Rule-Induction Algorithm generates a set of rules which are found to be meaningful and giving new knowledge. The results obtained from the implemented algorithm show relationships between related and unrelated parameters that can lead to emergence of new knowledge consistent with satellite design philosophy.

Keywords: Data Mining, Algorithm, Rule Induction, Satellite, Telemetry Parameter

INTRODUCTION

An artificial Satellite is an integrated system consisting of different subsystems (like Power, Telemetry & Telecommand, Attitude and Orbit Control, Sensors, Propulsion etc.) designed and developed to perform together for meeting the specific Mission requirements. After the satellite is injected in orbit each subsystem health and performance is monitored through a given set of parameters received from the satellite in the form of Telemetry data. This paper presents the results of applying Rule Induction Algorithm on a data set consisting of various parameters in the Telemetry data for deducing unknown patterns among them.

The health and performance of Satellites in orbit are monitored / evaluated by analysing Health-Keeping parameters. These parameters are collected by the Telemetry subsystem from all subsystems and beamed to the Ground station where online and offline analysis are carried out to determine the state of the spacecraft.

Archived spacecraft telemetry data contain a wealth of information about complex system behaviour. Recent developments in data mining techniques for anomaly detection make it possible to examine this archived data and extract embedded information to produce advanced system health monitoring applications. Such applications can aid mission controllers and engineering analysts in their task of ensuring that spacecraft systems under their watch are operating properly. In contrast to common individual parameter monitoring schemes, these “data driven” applications are capable of characterising and monitoring interactions between multiple spacecraft parameters and can provide additional insight and valuable decision support for controllers and engineers [2].

SATELLITE TELEMETRY DATA ANALYSIS USING DATA MINING TECHNIQUES AND TOOLS

Data Mining is the process of discovering interesting knowledge from large amounts of data stores in databases either from live or archived data.

Many researchers have used Spacecraft Telemetry Data to analyse the health of the Spacecraft and to predict its future performance. Data Mining techniques and tools have been regularly applied for analysing Satellite Telemetry data in order to find anomalous behaviour or predicting potential failures of subsystem elements. Commercial off-the-shelf tools like Orca and Inductive Monitoring System (IMS) have been applied to Shuttle and Space Station Telemetry data [2]. Techniques like Clustering have been applied to Satellite Data [3]. Data Mining Regression Algorithms have been used to adapt Telemetry measurements range in an automatic way [4].

Lance Self [1] predicted the use of Telemetry data to extract previously unknown relationships between various Telemetry parameters of different subsystems of the Spacecraft. In this work, he has mentioned the work done by contractors to Air Force Research Laboratory. A product called Model Quest Satellite Telemetry Anomaly Resolution (MQ STAR) developed by one of them used a mixture of statistical modeling approach and case based reasoning. Another product called ‘Mining Agent for Spacecraft Telemetry (MAST) took in learning parameter rules along with Telemetry Data. These rules were passed on to an interface agent which present these rules (candidates) to the user. Validated rules are then implemented by an interface engine. Events are stored in an event database, and validated rules are stored into a rules database. It is stated that each company has different approach to using Data Mining techniques for analyzing Satellite Telemetry Data.

David L.Iverson [2] has used a set of standard software tools to analyse the Telemetry data for mission operations and anomaly detection. Two data driven software tools namely Orca and Inductive Monitoring System (IMS) have been used for mission operations of Space Shuttle and International Space Station. Of them Orca uses a nearest-neighbour approach to search for unusual data points in multivariate data sets by calculating the distance of each data point from neighbouring points. The IMS tool uses the data mining technique Clustering to analyse archived spacecraft data and characterize nominal interactions between selected parameters. This characterization, or model, is compared with real time or archived system data to detect off nominal behaviour.

Mayu Sakurada et al. [3] have used a different approach for analyzing the TM data from Satellites. This work illustrates the possibility of integrating human and knowledge and artificial intelligence. Machine learning systems can handle big high dimensional data to extract some patterns or rules from the data. On the other hand, humans have domain knowledge, semantic rules or common sense relevant to the data that they have learnt from experience. In this work two models were built independently and later combined to deduce the classification status of satellite. The first model is built by humans based on human reasoning and the other by the system in a data-driven way. In order to combine the two models, the predictions of these two models are combined. The scores produced by multiplying both confidences are used as prediction of the combined model. Here both the models are made semantic so that rules are to be accessible both for humans and computers. Decision trees have been used with categorical features. Decision trees learn the rule sets which are understandable for humans. Categorical features indicate some status and are more informative than continuous features. First the Human-Drive Tree is constructed using the human selected features which are judged to be relevant to determine the class based on human knowledge, as well as the ordinary Data-driven Tree which is created using all the features.The focus of this work is on combining expert's knowledge and data – driven models.

Yu Gao et al., [4,5] presented a new unsupervised anomaly detection approach for spacecraft based on normal behaviour clustering. This is a straight forward approach wherein the Telemetry Data as received from the Spacecraft is sifted for abnormal values and the residual values are classified into clusters with closely similar features. This method takes as input a set of unlabeled historical telemetry data and automatically detects anomalies within the data. After these anomalous values are removed, the method constructs system normal behaviour model based on normal data clustering. Then at run-time it monitors the status of the spacecraft and detects any anomalies appearing in the real-time data by checking deviations from the normal behaviour model.

Tianshe Yang et al [6] have studied the Fault Detection and Prediction methods for in-orbit Satellites by Data Mining Technology. In this work a Deep Mining method of Satellite Telemetry is proposed. Parameter relationship at various instances has been mentioned as one of the features for Fault Detection and Prediction. The time interval of observation for a given parameter is studied with different methods.

APPYING RULE-INDUCTION METHOD FOR SATELLITE TELEMETRY DATA

A data mining algorithm is a set of heuristics and calculations that creates a data mining model from data. To create a model, the algorithm first analyzes the data provided, looking for specific types of patterns or trends. The algorithm then uses the results of this analysis to define the optimal parameters for creating the mining model. These parameters are then applied across the entire data set to extract hitherto unknown patterns..

The aim of the study has been to find the hidden relationships among the chosen set of telemetry parameters. Data mining algorithm is used in this exercise since conventional statistical analysis can provide meaningful relation only among inter related parameters that are within a defined range and where the database is dimensionally limited.

The data mining process for knowledge extraction from input data usually consists of a number of processes starting from Data preparation. In our present study, Data analysis through mining is treated as the major part.

An important step of the data mining process is to verify that the patterns produced on the data analysis stage occur in the wider dataset. For this purpose, the result validation process uses a test set of data on which the data mining algorithm was not trained. The learned patterns are applied to this test set, and the result is compared to the desired output.

The resulting "knowledge" extracted during the data mining process includes previously unknown patterns such as clusters of data samples, anomalous data structures, or dependencies between data features. These patterns form a model that describes the structure of the analyzed data.

In the present study emphasis is placed on finding the suitability of Rule Induction technique as compared to other Mining techniques. Hence test data and results validation have not been treated in a rigorous fashion.

Rule induction is one of the methods most used to extract knowledge from data. Although there are many other methods, like instanced based learning (e.g. k-nearest neighbour), statistical techniques (e.g. Naïve-Bayes classifier), neural networks and support vector machines, they do not normally return to the user, as the output of the system, comprehensible knowledge. Therefore, rule induction algorithm can be used for the present application because the representation of knowledge as if/then rules is very intuitive and easily understandable.

International Journal of Applied Engineering Research ISSN 0973-4562 Volume 13, Number 11 (2018) pp. 10124-10128

HIDDEN RELATIONS AMONG PARAMETERS IN SATELLITE TELEMETRY DATA

A. *Defining the Problem*

For geo-synchronous and deep-space missions the LAM engine in the satellite is fired as per requirement to inject into its intended orbit or to increase / increase its velocity. At the time of engine firing the values of different H-K parameters are likely to change. Mining the Telemetry data during such events can yield a wealth of information regarding the intended and/or unintended relationships between various parameters.

In our present study, a sample set of onboard telemetry parameters acquired in ground have been used as the data test bed. The parameters and their typical values are obtained from the onboard activity phases that have happened during LAM (Liquid Apogee Motor) firing.

The aim of the study has been to find the hidden relationships among the chosen set of telemetry parameters. Data mining algorithm is used in this activity since conventional statistical analysis is confined to providing meaningful relation only among inter related parameters that are within a defined range and are part of an essentially limited-dimensional database.

B. *Methodology*

Let A = {a,b,c,…x} be the dataset containing Telemetry parameters a to x with each of their values taken at any given instant of time during major, specific events of the on-orbit operations of the satellite. The objective is to find unknown and potential relationships between these parameters.

Here a to x can be any parameter like current, voltage, temperature etc. of a given subsystem. The exercise does not preclude any other similar parameter of any other subsystem. Parametric values are considered within the same event 'e' at different instances of time t1,t2,t3…tn where t1<t2<t3…<tn.. (a11,b11,c11 ….x11) refers to values for event e1 at time t1. The choice of event 'e' is made such that it is a major event during the on-orbit life of the satellite and where all major subsystems are in a state of dynamic performance, like during Engine firing.

If Ae1 corresponds to dataset during a given event e1, then

$$\text{Ae1} = \left| \begin{array}{l} \text{a11 b11 c11 …... x11} \\ \text{a12 b12 c12 …... x12} \\ \text{…………………......} \\ \text{a1n b1n c1n …...x1n} \end{array} \right| \qquad (1)$$

In order to find out the relationship of a1 with other items in Ae1, first its value at t1 is self -compared with its values at t2, t3, t4 …tn.

a11~a12~a13…~a1n gives the nature of a1's value changes within the duration of event e1.

Let *ae1* = | a11 ~ a1n | for all 'n' be denoted by an arbitrarily chosen value 0,1,2 such that

| a11 ~ a12 | = 0 if there is no value change

= 1 if there is an increase in value

= 2 if there is a decrease in value

If *ae1* is always of the same value i.e., either 0 or 1 or 2, it represents a vectoral change due to performance of the subsystem or as the result of an external performance. Any deviation in the trend can be due to either a break in data or an anomalous behaviour.

Similar functions *be1,ce1,…xe1* give the trend of the individual parameters during event e1.

Now the matrix in (1) can be represented as

Ae1 = {*ae1,be1,ce1,…xe1*} ….. (2)

The Rule-Induction algorithm treats these modified items to a step-wise comparison with every other item in the dataset and generates the output rules.

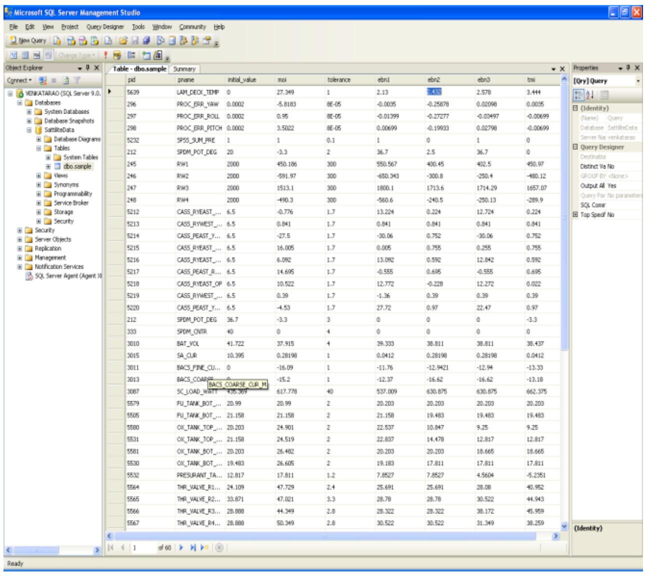

Figure 1. Table of parameters, their values during different events as obtained from archived satellite telemetry data

C. *Algorithm and implementation*

A part of the algorithm developed for the present application is reproduced below. Figures 2 to 4 show snap-shots of implementation and results obtained.

ALGORITHM

Input: Values of parameters from database.

Output: Set of rules

i← 0, j← 0, r←0, h1← total no of pids.

For j← 0 to h do

D[i][j]←in[j];

End for

For i← 0 to h1 do

D[i][j]← cur[j];

```
End for
For i← 0 to 1 do
        For j← 0 to h do
                If D[i][j] > D[0][j] Then
                        A[i][j]← 2
                End if
                Else if D[i][j] < D[0][j] Then
                        A[i][j]← 0;
                End if
                Else
                        A[i][j]← 1;
                End
        End for;
End for;
Return R;
```

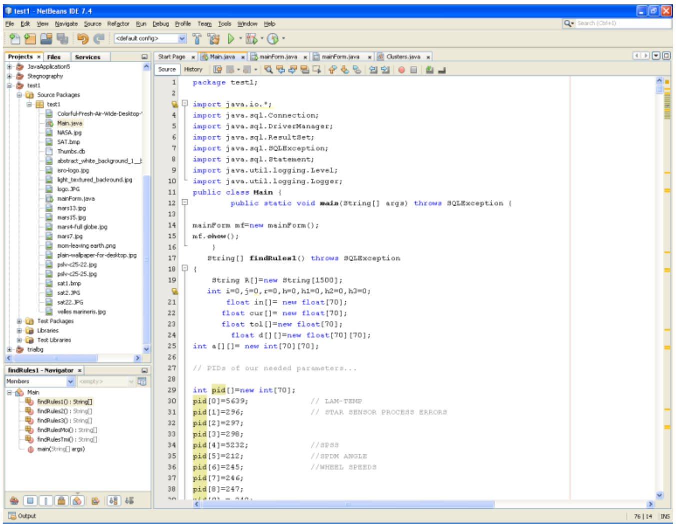

Figure 2. Implementation snap-shot of Rule Induction algorithm applied for Satellite Telemetry

RESULTS AND DISCUSSION

A sample Database with multi-dimensions has been created with Telemetry data from a Satellite functioning on-orbit. The data pertains to events when the LAM engine was fired during Earth burn phases, Trans-mars injection and Mars Orbit Insertion events.. Selected window covers an equalinterval of time during which the LAM engine is made on. Values of temperature, current, voltage, angle, speed and error changes of a randomly selected number of parameters in LAM Engine 'On' condition during all these events form the primary database.

Now a rule-set is generated based on the domain expertise giving the relation between different sets of related parameters. For this pilot study, a set of twenty parameters have been randomly chosen from the telemetry data and fed to the algorithm.

The design of the algorithm is such that it learns the relationship between these parameters, scans through the multi-dimensional database and generates a series of relationships hitherto unknown among various parameters within the same event set and between any two or more specified event sets.

For the initial study twenty parameters are chosen from the archived on-orbit Telemetry data. A pilot run has been implemented to train the algorithm and implementation (Fig.3). Based on validity of initial results, a test set data with sixty parameters has been created and fed to the algorithm (Fig.4).

It is found that out of the sixty-one rule-results obtained from among the set of twenty parameters, fifty-six have been found to be new knowledge and giving a new dimensional relation between the concerned parameters. Five of them have been found to be unconnected and improbable.

Typical outputs generated by applying Rule Induction technique :

1. RW1 increases with the increase in CASS_RYWEST_YN_OP
2. RW1 increases with the increase in CASS_RYEAST_YN_R_OP
3. PROC_ERR_PITCH increases with the increase in CASS_RYWEST_YN_OP
4. PROC_ERR_PITCH increases with the increase in CASS_RYEAST_YN_R_OP
5. SPSS_SUM_PRE increases with the increase in CASS_RYWEST_YN_OP
6. SPSS_SUM_PRE increases with the increase in CASS_RYEAST_YN_R_OP

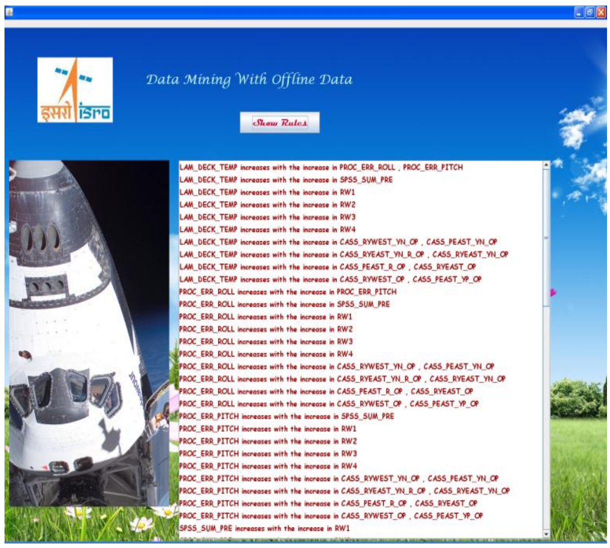

Figure 3. Snap-shot of results showing relational changes in parameters (implementation pilot–run)

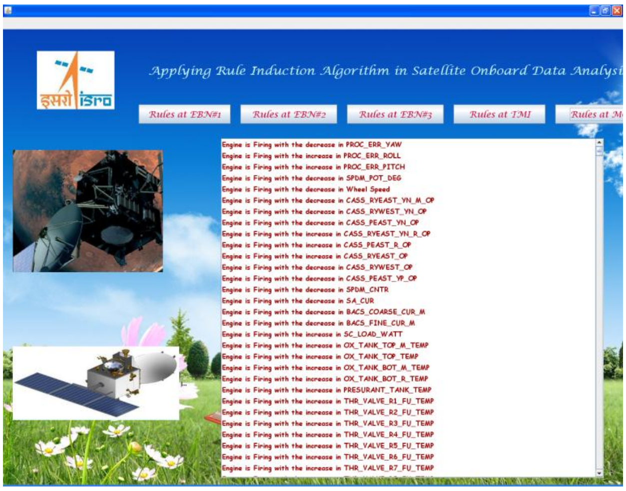

Figure 4. Snap-shot of results showing relational changes in LAM deck temperature and Processed-Error parameters during selected observation window (MOI)

CONCLUSIONS

Satellite telemetry received online as well as that assessed from archives are highly voluminous. It is found the Rule-Induction technique applied to Satellite Telemetry data has a clear edge over other mining techniques since it learns from the data and generates rules giving rise to meaningful and hitherto unknown knowledge from the uncovered patterns. This is a significant deviation from other researches carried out in applying Data Mining techniques where fault prediction is the main aim. At the primary level, the efficiency of the algorithm in giving correct results has been found to be better when the same dataset is tried for clustering and association mining techniques. The sample study carried out in the study can be extrapolated without any dilution of algorithm robustness to a very large multi-dimensional database.

ACKNOWLEDGEMENTS

The Authors would like to acknowledge the contribution of Ms.Saranya and Ms.Gunavathi, MCA students of Bharathiar University for their part in implementation of algorithm.

REFERENCES

[1] Lance Shelf, *"Use of Data Mining on Satellite Data Bases for Knowledge Extraction"*, FLAIRS-00 Proceedings Copyright 2000, American Association for Artificial Intelligence (www.aaai.org), 2000

[2] David L.Iverson, *"Data Mining Applications for Space Mission Operations Systems Health Monitoring"*, Space Ops 2008 Conference, AIAA 2008-3212, 2008

[3] Mayu Sakurada, *'Semantic Classification of Spacecraft Status: Integrating System Intelligence and Human Knowledge'*, Proceedings of the 2015 IEEE 9th International Conference on Semantic Computing (IEEE ICSC 2015), February 7-9,2015, Anaheim, California, 2015

[4] Yu Gao et al. *'An Unsupervised Anomaly Detection Approach for Spacecraft Based on normal behavior Clustering'* 2012 Fifth International Conference on Intelligent Computation Technology and Automation DOI 10.1109/ICICITA.2012.26, 2012

[5] Yu Gao et al.*' Fault Detection and Diagnosis for Spacecraft using Principal Component Analysis and Support Vector Machines'*, 2011 IEEE

[6] Tianshe Yang et al, *'Data Mining- Based Fault Detection and Prediction Methods for In-Orbit Satellite*', 2013 2nd International Conference on Measurement, Information and Control Harbin, China, IEEE, 2013