DEPENDABILITY IN SATELLITE SYSTEMS: AN ARCHITECTURE FOR SATELLITE TELEMETRY ANALYSIS

Denise Nunes Rotondi Azevedo[1], Ana Maria Ambrósio[2]

National Institute for Space Research (INPE/DSS)
S J Campos – SP, Brazil

denise@dss.inpe.br[1], ana@dss.inpe.br [2]

Abstract: *Safety, availability and reliability - attributes related to dependability - are very important issues in the context of Space Systems and Missions. In this paper we present a software architecture that integrates well-known Business Intelligent tools as Data warehouse, OLAP and Data Mining. We intend to take advantage of the largest amount of historical telemetry data which could help, in a "data driven" learning approach, diagnose and early detection of satellites anomalies and faults. We expect the integration of these tools in the proposed architecture to improve the process of telemetry analysis helping operators and experts in diagnosing and solving potential satellite problems before jeopardizing situations happen. The advantages of this architecture, challenges and research points are discussed as well.*

Keywords: *Data Mining, Anomaly Prediction, Telemetry Data*

1 Introduction

Artificial satellites normally render very important services to the society: communication, remote sensing, scientific experiments, etc. The complete or partial damage of a satellite implicates not only in a great financial loss, but also, the loss of essential, and sometimes, strategic unique services.

In this scenario, dependability plays a very important role in all the satellite life cycle: from the design up to the operation of such systems. Therefore, the early detection of anomalies and faults, diagnosis and prevention may promote reliability and availability of space systems and may extend their life time.

A satellite can be logically seen as a set of integrated subsystems (thermal, power supply, structure, payload, on-board computer, etc.). Aimed at monitoring the satellite and its health each subsystem has a set of sensors (thermistor, switch, battery depletion, etc.) for measuring the satellite states and conditions. These measurements are transmitted to the ground stations in each satellite pass and are called housekeeping telemetry data.

The telemetry data is analyzed and monitored in the ground station, mainly to assess if the values are not out of a pre-defined range. Values out of the limits normally indicate a potential anomaly or fault in the satellite or subsystem. However, the large amount of telemetry data (for China and Brazilian satellite -CBERS there are about 2000 telemetry-sensored measures, sampled in a short time interval) makes almost impossible a more careful and detailed analysis, for example, trend analysis, correlations between many sensor values and tendencies, etc.

Aiming at improving the TM analyses and the satellite monitoring process, this paper proposes an architecture that combines Business Intelligent tools as Data Warehouse, On-line Analytical Processing

(OLAP) and Visualization tools in an integrated way, to make the massive data analysis and health monitoring a more feasible task.

The architecture, also, incorporates Data Mining algorithms to discover new correlations patterns that can indicate anomalies and faults. This technique promotes a “data-driven” learning from the system and may help detect and diagnose anomalies in an automatic way.

This paper is organized as follows. Section 2 presents related works. Section 3 discusses the proposed architecture, and finally in Section 4 it is presented the conclusion and future work.

2 Related Works

Tools with the objective of monitoring telemetry data for anomalies and fault detection, mainly, using Data Mining techniques have been proposed in the literature. In (Yairi et al., 2004, Machida et al., 2006a and Machida et al., 2006b) data mining regression algorithms have been employed to adapt telemetry measurements range in an automatic way. In (Iverson, 2008) ORCA and IMS tool were used in the NASA projects. Orca uses nearest neighbor algorithms to find outliers, while IMS tool uses clustering techniques to generate a knowledge base of the normal operation for health monitoring. Most of the papers (Iverson, 2008, Hori at al., 2004, Machida et al., 2005 and Machida et al., 2006b) uses the approach of training the model for the normal operation, and consider as anomaly everything different from the patterns found.

VizTree tool (Nystrom et al., 2004) also combines data mining techniques with a specific visualization tool.

In our architecture, not only Data Mining tools are used. Instead, a set of complementary integrated tools work in a cooperative way to help operators and experts in the telemetry analysis and health monitoring activities, therefore promoting dependability in space missions.

3 Proposed Architecture

The operation of the CBERS satellite family by INPE/Brazil has produced about 15 Gbytes of Telemetry data. It is almost impossible to manually analyze this historical data. However, this data is a very rich source of information. This historical data may be used for discovering new patterns, correlation patterns and trends related to anomaly or fault occurrence, and also the eminence of these occurrences.

The proposed architecture, named Telemetry Analysis Architecture (TAA), integrates Automatic tools based on data mining techniques with Visualization tools. The automatic tools will be used for discovering new knowledge while Visualization tools will be used for helping operators and experts in the activities related to the telemetry and trends analyses and also in the analysis of the automatic extracted patterns and rules. The integration of these tools is quite important considering the amount of data.

Figure 1 illustrates the proposed architecture:

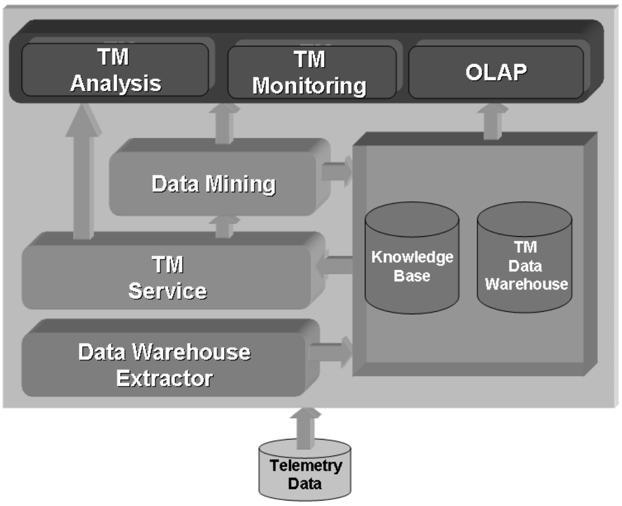

Figure 1. Telemetry analysis architecture.

The Telemetry Analysis Architecture is composed of the following integrated tools:

- Data Warehouse Extractor: automatically incorporates historical and incoming telemetries from satellites to a data warehouse and correlates them if it is possible with the historical data from other satellites (CBERS family), to constantly enhance the historical knowledge through similar telemetries in satellites series;
- Data mining based tools: predict anomalies and faults by identifying correlation rules between telemetries data and telemetries trends;
- TM Service: provide telemetry data in different formats for the other tools;
- On-line Analytical Processing (OLAP) tools: provide the operators and experts with the capability of making multidimensional queries on telemetry data, in any desired perspective: subsystem perspective, telemetry type perspective, etc.;
- TM Analysis: for viewing and analyzing large amount of telemetry data in a graphical and suitable way: This tool will also incorporate data mining results, so the experts or operators may analyze them;
- TM Monitoring: this tool will visually monitor the incoming telemetry data and advise operators about any problems using alarms.

In the next subsections it will be described in more detail what has been proposed for each tool.

3.1 Data Warehouse Extractor

There are some commercial off-the-shelf (COTS) Data Warehouse tools for extracting data from different sources in different formats. However, satellite telemetry data is a very specific data that is, in our case, stored in a framed and unprocessed format in relational data base, requiring a pre-processing. Aiming to solve it, the Data Warehouse Extractor shall extract telemetry data from different satellites, making the needed transformations and storing the data in the Data Warehouse database.

The proposed Telemetry Data Warehouse Star Model is shown in Fig. 2.

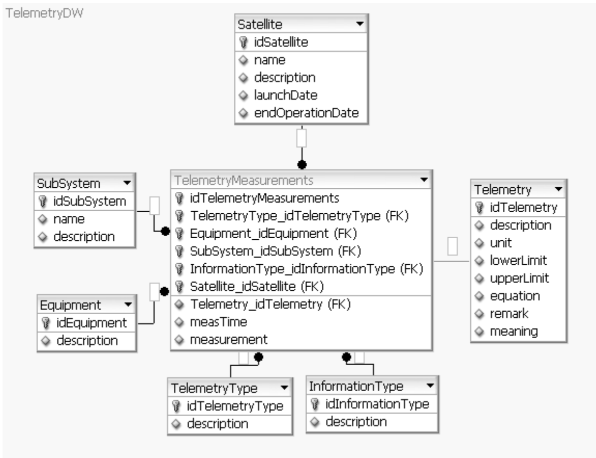

Figure 2. Telemetry star model.

3.1.1 Challenges and research points. A typical remote sensing satellite transmits to ground station, during the visibility period (about 15 minutes, two times a day), measurements of about 1000 sensors in a rate of 1.6 seconds. Data Warehouse Extractor design should take into account the following issues:

- Very huge data bases: a summarization method to store telemetry data in a suitable way may be necessary;
- Time-series sensor data: a careful relational database design has to be done, in order to represent data warehouse facts, namely, sensor measurements associated to the time of its occurrence. The Time and the sequence of the measurements over time is an important aspect.

3.2 Telemetry Service

In the proposed architecture a Telemetry Service has been designed to provide telemetry data in different formats for the other tools, such as Data Mining and Visualization tools (TM Analysis and TM Monitoring).

This component of the architecture has been intentionally designed as a Service, so it will be reutilized by other systems. In satellite operations the main objective of telemetry data is the satellite health monitoring but it may be a useful information for other applications like Satellite Simulators or for tests activities.

3.3 Data Mining Based Tool

Data Mining can be defined as automatic or semiautomatic discovering patterns in great amounts of data, where these patterns can be perceived as useful (Witten and Frank, 2005).

Using Data Mining techniques, unheard anomalies and faults patterns can be automatically discovered in the telemetry historical database, so the detection and diagnosis of new anomalies and faults in real time telemetry data may be improved. It may help in the satellite health analysis and it can lead to a more reliable and safer satellite operation.

The Data Mining tool will be the kernel of the proposed architecture. Further research is required before making a decision on existing algorithms that better fits the telemetry analysis needed. It will be also explored whether a single or a compounded algorithm shall be used in this tool. We intend to make experiments with some of the so far existing algorithms. Some points to be explored are:

- Detection of outliers in time-series telemetry data sensor. An outlier can be considered an anomaly (Iverson, 2008);
- The use of adaptive limits for more realistic limits checking (Yairi et al., 2004, Machida et al., 2006a and Machida et al., 2006b);
- The use of data mining algorithms, trained with normal operational data, for the establishment of normal correlations patterns between parameters. It will be generated a *Knowledge base* for monitoring new incoming data. Any incoming satellite telemetry data which doesn't match the expected patterns would be considered in an anomalous condition (Iverson, 2008, Hori at al., 2004, Machida et al., 2005 and Machida et al., 2006b);
- The use of data mining algorithms to discover new and unknown correlation patterns between sensors data of the same subsystem or inter subsystems that have led to an anomalous behavior or condition. Correlation between trends will be also analyzed.

The data mining tool will produce and update a *knowledge base* with normal and anomalous rules for satellite health monitoring and analysis. It will also produce formatted discovered knowledge to be analyzed and validated by experts and used as the base of OLAP queries.

3.3.1 Challenges and research points. Most of the research effort will be focused on this tool. A satellite is divided into subsystems and the knowledge is also spread among different expertise, making the understanding of the telemetry data a more complex issue.

The following subjects have to be addressed:

- The large numbers of attributes to be analyzed and selected in the data mining pre-processing phase;
- The mining of temporal series may need a combination of data mining techniques to produce some association rules;
- Trends, which are one of the most important aspects in the satellite anomalies, have to be considered in the research and experiments with data mining algorithms. The data mining association algorithms will have to address this characteristic of data.

3.4 Visualization Tools: TM Monitoring and TM Analysis Tool

The TM monitoring will be the tool used to effectively monitor the real time incoming telemetry data and the satellite health. This tool, using visual elements, will alert operators and experts about anomalies and faults. The alarms will take into account the telemetry range limits and the *knowledge base* generated by Data Mining tool to alert operators about anomalies, abnormal trends and critical faults.

The TM Analysis tool will provide experts with the capabilities to analyze data in graphical manner: trends, correlation between trends that have been discovered by Data Mining tool, etc. It will also be possible to analyze knowledge discovered by the data mining, association rules, etc. This tool will complement the OLAP COTS tool.

4 Conclusion and Future Work

Dependability is a key concept and issue in space systems and missions not only because the complexity of very expensive and time-consuming projects, but also because of the nature of the rendered services.

We will implement the proposed architecture in the context of INPE-CBERS satellite family. Nowadays, we have about 15 GBytes of CBERS telemetry data from the Cbers-1, Cbers-2 and Cbers-2B satellites, which will be used as historical data, mainly, for the mining activities.

Telemetries of thermal, power supply, on-board computer and other subsystems may be analyzed for anomaly tendencies. We intend to start our research using the Thermal subsystem as a case study.

The proposed architecture provides a set of integrated tools – both visual and automatic – to assess the telemetry data, helping operators and experts in the tasks of anomalies and faults detection and diagnoses aiming at preventing satellite fails and extending their life time. Another advantage of this approach is the use of the past failures knowledge in the conception of the future spacecrafts projects.

5 References

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

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

Machida, K., Fujimak, R., and Yairi, T., "An Approach to Spacecraft Anomaly Detection Problem Using Kernel Feature Space", In Proc. of the Eleventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD), 2005, pp. 401-410.

Machida, K., Fujimak, R. and Yairi, T., "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.

Machida, K., Fujimak, R., Yairi, T., Kawahara, Y. and Sato, Y., "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.

Nystrom, D. M., Keogh, E., Lankford, J.P., Lin, J., and Lonardi, S., "Visually Mining and Monitoring Massive Time Series", In Proc. of the Eleventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD), 2004.

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

Yari, T., Nakatsugawa, M., Hori, K., Nakasuka, S. and Machida, K., "Adaptive Limit Checking for Spacecraft Telemetry Data Using Regression Tree Learning", In Proc. of International Conference on Systems, Man and Cybernetics (IEEE-SMC), 2004, pp. 5130-5135.