Proceeding of the 6th International Symposium on Artificial Intelligence and Robotics & Automation in Space: i-SAIRAS 2001, Canadian Space Agency, St-Hubert, Quebec, Canada, June 18-22, 2001.

Fault Detection by Mining Association Rules from House-keeping Data

Takehisa Yairi, Yoshikiyo Kato and Koichi Hori

RCAST, University of Tokyo
4-6-1 Komaba, Meguro-ku Tokyo, JAPAN 153-8904
yairi@ai.rcast.u-tokyo.ac.jp

Keywords: fault detection, data mining, association rule, pattern clustering, house-keeping data

Abstract

This paper proposes a novel anomaly detection method for spacecraft systems based on data-mining techniques. This method automatically constructs a system behavior model in the form of a set of rules by applying *pattern clustering* and *association rule mining* to the time-series data obtained in the learning phase, then detects anomalies by checking the subsequent on-line data with the acquired rules. A major advantage of this approach is that it requires little *a priori* knowledge on the system.

1 Introduction

Fault detection is a key issue in the development of advanced spacecrafts. Although several detection techniques including limit-sensing, simulation and expert systems have been utilized for this purpose, they have often overlooked small anomalies in the house-keeping data (HK data) and some of them have led to fatal damages to the overall missions as a result.

One reason for the difficulty is that conventional fault detection methods generally require a tremendous a priori knowledge on the system behavior for each spacecraft, whereas that kind of knowledge is not always obtained easily beforehand. For example, a perfect dynamics model for simulation or a complete set of production rules for expert system is usually too expensive to prepare for each spacecraft. In addition, re-use of knowledge from past missions is also limited, because each spacecraft is usually more or less different from past ones.

Another reason is that these methods can grasp only limited aspects of overall spacecraft system behavior. For example, limit-sensing examines only upper and lower bounds of individual sensor values, and dynamics simulation can be performed on only several subsystems such as attitude control systems. To improve the fault detection reliability, therefore, it will be necessary to monitor and examine the spacecraft systems continuously from more various aspects.

In this paper, we propose a new fault detection method for spacecrafts based on data-mining techniques. This method first extracts a set of typical signal patterns or events from each time-series in the house-keeping data of a spacecraft by a clustering technique. Then it mines a set of association rules describing essential relationships among those time-series events. This set of rules can be regarded as a kind of "qualitative" model of the spacecraft. After that, it monitors the status of the spacecraft and detects any anomalies appearing in the house-keeping data by checking whether those association rules remain valid or not.

2 Conventional Approaches to Fault Detection Problem in Spacecrafts

Limit-sensing

Limit-sensing is the most basic and common technique of detecting anomalies in spacecraft systems. It constantly monitors the time-series of the house-keeping data and checks whether each sensor value is within the pre-defined upper and lower limits. Though it is very easy to apply this limit-sensing method to any type of spacecrafts because of its simplicity, there are some inevitable limitations.

First, it lacks in flexibility generally because the upper and lower limits of sensor values are supposed to be predefined in the design phases and fixed throughout the operation. Moreover, as those limit values usually contain wide margins, the checks tend to be too optimistic. As to this issue, a method which changes the limit values adaptively has been proposed recently[3].

Another problem is that the limit-sensing method regards the system to be normal if only the sensor values are within the ranges. In actual cases, however, anomalies can happen without exceeding those

Table 1: Comparison of Conventional and Proposed Fault Detection Methods

Method	**Limit Sensing**	**Simulation**	**Expert System**	**Proposed Method**
Representation Level	limit values (upper/lower limits)	mathematical models (differential equations)	production rules (if-then rules)	association rules (if-then rules)
Knowledge Source	experts (designers)	experts (designers)	experts (designers)	telemetry data (HK data)

limit values. In addition, limit-sensing monitors each of the sensor values individually. This implies that it may overlook some anomalies which can be detected only by looking at the relationships among more than one sensor values.

Simulation approach

Another traditional approach of detecting anomalies is the usage of simulations, in which mathematical (or sometimes qualitative) models of certain subsystems are utilized to simulate the system behavior and check the validity of the actual sensor values.

Simulation approach makes it possible to monitor the system in a detailed and strict way, compared with limit-sensing. However, both the preparation of the dynamics models and the execution of simulation using them are usually very expensive.

Expert systems

Several expert systems also have been proposed for this purpose, in which the knowledge for fault detection acquired from human experts is used. The knowledge is generally represented in the form of "if-then"production rules. This approach has an advantage that it can exploit a wide range of *know-hows* to detect anomalies in spacecrafts. However, knowledge acquisition from the human experts is usually a very costly and time-consuming process. What is still worse, re-use of the acquired knowledge (production rules) is generally difficult, because most spacecrafts are, more or less, specially designed and built and different from past ones.

In summary, there are two important issues in the conventional anomaly detection methods for spacecrafts, i.e.,

- How is the system behavior modeled and represented ?
- How is the necessary knowledge prepared (or acquired) ?

Tab.1 compares the three traditional methods described above and our approach from this viewpoint. As can be seen from this, the main purpose of our method is to automatically extract the qualitative relationships among the system elements in the spacecrafts from accumulated house-keeping data.

3 Application of Data Mining Techniques to Fault Detection

Data mining is one of the most active research fields nowadays, owing to several compound factors such as rapid progress of computers, emerging demands of knowledge extraction from large-scale information sources, and so on. The application target of data mining actually varies from biology, medical science to economics and social science.

In the meantime, spacecrafts in these days are also producing and sending to the ground a tremendous amount of house-keeping data (HK data) for monitoring a various kinds of system components, as well as mission data for the main purposes of the spacecrafts. The number of all time-series included in HK data reaches several hundreds to thousands in those spacecrafts. HK data is usually preserved in some storage devices after being firstly checked by limit-sensing or simulation mentioned above. However, the stored data is rarely used later, unless some faults occur.

It should be noticed here that, in many actual past fault cases, subsequent manual inspections of the stored data revealed that some symptoms had appeared long before the fatal faults occurred. Therefore, it is quite natural to think of applying the data mining techniques to the stored HK data to acquire automatically some useful knowledge for anomaly detection. More practically, a scenarion can be imagined that the data-mining (learning) processes is run in background during the initial phase in which human experts are monitoring a spacecraft, then the automatic anomaly detection system gradually re-

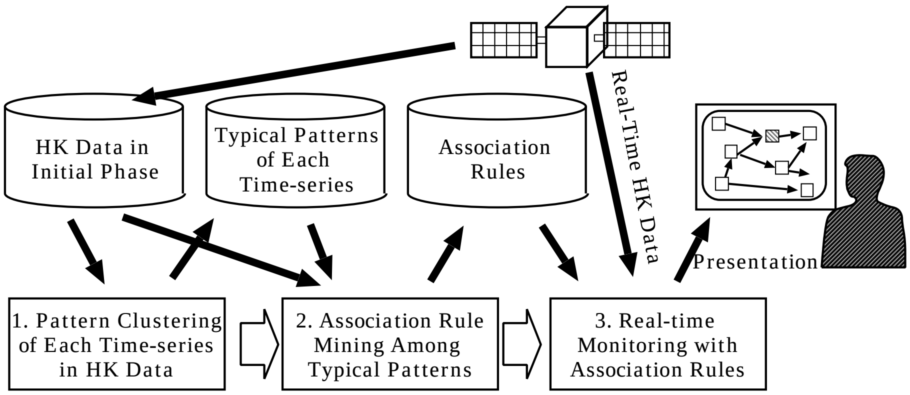

Figure 1: Overall Flow of the Proposed Method

places the experts as the knowledge is accumulated.

However, there are several difficulties in the actual application of data mining to the fault detection problem in spacecrafts.

First issue is the "imbalance of training data". When we learn to distinguish abnormal states of a system from normal states, it is desirable that both of the normal and abnormal examples are presented. In the cases of spacecrafts, however, that cannot be expected because it is almost unknown beforehand what faults may occur and what symptoms may appear due to them. This is a notable point that is completely different from the conventional application of data-mining techniques to mass-production industries.

Another important issue is the "presentation of the detection results". When the system gives alarms for some anomalies, human experts are supposed to reexamine them in detail. Therefore, it is demanded that the system should present the detection results in a fashion that human experts can easily understand them.

Based on these backgrounds, in the next section we introduce a new anomaly detection method for spacecrafts, which is grounded on two different kinds of data-mining techniques - time-series pattern clustering, and association rule mining.

4 Proposed Method

4.1 Overview

Basically, proposed method consists of the following three procedures:

1. Pattern clustering and event extraction for each time-series in the HK data
2. Mining association rules among the typical patterns (or events) in more than one time-series
3. Real-time monitoring of spacecraft system and anomaly detection using the acquired association rules

The former two procedures are performed on the data accumulated in the ground station during the initial operation phase right after the launch of the spacecraft. The last procedure, on the other hand, is applied to the real-time data telemetered from the spacecraft in order to detect any anomalies appearing in the HK data.

Figure 1 illustrates the outline of those three procedures. In the rest of this section, we describe each of the procedures in detail.

4.2 Pattern Clustering of Each Time-series in HK Data

The purpose of pattern clustering procedure is to extract a finite number of representative signal patterns from each continuous time-series in the HK data and label them. To be specific, this clustering procedure for each time-series can be described as follows (See also Fig.2):

1. A specified number of subsequences (we call them "instances") with a fixed length T_p are randomly sampled from the time-series.
2. Each instance or sample subsequence originally in the time domain is transformed or mapped into the feature vector space by calculating a set of feature values. Currently there are 9 features in total, including maximum and minimum values in the sequence, several coefficients of Fourier power spectrum, etc (Tab.2).

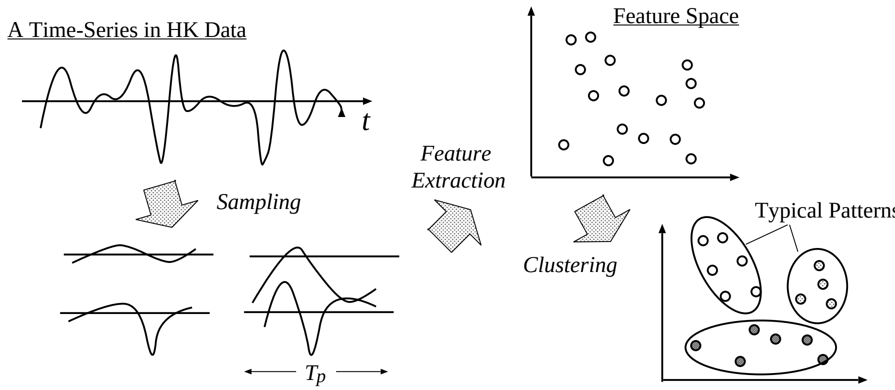

Figure 2: Pattern Clustering Procedure

3. Instances mapped in the feature vector space are clustered by *k-means* method. Instead of ordinary Euclidean distance measure, *standardized* Euclidean distance which is normalized by standard deviation is employed for calculating the similarity among instances.

4. Each cluster is labeled as a pattern class, which is supposed to correspond to a certain *event* in the time domain. For example, if a pattern cluster in the time-series of "output level of *Engine A*" is commonly characterized by an impulse, it might correspond to an event of "thrusting *Engine A*".

There are currently two major problems in this clustering procedure as follows:

- What features should be used in *Step 2* ?
- How should we decide the number of clusters in *Step 3* ?

A possible approach is to allow users (or experts) to decide these parameter interactively. As to the latter issue, automatic decision is also possible to some extent by employing information theoretic criteria such as *Akaike's Information Criteria*(AIC) and *Minimal Description Length*(MDL).

4.3 Association Rule Mining Among Patterns in Different Time-series

Association rule mining[1] is one of the most popular data mining techniques these days. The purpose of the association rule mining is to find a set of interesting or useful rules when a database or a set of *transactions* is given.

The basic framework of the association rule mining is formalized as follows[1, 5]: Let $\mathcal{I} = \{I_1, I_2, \cdots, I_m\}$ be a set of *items* or binary attributes. Let $\mathcal{D}$ be a database or a set of *transactions*. Each transaction is represented by a m-dimensional binary vector specifying which items are contained. For example, a transaction expressed by $[1, 0, 1, \cdots, 1]$ implies that it contains I_1, I_3, I_m. An association rule has the form of $X \longrightarrow Y$, where X and Y are sets of some items ($X \subset \mathcal{I}, Y \subset \mathcal{I}$) and disjoint ($X \cap Y = \emptyset$). The meaning of this association rule is, "if a transaction contains an item set X, then it is likely to contain Y too". The *confidence* of this rule is the percentage of transactions containing Y (and also X) among the transactions containing X. On the other hand, the *support* of this rule is the percentage of transactions containing both X and Y in D. Now, the problem of association rule mining is to find all rules that satisfy minimum support and confidence which are specified by users.

Table 2: List of Features

No.	Feature description
1	*Maximum* value
2	*Minimum* value
3	*Mean* value
4	*Standard deviation*
5	1st coefficient of FFT power spectrum
6	2nd coefficient of FFT power spectrum
7	3rd coefficient of FFT power spectrum
8	4th coefficient of FFT power spectrum
9	5th coefficient of FFT power spectrum

In our problem, we regard the typical patterns obtained by the clustering procedure as the *items*. Also, we divide all the time-series of HK data into segments with a fixed length of T_s ($> T_p$), and regard each time

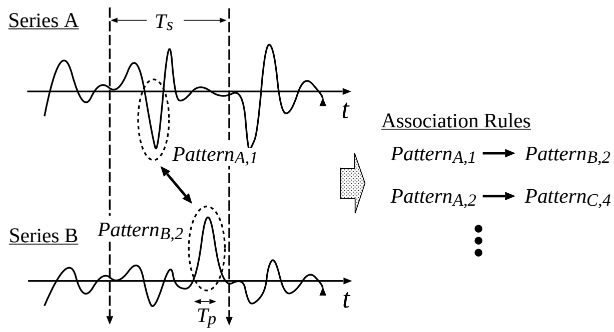

Figure 3: Association Rule Mining Among Patterns

segment as a *transaction* (Fig.**??**). By this interpretation, we can apply the framework of association rule mining mentioned above to acquire knowledge about the relationship among the typical patterns in different time-series of HK data. For example, a rule $Pattern_{A,1} \longrightarrow Pattern_{B,2}$ has the meaning of "if $Pattern_{A,1}$ appears in a time segment of *series A*, then $Pattern_{B,2}$ is also likely to appear in the same segment of *series B* ". In addition to the traditional measures - *confidence* and *support*, we employ *J-measure* [4, 2] as another measure of "interestingness" of each rule. J-measure for a rule $X \longrightarrow Y$ ("If X Then Y") is defined as:

$$J(Y : X) = Pr(X) \cdot [Pr(Y|X) \cdot \log \frac{Pr(Y|X)}{Pr(Y)} + Pr(\neg Y|X) \cdot \log \frac{Pr(\neg Y|X)}{Pr(\neg Y)}]$$

where $Pr(X)$ is the probability that the proposition X holds in the database, and $Pr(Y|X)$ is the conditional probability that Y holds given that X holds. Roughly speaking, *J-measure* is the average amount of information contained in a rule, and can be used to measure *how important the rule is*.

A set of discovered association rules can be regarded as a kind of model that describes the nominal behavior of the spacecraft system.

4.4 Real-time Monitoring with Acquired Association Rules

As explained above, each association rule is supposed to describe a local aspect of the system behavior in normal mode. Therefore, by monitoring and checking continuously whether these rules are valid or not, it is possible to detect some changes from the normal behavior. More strictly, some statistical test such as χ^2-test is utilized to determine if the changes of the *confidence* and *support* values of a rule are significant or not.

One advantage of using a set of association rules for anomaly detection is that it can detect one fault in more than one ways by monitoring the validity of multiple rules which are related to one element. For example, when a fault occurs in a reaction wheel for the attitude control, it will not only cause some changes in the cause-effect relationship between the control input to the wheel and the actual angular velocity, but also affect other time-series in HK data such as the amount of the electric current in the motor, the total power consumption, and so on. If an appropriate set of association rules is obtained in the learning phase, this kind of faults should be found by detecting significant changes in the *confidence* and *support* values of more than one relevant association rules. This property of association rules, or *redundancy* is expected to increase the reliability of the fault detection.

Table 3: Description of 6 Time-series in HK Data

Series ID	Description
S_1	Output of the thruster 1 (roll)
S_2	Output of the thruster 2 (pitch)
S_3	Output of the thruster 3 (yaw)
S_4	Angular acceleration around *roll* axis
S_5	Angular acceleration around *pitch* axis
S_6	Angular acceleration around *yaw* axis

5 Experiment

In order to examine the validity of the proposed method, we applied it to a part of past house-keeping data of ETS-VII (Engineering Test Satellite VII) provided by NASDA. It has been reported that this satellite suffered a fault in a reaction wheel on Dec 25, 1999.

In this experiment, we chose six time-series of two groups in HK data (Tab.3). They are *outputs of three thrusters for attitude control about roll, pitch and yaw axes* (S_1,S_2,S_3) and *angular acceleration around the three axes* (S_4,S_5,S_6). The sampling period for each time-series is equally 0.083 (1/12) sec.

Pattern Clustering of Each Time-series

Based on the pattern clustering method described in 4.2, typical patterns were extracted from each time-series as follows.

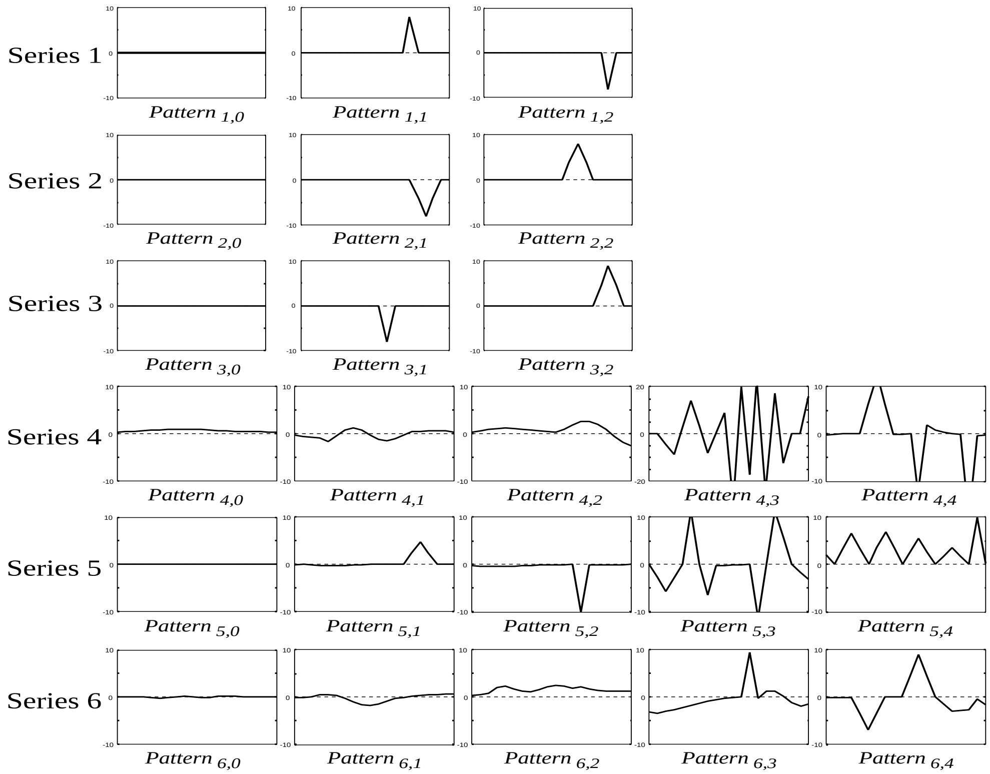

Figure 4: Representative Patterns of Each Time-series Obtained by Pattern Clustering

First, a set of instances was formed by sampling 5000 subsequences with the length of $T_p = 1.67$ sec (consisting of 20 data points) randomly from the time-series. After that, feature values were computed for each element of the instance set. Then, the clustering procedure was performed by applying *k-means* method to the instance set in the feature vector space. The numbers of the clusters were set to 3 in series S_1,S_2,S_3, and 5 in series S_4,S_5,S_6.

Fig.4 illustrates the representative instances (patterns) of the obtained clusters in each series.

Any of the listed rules can be interpreted as "If one thruster is turned on, a certain pattern is likely to appear in the time-series of the angular acceleration about the corresponding axis", or inversely, "If a certain pattern appears in a time-series of the angular acceleration, it is likely that the corresponding thruster engine was turned on". This result is reasonable and consistent with our intuition.

Association Rule Mining Among Patterns

Next, we applied the association rule mining to the obtained patterns in the 6 time-series. Though negation ($\neg$) is allowed both in the pre-condition part (X) and result part (Y), conjunction ($\wedge$) and disjunction ($\vee$) of more than one patterns are not considered this time.

Tab.4 shows the top eight association rules in terms of the magnitude of *J-measure* value (6th column). This table also lists the *confidence* (4th column) and *support* (5th column) values of each rule.

Anomaly Detection Using Association Rules

Finally, we investigated how the validity of those association rules changed after the fault occurred. The last column of Tab.4 shows the *confidence* values of the eight rules *before* and *after* the occurrence of the fault. In most of the rules, the confidence values decreased after the fault. Especially, declines of the values in *Rule 1,4,7* are remarkable, which is considered to be reflecting the change of the system behavior induced by the fault.

Table 4: Obtained Association Rules and Confidence After Fault

Rule no.	Condition.	Result.	Confidence	Support	J-measure	Confidence *after* fault
1	$P_{5,2}$	$P_{2,1}$	0.882	0.0312	0.0693	0.462
2	$P_{5,1}$	$P_{2,2}$	0.684	0.0271	0.0571	1.000
3	$P_{2,1}$	$P_{5,2}$	0.429	0.0312	0.0561	0.353
4	$P_{3,1}$	$\neg P_{6,0}$	0.813	0.0271	0.0553	0.400
5	$P_{2,2}$	$P_{5,1}$	0.542	0.0271	0.0539	0.431
6	$P_{3,1}$	$P_{6,1}$	0.813	0.0271	0.0433	0.600
7	$P_{6,1}$	$P_{3,1}$	0.236	0.0271	0.0324	0.043
8	$P_{1,1}$	$P_{4,2}$	0.529	0.0187	0.0311	0.563

6 Conclusion

This paper proposed a new anomaly detection method for spacecrafts based on two different kinds of data-mining techniques - *time-series pattern clustering* and *association rule mining*.

In this method, typical temporal patterns are extracted from each time-series of the house-keeping data which has been accumulated in the initial operation phase. Then, qualitative cause-effect relationships among those patterns in different time-series are explored and obtained in the form of association rules. The set of association rules can be regarded as a collective model of the spacecraft, and can be used for detecting whether the system behavior is normal or not.

This approach has two notable features, compared with the traditional fault detection methods. Firstly, it requires little *a priori* knowledge on the spacecraft system, as it obtains it automatically from the house-keeping data. Therefore, this method can be applied to various kinds of spacecrafts at relatively small costs. Second, it models the behavior of spacecrafts by a set of association rules, which is different from the representation by either differential equations or limit values. As a result, it is expected to detect some sorts of anomalies which have been usually overlooked by the conventional methods.

One of the most significant issues in future is the design of an effective interaction between the proposed method and users. This is because the proposed method contains a variety of parameters, and it is not always easy to decide them properly in advance. It is also important to present the results of pattern clustering, association rule mining and fault detection in such a way that users can understand and justify them easily.

In addition, we are going to examine the effectiveness of the proposed method in more practical situations with larger number of time-series of HK data.

Acknowledgment

We thank Yohsuke Fukushima and Naoki Ishihama for providing us with the data of ETS-VII and a lot of insightful comments.

References

[1] R. Agrawal, T. Imielinski, and A. Swami. Mining association rules between sets of items in large databases. In *Proceedings of the 1993 ACM SIGMOD Conference*, 1993.

[2] G. Das, K. Lin, H. Mannila, G. Renganathan, and P. Smyth. Rule discovery for time series. In *Proceedings of the Fourth nternational Conference on Knowledge Discovery and Data Mining (KDD-1998)*, pages 16–22, 1998.

[3] D. DeCoste. Automated learning and monitoring of limit functions. In *Proceedings of International Symposium on Artificial Intelligence, Robotics and Automation in Space (i-SAIRAS*, pages 287–292, 1997.

[4] P. Smyth and R. Goodman. An information theoretic approach to rule induction from databases. *IEEE transactions on knowledge and data engineering*, 4(4):301–316, 1992.

[5] R. Srikant and R. Agrawal. Mining generalized association rules. In *Proc. of the 21st Int'l Conference on Very Large Databases*, 1995.