

An Anomaly Detection Method for Spacecraft Using ICA Technology

Xu Binglin, Li Zhanhuai

School of Computer Science, Northwestern Polytechnical University, Xi'an, China

Abstract - Anomalies, outlier and fault means that the received data do not fit the common pattern, and may often be occur in spacecraft missions because of the high complexity and rugged operating environment of spacecraft. In order to reveal the potential faults or anomalies many methods and technology have been researched and achieved fruitful results. In this paper, we proposed an anomaly detection method for Spacecraft based on ICA technology. The independent component analysis (ICA) is recently developed in order to extract possible independent component and reveal the hidden anomaly pattern and can be used in the anomaly detection for spacecraft. By Applying ICA technology on spacecraft telemetry data, we can reveal the hidden anomalies in telemetry data which can not be accomplished by the ordinary method. The fastICA algorithm is used when doing ICA on real spacecraft telemetry. The simulation result and the application on real telemetry data prove the effectiveness of the method.

Index Terms - Data Process, Spacecraft Health Monitoring, Independent Component Analysis, Fault Diagnosis

1. Introduction

It's obvious that a space mission requires the highest level of safety and reliability. On the other hand, since the scale of space systems are so large and complex, no matter how carefully they are designed and produced, it is still very difficult to completely eliminate the possibility of malfunction in space systems, which has been demonstrated by the history of space mission anomalies and malfunctions[1][2]. Therefore, various efforts in the design, development, test and operation phases are required to ensure the highest level of safety and reliability. It is no doubt that the development of advanced anomaly detection and fault diagnosis technologies also plays an important role for this purpose. Actually, several attempts to develop such advanced anomaly detection and fault diagnosis methods by applying various information technologies and artificial intelligence have been made[3][4]. Conventionally, anomaly detection methods based on prior expert knowledge and deductive reasoning process, such as expert systems and model-based reasoning, have been principally studied for this purpose[4]. Though these knowledge-intensive approaches have been proved to be much better than the classical limit checking method, they are still costly and time-consuming to prepare the required knowledge or model. On the other hand, in recent years, deductive reasoning techniques, such as data mining or machine learning technologies, have drawn much attention as alternative approach to the anomaly detection problems in various application fields [2].

The ICA aims at identifying a set of independent unobservable variables that are supposed to generate the data

set of interest. ICA is a computational method for separating a multivariate signal into additive subcomponents supposing the mutual statistical independence of the non-Gaussian source signals. It is a special case of blind source separation and has been successfully applied in a variety of fields, such as speech, radar, signal processing and time series[5]. ICA defines a generative model for the observed multivariate data, which is typically given as a large database of samples. In the model, the data variables are assumed to be linear mixtures of some unknown latent variables, and the mixing system is also unknown. The latent variables are assumed nongaussian and mutually independent and they are called the independent components of the observed data. These independent components, also called sources or factors, can be found by ICA. This paper proposes an anomaly detection method of spacecraft by using the ICA technology in telemetry data. This method can identify the locations of multiple anomalies in the telemetry data. This paper first introduces the basic theory of the ICA, proposes the method model, and finally provides the experiment result and conclusion.

2. Theory of ICA

ICA comes from blind signal separation (BSS), which means the separation of a set of signals from a set of mixed signals, without the aid of information about the source signals or the mixing process [6]. In many scientific and engineer system, an unknown mix matrix is postulated to linearly transform the unobservable signals to a set of observable mixed ones. The model can be expressed by fig.1 and formula (1). Because both the signal variable $S(t)$ and the mix matrix A are unknown, BSS technology can be only used to estimate them from the data.

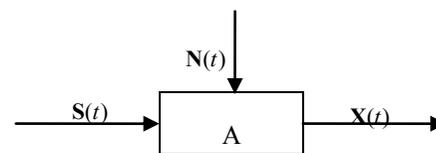


Fig. 1: Signals and Mix System Model

$$\mathbf{X}(t) = \mathbf{A}\mathbf{S}(t) + \mathbf{N}(t) \quad (1)$$

Where, $\mathbf{S}(t)$ refers to the source signals which are unobservable, \mathbf{A} refers to the mix matrix, $\mathbf{N}(t)$ refers to the gaussian white noise and $\mathbf{X}(t)$ refers to the received data.

ICA is a statistical and computational technique revealing the mix matrix and source signals in BSS. In the model, the data variables are assumed to be linear mixtures of the unobserved source signals. Since they are nongaussian and mutually independent, they are called the independent components of the observed data. These independent components or sources signals can be revealed by ICA when these classic methods fail completely.

ICA has evolved into many forms, include linear ICA, Nonlinear ICA etc. In this paper, we just research on the linear ICA, which assumes that the received data is a linear mixture of the unobserved source signals and the source signals do not correlate with each other, that is, the signals may be mutually statistically independent or decorrelated. Because anomalies can also be looked on as one of the source signals and they are independent to the source signals at most of time, the ICA can be used to separate the anomalies.

Before the application of an ICA algorithm in the spacecraft telemetry data, it is usually very useful to perform some preprocessing [5], including:

(1) Centering

The most basic and necessary preprocessing procedure is to center \mathbf{x} to make it a zero-mean variable. The method is to subtract \mathbf{x} by its mean vector $E(\mathbf{x})$. This preprocessing implies that \mathbf{S} is zero-mean as well and can make the ICA algorithms simple. After estimating the mix matrix \mathbf{A} with centered data, we can complete the estimation by adding the mean vector of \mathbf{S} back to the centered estimates of \mathbf{S} . The mean vector of \mathbf{S} is given by $\mathbf{A}^{-1}\mathbf{m}$, where \mathbf{m} is the mean that was subtracted in the preprocessing.

(2) Whitening

Another useful preprocessing procedure in ICA is to firstly whiten the observed variables. Before the application of the ICA algorithm (and after centering), we transform the observed vector \mathbf{x} linearly so that we can obtain a new vector $\hat{\mathbf{x}}$ which is white. Its components are uncorrelated and their variances equal unity. In other words, the covariance matrix of $\hat{\mathbf{x}}$ equals the unit matrix \mathbf{I} .

One popular method for whitening is to use the eigenvalue decomposition of the covariance matrix $E(\mathbf{xx}^T) = \mathbf{E}\mathbf{D}\mathbf{E}^T$, where \mathbf{E} is the orthogonal matrix of eigenvectors and \mathbf{D} is the diagonal matrix of its eigenvalues. Whitening can now be given by:

$$\hat{\mathbf{x}} = \mathbf{E}\mathbf{D}^{1/2}\mathbf{E}^T\mathbf{x} \quad (2)$$

where the matrix $\mathbf{D}^{1/2}$ is computed as $\mathbf{D}^{1/2} = \text{diag}(d_1^{1/2}, d_2^{1/2}, \dots, d_n^{1/2})$, and now it is easy to check that $E(\hat{\mathbf{x}}\hat{\mathbf{x}}^T) = \mathbf{I}$.

Whitening preprocessing also transforms the mix matrix into a new one $\hat{\mathbf{A}}$. With formula (1), (2) and zero gaussian

white noise, we can get $\hat{\mathbf{x}} = \mathbf{E}\mathbf{D}^{1/2}\mathbf{E}^T\mathbf{A}\mathbf{s} = \hat{\mathbf{A}}\mathbf{s}$. With $E(\hat{\mathbf{x}}\hat{\mathbf{x}}^T) = \mathbf{I}$, the new mix matrix $\hat{\mathbf{A}}$ is also orthogonal, which can be seen from formula (3).

$$E(\hat{\mathbf{x}}\hat{\mathbf{x}}^T) = \hat{\mathbf{A}}E(\mathbf{ss}^T)\hat{\mathbf{A}}^T = \hat{\mathbf{A}}\hat{\mathbf{A}}^T = \mathbf{I} \quad (3)$$

3. Anomaly Detection Model

Let $\{\Delta_t\}$ denote the sequence describing the anomaly pattern, that is, $\Delta_t = 1$ when there is an outlying observation at time $t = t_0$, and $\Delta_t = 0$ otherwise. A level change at time $t \in [t_m, t_l]$ is represented by the sequence, $\Delta_t = 1$ at time $t \in [t_m, t_l]$ and 0 otherwise.

Let $\mathbf{x}_t = (\mathbf{x}_{1t} \cdots \mathbf{x}_{ht})^T, t = 1, \dots, n$ denote the observed telemetry data and $\boldsymbol{\varepsilon}_t = (\boldsymbol{\varepsilon}_{1t} \cdots \boldsymbol{\varepsilon}_{ht})^T$ denote a sequence of independent identically distributed vector Gaussian variables with mean 0 and covariance matrix $\sum_{\varepsilon} = \text{diag}\{\sigma_1^2 \cdots \sigma_h^2\}$. We suppose that the observed time series is obtained by a linear transformation of a k individual signal $\boldsymbol{\varepsilon}_t = (\boldsymbol{\varepsilon}_{1t} \cdots \boldsymbol{\varepsilon}_{kt})^T$, and the mix matrix is denoted by \mathbf{A} . Then the observed data with anomalies can be written as:

$$\mathbf{x}_t = \mathbf{A}\mathbf{s}_t + \boldsymbol{\omega}\Delta_t + \boldsymbol{\varepsilon}_t \quad (4)$$

Where $\boldsymbol{\omega} = (\omega_1 \cdots \omega_h)^T$ is the anomaly magnitude. The model (4) is known as dynamic factor model and has been often used in time series analysis and econometric studies with $\boldsymbol{\omega} = \mathbf{0}$. It has recently attracted a new interest as a framework for forecasting from a large number of predictors[7] and for classifying outliers in innovations[8].

After data preprocessing, we can assume a new mix matrix without losing the system information with property $\mathbf{A}\mathbf{A}^T = \mathbf{I}$ (formula 3). We can also assume without loss of generality that the k column vectors $\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_k\}$ of the mix matrix \mathbf{A} are linearly independent. $\boldsymbol{\omega}$ can be written as $\boldsymbol{\omega} = \mathbf{A}\boldsymbol{\alpha} + (\boldsymbol{\omega} - \mathbf{A}\boldsymbol{\alpha})$, where $\boldsymbol{\alpha} = (\alpha_1 \cdots \alpha_k)^T$ and the vector $\boldsymbol{\omega} - \mathbf{A}\boldsymbol{\alpha}$ can be orthogonal to each \mathbf{a}_j through computing different coefficient vectors $\boldsymbol{\alpha}$. Now on defining the unit norm vector $\mathbf{b}_1 = (\boldsymbol{\omega} - \mathbf{A}\boldsymbol{\alpha}) / \|\boldsymbol{\omega} - \mathbf{A}\boldsymbol{\alpha}\|$, we may consider the $k+1$ vectors $\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_k, \mathbf{b}_1\}$ and complete an orthonormal basis of space R^h by further choosing $h-k-1$ orthonormal vectors $\{\mathbf{b}_2, \dots, \mathbf{b}_{h-k}\}$. The resulting basis is denoted by $\mathbf{C} = [\mathbf{A}, \mathbf{B}] = [\mathbf{a}_1 \cdots \mathbf{a}_k, \mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_{h-k}]$. Since $\boldsymbol{\varepsilon}_t \in R^h$, it may be projected onto the space spanned by the orthonormal

vectors. Let us further define the sequence $\boldsymbol{\eta}_t$ and the relationship $\boldsymbol{\varepsilon}_t = \mathbf{C}\boldsymbol{\eta}_t$. The elements of $\boldsymbol{\eta}_t$ are mutually independent because this property is assumed to hold for the sequence $\boldsymbol{\varepsilon}_t$. Then model (1) can be written in the following equivalent form:

$$\mathbf{x}_t = \mathbf{A}s_t + [\mathbf{A}, \mathbf{b}_1] \begin{pmatrix} \boldsymbol{\alpha} \\ |\boldsymbol{\omega} - \mathbf{A}\boldsymbol{\alpha}| \end{pmatrix} \Delta_t + [\mathbf{A}, \mathbf{B}]\boldsymbol{\eta}_t, \quad (5)$$

$$= [\mathbf{A}, \mathbf{B}] \left\{ \begin{pmatrix} s_{1t} + \alpha_1 \Delta_t \\ \vdots \\ s_{kt} + \alpha_k \Delta_t \\ |\boldsymbol{\omega} - \mathbf{A}\boldsymbol{\alpha}| \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \begin{pmatrix} \eta_{1t} \\ \vdots \\ \eta_{kt} \\ \eta_{k+1,t} \\ \eta_{k+2,t} \\ \vdots \\ \eta_{ht} \end{pmatrix} \right\}$$

where \mathbf{x}_t is written as a linear combination of independent series, the $(k+1)$ th source is proportional to Δ_t plus a noise term, the first k components are the origin independent components perturbed by outliers, and the remaining $h-k-1$ sources are purely Gaussian noise. This formula is just the linear noisy ICA model, therefore ICA may extract the mix matrix $[\mathbf{A}, \mathbf{B}]$ and the components can be characterized as follows:

- (1) The component $|\boldsymbol{\omega} - \mathbf{A}\boldsymbol{\alpha}| \Delta_t + \eta_{k+1,t}$ only contains the term closely related to the anomaly pattern, plus a noise term.
- (2) k components $s_{jt} + \alpha_j \Delta_t + \eta_{jt}$, $j = 1, 2, \dots, k$ include the original factor series, a possible term due to the anomaly and a random error.
- (3) The remaining $h-k-1$ η_{jt} , $j = 2, \dots, h-k$ are purely random noise.

So after doing ICA, the component $|\boldsymbol{\omega} - \mathbf{A}\boldsymbol{\alpha}| \Delta_t + \eta_{k+1,t}$ may give an indication of the behavior of the anomaly.

4. Application and result

In order to test the effective of proposed method, we apply the method on the real satellite telemetry data received from a geosynchronous satellite. The data come from power supply subsystem, including the voltage of accumulator-A, accumulator-B, main power bus and SADA bus. The original data set is shown in fig.2(a). The ICA algorithm used in this paper is fastICA, which can be found in Hyvarinen's paper[5].

We set two kinds of anomalies, one is the sudden change, which means that the data is change suddenly and resume to the normal state, although this circumstances perhaps mean

only a disturbance but should also attract attention, and the other is continuous change at a longer time and should be taken seriously before a big malfunction is happen. We explain the experiment as below.

(1) $\boldsymbol{\omega} = -3$ at time 50 and $\boldsymbol{\omega} = 3$ at time 100, this is sudden change type at point time as shown in fig.3(a). Although the data resume normal quickly, there may exist some important malfunction and shall attract more attention. Fig.3(b) shows the result after performing ICA with the anomalies.

(2) $\boldsymbol{\omega} = -3$ at time 50-60 and $\boldsymbol{\omega} = 3$ at time 100-110, this is continuous sudden change type as shown in fig.4(a) and shall be revealed as soon as possible. Fig.4(b) shows the result after performing ICA with the anomalies.

In order to contrast with normal status, we also perform ICA on the origin telemetry without any anomalies, and the result is shown in fig.2(b).

As the figures shown below, the origin telemetry data values vary from 30 to 40, a change level at 3 is inappreciable and is difficult to be revealed by ordinary method. The proposed method is able to reveal the associate anomalies as shown in the result figure even the change value is exiguous. The first components of the ICA result shown in figure can be the evidence that anomalies really occurs.

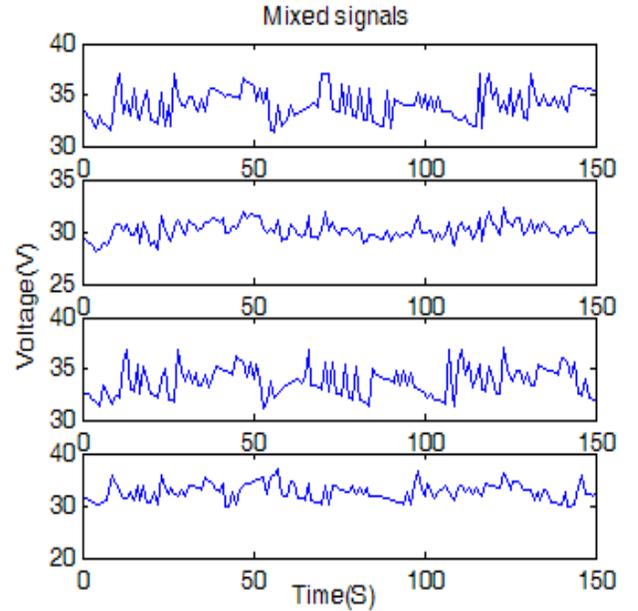


Fig 2: (a) Origin telemetry without anomaly

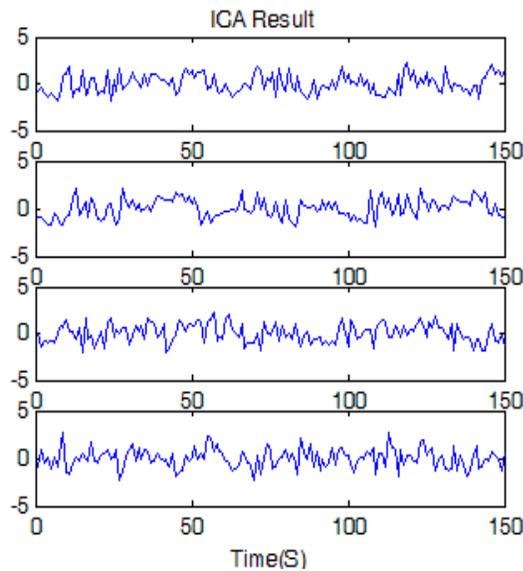


Fig. 2: (b) Result after performing ICA

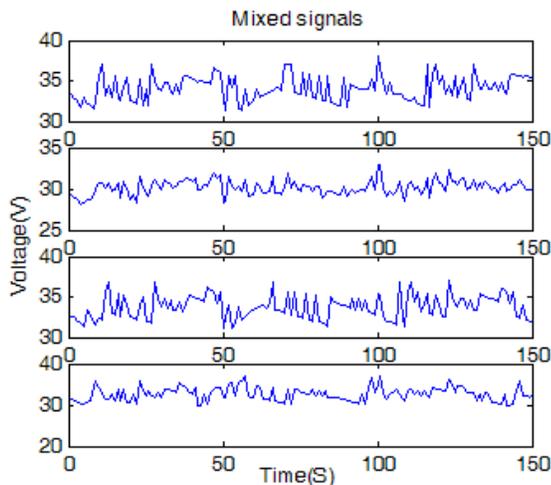


Fig. 3: (a) Origin telemetry with anomaly (1)

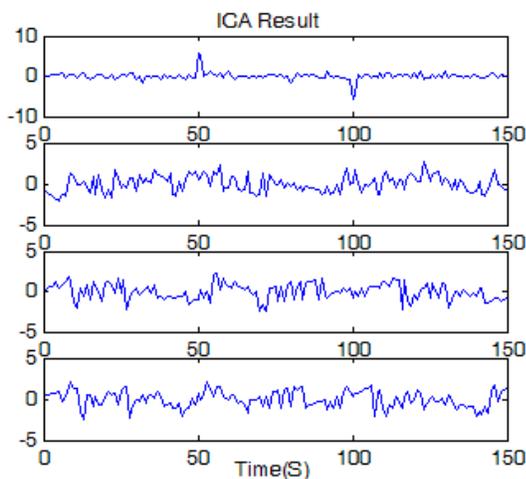


Fig. 3: (b) Result after performing ICA

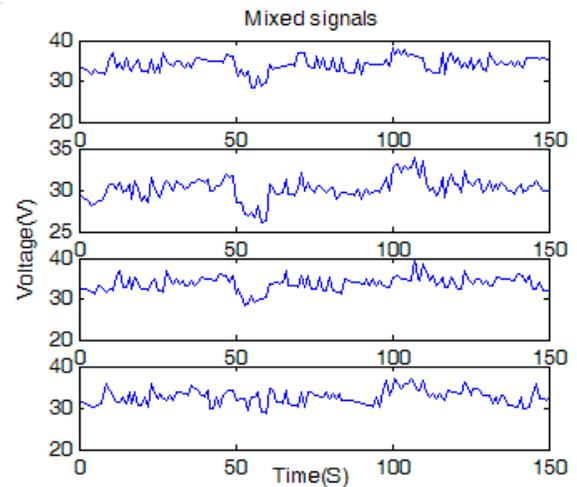


Fig. 4: (a) Origin telemetry with anomaly (2)

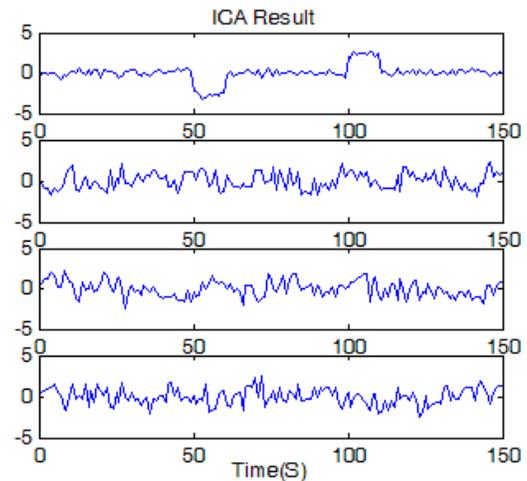


Fig. 4: (b) Result after performing ICA

The experiment proves that the proposed method can be used in anomaly detection successfully. But as shown before, the anomalies in the experiment are simulated by adding artificial components, the effectiveness of the method in real anomaly telemetry data should also be checked in future work. Because of the noise and disturbance of the space environment, how to reduce the mis-detection rate automatically should also be resolved in the future.

5. Conclusion

This paper proposes an ICA-based method to reveal potential anomalies in spacecraft telemetry data. The method can reveal hidden anomalies in telemetry data which cannot be accomplished by using ordinary methods. This paper applies the ICA-based method to real telemetry data, simulating sudden and consecutive level changes. The results show that simulated anomalies can be revealed successfully by using the fastICA algorithm. Before the method is widely used in fault diagnosis and anomaly detection in

spacecraft practicality, the telemetry data classification and the automatic analysis and detection of anomalies in the ICA component shall be solved in the future.

Acknowledgements

I would like to show my deepest gratitude to my supervisor, Dr. Li Zhanhuai, a respectable, responsible and resourceful scholar, who has provided me with valuable guidance in every stage of the writing of this thesis.

Reference

- [1] Martin, R. A., Schwabacher, M., Oza, N., and Srivastava, A., Comparison of Unsupervised Anomaly Detection Methods for Systems Health Management Using Space Shuttle Main Engine Data, Proceedings of the JANNAF Propulsion Meeting, JANNAF, Denver, 2007.
- [2] VARUN CHANDOLA, ARINDAM BANERJEE, VIPIN KUMAR et al. Anomaly Detection: A Survey . ACM computing surveys, 2009, 41(3) :15.1-15.58.
- [3] Enrica Vecchio, Beatrice Lazzerini, Steve Foley etc, Spacecraft fault analysis using data mining techniques, Proc. of ISAIRAS 2005 Conference , Munich, Germany, 2005.
- [4] R. Fujimaki, T. Yairi, and K. Machida., An anomaly detection method for spacecraft using relevance vector learning, Proc. of the 9th Pacific-Asia Conference on Knowledge Discovery and Data Mining , 2005, 785–790.
- [5] Enrica Vecchio, Beatrice Lazzerini, Steve Foley, Spacecraft Fault Analysis Using Data Mining Techniques, Proc of ISAIRAS Conference, 2005, 5-8
- [6] Hyvarinen, A., Oja, E., Independent component analysis: algorithms and applications, Neural Networks , 2000, 13, 411–430.
- [7] Weiguo Huang, Shuyou Wu, Research on Blind Source Separation for Machine Vibrations, Wireless Sensor Network , 2009, 1, 453-457
- [8] Stock, J. H., Watson, M. W., Forecasting using principal components from a large number of predictors, Journal of the American Statistical Association, 2002,97, 1167–179.
- [9] Georgiev, I., A factor model for innovational outliers in multivariate time series, First Italian Congress of Econometrics and Empirical Economics , Venice, January 2005, 24–25.