Research on Anomaly Detection Method for Satellite Power Supply Based on Bayesian Model

Hui Li¹,²,*, Jing He¹, Fuqiang Cheng²

¹Information and Navigation College, Air Force Engineering University, Xi’an China
²Xi’an Satellite Control Center, Xi’an, China

*Corresponding author email: mykafu0517@163.com

Abstract. Satellite power system is an important system of satellites and it has a significant impact on safe and reliable operations of satellites. Researches on anomaly detection method of satellite power systems can improve the reliability of normal satellite operations. Fixed top and bottom limitations for telemetry parameters are often used in engineering to detect anomalies of telemetry parameters, but the detection thresholds do not change with time and can neither reflect the dynamic tendency of telemetry parameters nor anomalies that occur during dynamic changes of telemetry parameters. In order to solve this problem, this paper puts forward a dynamic detection method for finding anomalous satellite power supply telemetry parameters by comparing predicted values and real-time telemetry data by using dynamic characteristics of satellite power supply telemetry parameters and the Bayesian Model. This method can be used to verify telemetry data of voltage parameters of the storage battery of satellite actually in orbit. The results show that this method can effectively detect anomalies of satellite power supply.

1. Introduction

Power system is an important system of satellites and it has a significant impact on safe and reliable operations of satellites. It is one of the key systems that determine whether a satellite can fulfill its mission. According to the statistics [1], power supply system has caused more than 50% of satellite failures of high-orbit satellites in the United States in the last decade. Satellite power supply telemetry data are important embodiments of running state of satellite power supply. Timely detection of anomalous change data of satellite power system can improve automatic detection and identification capabilities of satellite faults.

In engineering, fixed upper and lower limits of telemetry parameters are often used to detect anomalous values in satellite telemetry data. The detection threshold does not change with time and cannot reflect dynamic trends of telemetry parameters, so it is difficult to capture anomalies in the dynamic change process of telemetry parameters. Therefore, we should build an anomaly detection method for satellite power supply telemetry parameters which can meet dynamic demands. In order to solve this problem, many domestic and foreign experts in this field have also carried out relevant studies. Literature [2] models circuit data of satellite and plane power supply systems through Bayesian network and verifies the feasibility of probability technique in fault diagnosis. Literature [3] demonstrates the application of Bayesian network in satellite fault diagnosis. Literature [4] uses the principle that
telemetry parameters of adjacent cycles are equal or close to build an Auto Regression Moving Average model so as to obtain predicted values of telemetry parameters. It conducts anomaly detection by comparing the measured value and the predicted value. Literature [5] puts forward a neural network method for Particle Swarm Optimization. The neural network optimized by particle swarm algorithm is used to approximation and modelling of key telemetry parameters of the satellite. After that, the predicted time sequence obtained and the measured values are compared to realize anomaly detection. Researches find that the ARMA model can be established to learn and predict stable change parameters or parameters near a constant value in a historical period. However, the satellite power system has a wide variety of parameters, including many kinds of non-stationary parameters. Neural network can well predict Time Sequence of telemetry parameters, but it needs to constantly improve the weight to achieve the goal of accurate prediction. On the basis of that, this paper puts forwards a method based on Bayesian network which can simply and effectively realize anomaly detection of satellite power supply telemetry parameters dynamically. This method has been used for many times in many tests and this verifies the effectiveness of anomaly detection in satellite power supply parameters.

2. Anomaly detection strategies based on Bayesian Model

2.1. Basic principles

Bayesian classification is an independent assumption based on Bayesian theorem and characteristic conditions [6]. A comparative study of classification algorithms shows that naive Bayesian classification is comparable to decision tree and neural network classifier selected. Bayesian classification algorithms also show high accuracy and high speed for large data.

Suppose X is a DataSet and H is some hypothesis, for example DataSet X belongs to a certain class C, then the probability of H is as follows under the given DataSet conditions:

\[ P(H \mid X) = \frac{P(HX)}{P(X)} \]  

In the formula, \( P(H) \) means prior probability, \( P(H \mid X) \) means posterior probability. Bayesian theorem is given directly without proving it:

\[ P(H \mid X) = \frac{P(X \mid H)P(H)}{P(X)} \]  

Next, let’s study naive Bayesian classification. Suppose \( X \) is the \( n \)-dimension feature vector of input space \( X \subseteq \mathbb{R}^n \), \( Y \) a random variable of the output space and \( P(X,Y) \) is the joint probability distribution of X and Y. The training sample \( T = \{(x_1,y_1),(x_2,y_2),...,(x_n,y_n)\} \) is the independent and identically distributed \( P(X,Y) \).

Naive Bayesian classification learns the joint probability distribution \( P(X,Y) \) through training samples. We mainly learn the joint probability distribution by prior probability \( P(Y = c_k), k = 1, 2, ..., K \).

According to the assumption of conditional independence, posterior probability can be calculated according to Bayesian theorem [7]. The conditional independent hypothesis is:

\[ P(X = x \mid Y = c_k) = \prod_{i=1}^{n} P(X^{(i)} = x^{(i)} \mid Y = c_k) \]  

The posterior probability is:
\[ P(Y = c_k \mid X = x) = \frac{P(X = x \mid Y = c_k)P(Y = c_k)}{\sum_k P(X = x \mid Y = c_k)P(Y = c_k)} \]  \tag{4}

Substitute equation (3) into equation (4) to get:

\[ P(Y = c_k \mid X = x) = \frac{P(Y = c_k) \prod_{i=1}^{n} P(X^{(i)} = x^{(i)} \mid Y = c_k)}{\sum_k P(Y = c_k) \prod_{i=1}^{n} P(X^{(i)} = x^{(i)} \mid Y = c_k)} \]  \tag{5}

In posterior probability, because the denominator is the sum of all \( c_k \), the Bayesian classifier can be expressed as:

\[ y = \arg \max_{c_k} P(Y = c_k) = P(Y = c_k) \prod_{j} P(X^{(i)} = x^{(i)} \mid Y = c_k) \]  \tag{6}

The naive Bayesian classification algorithm is given below.

In training samples \( T = \{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\} \), \( x_j = (x_j^{(1)}, x_j^{(2)}, \ldots, x_j^{(n)})^T \), \( x_j^{(i)} \) means the j feature \( x_j^{(i)} \in \{a_{j_1}, a_{j_2}, \ldots a_{j_{S_j}}\} \) of sample \( i \); \( a_j \) refers to possible value \( l \) of feature \( j \) and \( j = 1, 2, \ldots, n \), \( l = 1, 2, \ldots, S_j \), \( y \in \{c_1, c_2, \ldots, c_k\} \). The prior probability and the conditional probability shall be calculated firstly:

\[ P(Y = c_k) = \frac{\sum_{i=1}^{N} I(y_i = c_k)}{N}, k = 1, 2, \ldots, K \]

\[ P(X^{(i)} = a_{jl} \mid Y = c_k) = \frac{\sum_{i=1}^{N} I(X^{(i)} = a_{jl}, Y = c_k)}{\sum_{i=1}^{N} I(Y = c_k)} \]  \tag{7}

\[ j = 1, 2, \ldots, n, l = 1, 2, \ldots, S_j, k = 1, 2, \ldots, K \]

With the given characteristics \( x = [x^{(1)}, x^{(2)}, \ldots, x^{(n)})^T \), we will calculate

\[ P(X = x \mid Y = c_k) = \prod_{i=1}^{n} P(X^{(i)} = x^{(i)} \mid Y = c_k) \]  \tag{8}

Finally, the category of feature \( x \) will be determined

\[ y = \arg \max_{c_k} P(Y = c_k) \prod_{j} P(X^{(i)} = x^{(i)} \mid Y = c_k) \]  \tag{9}
2.2. Dynamic detection strategy for satellite power supply telemetry data thresholds based on Bayesian Model

Satellite power supply telemetry data are arranged in chronological order [8-11]. Therefore, telemetry parameters should be used to construct telemetry data time series. The anomaly detection strategy based on Bayesian Model is used to establish sample sequence data samples by selecting time sequence of satellite power system parameters. Then, the sample set composed of sample sequence input vector and output value is obtained. The Bayesian classification model is used to predict output values of the input vector and then, the Lagrange interpolation algorithm is used to build the interpolating function of predicted values and the predictive continuous function with time as independent variable is obtained. Dynamic anomaly detection was conducted on the data according to the difference degree between the predicted result and the actual value so as to establish the dynamic threshold of telemetry data and realize the anomaly detection of satellite power supply telemetry data.

The time sequence point of satellite power supply telemetry parameters is \( \{ t, s(t) \} \), in which it means denote time; \( t \in [t_s, t_e] \), \( s(t) \) represents parameter values at the corresponding time.

1) Timing alignment and generation of equally spaced sample sequence

Give the starting time of sample sequence \( t_s \) and build the uniformly-spaced sample sequence \( \{ (t, x(t)) \mid t = t_s + p \cdot t_d, p = 0, 1, 2, \ldots P - 1 \} \) according to regular intervals \( t_d \), \( t_d < t_e - t_s \). \( X(t) \) is the sample value at the corresponding time and no value is assigned at its initialization; \( P \) means the length of the sample sequence. \( P = \text{floor} \left( \frac{t_e - t_s}{t_d} \right) \), floor function means we round down.

According to the time scale of sample sequence \( x(t) \), the moment closest to the sample timings will be found in the original sequences \( t \) records and the data values at that moment is extracted as the sample value. After that, the sample sequence \( \{ x(t_s + p \cdot t_d) \mid p = 0, 1, 2, \ldots P - 1 \} \) with values assigned will be obtained.

2) Model sample preparation

As for one dimensional nonlinear time series \( x(t), t \in [1, n] \), if we want to predict values of \( x(n+1) \) and \( x(n+2) \ldots \), first, we need to construct structural forms of the prediction model, which means the input pattern of sequential inputs (input nodes) of many known sequence values and the output pattern of expected outputs (output nodes).

Take sequence \( \{x(0), x(1) \ldots x(n)\} \) as an example; the number of input nodes of the machine learning model is recorded as \( L_i \) and the number of output nodes is \( L_o \), the sample interval is selected as \( d_s \). The sample selection process is shown as Figure 1.

![Figure 1. Sample generation of time series](image)

From the above process, the input array of model sample can be obtained:
Output (target) array of the model sample:

\[
Y = \begin{bmatrix}
Y_0 \\
\vdots \\
Y_m \\
\vdots \\
Y_p
\end{bmatrix} = \begin{bmatrix}
x(L_1 + L_o) & \ldots & x(L_1 + L_o + ds - 1) \\
\vdots & \ddots & \vdots \\
x(m \cdot (L_i + L_o + ds) + L_i) & \ldots & x(m \cdot (L_i + L_o + ds) + L_i + L_o - 1) \\
\vdots & \ddots & \vdots \\
x(P \cdot (L_i + L_o + ds) + L_i) & \ldots & x(P \cdot (L_i + L_o + ds) + L_i + L_o - 1)
\end{bmatrix}, \quad (11)
\]

Here, the length of the input vector, which is, the number of elements, is set as Lin; the number of output nodes is 1. Lin should be less than P, then, \([x(t_i), x(t_i + t_d), \ldots, x(t_i + (L_{in} - 1) \cdot t_d)]^T\) is the first input vector and recorded as X1; \(x(t_i + L_{in} \cdot t_d)\) is the first output value, which is denoted as Y1.

The sample selection interval is 1, and so on. We can get the input vector m of Xm as \([x(t_i + (m - 1) \cdot t_d), \ldots, x(t_i + (m + L_{in} - 2) \cdot t_d)]^T\) and the m output value of Ym as \(x(t_i + (m + L_{in} - 1) \cdot t_d), m \in \{1, 2, \ldots, P - L_{in}\}\) and the total number of samples is N, then N=P−Lin.

Establish the training sample set \(T = \{(X_1, Y_1), (X_2, Y_2), \ldots, (X_N, Y_N)\}\) to prepare data for Bayesian modeling.

(3) Bayesian modeling

If the input vector of Sample I is \(X_i = \left[x_i^{(1)}, x_i^{(2)}, \ldots, x_i^{(L_m)}\right]^T\), \(x_i^{(j)}\) means element j of input vector I, \(x_i^{(j)} \in \{a_1, a_2, \ldots, a_s\}\), \(s = 1, 2, \ldots, S\). The set of as represents the possible values of x (j), and its number is S. The output value is \(Y \in \{c_1, c_2, \ldots, c_k, \ldots, c_K\}\). The set of ck represents the possible values of Y and its number is K.

Calculate the prior probability of all possible values in the output value Y.

\[
P(Y = c_k) = \frac{\sum_{i=1}^{N} I(Y_i = c_k)}{N}, k = 1, 2, \ldots, K, \quad (12)
\]

Among which, \(I(Y_i = c_k)\) means that if Yi is ck, 1 will be used, or else, 0 will be used.

Calculate the posterior probability of all possible values of all elements in X sequence when considering all possible values of Y. When Y=ck, element j of X can meet the following formula:
\[
P(x^{(j)} = a_s \mid Y = c_k) = \sum_{i=1}^{N} I(x_i^{(j)} = a_s, Y_i = c_k) = \frac{\sum_{i=1}^{N} I(Y_i = c_k)}{\sum_{i=1}^{N} I(Y_i = c_k)} \quad (13)
\]

Next, calculate predicted values of input vectors in turn. Take the input sequence \( i X_i \) as an example. First, calculate the probability of \( X = X_i \) under different values of \( Y \), that is

\[
P(X = X_i \mid Y = c_k) = L_{in} \prod_{j=1}^{L_{in}} P(x^{(j)} = x_i^{(j)} \mid Y = c_k) \quad (14)
\]

Then, determine the prediction result \( \hat{y}_i \) of input vector \( X_i \) and the probability maximization is the minimization of structural risks.

\[
\hat{y}_i = \underset{c_k}{\text{arg max}} \ P(Y = c_k) \prod_{j=1}^{L_{in}} P(x^{(j)} = x_i^{(j)} \mid Y = c_k) \quad (15)
\]

The prediction result \( \hat{y}_i \) means the predicted value at \( t = t_s + (i + L_{in} - 1) \cdot t_d \). Do the previous step and we can get the corresponding predicted value at time \( \{ t = t_s + q \cdot t_d \mid q = L_{in}, L_{in} + 1, \ldots, P - 1 \} \), which is noted as \( \{ \hat{x}(t_s + q \cdot t_d) \mid q = L_{in}, L_{in} + 1, \ldots, P - 1 \} \). Due to the limitation of the algorithm, the value at \( t < t_s + L_{in} \cdot t_d \) cannot be detected.

(4) Lagrange interpolation of the predicted value

The model evaluation function is established and the RMSE of root mean square error is defined as:

\[
RMSE = \sqrt{\frac{1}{P - L_{in}} \sum_{j=L_{in}}^{P-1} \left( x(t_s + j \cdot t_d) - \hat{x}(t_s + j \cdot t_d) \right)^2} \quad (16)
\]

It represents the dispersion degree of prediction error, which is also called as standard error. The best case is RMSE=0, which is one of the comprehensive indexes of error analysis.

Do Lagrange interpolation to prediction sequence \( \hat{x} \) and the interpolation polynomial of \( t \) \( t \in [t_s + L_{in} \cdot t_d, t_s + (P - 1) \cdot t_d] \) is \( \hat{L}(t) \):

\[
\hat{L}(t) = \sum_{L_{in}}^{P-1} \left( \hat{x}(t_s + j \cdot t_d) \cdot \prod_{j=L_{in}, j \neq i}^{P-1} \frac{t - (t_s + j \cdot t_d)}{(i - j) \cdot t_d} \right) \quad (17)
\]
\( \hat{L}(t) \) Represents the continuous function of the predicted value.

(5) Dynamic threshold generation. If the training error of the model is RMSE, the detection threshold will be set as \( \kappa \). If \( \kappa > 0 \), the upper limitation \( L_u(t) \) and the lower limitation \( L_l(t) \) of the detection threshold are set as:

\[
L_u(t) = \hat{L}(t) + \kappa \cdot \text{RMSE} \\
L_l(t) = \hat{L}(t) - \kappa \cdot \text{RMSE}
\]  

Detect anomalous points. For any point in the time sequence \( s(t) \), \( t \in [t_s + L_{\text{in}} \cdot t_d, t_s + (P - 1) \cdot t_d] \), the condition for judging whether it is anomalous or not is that if the deviation degree between the predicted value and the actual value exceeds a certain multiple of the model RMSE, it should be predicted that the actual value is anomalous, that is

\[
\begin{align*}
\text{Normal value:} & \quad L_l(t) \leq s(t) \leq L_u(t) \\
\text{Abnormal value:} & \quad s(t) < L_l(t) \text{ or } s(t) > L_u(t)
\end{align*}
\]  

3. Experimental Tests

By analyzing satellite power supply parameters, telemetry data from 2018-07-01 to 2018-07-31 were selected to verify the voltage parameters of the 70Ah battery of a certain satellite. The specific flow chart is shown in Figure 2.

![Figure 2. Anomaly detection flow chart of the satellite power system](image-url)
(2) The length of the input vector is set as 100, and the sample selection interval is 1. The sample set T is established according to the method in step 2.2 (2) and a total of 8828 samples are obtained.

(3) The probability values of different output values are calculated according to step 2.2 (3). The conditional probability of each element in the input vector under different numerical conditions is calculated. On this basis, the predicted value of each input vector can be calculated to form a "predicted value" sequence \( \{ \hat{x}(t) \} \).

(4) Do Lagrange's interpolation to the predicted value according to step 2.2 (4). The RMSE of the predicted results is calculated, and the detection threshold is set $\kappa=2$. Upper and lower detection thresholds were generated according to step 2.2 (5).

(5) The time sequence of the original telemetry data are tested according to formula (19). The detection effect of sometime intervals is shown in Figure 3 and Figure 4. The actual telemetry values are shown in black what the yellow and purple lines show is the dynamic upper and lower limitations generated. The blue line in Figure 5 shows the model residuals, that is, the deviation between the predicted value and the actual value and the red line shows the upper and lower line of the threshold. It can be seen that the dynamic threshold reflects changes of the actual value change well and can cover parameter trend well. Actual values of parameters of Figure 3 and Figure 4 present the variation trend of the approximate sine function, but in Figure 3, from around 21:00 of 7.06 to around 22:00, the actual value becomes constant, which obviously does not conform to the variation law of the parameters. The value of this region is marked as an exception (red dot). Figure 4 shows that the actual value abruptly changes at about 10:48 of 7.15, which obviously does not conform to the law of parameter change. The value of this point is marked as anomalous (red dot). Figure 5 shows model residuals, which corresponds to Figure 4. At about 10:48 of 7.15, the model residual exceeds the set threshold and the actual value corresponding to this dot is marked as anomalous.

**Table 1. Telemetry parameters data of a satellite**

| Time 2018-07-06 | Parameter values | Time 2018-07-06 | Parameter values | Time 2018-07-06 | Parameter values |
|-----------------|------------------|-----------------|------------------|-----------------|------------------|
| 08:00:01.230    | 26.25            | 09:12:01.230    | 26.35            | 10:24:01.230    | 24.1             |
| 08:12:01.230    | 25.15            | 09:24:01.230    | 26.95            | 10:36:01.230    | 23.6             |
| 08:24:01.230    | 24.65            | 09:36:01.230    | 25.35            | 10:48:01.230    | 25.45            |
| 08:36:01.230    | 23.8             | 09:48:01.230    | 26.35            | 11:00:01.230    | 26.15            |
| 08:48:01.230    | 24.95            | 10:00:01.230    | 25.35            | 11:12:01.230    | 26.65            |
| 09:00:01.230    | 25.95            | 10:12:01.230    | 24.7             | ……             | …… |

**Figure 3. A comparison diagram of predicted values and actual values**
4. Conclusion
By analyzing defects of the static threshold method currently adopted in anomaly detection and dynamic characteristics of satellite power system telemetry data, this paper puts forward an anomaly detection method based on dynamic changes of telemetry data and gives its theoretical bases. Moreover, the anomaly detection effect is verified according to voltage parameter data of a satellite's power supply system. The results show that this method can detect anomalous conditions of satellite power supply telemetry data in real time according to dynamic changes of parameters. The method can do anomaly detection of data effectively and avoid failure in reports, which is conducive to finding anomalous data from time sequence of satellite power supply telemetry parameters and assisting satellite power system in troubleshooting and positioning parameters when doing anomalous analysis.

References
[1] Bingxin, Y.; Jinyuan, G.; and Tao, H. Introduction to the Latest High Reliability Space Power Standard in USA, Aerospace Standardization, 2013, 1, 38 - 41.
[2] Mengshoel, B.O.J.; Darwiche, A.; Cascio, K.; et al. Diagnosing Faults in Electrical Power Systems of Spacecraft and Aircraft. National Conference on Innovative Applications of Artificial Intelligence, AAAI Press, 2008, 620 - 630.
[3] Abhinav, S.; Indranil, R.; and Celaya, J. Requirements specifications for prognostics: an overview, American Institute of Aeronautics and Astronautics, 2004, 1 - 6.
[4] Weizheng, L.; Qiao, Meng. Satellite Anomaly Detection Method Based on Dynamic Characteristics of Telemetry Data, Chinese Journal of Space Science, 2014, 34 (2), 201 - 207.
[5] Hongzheng, F.; Hui, S.; Liming, H. Satellite Fault Prediction Method Based on Particle Swarm Optimization Neural Network, Computer Measurement and Control, 2013, 21 (7), 1730 - 1733, 1745.
[6] Campos, L. M. D.; Juan, M. Fernández-Luna; and Huete, J. F. Clustering terms in the Bayesian network retrieval model: a new approach with two term-layers, Applied Soft Computing, 2004, 4 (2), 149 – 158.
[7] Chundi, M.; Jianbin, D.; Jun, Y. Bayesian Networks for Data Mining, Journal of Software, 2000, 11 (5), 660 - 666.
[8] Xiyuan, P.; Jingyue, P.; and Yu, P. Review on Anomaly Detection of Spacecraft Telemetry Data, Chinese Journal of Scientific Instrument, 2016, 37 (9), 1923 - 1945.
[9] Ying, D.; Sihu, L.; Shaolin, H. Anomaly Detection Algorithm for Subsequence of Telemetry
Pseudo-periodic Time Sequence, Chinese Space Science and Technology, 2016, 36 (5), 18 - 24.

[10] Dong, Z.; Shaolin, H.; and Ningli, Y. Research on Anomaly Detection Algorithm of Spacecraft Downlink Data, Journal of Aircraft Measurement and Control, 2013, 32 (2), 123 - 126.

[11] Gang, Z.; Yanjun, L. Spacecraft Fault Diagnosis Method based on Time Series Data Mining, Journal of Aircraft Measurement and Control, 2010. 23 (3), 1 - 5.

[12] Zhihua, Z. Machine Learning, Beijing: Tsinghua University Press, 2016, 147 - 156.