



SPACECRAFT TEST DATA INTEGRATION MANAGEMENT TECHNOLOGY BASED ON BIG DATA PLATFORM

NANQI GONG*

Abstract. In this paper, a general test platform for spacecraft data management is designed and constructed. This paper introduces a portable software development environment based on LUA. The technology of space environment data management, comprehensive analysis, parameter correction and visual display of spacecraft is realized. The relationship between continuity, mixed dispersion, variation and indication of remote sensing data is studied. This project uses the integrated Long Short Term Memory network (LSTM) technology to detect anomalies in satellite remote sensing observation data. Give full play to the advantages of laser scanning tunneling microscope in the nonlinear field. The combination of this method and the matrix method can improve the adaptive ability of spacecraft in an operation state to better identify abnormal information in remote sensing data. Experiments show that the algorithm can significantly improve the anomaly detection rate of the system. The system can monitor the front test device and record the data. The method can be connected with the space vehicle's central control and automatic test system. The comprehensive management of the integrated test system of space vehicles is realized.

Key words: Spacecraft; Front-end test equipment; Telemetry; Anomaly detection; LSTM

1. Introduction. The spacecraft data processing subsystem (DMS), as the "brain" of spacecraft, plays an essential role in space information processing. DMS undertakes remote sensing measurement in aviation systems, such as remote sensing data collection and transmission, remote control command receiving and execution, flight plan maintenance and space flight routine work. Each subsystem on the spacecraft is subject to distributed control management, so the stability and reliability of distributed control management are crucial to the successful operation of the spacecraft. DMS functional testing is significant. Spacecraft face a highly complex space environment when they fly in space [1]. It involves all kinds of particle radiation, electromagnetic radiation, atomic oxygen, temperature changes, etc. Real-time monitoring of the space environment and obtaining relevant space environment information have practical application value and prospects. It provides a good condition and basis for the country's high-life and high-quality artificial satellites. At the same time, the spacecraft data processing subsystem will also provide necessary data support for space weather prediction and scientific research [2]. Promoting the international exchange of space environment detection technology and data in China is of great significance.

Scientific space environment monitoring is significant to studying astrophysics and planetary evolution. There has been a long history in the field of automatic detection of space vehicles. The software and hardware products on the space vehicle have formed a relatively complete automatic detection method library [3]. The system can realize the function check and reliability test of hardware equipment and software. Some scholars have proposed a ground test device for space vehicles based on artificial intelligence. This system has an automatic discriminating ability similar to conditional criteria. The system is applied to the automatic verification test of the power supply and distribution equipment of space vehicles, which can improve the verification test effect of the equipment. In recent years, the automatic detection technology of spacecraft has been dramatically developed in China. The method has been successfully used in space fields such as human-crewed spaceflight and the Beidou satellite. However, the current automated test system design at home and abroad mainly focuses on the execution of spacecraft test sequences. Most related systems are configured at the back-end level of the EGSE system. Little attention has been paid to the automatic operation of the front-end test unit. Although some automatic detection systems have detection devices, they can only complete the command control of the detection devices. It cannot obtain the working condition of the device under test in time. The device has

*The University of Hong Kong, Hong Kong (vanessagong2001@gmail.com)

a good coupling relationship with the tested space vehicle. In addition, the system is highly dependent on hardware. The development of a new standardized intelligent test platform can fully meet the testing needs of space agencies. The system platform designed in this paper abstracts the test business of the spacecraft data management subsystem, thus completing the standardization of the organization structure, test flow and communication protocol of the spacecraft data management subsystem test system [4]. At the same time, introducing the LUA scripting language makes the test program customization more intelligent and standardized. The test bed is also a reconfigurable system. The system can be tested under the imperfect spacecraft data processing subsystem architecture.

2. Design of automatic management system.

2.1. System Software Architecture. Based on the mature integrated automatic detection technology, an automatic detection technology for spacecraft front-end detection devices is proposed. This system uses Visual Studio software. The software is designed according to the modularization idea [5]. Divide the system functions to be completed into several functional modules. The data interface between each module is straightforward. It has good expansibility in device type and function control. Figure 2.1 shows the architecture of the entire system. It is mainly divided into six parts: 1) The network communication part completes the network interface function of the system. A network connection is provided between the managed front-end test device and the master control master test processor (MTP). Receive the health information sent by the managed device and process the content of the information according to the agreement. Send the control command to the device and get a reply. 2) The data analysis function module retrieves packets sequentially from a valid packet cache queue. According to the designed packet format, the spacecraft identifier, the local time of the sending time of the sender, the device identifier of the sender, the device identifier of the receiver, the information identifier, the cumulative count of the packet sending of the sender, and the data content are analyzed. The identification of this information determines the status of this level. And write the received data into a log file. 3) The instrument control module is dynamically maintained as a data link table. Managed devices can be added or removed via the system's software interface or profile. Device information includes device name, device number, control instruction name, control instruction number, IP address, port number, etc. The network communication module receives the control indicator and analyzes the indicated content. The device generates an indication response according to the communication protocol and returns a message to the network communication module. 4) The program's sequence implementation part reads the instruction sequence selected by the human-machine interface. According to the command number and the device ID, it is sent to the device management module as the corresponding command. You can set the time interval for instruction execution, the number of loops, and boundary conditions. 5) The data file module records the diary data generated by the system in its operation and the data received by the managed device. At the same time, it makes an instant recording. 6) Display/operation module displays the working status of the system. Various types of status information transmitted by the managed device respond to the tester's input. It is a kind of management system software with a man-machine dialogue function.

2.2. System module design.

2.2.1. Product structure module. According to the four levels of whole, cabin, cabin board and equipment, the prototype is constructed to realize the whole structure of the product. The process prototype is summarized from the system's overall structure, final assembly design pattern, subsystem design pattern and related design documents. The system was produced during the integration process. The creation process is shown in Figure 2.2.

2.2.2. Spacecraft data transmission module. The same information is often recorded in multiple places during the AIT process. Such duplication of data leads to "data contradictions." This situation has an impact on the tracking of the process. Therefore, the system should follow the principle of consistency of data sources. Ensure all data is transferred from existing systems or entered via files etc. According to different information systems, the corresponding communication mode and interface are designed in this paper [6]. The system is integrated through network service, ETL, intermediate table synchronization and other methods to determine whether the transmitted data is complete. The data management system itself only

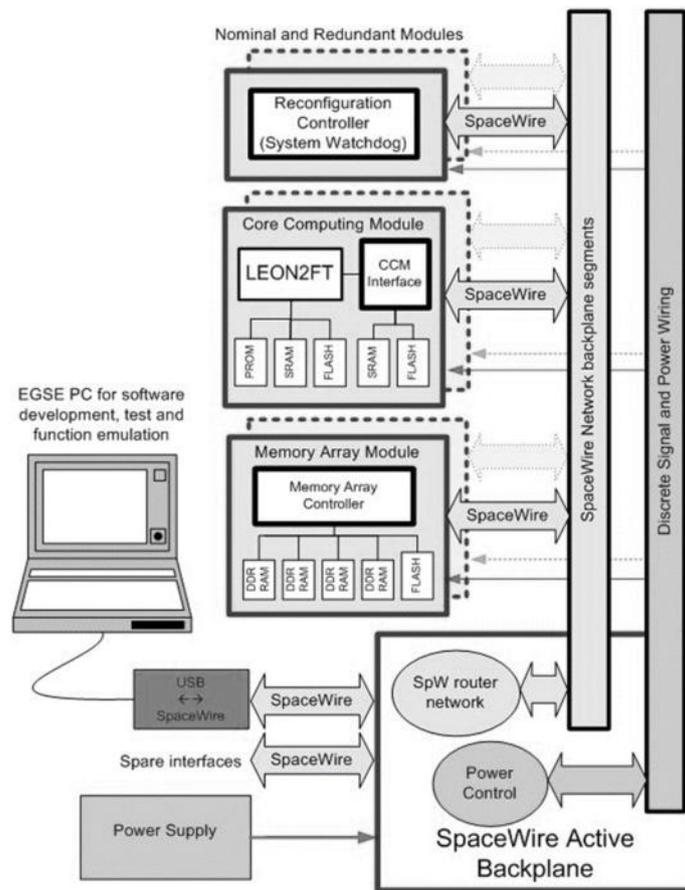


Fig. 2.1: Software module and interface of the front-end detection device AIS.

transfers data from other data systems. The data outside the work network is entered into the corresponding information system through the physical media. Then use the data interface of the office network to transmit this information. See inter-system data integration relationship Figure 2.3 (image cited by the National Aeronautics and Space Administration).

2.2.3. Space data reconstruction and contact module. The data reconstruction and association module reconstructs process-based data. Connect it to the "Product Architecture Framework" to complete the data content filling. Figure 2.4 shows the process flow.

The information in the message display system comes from different systems. Benchmarking and versioning the data in the process is necessary to ensure the integrity and accuracy of AIT process data [7]. Keep the same version of data for each dimension. This can effectively prevent the query error caused by the difference in data.

2.2.4. Data Application Module. The data module is used for statistics and a summary of the forms of each module. It includes all biased data, custom queries and data mining. The retrievability and scalability of AIT data are realized through the configuration and management of big data such as Impala. Thus, the vertical comparison between the same model and the horizontal comparison between the different models is achieved. For example, it can quickly compare the execution records of each disassembly and site photos in multiple disassembly operations of an important part. The assembly situation and data of the same type of equipment on the same batch of satellites are easily compared.

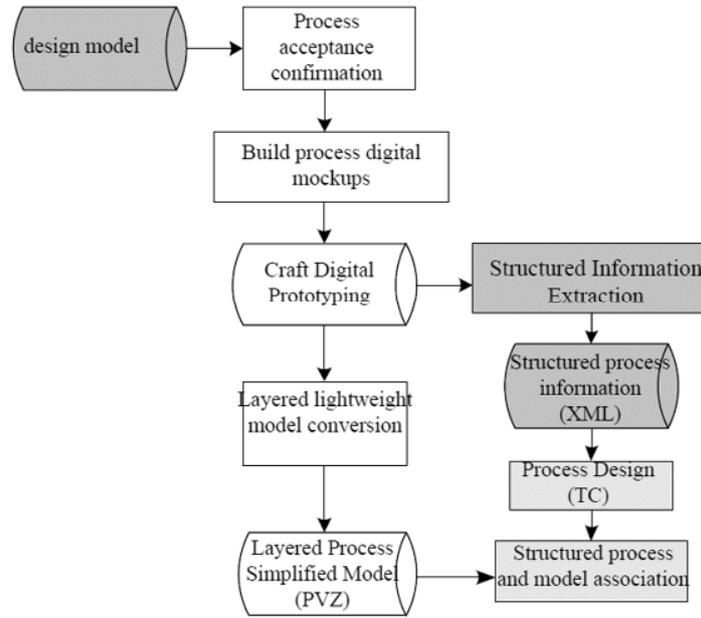


Fig. 2.2: Data generation process based on product structure.

3. Anomaly detection of spacecraft remote sensing data based on LSTM. LSTM model neglects the modal characteristics of remote control commands when establishing remote control commands. This will affect the accuracy of the forecast. This dramatically impacts some problems existing in satellite remote sensing data. This project plans to conduct research from three aspects: 1) Feature extraction of control commands. 2) Multi-LSTM model training. 3) Anomaly detection method based on remote sensing data. This paper presents an algorithm for anomaly detection of remote sensing data using an integrated LSTM prediction model (Figure.3.1).

3.1. Excavation of Control command pattern. The algorithm is performed before the LSTM prediction model is started. The first is the preprocessing of these training samples. The method of cluster analysis is used to classify them [8]. Among them, the preprocessing of training data and the mining of remote indication control mode are the focus of this paper.

3.1.1. Training data preprocessing. The training set R can be regarded as a set of n dimensional vectors at time t . $r^t = \{r_1^t, r_2^t, \dots, r_n^t\}$. Where $\{r_1^t, r_2^t, \dots, r_{n-1}^t\}$ is the multi-dimensional remote control command at time t . $\{r_n^t\}$ is the telemetry data at time t . Then the expression R for the training set of telemetry data containing a subsystem is:

$$R = \left\{ \left[\begin{array}{c} r_1^1 \\ \vdots \\ r_{n-1}^1 \\ r_n^1 \end{array} \right], \left[\begin{array}{c} r_1^2 \\ \vdots \\ r_{n-1}^2 \\ r_n^2 \end{array} \right], \dots, \left[\begin{array}{c} r_1^t \\ \vdots \\ r_{n-1}^t \\ r_n^t \end{array} \right], \dots \right\} \quad (3.1)$$

Then R is reconstructed. Because the telemetry data at the next moment is related to the remote command at the next moment, and the telemetry data at the previous time is related to the corresponding remote command [9]. Therefore, when constructing the input, \tilde{R} is divided into multiple submatrices R_m^* containing continuous-time vectors. Each submatrix contains telemetry data and global remote control instructions within s_t time. Input \tilde{R} for reconstruction is represented as follows

$$\tilde{R} = \{\{R_1^*\}, \{R_2^*\}, \dots, \{R_m^*\}, \dots\} \quad (3.2)$$

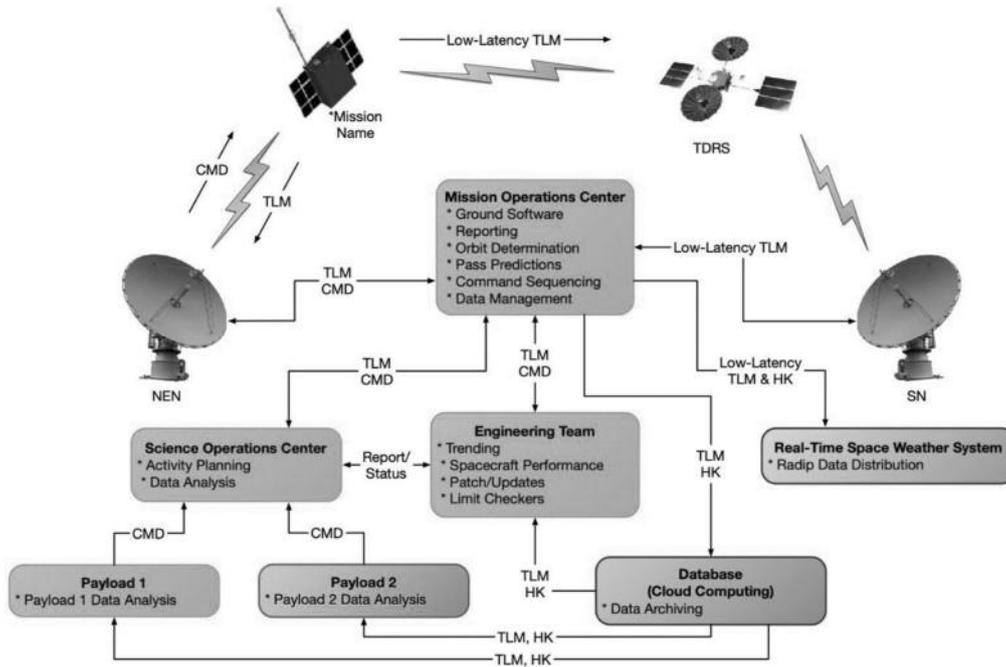


Fig. 2.3: Data integration relationships between different systems.

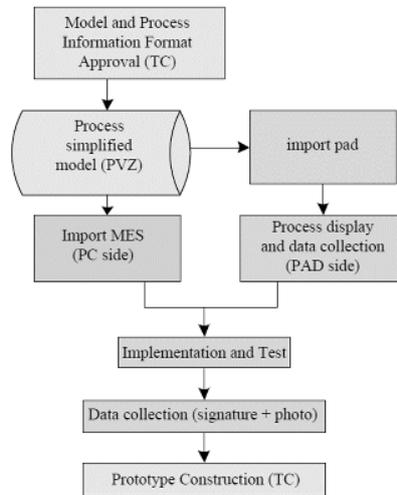


Fig. 2.4: Data reconstruction and association process.

Among them:

$$R_m^* = \left\{ \begin{bmatrix} r_1^m \\ r_2^m \\ \vdots \\ r_n^m \end{bmatrix}, \begin{bmatrix} r_1^{m+2} \\ r_2^{m+2} \\ \vdots \\ r_n^{m+2} \end{bmatrix}, \dots, \begin{bmatrix} r_1^{s_t+m} \\ r_2^{s_t+m} \\ \vdots \\ r_n^{s_t+m} \end{bmatrix} \right\} \quad (3.3)$$

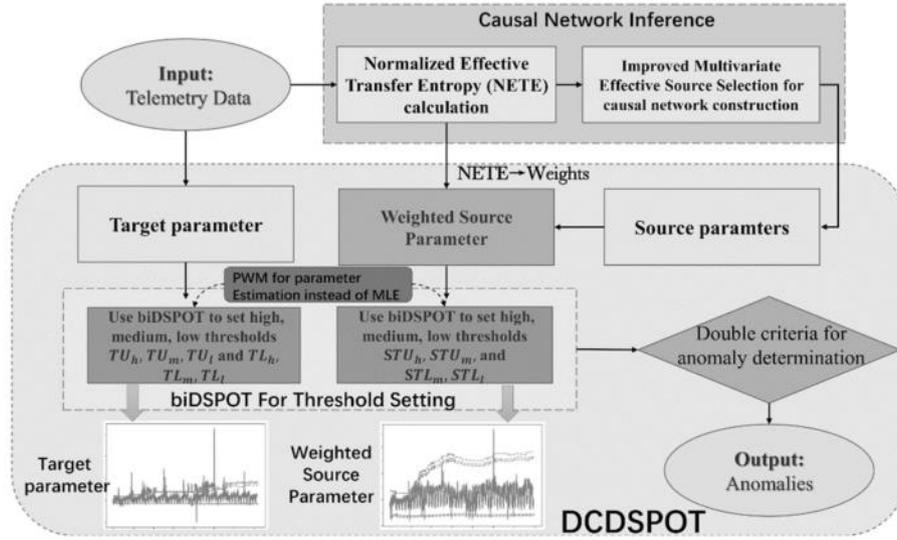


Fig. 3.1: Anomaly detection method of remote sensing data integrated with LSTM prediction model.

3.1.2. Remote command control mode mining. Take input \tilde{R} as an example to extract the remote control instruction D_m^* in each submatrix R_m^* . In this way, the control instruction matrix D^* corresponding to \tilde{R} is obtained.

$$D_m^* = \left\{ \left[\begin{array}{c} r_1^m \\ r_2^m \\ \vdots \\ r_{n-1}^m \end{array} \right], \left[\begin{array}{c} r_1^{m+1} \\ r_2^{m+1} \\ \vdots \\ r_{n-1}^{m+1} \end{array} \right], \dots, \left[\begin{array}{c} r_1^{s_t+m} \\ r_2^{s_t+m} \\ \vdots \\ r_{n-1}^{s_t+m} \end{array} \right] \right\} \quad (3.4)$$

By calculating the S_2 norm of D^* submatrix. In this way, a set of control mode feature vector $\|D^*\|_2$ is obtained.

$$\|D^*\|_2 = \{\|D_1^*\|_2, \|D_2^*\|_2, \dots, \|D_m^*\|_2, \dots\} \quad (3.5)$$

$$\|D_m^*\|_2 = \sqrt{\eta_{\max}(D_m^{*T} D_m^*)} \quad (3.6)$$

Reorder the values in the control mode feature vector $\|D^*\|_2$. The median in $\|D^*\|_2$ is taken as the threshold by dividing it into two sub-training sets. The similar sequence of centralized remote instructions divides the training set into 2 classes.

3.2. Multi-LSTM prediction model training. LSTM prediction model respectively trained the two sub-training sets obtained by pattern mining and clustering. For example, the reconstructed training matrix R is used as input for telemetry data prediction. Let $t = s_t$, the prediction process is as follows:

$$R_m^* = \left\{ \left[\begin{array}{c} r_1^m \\ r_2^m \\ \vdots \\ r_n^m \end{array} \right], \dots, \left[\begin{array}{c} r_1^{m+t} \\ r_2^{m+t} \\ \vdots \\ r_n^{m+t} \end{array} \right] \right\} \rightarrow \left[\begin{array}{c} r_1^{m+t+1} \\ r_2^{m+t+1} \\ \vdots \\ \hat{f}(x)^{m+t+1} \end{array} \right] \quad (3.7)$$

$\hat{f}(x)^{m+t+1}$ is the predicted value at time $m+t+1$. When constructing the prediction model, the absolute error function (MAE) is the loss function of the LSTM prediction model:

$$MAE(f(x), \hat{f}(x)) = \frac{\sum_{i=1}^m |f(x)_i, \hat{f}(x)_i|}{m} \quad (3.8)$$

m is the number of total predicted values. $f(x)_i$ is the i true value.

The optimal solution of LSTM prediction model is Adam. The sensitivity of the method to non-normal changes is improved [10]. To avoid overfitting, the memory modules of each hidden layer cannot enter the LSTM learning. This approach is also known as dropout. LSTM prediction model P and LSTM prediction model Q are obtained after all the sub-training sets are trained.

3.3. Telemetry data anomaly detection model. After the training of the LSTM prediction model P and LSTM prediction model Q . The test data will be separately passed through each prediction model. Two prediction sequences were obtained [11]. The final prediction sequence is obtained by integrating two groups of prediction sequences according to specific weights.

The test data is reassembled using training set \tilde{R} as an example [12]. The reconstructed test set was simultaneously predicted by LSTM prediction model P and LSTM prediction model Q to obtain two sets of prediction sequences. Telemetry data prediction set $P, \hat{f}(x)_P = \{\hat{f}(x)_P^1, \hat{f}(x)_P^2, \dots, \hat{f}(x)_P^t, \dots\}$ and telemetry data prediction set $Q, \hat{f}(x)_Q = \{\hat{f}(x)_Q^1, \hat{f}(x)_Q^2, \dots, \hat{f}(x)_Q^t, \dots\}$ respectively. The prediction sequence \hat{Y} is obtained by integrating it with the weight matrix ζ .

$$\hat{f}(x) = \begin{Bmatrix} \hat{f}(x)_P^1 & \hat{f}(x)_Q^1 & \dots & \hat{f}(x)_P^t & \dots \\ \hat{f}(x)_P^2 & \hat{f}(x)_Q^2 & \dots & \hat{f}(x)_P^t & \dots \\ \dots & \dots & \dots & \dots & \dots \\ \hat{f}(x)_P^t & \hat{f}(x)_Q^t & \dots & \hat{f}(x)_P^t & \dots \\ \dots & \dots & \dots & \dots & \dots \end{Bmatrix} \quad (3.9)$$

R_t^* as some input at time t . The corresponding remote control instruction matrix D_t^* obtains the predicted value $\hat{f}(x)_P^{t+s_t+1}$ at time $t+s_t+1$ through the prediction model P . D_t^* obtains the predicted value $\hat{f}(x)_Q^{t+s_t+1}$ at time $t+s_t+1$ through the prediction model Q . If D_t^* corresponds to $S2$ norm $\|D_t^*\|_2$ belongs to the training set P , then:

$$\hat{f}(x)^{t+s_t+1} = \zeta^{t+s_t+1} \hat{f}(x)_P^{t+s_t+1} + (1 - \zeta^{t+s_t+1}) \hat{f}(x)_Q^{t+s_t+1} \quad (3.10)$$

If D_t^* corresponds to $S2$ norm $\|D_t^*\|_2$ belongs to training set Q .

4. Experimental verification and result analysis. The test data used in this paper are from the NASA Report published by NASA. It contains Acoustic Memory Activities and Cryptography (SMAP) and the Marten Stein Science Laboratory (MSL). Finally, the remote sensing data of the fusion LSTM model is analyzed, and the results are compared with those of the traditional LSTM model.

4.1. Experimental Settings. The training and calibration data include the telemetry data and the corresponding telemetry indication of each channel of 12 subsystems in SMAP and MSL. Ninety-four data sets in total. There were 47 groups (50%) with SMAP. There are 47 groups (50%) that are MSL. Each training set contains 11 remote control commands [13]. The number of remote sensing data sets varies from 300 to 4,000. The specific total number of training data, abnormal sequences, test telemetry channels and detection data are shown in Table 1.

4.2. Comprehensive LSTM tests the anomaly detection algorithm in remote sensing data. The abnormal fragments are compared with the actual fragments. Results, such as the exception detection component, are illustrated in Figure 4.1. As you can see from Figure 6, the exceptions found are essentially the same as those flagged [14]. Due to error redundancy q , the difference between the number of anomalies detected in this paper and the number of real anomalies is about ± 100 .

Table 4.1: Experimental data.

| | SMAP | MSL | Total |
|--|------------|------------|------------|
| Anomalous sequential total | 72 | 38 | 109 |
| The sum of dotted anomalous series | 48(66.67%) | 22(57.89%) | 81(74.31%) |
| Background exception value total | 24(33.33%) | 16(42.11%) | 28(25.69%) |
| Remote channel total | 57 | 28 | 49 |
| Total number of remote sensing data detected | 447641 | 69489 | 517129 |

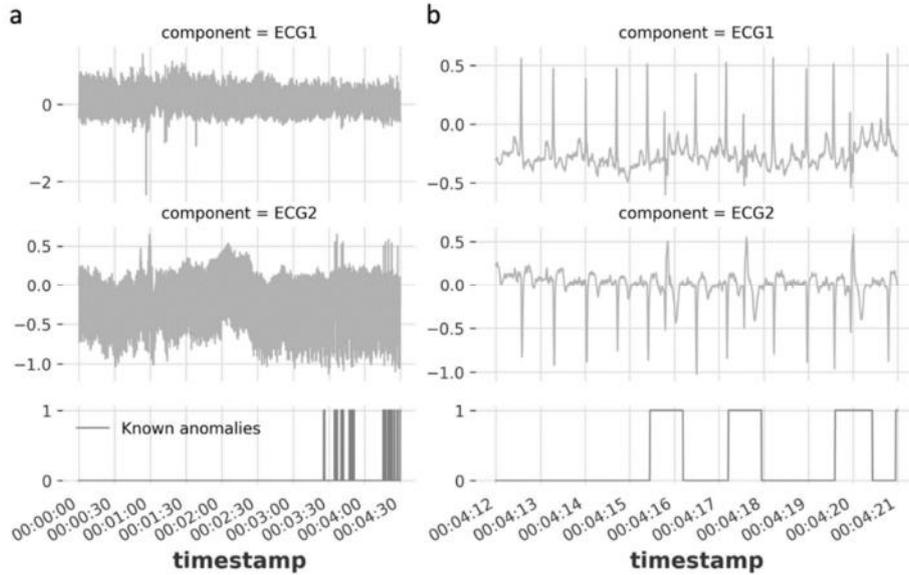


Fig. 4.1: Comparison between the actual anomalous sequence and anomalous detection results.

4.3. Comparison of actual anomalous sequence and anomalous sequence detection results. The ability of LSTM prediction is constantly improving, and the sensitivity to abnormal signals is also constantly improving. The change rule is evident for some remote sensing data, and remote control commands do not affect the remote sensing data. Therefore, anomaly detection in this case still needs further improvement [15]. The anomalies and background anomalies in two satellite remote sensing data types are examined based on ten experiments. FIG. 4.2 shows the performance comparison between an integrated LSTM-based anomaly detection method and an LSTM-based anomaly detection method in point anomaly detection and context anomaly detection for SMAP and MSL2 spacecraft.

A new method for anomaly detection of satellite remote sensing data using integrated LSTM is proposed. Compared with the LSTM algorithm, the single point anomaly detection rate of the SMAP algorithm is increased from 87.5% to 98.96%. The anomaly detection rate of SMAP applied to text content increased from 80.21% to 91.67%. The anomaly detection rate of MSL text was increased from 67.71% to 79.17%. A multimodal LSTM model based on the constraint model is proposed and applied to remote sensing data prediction. This can improve the recognition rate of abnormal events. Compared with the LSTM algorithm, this algorithm's overall point anomaly detection rate is increased from 73.96% to 81.25%. The abnormal detection rate of the whole environment increased from 75.00% to 84.38%.

5. Conclusion. This paper introduces an automatic control and management method for space vehicle test equipment. The monitoring and control of the test equipment are realized. This interface matches the

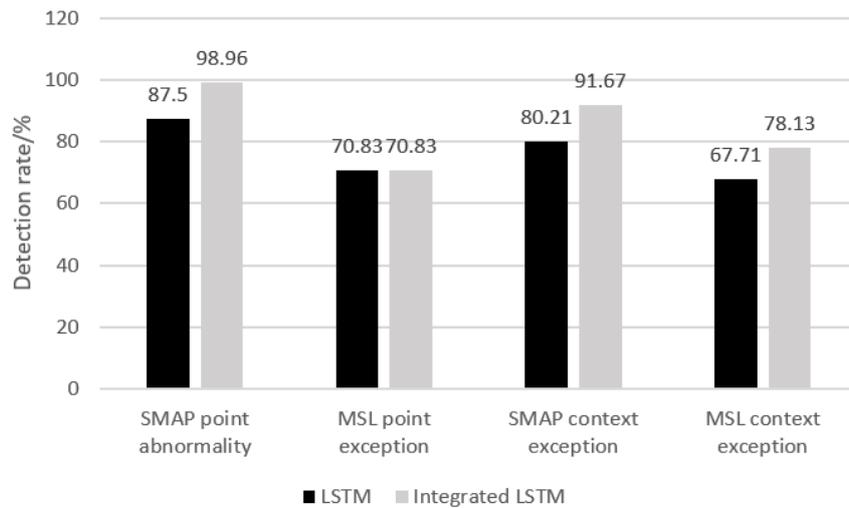


Fig. 4.2: Detection rates of different anomalies by two methods.

central control ATS. The system can automate and remotely manage the front-end test equipment of spacecraft. A spacecraft remote sensing measurement system is proposed, and experiments verify its effectiveness. An anomaly detection method for spacecraft telemetry data based on an integrated LSTM model is proposed. A dynamic threshold is established, and the difference between prediction and measurement is analyzed. The classification of remote measurement data is realized by mining the control mode of the remote command. This method improves the prediction level of remote sensing data by the LSTM model. The application of this system can effectively improve the management efficiency of front-end test equipment and promote the implementation of spacecraft automated tests.

Acknowledgments. The National Natural Science.

REFERENCES

- [1] Shi, H. B., Huang, D., Wang, L., Wu, M. Y., Xu, Y. C., Zeng, B. E., & Pang, C. (2021). An information integration approach to spacecraft fault diagnosis. *Enterprise information systems*, 15(8), 1128-1161.
- [2] Häseker, J. S., Aksteiner, N., & Ofenloch, A. (2021). From Behavioural to In-Depth Modelling of a Scalable Spacecraft Power System in SystemC-AMS using Continuous Integration Techniques. *Transactions of the Japan Society for Aeronautical and Space Sciences, Aerospace Technology Japan*, 19(5), 709-718.
- [3] Zheng, K., Zhang, Z., Chen, Y., & Wu, J. (2021). Blockchain adoption for information sharing: risk decision-making in spacecraft supply chain. *Enterprise Information Systems*, 15(8), 1070-1091.
- [4] Shen, B., Da Xu, L., Cai, H., Yu, H., Hu, P., Jiang, L., & Guo, J. (2022). Enhancing context-aware reactive planning for unexpected situations of on-orbit spacecraft. *IEEE Transactions on Aerospace and Electronic Systems*, 58(6), 4965-4983.
- [5] Zhang, C., Wu, J., Huang, Y., Jiang, Y., Dai, M. Z., & Wang, M. (2021). Constructive schemes to spacecraft attitude control with low communication frequency using sampled-data and encryption approaches. *Aircraft Engineering and Aerospace Technology*, 93(2), 267-274.
- [6] Lin, L. I., Li, Y. U. A. N., Li, W. A. N. G., Zheng, R., Yanpeng, W. U., & Xiaoyan, W. A. N. G. (2021). Recent advances in precision measurement & pointing control of spacecraft. *Chinese Journal of Aeronautics*, 34(10), 191-209.
- [7] Álvarez, J. M., & Roibás-Millán, E. (2021). Agile methodologies applied to Integrated Concurrent Engineering for spacecraft design. *Research in Engineering Design*, 32(4), 431-450.
- [8] Genova, A., & Petricca, F. (2021). Deep-space navigation with intersatellite radio tracking. *Journal of Guidance, Control, and Dynamics*, 44(5), 1068-1079.
- [9] Yin, Z., Cai, Y., Ren, Y., Wang, W., & Chen, X. (2020). A high precision attitude measurement method for spacecraft based on magnetically suspended rotor tilt modulation. *IEEE Sensors Journal*, 20(24), 14882-14891.
- [10] Schmidt, L., Gamper, E., Feuerle, T., & Stoll, E. (2020). Evaluation of the impact on European air traffic by uncontrolled reentering spacecraft with initial state uncertainty. *CEAS Aeronautical Journal*, 11(2), 401-416.

- [11] Zhu, W., Du, H., & Li, S. (2022). On event-triggered nonsmooth attitude tracking controller for a rigid spacecraft. *International Journal of Robust and Nonlinear Control*, 32(2), 900-916.
- [12] Pehrson, N. A., Ames, D. C., Smith, S. P., Magleby, S. P., & Arya, M. (2020). Self-deployable, self-stiffening, and retractable origami-based arrays for spacecraft. *AIAA Journal*, 58(7), 3221-3228.
- [13] Mote, M., Egerstedt, M., Feron, E., Bylard, A., & Pavone, M. (2020). Collision-inclusive trajectory optimization for free-flying spacecraft. *Journal of Guidance, Control, and Dynamics*, 43(7), 1247-1258.
- [14] Hovell, K., & Ulrich, S. (2021). Deep reinforcement learning for spacecraft proximity operations guidance. *Journal of spacecraft and rockets*, 58(2), 254-264.
- [15] Yu, X., Zhu, Y., Qiao, J., & Guo, L. (2021). Antidisturbance controllability analysis and enhanced antidisturbance controller design with application to flexible spacecraft. *IEEE Transactions on Aerospace and Electronic Systems*, 57(5), 3393-3404.

Edited by: Zhigao Zheng

Special Issue on: Graph Powered Big Aerospace Data Processing

Received: Jul 14, 2023

Accepted: Aug 25, 2023