The Design and Application of Space Target Monitoring Data Management System based on Big Data

Hongyan Chen[a], Sihong Wang, Junwei Wan

Beijing Institute of Tracking and Telecommunications Technology Beijing, China

[a]child_smile@163.com

Keywords: Observation data, Monitoring station, Space target recognition, Large data, Hadoop.

Abstract: Aiming at the low efficiency of data management, inadequate data analysis and processing ability, incomplete utilization of observation data and low contribution to the improvement of system service performance in the monitoring station(Abbreviated as monitoring station) of space target detection and recognition test system, Hadoop and other big data technology are used to develop a platform with functions of data storage management, analysis and mining, quality evaluation, etc. The active controlled early warning system of system equipment performance status is established, which solves the difficult problem of storage and management of massive observation data. At the same time, using the results of large data analysis, mining and exploring the correlation and influence rules of monitoring data, to provide reliable support for improving the service performance of space target monitoring system.

1. Introduction

The main task of space target monitoring station is to collect, process and transmit monitoring data of space target recognition by using various equipment, and ensure the long-term continuous and stable operation of system equipment. The data collected and managed by the monitoring station include orbit and situation data, supporting data (earth rotation, solar radiation, geomagnetic index, event analysis), data that have occurred (launch, meteorite, collision, disintegration), future data (collision warning, meteorite forecast) and equipment working condition data such as atomic clock and satellite communication system. With the steady development of China's aerospace industry, the information data acquisition and processing ability of the space target detection and recognition test system is gradually improved, and the system will gradually form a practical space target surveillance information support ability. The data quality of the monitoring station is the most direct and accurate reflection of the overall performance, service level and design index of the space target detection and recognition test system. For various reasons, there are still some problems in the management and depth application of observation data of monitoring stations, which need to be solved.

2. Current Situation and Problems

2.1 Data Management Methods Lag Behind

At present, the monitoring station receives real-time space target detection and recognition data, and generates hundreds of thousands of lines of data records every day. The amount of data reaches tens of gigabytes, which generates TB-level data every year. With the development and construction of space target detection and recognition experimental information system, the number of space targets is increasing, and the number of observation data produced every year is doubling. At present, single computer acquisition, cataloguing, short-term hard disk storage and long-term CD-ROM recording are used to manage data. To analyze the long-term change of a characteristic data (one year or several years), it is necessary to read the original observation data of CD-ROM, decompression and decompression every day in order to carry out the corresponding data analysis,

which is time-consuming, laborious and inefficient. In addition, the damage and loss of CD-ROM bring great hidden dangers to the security of data, which is not conducive to the long-term storage and management of data.

2.2 Data Utilization Efficiency is not High

All kinds of data of monitoring station are mainly used to complete the basic data such as the success rate of real-time original observation data acquisition, the success rate of satellite communication data transmission, the consistency and rationality of measurement data, track and situation data, and simple data processing such as pseudo-range measurement accuracy analysis, carrier cycle slip analysis, single-double difference analysis of measurement value. All kinds of data of monitoring station are independent and lack of cross-correlation. At present, there is no effective means to excavate the effective information contained in the massive observation data. It is difficult to analyze and process the parameters of the monitoring station for a longer period, which cannot provide support for system state assessment, fault prediction and performance analysis.

2.3 Few Studies have been Done on Analytical Methods

At present, in the field of space target recognition test information, there are more studies on service performance monitoring and evaluation at the user level, less on the daily operation management of monitoring stations, and there are problems of emphasizing ex post monitoring and neglecting pre-warning in the operation management of monitoring stations. Especially from the level of monitoring station observation data analysis system performance, the use of large data technology for in-depth analysis and mining of monitoring station observation data, the establishment of a platform for management, analysis and evaluation of research is not much. At the same time, multi-dimensional navigation data, such as multi-station data, differential enhancement correction data, user terminal data and so on, are synthetically used to mine the closed-loop structure relationship between monitoring station and user terminal data, and to study and explore the correlation and influence law of multi-dimensional data. At present, there are few such studies.

In view of the above practical problems, it is necessary to establish a management system, which combines the operation management of the system with the daily monitoring and evaluation organically, facilitates the benign interaction of trend analysis, forecasting problems and timely treatment, and realizes the deep coupling between operation and service. Complete the collection, transmission, storage, management, real-time analysis, display, early warning and alarm of monitoring station data, and realize the efficient safety management of data. Using the observation data of each monitoring station, the operation status and rules of mining system and equipment are analyzed, hidden dangers are found, and the prediction and evaluation of the operation status of monitoring station are provided. The data support is provided for the stable operation, evaluation and early warning of monitoring station and space target recognition test information system.

3. Overall Framework of the System

According to the sequence of data acquisition and use, the flow from left to right is divided into data source, data management, data mining, quality assessment and data access. The frame structure is shown in Figure 1. Data source obtains data from monitoring station and user terminal, obtains original observation data through data decoupling software, and stores the data into management subsystem. Through data mining, the data of monitoring station and user terminal are analyzed in full. The relationship between each measurement value and working condition information is obtained, and the evaluation model of monitoring station system status and user positioning performance is established. It objectively reflects the changes of service performance of monitoring stations and systems.

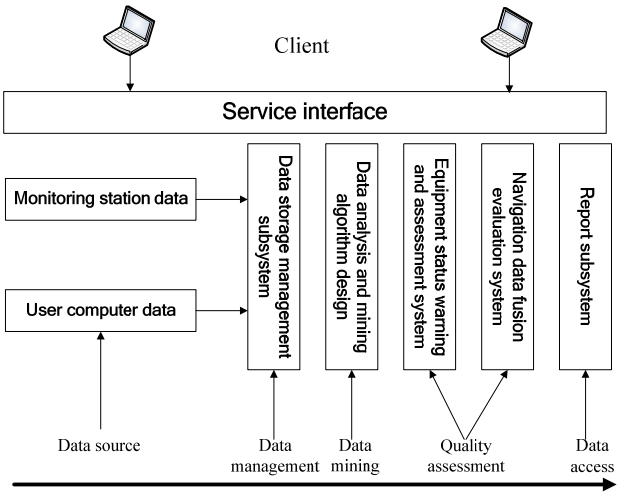

Fig 1. Schematic diagram of the overall structure of the system

3.1 Data Source Acquisition

The data of monitoring station include multi-channel pseudorange measurement, multi-channel carrier phase observation, multi-channel carrier-to-noise ratio observation, multi-frequency Doppler observation and interference level monitoring, orbit and situation data, meteorological data, working condition data of space target recognition system, network communication, atomic clock system, data preprocessing and so on. User data includes pseudorange value, ephemeris, loop locking, differential correction information, positioning results, positioning success rate and other information.

3.2 Data Management

Before storing the data, the collected data is preprocessed. According to gross error theory, prior information of data characteristics and data processing requirements, gross errors are eliminated to remove invalid data. At the same time, the rationality of the original data, the consistency of multi-channel data and the integrity of the target signal are checked. The pre-processed observation data is stored in the HBase non-relational database instead of using CD-ROM to save data, which improves the efficiency of data storage and reading and writing, and achieves random real-time access to large-scale data.

3.3 Data Mining

Using R language and Hadoop platform to mine the data of the monitoring station and the user, the hidden relationship between various kinds of observation data is found, such as the correlation between carrier-to-noise ratio and pseudo-range measurement accuracy, orbit angle and pseudo-range measurement accuracy, Doppler value and carrier phase, and the influence of the change of electromagnetic environment around the monitoring station on the accuracy of measurement data is explored.

3.4 Quality Assessment and Data Access

According to the characteristic data of the fault case, a model is established to evaluate and warn the operation status of the monitoring station. According to the specific business needs, the related data analysis results are generated by combining charts and graphs, which is convenient for query and statistical analysis.

4. Implementation of Key Technologies

4.1 Design and Implementation of Data Management Platform

The storage management, analysis and processing of massive data generated by monitoring stations need a platform with strong scalability, easy maintenance, scalability and powerful computing and processing capabilities. According to the requirement of business data processing, Hadoop large data processing platform is chosen to build to meet the needs of security, fault tolerance, scalability and efficient reading and writing of business data. At the same time, the establishment of HBase non-relational database based on HDFS distributed file storage system not only has the advantages of HDFS, but also achieves real-time random reading and writing access to business data, and enables the data statistics office. It's very simple. The network topology of the monitoring station's large data platform is shown in Figure 2.

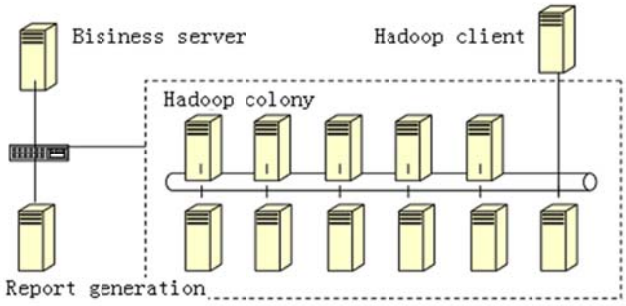

Fig 2. Network Topology Map of Data Management Platform

On the basis of meeting the requirement of data storage of early monitoring stations, considering that the annual data of monitoring stations is in the TB level, the cluster can meet the requirement of the next three years without adding nodes. The default backup strategy of HDFS (i.e. the number of replicas is 3) is selected, and the initial storage capacity of the cluster is calculated to be at least 120 TB. Based on this and considering Hadoop's computing requirements, the hardware structure is as follows: 6 PC servers (memory: 256GB, hard disk: 20TB), including 1 NameNode node, 1 Secondary NameNode node (combined with a slave node), 5 slave nodes, cluster storage capacity of 120 TB, meeting the needs of project data processing, and can act according to needs. State expansion. Ten thousand Mega switches are selected to meet the requirement of network bandwidth performance of cluster.

In the selection of database software, NoSQL database is used for data storage, and column-oriented HBase data model is used for database model. The data is stored by column, which facilitates the storage of structured and semi-structured data and data compression. It has great I/O advantages for queries against a column or columns. These characteristics just satisfy the long-term analysis and research of some characteristic data of monitoring stations. In the aspect of database design, different data tables are divided and stored according to the functions of monitoring station equipment and types. The data of different types and equipment are stored mainly in the following tables: original observation data table of monitoring receiver, data preprocessing package data table, satellite antenna tracking data table, satellite receiving level data table, satellite baseband data table and meteorological observation data table.

4.2 Design and Implementation of Data Analysis and Mining

Data analysis and mining design is mainly composed of six modules: carrier phase cycle slip detection and evaluation module, data quality comprehensive analysis module, multi-path effect evaluation module, ionospheric monitoring and analysis evaluation module, satellite orbit extrapolation performance evaluation module, monitoring receiver integrity monitoring module. As shown in Figure 3.

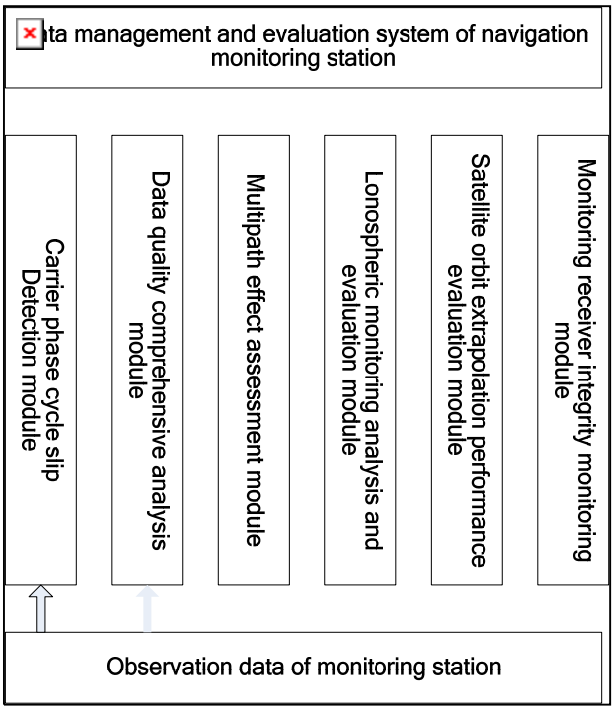

Fig 3. Data Analysis and Mining Function Module

The function of carrier phase cycle slip detection and evaluation module is to accurately detect and statistically monitor the cycle slip of carrier phase observation data of the receiver, and to evaluate the performance of carrier phase cycle slip detection of the receiver. Because the monitoring receiver is static, polynomial fitting method is suitable.

According to the theory of data quality analysis, the function of data quality comprehensive analysis module is to establish the standard model of navigation signal quality analysis of monitoring station, and to analyze and statistic the quality of observation data comprehensively. The single-difference and double-difference combination analysis of pseudo-range observation data and phase observation data is carried out to extract the change information, and the quality of satellite ranging signal, carrier phase signal and the stability of monitoring receiver are evaluated.

The function of multi-path effect evaluation module is to use carrier phase data and pseudo-range observation data to make linear combination, extract the multi-path variation on pseudo-range observation data, and smooth pseudo-range with carrier phase. Finally, according to the characteristics of multi-path effect, the observation environment of monitoring station and the ability of receiver to resist multi-path effect are evaluated.

The function of the ionospheric monitoring analysis and evaluation module is to calculate the time series of the ionosphere by using the dual-frequency observation data, to evaluate the stability of the multi-frequency data to the ionospheric monitoring ability, and to evaluate the impact of the monitoring ionosphere on the ranging quality.

The function of the target orbit extrapolation performance evaluation module is to evaluate the precise dynamic model extrapolating the target orbit technology based on the target ephemeris calculation. It mainly evaluates the cumulative effect of errors with the extension of extrapolation time.

The function of the integrity monitoring module of the monitoring receiver is to detect and identify satellite faults by using redundant observations of the receiver itself. It does not need the assistance of external equipment, costs less and is easy to implement. It is a widely used integrity monitoring method at present. Consistency checking is carried out by calculating deviation.

4.3 Design and Implementation of Equipment Status Early Warning and Assessment Sub-system

According to the large amount of observation data obtained by monitoring stations, the evaluation and analysis model of business data is studied. Through the analysis and mining of all kinds of data, the correlation between various parameters and indicators is found, and the trend and rules of system operation are analyzed and extracted to realize the state prediction and early warning evaluation of system equipment. By means of analysis methods and means of measurement quality analysis of observation data, ionospheric analysis of multi-frequency observation data, multi-path detection and analysis, detection and repair of carrier phase cycle slip, and correlation between carrier-to-noise ratio of navigation signal, accuracy of multi-path analysis value and pseudo-range measurement value, carrier phase cycle slip and other measurement values, the long-period data analysis and mining are carried out, and the reference of fault cases is given. The operation model of monitoring station is established by numerical analysis and comparison, which provides data support for early warning and evaluation of monitoring station system.

The system fault alarm is mainly realized from the following aspects: system integrated state display, subsystem detailed state real-time monitoring, performance index real-time monitoring, navigation signal real-time monitoring, satellite orbit prediction display, alarm threshold design and so on. At the same time, it establishes the typical fault database of the system, establishes the logical tree relationship between the possible fault locations and nodes, analyses the interrelations among them, infers the running state and health status of the system equipment from the changes of real-time indicators, predicts the possible fault locations of the subsystem integrated state, maintains and repairs or replaces spare parts in advance according to the possible faults, and makes a plan. In order to predict the equipment failure in advance and early warning, the active and controlled early warning function is realized.

In addition, we can use the quality and positioning accuracy of the received signal from the user to deduce the operation status of the monitoring station, establish a closed-loop structure from the system to the user, open the two-way feedback link of the state and performance of the system equipment, and verify each other from the two dimensions of system data and user data, so as to achieve the purpose of early warning and evaluation, and improve the quality and efficiency of operation and maintenance management of the monitoring station.

4.4 Design and Implementation of Navigation Data Fusion Evaluation System

This part mainly studies the integration of the operation and maintenance data of the monitoring station system and the location data of the user, and establishes the closed-loop network structure among the monitoring station, differential enhanced receiving terminal and user receiving terminal. On the basis of monitoring station data management platform, differential enhanced data receiving subsystem and user receiving terminal data processing subsystem are established. Using large data technology to fuse analysis and explore the impact and regularity of multi-dimensional data such as monitoring station, differential enhancement correction, user receiving terminal data and so on.

5. Summary

The data management system of monitoring station is designed and implemented, which changes the backward mode of using CD-ROM to record and save data. At the same time, it deeply analyzed and excavated the effective information contained in the massive navigation data, analyzed the correlation and coupling between the various data of the monitoring station, developed the observation data analysis and evaluation model, realized the real-time analysis and evaluation of the data quality of the monitoring station, and provided an effective judgment basis for the system status judgment. Research and explore the operation and maintenance model of monitoring station of space target detection and recognition test system, construct active and controlled equipment prediction analysis and fault early warning system, real-time monitor the overall operation status and performance of the system, change the traditional passive working mode of "firefighting" and make

emergency repairs as pre-analysis and prediction, so as to minimize the occurrence of major system failures and greatly enhance the performance of the system. Monitoring the quality and efficiency of operation and maintenance management. This project can be used for single monitoring station, combined with multi-station data, differential enhancement correction data and user terminal data, to realize the analysis and mining of all data from large-scale system to user-end closed-loop structure, and to provide a reliable integrated processing platform for improving the service performance of space target detection and recognition test system.

References

[1] Junwei WAN, Hongyan CHEN, Jing ZHAO. Application verification of cloud computing technology in real time space test mission area [J]. Netinfo Security,2017(05):63-68.

[2] Hui ZHAO, Junwei WAN etc. Application Research of autonomous control technology in experimental task domain [J]. Journal of Spacecraft TT&C Technology,2015,34(02):109-114.

[3] Hongyan CHEN, Junwei WAN. Research on Application Model of Domestic Basic Software in High-Density Information Integrated Application System [J]. Modern Electronics Technique, 2018, 41(21):84-87+91.

[4] Hongyan Chen, Junwei Wan. Research and Application of Web-based Performance Optimization Scheme for Enterprise Websites [J]. Internet of Things Technologies,2018,8(02):67-69+73.

[5] Hongyan Chen, Junwei Wan, Qi Wang. Design and Implementation of High-Performance Sorting Algorithms for Large Data [J]. Journal of Spacecraft TT&C Technology,2015,34(02):120-127.

[6] Junwei WAN, Hui ZHAO etc. The development status and application analysis of autonomous controllable information technology [J]. Journal of Spacecraft TT&C Technology, 2015, 34 (04) :318-324.

[7] Hongyan CHEN, Junwei WAN, Hongwei QI. Research on cloud smart office integrated management system solution based on Internet of things [J/OL]. Modern electronic technology, 2018 (10):85-89[2018-05-30]. https://doi.org/10.16652/j.issn.1004-373x.2018.10.022.