The Dynamic Monitoring System of Document Bibliography Information Based on Wireless Communication

Ri-han Wu(✉)

School of Chinese Language and Literature, Northwest Minzu University, Lanzhou 730030, China
wurihan21322@yeah.net

Abstract. The dynamic monitoring of changes in the bibliographic information is an important task for insight into the development of informatization in the field of science and technology. Based on this, a wireless communication-based dynamic monitoring system for document bibliographic information is designed. By optimizing the system hardware equipment, the system operation efficiency is improved, and the system software functions and operating procedures are further optimized by combining wireless communication technology. Through the collection of information and knowledge extraction, information analysis and other technologies can effectively reveal the important dynamic information of the target scientific research institution in terms of strategic planning, research layout, important research progress, etc., and in-depth monitoring of the dynamic changes of literature bibliographic information. Finally, it is confirmed by experiments that the dynamic monitoring system of bibliographic information based on wireless communication has high effectiveness in the actual application process and fully meets the research requirements.

Keywords: Wireless communication · Document bibliography · Information dynamic monitoring

1 Introduction

Dynamic monitoring of bibliographic information can effectively query the changes of science and technology in real time. The national science library requires each strategic information research team to continuously track and accumulate the changes of literature information in their respective fields, so as to effectively provide multi-level literature information services and grasp the major trends and strategic trends of scientific and technological development. Literature [1] designed a Scrapy-based network space information dynamic monitoring system, with the help of the information collection layer in the system, from various information sources and Scrapy web crawlers in the network, the required network space information was collected, and the information collection layer was constructed based on the Scrapy framework Establish the

S. Liu and X. Ma (Eds.): ADHIP 2021, LNICST 416, pp. 41–55, 2022.
https://doi.org/10.1007/978-3-030-94551-0_4

request response to the web page, analyze the web page layout and cyclically capture the required cyberspace information. Use regular expressions to remove the captured cyberspace information as formatted data, and store the collected data in the cyberspace information processing module database In. The network information processing module uses reasonable methods to purify the webpage where the information is located, and transfers the content extracted from the database to the network spatial information analysis module. The analysis module uses text clustering, feature selection and other technologies to explore the spatial information in the network and form a topic List and set up keyword information, and establish a cyberspace information lexicon in related fields. Use WM algorithm to process homophones, split words and interference symbols, to achieve accurate matching of required cyberspace information, and achieve the purpose of dynamic monitoring of cyberspace information.

However, the above-mentioned systems have low efficiency in network information monitoring. Based on this, combined with wireless communication technology, the literature bibliographic information dynamic monitoring system is optimized to optimize the functions of the automatic information monitoring service system, simplify the dynamic information monitoring process, and analyze through simulation experiments. The validity of the system in this paper is verified, and the problems existing in the traditional system are solved.

2 Literature Bibliographic Information Dynamic Monitoring System

2.1 The Hardware Configuration of the Dynamic Monitoring System for Bibliographic Information

The information collection layer of the document bibliographic information dynamic monitoring system mainly realizes the regular collection and harvesting of target resources. In the document bibliographic information dynamic monitoring system, it is necessary to build a series of distributed network directional collectors to achieve accurate collection of target resources [2]. The monitoring calculation and analysis layer is the main functional layer of the monitoring system. For each information resource collected by the network directional collector, the monitoring system first identifies new resources, and from the collected resources, it identifies new resources that were not in the original system [3]. For the identified new HTML page, the structure of this HTML page is analyzed, and the main text content of this page is extracted [4]. If this resource is not an HTML page, but a rich document resource such as PDF and WORD, the monitoring system will automatically realize the analysis of this rich document resource and extract the corresponding text content. After that, the monitoring system then implements title and abstract extraction, monitoring target object and object relationship extraction, and domain term extraction for this resource, transforming the original free text into structured data with certain semantic support, and incorporating it into the corresponding semantic knowledge base Stored in [5]. For these resources that have been expressed by structured monitoring objects and object relationships, the monitoring system can adopt corresponding wireless communication methods to realize the intelligence value judgment of these resources to reveal important intelligence resources; through automatic

classification tools, the intelligence type can be realized. Recognition; through semantic similarity calculation, to reveal the clustering and distribution relationship between intelligence. The monitoring service provider layer is the layer that provides automatic monitoring services to strategic intelligence personnel [6]. For the new resources collected from the network, through various processing of the monitoring calculation and analysis layer, a specific service function is formed at the monitoring service providing layer, including the disclosure of important information content, the tracking of key monitoring targets, the disclosure of hot topics and hot objects, The overall situation of the organization is revealed, etc., through the friendly user interface, it serves the strategic intelligence monitoring [7]. Based on the business level of the above-mentioned network technology information automatic monitoring system, the corresponding system hardware structure framework is proposed, as shown in the figure:

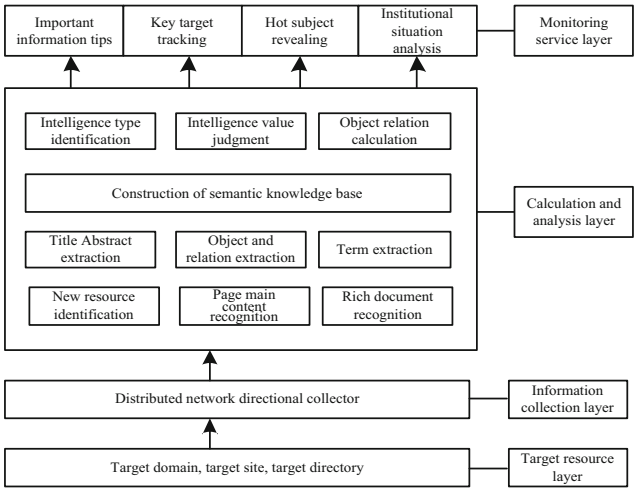

Fig. 1. Hierarchical structure of information automatic monitoring system

The main function of the monitoring information collection subsystem (as shown in Fig. 1) is to discover and collect new content from the target website in real time and accurately. The specific technical realization mainly has the following functional characteristics:

Distributed collection architecture realizes a scalable monitoring collection platform [8]. The central control server controls the collection strategy and collection cycle, and distributes the collection tasks to each collection node in a unified format, and the central control server performs load balancing control on each collection node at the same time.

Fine-grained collection methods are used to accurately collect target resources [9]. Different from the traditional full-site collection system, the monitoring subsystem starts from the particularity of the application, refines the monitoring collection resources from the site to the catalog, and conducts targeted collection according to different catalog page formats and URL characteristics.

Adopting efficient and accurate new resource identification technology. In addition to using the traditional MD5 code to check the text and extract page fingerprints, the monitoring collection subsystem also uses the scientific research objects extracted from the network information resources to calculate the similarity of the pages to realize the identification of new resources.

In network monitoring, URLs of monitored institutions and directories will change due to factors such as website revisions or adjustments to related modules. At this time, timely early warning and feedback mechanisms are particularly important, which can help strategic intelligence personnel to monitor the monitoring in a timely manner. The target is adjusted [10]. In the WSN data processing system, the upper application system mainly provides user services including model deployment, data playback, and situation analysis. These services are based on the validity and accuracy of the data in the data processing system. The core functions of data processing mainly include data fusion, data compression, data association, data completion, data compensation, data storage and query, etc. These core functions provide basic functional support for upper-layer applications. Among them, data fusion and data compression are necessary data processing to save hardware resources during data transmission; data association is to restore the data after fusion and compression.; Data completion and compensation is a kind of data post-processing to ensure the accuracy and effectiveness of the data in order to ensure that the data is accurate and effective; data storage and query are basic data services. According to requirements, the software in the data processing system adopts a modular design method, which integrates deployment model generation, data post-processing and data query and other related modules to encapsulate the large and complex data processing process based on WSN to form an organic whole. The software structure of the WSN-based dynamic monitoring data processing system is shown in the figure. The data processing system is composed of four parts: a system manager, a database,

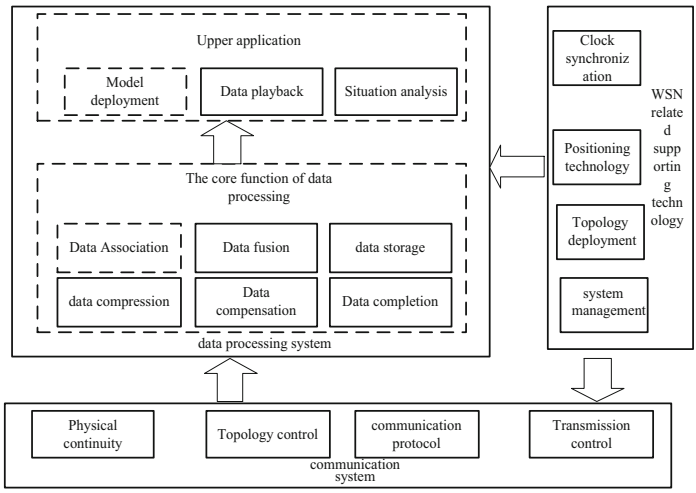

Fig. 2. The frame structure of the dynamic monitoring data processing system

a functional component and a user interaction interface, and each part completes the transmission and management of information through a public interface (Fig. 2).

The system structure design adopts a design model similar to MVC, which separates system management, functional components and user interaction interfaces, effectively guarantees the integration between modules, and provides convenience for system development and later maintenance. The functional components of the system adopt a modular design, which is divided into a deployment model generation module, a data post-processing module, and a data query module to realize data processing at different stages.

2.2 Literature Bibliographic Information Dynamic Monitoring System Software Function

The data processing process in WSN nodes and gateways is called data pre-processing. Compared with data pre-processing, the processing of data after the data is transmitted to the gateway is called data post-processing in this article. Each module is divided into more specific sub-modules according to its function to achieve high cohesion of the internal functions of the module. The model deployment generation module is composed of model data generation and model parameter setting modules; the data post-processing module is composed of data analysis and association, data compensation and data completion modules; the data query module is composed of integrated query and data playback modules.

The database of the system mainly includes original database, process database, result database, model database, algorithm database and rule database. Among them, the original database stores the original data transmitted by WSN, the process database and the effective database store some data in the data processing process and the effective data after analysis, and the model library stores the basic model and algorithm library of the large deformation flexible body of the airdrop equipment. Stored are the algorithms used in the data processing process, and the rule base is the rules used in the data parsing process.

The user interaction interface component of the system is to provide the user's operation interface for the user to use. When users use software, they pay more attention to the data in their domain and do not care about the process and logic of the software itself. Therefore, the user interface encapsulates the three functional modules, and only provides users with the data and processing processes they care about (Fig. 3).

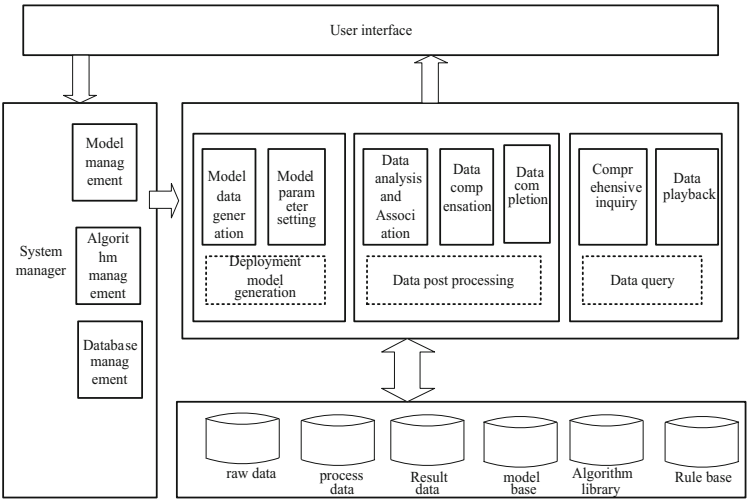

Fig. 3. Software structure of dynamic monitoring data processing system

The three parts of database, functional components and user interaction interface complete the interaction between each function through the control and coordination of the system management module. The user operates the system through the interaction interface and feeds it back to the system manager. The system manager determines the needs according to the user's operation. After obtaining the instructions of the manager, which functional component to use, which algorithm to choose, and which data is required, the functional component selects the relevant functional module and database as needed to complete the operation, and presents the operation result on the user interaction interface. The following will design and introduce the workflow of the data processing system according to the interactive process. The data post-processing module receives the monitoring data transmitted from the communication interface, and imports these data into each data table built to complete the storage of the local original data. Analyze and associate the original data, establish the internal index of the module, load the original data into the module through the index information, perform operations such as denoising and removing anomalies on the data, and store the parsed effective data. Call the core algorithm to effectively compensate and complete the data, and store the process data for processing and analysis. Analyze the compensated and supplemented data, and store the analysis results.

The data query module can replay the analyzed and processed raw data, help users understand the various states of the large-deformation flexible body of the airdrop and airborne equipment, and provide data support for the development of the large-deformation flexible body (Fig. 4).

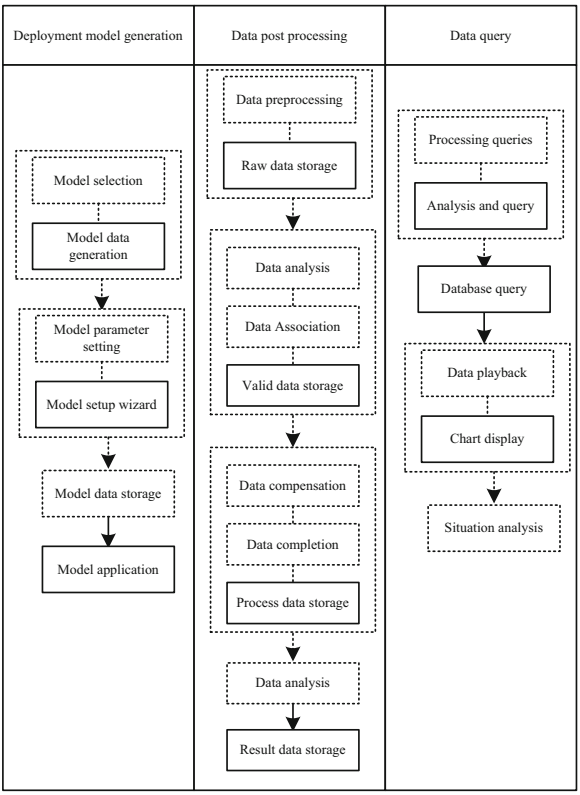

Fig. 4. The flow of the dynamic monitoring data processing system

When a certain data is not updated after multiple searches, it means that it has fallen into a local optimum, and a feasible solution is randomly searched for as the new dependent data, and a new search is restarted. Through the division of labor and cooperation, a better information source can be obtained. The specific algorithm is as follows.

$$\min\left(f(x) : x \in S \subset R^d\right) \tag{1}$$

Where $x = (x_1, x_2, \ldots, x_d)$ is the variable to be optimized, f(x) is the objective optimization function, and S is the solution space, then:

$$S = \{\left(x_j^{\min}, x_j^{\max}\right) \Big| j = 1, 2, \ldots, d)\} \tag{2}$$

Further calculations can be:

$$v = r(-1, 1) \times \left(x_{ij} - x_{\text{neighberw},j}\right) \tag{3}$$

$$x_{ij}^{\text{new}} = x_{ij} + v \tag{4}$$

After each round of searching, observe the information to judge the income of the information source, and make selections based on probability to speed up the convergence of the algorithm.

$$h_j = g\left(\sum_{i=1}^{I-1} w_{ij} t_i + S x_{ij}^{\text{new}} - v\right) \tag{5}$$

Among them, w_{ij} is the connection weight between the i-th neuron in the input layer and the j-th neuron in the hidden layer, a is the threshold of the j-th hidden layer neuron, if h is the j-th neuron in the hidden layer The output of the excitation function is as follows:

$$g(x) = \frac{1}{1 + he^{-x}} \tag{6}$$

Then the output of wireless communication is:

$$p = \sum_{j=1}^{J} h_j c_j - b \tag{7}$$

The mean square error of data compensation wireless communication is:

$$E = \frac{1}{2} \sum_{k}^{s} e_k^2 \tag{8}$$

According to the gradient descent method in the feedback adjustment mechanism of wireless communication, the specific calculation process is as follows:

$$\Delta w_{\text{neighbour},j} = \begin{cases} \delta H_j \left(1 - H_j\right) t_{\text{neighbour}}\, c_j e \\ \delta H_{j-I^*J} e \\ \delta H_{j-(I+1)^*J} \left(1 - H_{j-(I+1)^*J}\right) c_{j-(I+1)^*J} e \\ \delta e \end{cases} \tag{9}$$

Based on the above algorithm, further use the document ontology storage query method to store the constructed ontology document. The specific steps are as follows (Fig. 5):

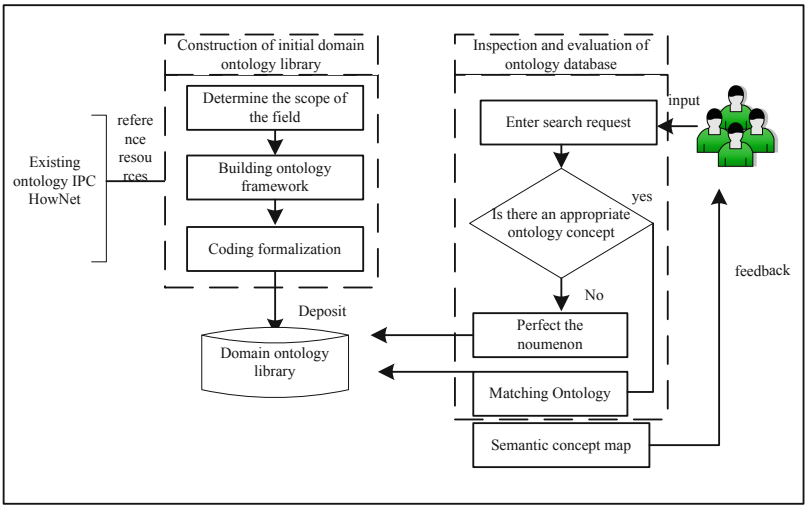

Fig. 5. The flow chart of the ontology construction of the document bibliographic information dynamic monitoring system

2.3 Realization of Dynamic Monitoring of Bibliographic Information

With the computerization of information management, people's use and reliance on databases has become more and more common. Although the database building software is becoming more intelligent and easier to use, many organizations can build a database system that meets its characteristics. However, the design of the database table is very important at the beginning of the development. If you do not pay attention to it, it will cause the later stage of the system development. There were some unsatisfactory points in the created table, and finally the table had to be modified. The program made before must also be modified due to the modification of the table structure. Therefore, when building a database, you must pay attention to the following:

(1) The data is complete, and the fields of the data table can fully reflect the document information required by the user.
(2) Reduce data redundancy, set reasonable field lengths, and avoid waste of storage space; at the same time, eliminate some unnecessary fields to prevent field duplication.
(3) Regularization of field naming. When creating a table, the name of the data table should not conflict with the keyword name in the table. Choose the appropriate field type. Based on this, the document information collection model is constructed as follows (Fig. 6):

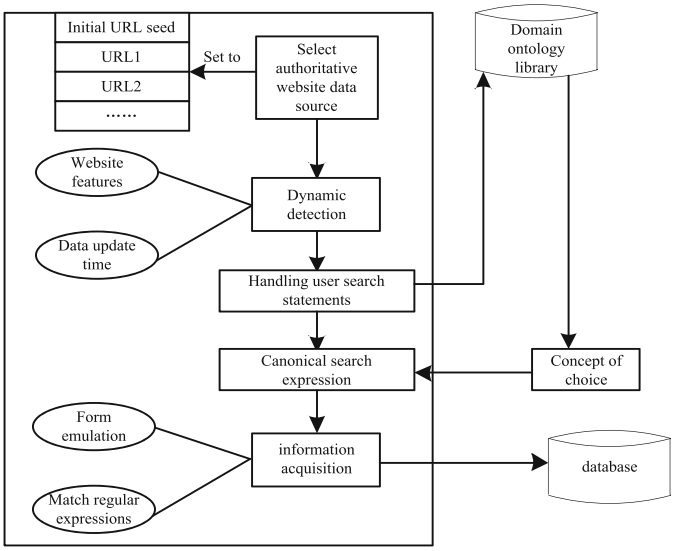

Fig. 6. Document information collection model

After the document database is formed, it needs to be maintained and managed, because in the document information collection process, due to network speed or other problems, the document information collection procedure may be interrupted, resulting in incomplete data records in the document database. If it is not timely Dealing with these incomplete data, the user's literature information analysis results will lack objectivity

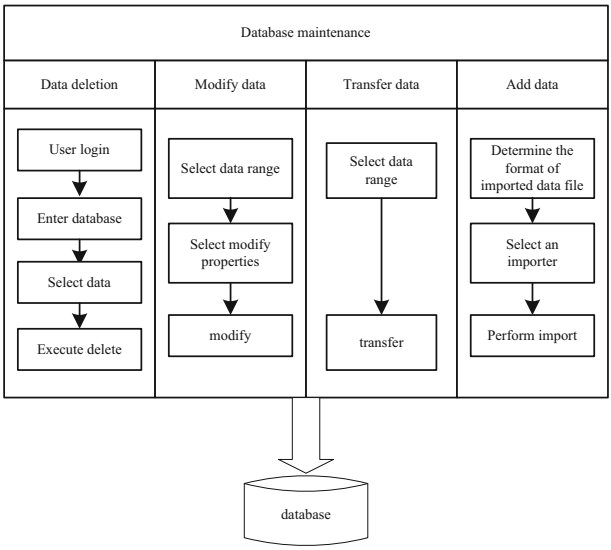

Fig. 7. Document information management model

and scientificity, so it is necessary to maintain the literature database from time to time. The document information management model is shown in the figure (Fig. 7).

According to the above-mentioned table building rules, design the document data table and the mapping relationship between it suitable for this system, and create the document database. The database should include data deletion, modification, and transfer. The system front-end development tool chooses PHPS. The PHPS development language is chosen among the many development tools mainly because the PHPS development language has advantages that other development tools do not have, such as fast speed, good openness and scalability, and support for a variety of mainstream and non-standard development tools. Mainstream databases, object-oriented programming, rich functions, etc. The realization principle of the user login module is to compare whether the user name and password entered by the user match the two fields of name and password in the login table in the literature data. If they match, enter the main module of the system. The flowchart of the user login module is shown in the figure (Fig. 8):

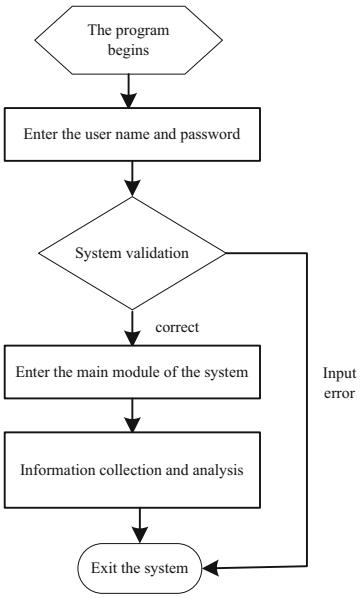

Fig. 8. Flow chart of user login module

Wireless sensor network node is mainly composed of four parts: sensor module, processing module, wireless communication module and energy supply module. Through the physical connection of Q-type sensor, the force information of large deformation flexible body and other measured target parameters are detected, and the circuit is adjusted. After a/D conversion, the sensing signal is processed and amplified, analyzed and transmitted by the processor, and stored by the memory. The wireless communication module is mainly responsible for the communication with other nodes, mainly including ZigBee

wireless network, MAC layer and wireless transceiver. The energy supply module generally uses micro dry batteries to provide energy for the above three modules to ensure the normal operation of the node (Fig. 9).

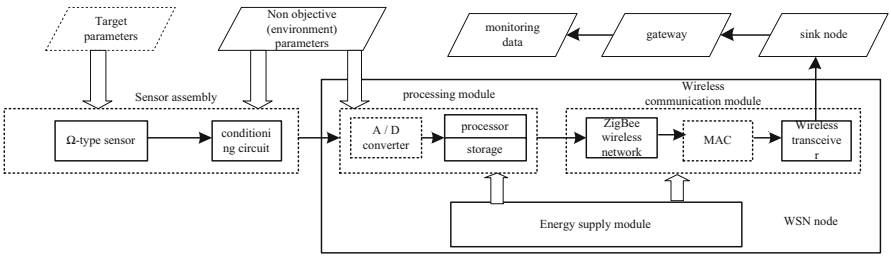

Fig. 9. WSN node and working mechanism structure

As can be seen in the figure, the sensor component will be affected by non-target parameters such as the environment during the data collection process. Therefore, the actual force of the measured object will deviate from the monitoring data obtained by the gateway node, and a certain amount must be used. The method to eliminate these deviations, this chapter gives the corresponding compensation model and algorithm to eliminate the deviation. After the data undergoes A/D conversion, the analog signal becomes a digital signal. The digital signal generally does not produce deviations, but abnormalities or missing during high-frequency transmission and processing. The processing of missing data is in Chapter 4. The corresponding model and algorithm are given.

3 Analysis of Results

In order to verify the actual application effect of the document bibliographic information dynamic monitoring system based on wireless communication, experimental testing was carried out. The experimental data processing system was developed in the VS2010 integrated development environment using C++ language, and the hardware and software environment is as follows:

Hardware environment:
CPU: Intel Core i54210M2.6 GHz/3200 MHz/3 MB
Memory: 4GDDR3L1600MHz*2
Hard Disk: SOOGHDDSATA
Communication interface: RS-232 interface

Software Environment:
Operating system: Windows7NT6.1 × 64
Database: Oracle11g
Integrated development environment: VisualStudio2010.

In order to facilitate users to observe the processing progress of data compensation, the processing progress bar is at the bottom of the interface. The figure shows the status of data processing (Fig. 10).

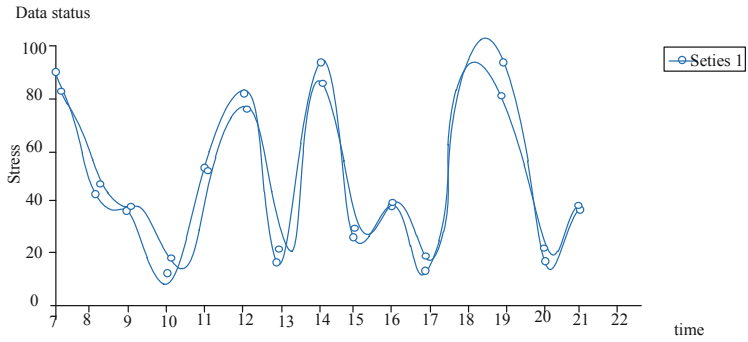

Fig. 10. System data status detection

The interface of data completion is similar to this interface, you can select the corresponding parameters, and you can view the corresponding processing progress (Table 1).

Table 1. Frequency table of main and subordinate classification numbers

Classification number	Frequency	Classification number	Frequency	Classification number	Frequency
A41D	5796	A61M	347	F21V	114
A41B	2212	A61H	267	A45C	100
A44B	1792	H05B	262	D06M	98
A41C	1294	A41G	261	A43B	91
A41H	654	A61K	186	A61L	91
A61F	510	A62B	146	B29C	85
A41F	479	B32B	134	D04B	79
A61N	408	A42B	117	A63B	77

In this system, in order to prevent suspended animation during data processing, the data processing system uses multi-threading technology. Several functions that require high computer resources such as CPU and memory are data association, data compensation, and data completion. When processing a large amount of data, it is necessary to test the time and space spent on these functions. At the same time, the growth trend of them can also be used to determine whether the data processing volume is within the acceptable range of the user. The figure below shows the data processing time of the system and related important algorithms (Fig. 11).

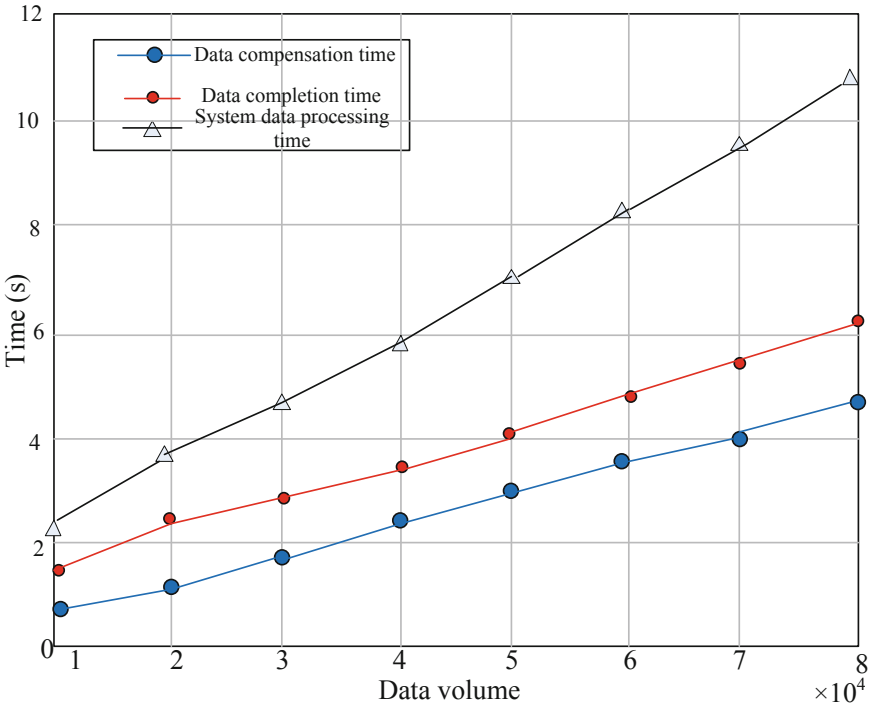

Fig. 11. System document detection time-consuming

The completion algorithm is tested when the data is missing at least 1%. It can be seen from the figure that as the amount of data increases, the running time of the system increases linearly, and the running time of the data compensation algorithm is significantly less than that of the data completion algorithm. The figure shows the running time of the system from 10,000 data to 80,000 data. In an experiment, the effective data monitoring time of the drop experiment and wind tunnel experiment is within a few seconds. If the data collection per second is 1000 times, The data volume of an experiment is basically within the scope of this performance test. Therefore, from the operating time of the entire system, the user is basically acceptable. This proves that the system in this paper has high monitoring accuracy and efficiency in the actual application process, and fully meets the research requirements.

4 Conclusion and Outlook

At present, digital resources have become an indispensable and important resource for social development, and the rational use and scientific analysis of network literature information resources can form literature information with high technical and commercial value, which is helpful to analyze the technology distribution situation and master related technologies. Historical origin, current conditions and future trends, so as to make a scientific and reasonable positioning for technological innovation activities, and

provide important direction guidance and decision-making assistance for enterprises to improve independent innovation capabilities, optimize implementation effects, and increase development speed. The scientific analysis results of literature information rely on comprehensive and authoritative information sources and a reasonable analysis index system. However, the system in this paper did not consider the security of network literature information when designing it. In the following research, we will focus on information encryption.

FundProjects. This paper is supported by the Postgraduate Research and Innovation Project of Northwest Minzu University , the name of the project is "A study on the names and descriptions of the things of mongolian <qin ding li fan yuan ze li>" (project number is Yxm202009), which is the phased achievement of this project.

References

1. Li, C., Jian, M.: Scrapy-based agricultural network spatial information dynamic monitoring algorithm. J. Shandong Agric. Univ. (Nat. Sci. Ed.) **51**(02), 64–69 (2020)
2. Shen, S., Zhu, D., Rousseau, R., et al.: A refined method for computing bibliographic coupling strengths. J. Informet. **13**(2), 605–615 (2019)
3. Yang, J., Wen, J., Wang, Y., et al.: Fog-based marine environmental information monitoring toward ocean of things. IEEE Internet Things J. **7**(5), 4238–4247 (2020)
4. Shahshahani, A., Zilic, Z., Bhadra, S.: an ultrasound based biomedical system for continuous cardiopulmonary monitoring: a single sensor for multiple information. IEEE Trans. Biomed. Eng. **7**, 1–10 (2019)
5. Barua, A., Zhang, Z., Al-Turjman, F., et al.: Cognitive intelligence for monitoring fractured post-surgery ankle activity using channel information. IEEE Access **99**, 1 (2020)
6. Jiang, B., Luo, Y., Lu, Q.: Maximized mutual information analysis based on stochastic representation for process monitoring. IEEE Trans. Ind. Inf. **15**(3), 1579–1587 (2019)
7. Kodai, S., Tatsuya, O., Yuichi, O., Masayuki, M.: Traffic engineering and traffic monitoring in the case of incomplete information. IEICE Trans. Commun. **E102.B**(1), 111–121 (2019)
8. Lu, Z., Wang, M., Dai, W., et al.: In-process complex machining condition monitoring based on deep forest and process information fusion. Int. J. Adv. Manuf. Technol. **104**, 5–8 (2019)
9. Liu, S., Fu, W., He, L., Zhou, J., Ma, M.: Distribution of primary additional errors in fractal encoding method. Multimedia Tools Appl. **76**(4), 5787–5802 (2014). https://doi.org/10.1007/s11042-014-2408-1
10. Liu, S., Bai, W., Liu, G., et al.: Parallel fractal compression method for big video data. Complexity **2018**, 2016976–2016982 (2018)