

# Design of University Library and Information Management System Based on Big Data Fusion

Zhang Jia<sup>1, a</sup>

<sup>1</sup>Huali College Guangdong University of Technology, 511325 Guangdong Guangzhou, China

**Abstract:** In order to improve the ability of library and information management in colleges and universities, and improve intelligent retrieval level of books, a design method of library information management system is proposed based on big data fusion. The phase space reconstruction technology is used to reconstruct the feature of library and information. The feature quantity of semantic concept set of library information is extracted, and the classification storage and information retrieval of library information is carried out by fuzzy clustering method. The adaptive training method is used for feature fusion, and big data fusion of library and information is realized in high dimensional feature space. The data processing center is set up under the Linux kernel environment, the application program of the university library information management system is developed under the Linux kernel, and the VXI bus technology is used to transmit and schedule the university library information management information and data. Realize the software development and design of the school library information management system. The test results show that the design of university library information management system with this method has good information storage and scheduling ability, and it improves the performance of library information retrieval. In the information recall rate and recall rate and other indicators performance has an advantage.

## 1 INTRODUCTIONS

With the continuous development and progress of social economy, people's demand for knowledge and culture is increasing. As an important place and intelligence department for document collection, document collation and document collection, libraries provide readers with a large amount of information cultural resources. In order to meet the reader's reading and document collection needs to provide an instrumental role<sup>[1]</sup>. As an important department to deal with information and provide information services, library plays an important role in the accumulation and dissemination of human knowledge. With the development of computer technology and the popularization of Internet, the management mode of university library is also developing towards the direction of information, technology and network. The traditional library management mode is difficult to adapt to and meet the needs of the times, which affects the development of library information management and service function. It is necessary to set up a library and information management system supported by information technology and computer technology, and combine information, management and system organically in order to maximize library function, standardize management and expand service. It is of great significance to realize the intelligent management of library and information and to study the optimization

design of library information management system in colleges and universities<sup>[2]</sup>.

The design of library information management system is constructed based on big data information processing and database optimization design. The intelligent design of library information management system is realized. In the traditional methods, the methods of nonlinear time series analysis, statistical sequence analysis and piecewise regression analysis are mainly used in the research of library information configuration and access control model<sup>[3]</sup>. Semantic fusion methods and so on, the above methods first carry on the non-linear time series information flow fitting to the library information resources, through the library information resources storage structure optimization design to improve the throughput of library and information resources, the balanced control and optimal allocation of library and information is achieved, some research results have been achieved<sup>[4]</sup>. In reference [5], a library information resources equilibrium control and configuration model is proposed based on adaptive screening of distributed structure, and phase space reconstruction method is used to reorganize the structure of storage space to improve the ability of information management and retrieval in library, however, there are some problems in the method of information retrieval in library, such as too much calculation and low precision. In reference [6], a method for optimizing the retrieval of university library and information is proposed based on the fusion clustering of relevance semantics, and the

distributed structure model of university library information retrieval node is constructed by using directed graph model. In order to improve the retrieval performance and the ability of information management, the piecewise fusion fuzzy clustering method is used to cluster library information. The global convergence of this method is bad.

In view of the above problems, a design method of library and information management system is proposed based on big data. This paper uses phase space reconstruction technology to reconstruct library and information features, extracts feature quantity of semantic concept set of library and information, and realizes big data fusion of library and information in high-dimensional feature space. The VXI bus technology is used to transmit and schedule the information and data of university library and information management, and the software development and design of the school library information management system is realized. Finally, the test and analysis are carried out to show the superior performance of this method in improving the management ability of library and information.

## 2 LIBRARY AND INFORMATION FEATURE EXTRACTION AND INFORMATION PREPROCESSING

### 2.1. Reconstruction of Library and Information distributed structure

In order to realize the library information retrieval and information fusion based on big data, this paper adopts big data information analysis method to carry on the library information retrieval, combines the data clustering method to carry on the book information optimization dispatch and the information fusion<sup>[7]</sup>. The distributed structure reconstruction and data structure analysis of library and information are realized. The two subspaces of library and information distribution in phase space are  $S$  and  $Q$  respectively, which are composed of the solution vector  $\{s_1, s_2, \dots, s_n\}$  of fuzzy training set. The conditional probability of the clustering attribute of library and information is  $P_s(s_i)$ ,  $i = 1, 2, \dots, n$ ,  $Q$  is composed of the solution vector  $\{q_1, q_2, \dots, q_n\}$  of the fuzzy test set of the book retrieval information, and the detection probability of the corresponding library information fusion is  $P_q(q_j)$ ,  $j = 1, 2, \dots, n$ . In the library information management system, the phase space reconstruction model of library information is constructed as follows:

$$H(S) = -\sum_{i=1}^n P_s(s_i) \log_2 P_s(s_i) \quad (1)$$

$$H(Q) = -\sum_{j=1}^n P_q(q_j) \log_2 P_q(q_j) \quad (2)$$

Where,  $P_s(s_i)$  denotes the probability that the semantic concept set  $s_i$  of library information appears

in the chaotic region  $S$ . similarly,  $P_q(q_j)$  denotes the probability that the ontology feature concept set  $q_j$  of library information appears in the chaotic region  $Q$ . In phase space  $S$ , the average mutual information satisfying the condition of library and information clustering is expressed as follows:

$$I(Q, S) = H(Q) - H(Q|S) \quad (3)$$

Where

$$H(Q|s_i) = -\sum_j \left[ \frac{P_{sq}(s_i, q_j)}{P_s(s_i)} \right] \log_2 \left[ \frac{P_{sq}(s_i, q_j)}{P_s(s_i)} \right] \quad (4)$$

It represents the classification attribute set satisfying condition  $P$ , and the method of average mutual information fusion is taken to obtain the distributed reconfiguration structure of library and information retrieval:

$$X(n) = \{x(n), x(n + \tau), \dots, x(n + (m-1)\tau)\} \quad n = 1, 2, \dots, N \quad (5)$$

The adaptive training method is used for feature fusion. The distribution locus of library and information retrieval in high dimensional feature space is shown as follows:

$$X = [s_1, s_2, \dots, s_K] = \begin{bmatrix} x_1 & x_2 & \dots & x_K \\ x_{1+\tau} & x_{2+\tau} & \dots & x_{K+\tau} \\ \dots & \dots & \dots & \dots \\ x_{1+(m-1)\tau} & x_{2+(m-1)\tau} & \dots & x_{K+(m-1)\tau} \end{bmatrix} \quad (6)$$

Where  $K = N - (m-1)\tau$ , it represents the semantic feature concept set of book information retrieval,  $\tau$  is embedded dimension,  $m$  is the layer number of semantic ontology attribute of book information retrieval,  $s_i = (x_i, x_{i+\tau}, \dots, x_{i+(m-1)\tau})^T$  is called feature vector set in phase space.

### 2.2. Feature extraction of semantic concept set of library and information

On the basis of the reconstruction of the phase space of Library and information, the feature quantity of the semantic concept set of books and information is extracted<sup>[8-10]</sup>. The fuzzy clustering method is used to carry out the classified storage and information retrieval of the books and information, and the statistic of the library and information fusion is:

$$R_{mn} = \|X_{\eta(n)} - X_n\|_2^{(m)} = \min_{j=N_0, \dots, N, j \neq n} \|X_n - X_j\|_2 = \sqrt{\sum_{l=0}^{m-1} (x_{\eta(n)+l\tau} - x_{n+l\tau})^2} \quad (7)$$

Adaptive learning is taken in the reconstructed phase space, the embedded dimension of the phase space is increased from  $m$  to  $m+1$ , and the semantic feature concept set of library and information fusion is obtained as follows:

$$R_{(m+1)n} = \|X_{\eta(n)} - X_n\|_2^{(m+1)} = \sqrt{\sum_{l=0}^m (x_{\eta(n)+l\tau} - x_{n+l\tau})^2} \quad (8)$$

$$= \sqrt{\left[\|X_{\eta(n)} - X_n\|_2^{(m)}\right]^2 + [x_{\eta(n)+m\tau} - x_{n+m\tau}]^2}$$

If  $R_{(m+1)n}$  is much larger than  $R_{mn}$ , it can be used as a fuzzy clustering center, and it can be used as a fuzzy clustering center. According to this property,  $X_{\eta(n)}$  is the nearest neighbor of  $X_n$ . The judgment criteria of library and information retrieval are obtained as follows:

Norm (1):

$$\sqrt{\frac{R_{(m+1)n}^2 - R_{mn}^2}{R_{(m+1)n}^2}} = \frac{|x_{\eta(n)+m\tau} - x_{n+m\tau}|}{R_{(m+1)n}} \geq R_{tol} \quad (9)$$

Norm (2):

$$\frac{R_{(m+1)n}}{\sqrt{\frac{1}{N} \sum_{k=1}^N \left[ x_k - \frac{1}{N} \sum_{k=1}^N x_k \right]^2}} > A_{tol} \quad (10)$$

Where,  $R_{tol}=15$ ,  $A_{tol}=2$ .

According to the result of feature quantity extraction from the semantic concept set of library and information, this paper combines big data's information fusion method to carry out information management and optimization retrieval<sup>[11]</sup>.

### 3 OPTIMIZATION OF LIBRARY AND INFORMATION RETRIEVAL IN COLLEGES AND UNIVERSITIES

Based on the reconstruction of library and information features using phase space reconstruction technology, the optimization design of library and information retrieval in colleges and universities is carried out. In this paper, the optimal design of library and information management system based on big data's fusion is put forward, and fuzzy clustering is adopted<sup>[12]</sup>. Class method is used to store library and information, and the feature classification set is obtained as follows:

$$J(\bar{X}_j) = \frac{y_j^T S_b y_j}{\lambda_j}, \quad j = 1, 2, \dots, l \quad (11)$$

Clustering in descending order according to the size of the distribution of characteristic attributes:  $J(\bar{X}_1) \geq J(\bar{X}_2) \geq \dots \geq J(\bar{X}_l)$ , taking the minimum eigenvalue as the clustering center, the data point  $\bar{X}_i$  is obtained. Band the feature distribution vector of information retrieval is obtained as:

$$W = [y_1, y_2, \dots, y_d] \quad (12)$$

Optimized information classification set is:

$$F_j = \sum_{k=1}^n X_{kj}, \quad Q_j = \sum_{k=1}^n (X_{kj})^2 \quad (13)$$

Thus, the big data information fusion of library and information management is realized, and the performance of library information management is analyzed with recall rate, load and time cost as test

index<sup>[13]</sup>.

### 4 SOFTWARE DEVELOPMENT AND DESIGN OF THE SYSTEM

On the basis of the design of intelligent information retrieval and data fusion algorithm in university library and information management, the software development and design of library information management system are carried out, and the data processing center is established under the environment of Linux kernel. The application program of university library and information management system is developed under the Linux kernel, and the data transmission and scheduling of university library information management information is carried out by using VXI bus technology<sup>[14]</sup>. The software design of library information management system mainly includes program compiling module and bus driver module. The data cache is designed by MySQL, and the library information is read and received in real time by combining TCP/IP server and A / D conversion protocol. Send control, call free\_irq() function to realize the cross-compilation of library information management program. This paper adopts the design method of three-layer network system, adopts radio frequency tag identification technology and information scanning technology to collect book resource information in the perceptual layer of book information, and inputs the collected book resource information into the middle layer<sup>[15]</sup>. Adaptive processing, using the above designed library and information resources mining algorithm, resource optimization fusion and integration processing, in the network layer of adaptive transceiver conversion control, according to the above design, get the construction of this paper. The software development flow of college library and information management system based on big data is shown in figure 1.

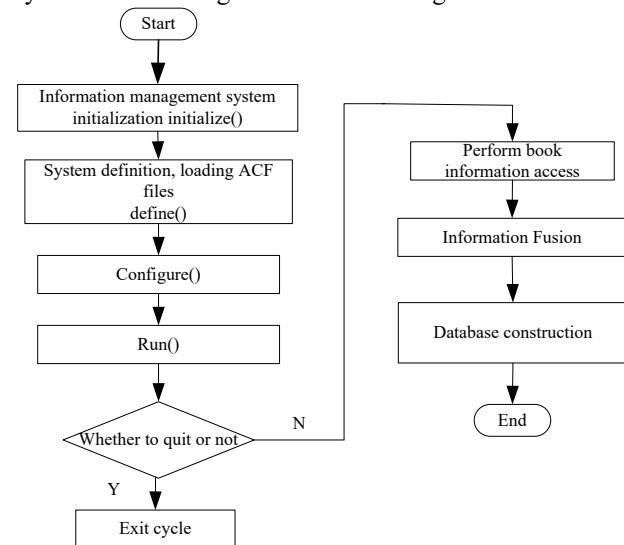
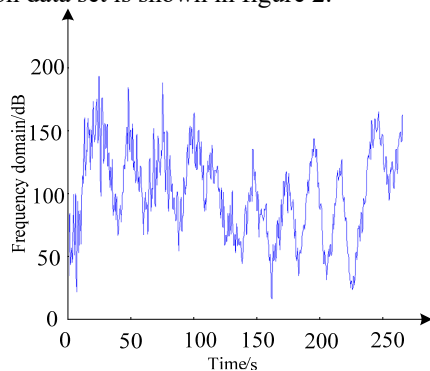


Figure 1 Software development and design flow of the system

### 5 EXPERIMENTAL TEST ANALYSIS

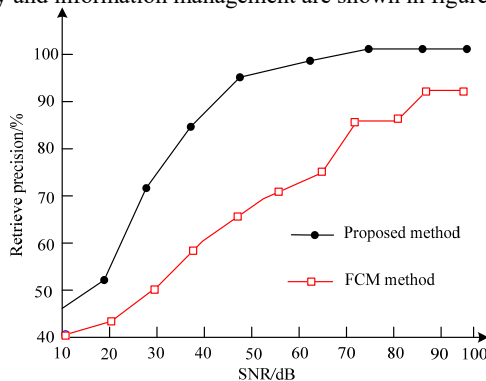
In order to test the superior performance of the system

designed in this paper in the realization of library and information management in colleges and universities, the system is debugged and analyzed. The experiment is built on the embedded Linux kernel platform, and the command source install is compiled. Install drivers for data loading and program loading, test the accuracy of library and information retrieval at the main interface of the information management system, the interval of sampling time for library and information is 2.5s, the scale of test set is 1200 Gbit, the sample length of big data training set is 1024, the embedded dimension of phase space reconstruction of information retrieval is  $m=4$ , and the data clustering distribution interval is  $[0,1000]$ . According to the above simulation environment and parameter setting, the university library information retrieval is carried out, and the original collection data set is shown in figure 2.



**Figure 2** Time-domain sampling waveform of library information

The sampling waveform of figure 2 is taken as input, the performance test of information management function is carried out, and the results of comparison of the precision of library and information management are shown in figure 3.



**Figure 3** Comparison of checking performance

Figure 3 shows that the design of University Library and information management system is carried out by this method, and the accuracy rate of information retrieval is higher and the performance is better.

## 6 CONCLUSIONS

In this paper, a design method of university library and information management system based on big data fusion is proposed. This paper adopts phase space reconstruction technology to reconstruct library and

information features, extracts feature quantity of semantic concept set of library and information, uses fuzzy clustering method to classify, stores and retrieves library information, and adopts adaptive training method. The data processing center is established under the Linux kernel environment, the application program of the university library information management system is developed under the Linux kernel, and the VXI bus technology is used to transmit and schedule the university library information management information and data, so as to realize the school. Software development and design of library and information management system are obtained. The test results show that this method has good information storage and scheduling ability, improves the performance of library and information retrieval, and has good performance in information recall and recall rate. The research results show that the good application value of the design system.

## REFERENCES

1. LING Conggang, WANG Hongzhang. Optimization research on differential evolution algorithm and its application in clustering analysis[J]. Modern electronic technology, 2016, 39 (13): 103-107.
2. MI Jie, ZHANG Peng, YU Haipeng. Large Data Clustering Algorithm Based on Particle Swarm Differential Perturbation Optimization[J]. Journal of Henan University of Engineering (Natural Science Edition), 2016, 28(1):63-68.
3. XING Shuning, LIU Fang'ai, ZHAO Xiaohui. Parallel high utility pattern mining algorithm based on cluster partition. Journal of Computer Applications, 2016, 36(8): 2202-2206.
4. Palomares I, Martinez L, Herrera F. A consensus model to detect and manage non-cooperative behaviors in large scale group decision making[J]. IEEE Trans on Fuzzy System, 2014, 22(3): 516-530.
5. LI Ziyang, YU Jiong, BIAN Chen, WANG Yuefei, LU Liang. Dynamic data stream load balancing strategy based on load awareness. Journal of Computer Applications, 2017, 37(10): 2760-2766.
6. Sun Li-juan, Chen Xiao-dong, Han Chong, Guo Jian. New Fuzzy-Clustering Algorithm for Data Stream[J]. JEIT, 2015, 37(7): 1620-1625.
7. XING Changzheng, LIU Jian. Evolutionary data stream clustering algorithm based on integration of affinity propagation and density. Journal of Computer Applications, 2015, 35(7): 1927-1932.
8. HOU Sen, LUO Xing-guo, SONG Ke. A Maximum Entropy Weighted Trust-Analysis Algorithm Based on Sources Clustering[J]. Chinese Journal of Electronics, 2015, 43(5): 993-999.
9. BI Anqi, WANG Shitong. Transfer Affinity Propagation Clustering Algorithm Based on Kullback-Leiber Distance[J]. Journal of electronics and information, 2016, 38(8): 2076-2084.

10. LIU Jun, LIU Yu, HE You, SUN Shun. Joint Probabilistic Data Association Algorithm Based on All-neighbor Fuzzy Clustering in Clutter[J]. *Journal of electronics and information*, 2016, 38(6): 1438-1445.
11. WU Hong-hua, MU Yong, QU Zhong-feng, DENG Li-xia. Similarity and nearness relational degree based on panel data[J]. *Control and Decision*, 2016, 31(03): 555-558.
12. PARVIN H, MIRNABIBABOLI M, ALINEJAD-ROKNY H. Proposing a classifier ensemble framework based on classifier selection and decision tree[J]. *Engineering Applications of Artificial Intelligence*, 2015, 37:34-42.
13. DURAO F, DOLOG P. Improving tag-based recommendation with the collaborative value of wiki pages for knowledge sharing[J]. *Journal of Ambient Intelligence Humanized Computing*, 2014, 5(1):21-38
14. GUO L, MA J, CHEN Z M, et al. Incorporating item relations for social recommendation[J]. *Chinese Journal of Computers*, 2014, 37(1):219-228.
15. XIONG Xinglong, YANG Lixiang, MA Yuzhao, ZHUANG Zibo. Alerting algorithm of low-level wind shear based on fuzzy C-means. *Journal of Computer Applications*, 2018, 38(3): 655-660.