

Research Article

Wireless Network Communication in the XML Metadata Storage of Wushu Historical Archives

Xin Liu 

Department of Arts and Sports, Henan Technical College of Construction, Zhengzhou, 450064 Henan, China

Correspondence should be addressed to Xin Liu; liuxin@hnjs.edu.cn

Received 4 August 2021; Revised 28 September 2021; Accepted 6 October 2021; Published 8 November 2021

Academic Editor: Zhihan Lv

Copyright © 2021 Xin Liu. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The wireless communication network has a huge promotion effect on data processing. As an important standard for internet data transmission, XML markup language is currently widely used in the internet. The rapid development of XML has brought fresh blood to the research of database. This article is aimed at studying the application of wireless network communication in the storage of the XML metadata database of Wushu Historical Archives. One part uses data load-balancing algorithms and two-way path constraint algorithms to conduct research on the metadata storage of XML databases, such as document size, document loading, and document query. Combining wireless technology and comparing with SGML and HTML, it highlights the flexibility and importance of XML database. The other part analyzes the XML database of Wushu Historical Archives itself and uses questionnaire surveys combined with the experience of 30 users randomly checked to provide constructive suggestions for the construction of XML database, such as strengthening the operation page design, text, and graphics. The experimental results show that the time consumed by the XML database in the case of bidirectional path indexing is 0.37 s, and the time consumed in the case of unidirectional path indexing is 0.27 s. This data shows that the application of wireless network communication technology has played an important role in the easy update and flexibility of the XML database. The data load capacity presented by the two path indexes is inconsistent; the current experience of XML database mainly focuses on the copy of the database quantity.

1. Introduction

With the rapid development of communication technology, especially the popularization and application of internet technology, people can receive and send information through the internet. A common problem in the process of information exchange is the diversity of data formats, which greatly hinder the exchange of information and resource sharing [1]. With its compact size, personalized description, and structural features, it provides a new solution for diversified data exchange and has become a standard format for agent and data exchange. How to effectively manage large-scale XML database data has become a very difficult new problem in current databases. In addition, with the expansion of XML data and the increasing development of self-description, more and more developers regard it as the main form of network data transmission. At the same time, the demand for massive

XML data storage has led to the birth of a new type of XML database. How to improve the query performance of XML database is particularly important under the premise of massive data. XML file is essentially a structured carrier for storing information, and its processing capacity for its own database is limited [2]. The advantage is that it is semistructured and can customize the schema. Compared with the relational two-dimensional table, it does not need to follow the first normal form. Heterogeneous data integration can be achieved through schema. Converting XML data into relational data, and then relying on the powerful functions of relational databases to process the data, will bring considerable processing effects [3].

As a component of the data system, historical data can ensure the stability and normal operation of the system [4]. With the growth of research in big data-related fields, research on historical data storage has become an important topic. From traditional relational databases to nonmodern

databases, from distributed databases to distributed databases, from massive data to large databases, the current technical research-database equipment is based on compressed field data and actual scene data. The first research direction is the development of compression algorithms. However, due to the shortcomings of large databases in related databases, it cannot meet the system requirements of large data with simultaneous high read and write, high-performance access, and large-capacity storage [5].

For the development of wireless sensor network technology, high memory, storage, processing, and communication capabilities make it an inevitable trend to reduce the energy consumption used in data transmission. Jawhar et al. adopted a model to reduce hidden terminals and conflicts in the multihop data routing method to establish a model for data collection, and the experimental results obtained an effective method for delay tolerant data traffic [6]. Gafurova et al. discussed the presentation characteristics of various electronic collection document metadata, which are related to the format used and changes in the composition and completeness of the collection. The metadata that runs through the entire publication of the corresponding scientific journal shows and describes the software tools used to manage scientific content and the automatic integration methods used to organize the mathematical document repository and other information systems [7, 8]. Adams introduced XML (extensible markup language) coding as a way to describe the intelligent structure of metadata records, which makes it possible to develop tools to manipulate, merge, and analyze metadata records. In XML, there are content and markup. Tags are added to disclose the structure of the content, including elements, attributes, comments, and processing instructions. Similar to HTML tags, XML tags are separated from parenthesized content. The elements are in parentheses, and the attributes are the values in quotes. Elements show the hierarchy of structures through nesting. In XML, developers can define all the element and attribute names required by the metadata content [9]. Olmedo's legislative semantic web based on structured document (XML), metadata (RDF), and ontology (OWL) technology emerged as the best platform model for the presentation, organization, and distribution of legal content, which is an effective response to the phenomenon of pollution caused by legislation. The results require the need for better and easier access to legislation [10]. Guccio et al. solve this problem by studying the efficiency of the Italian Public Historical Archives (PHA) from 2011 to 2012, use data envelopment analysis (DEA), and implement one-stage and two-stage methods to investigate the differences and possibilities of PHA in managing it, the performance of conflicting functions [11]. In his experiment, Michalak used the example of two teachers trying to teach the basic file practice of summer bridge course (boot camp) students, through the use of information literacy skills (ILA) and students' perceptions of their information skills-questionnaire (SPIL-Q) tools to quantitatively measure their information literacy skills (citation). In the SPIL-Q after training, the average perception confidence of the boot camp students through the SPIL-Q instrument to assess IL skills increased from

4.00 to 4.77 (+19.2%) [12]. Silovsky et al. based their study on manually annotated archive partitions. We have compiled a database covering a span of more than 30 years for experimental research. The score shows changes in the target (real) speaker test introduced by the aging effect, the signal-to-noise ratio value, or the variable number of registration and test data. The score of the speaker detection test is based on the vector paradigm and probabilistic linear discriminant analysis system evaluation. They also evaluated the performance of the system using an evaluation database, which contains contemporary recordings collected over a time span of approximately 4 years [13]. The above-mentioned research on data processing and data storage has undoubtedly played a great role in promoting the exploration of the field of database. However, after careful study, it can be found that their research methods are relatively speaking. There are shortcomings such as lack of novelty, too single data, and lack of rigorous construction of data index processing models.

The reliability of data is the top priority in the field of data storage. This paper uses dynamic and static data storage algorithms to analyze and research the application of wireless network communication in the XML metadata storage of Wushu Historical Archives. The original problem was decomposed into several subproblems, then the smallest subproblem was solved first from the bottom up, and the results were stored in a table. When solving a large subproblem, the solution of the small subproblem was directly queried from the table to avoid repeated calculations to improve algorithm efficiency. The logical model of XML data and the two-way path constraint model are constructed through the data load-balancing algorithm to analyze the data volume of the real-time monitoring database inside the operation and maintenance system and find ways to improve the data storage efficiency. The polling method, random method, and minimum average method of load-balancing algorithm were used to select the document size and calculate the time of document loading. The similarity of XML data was combined to build an ordered tag tree, and the bidirectional path constraint algorithm was used to calculate it; aiming at the characteristics of the big data of the martial arts historical archives and the user's needs to access behavior, we are determined to research and implement an effective big data system database organization system to meet the needs of secure distribution of massive data in digital libraries [14].

2. Data Storage Method of Wushu Historical Archives

2.1. Wireless Network Communication and Martial Arts Historical Materials Museum. With the strong development of wireless communication technology and the increase of wireless business doubling speed, users have put forward higher requirements for the resources provided by wireless networks, that is, smart and close to life [15]. How to improve the performance of wireless networks under the premise of ensuring user integrity and intelligent benefits has always been one of the most important research

guidelines for wireless communication networks. In today's society, the wireless information network community is slowly spreading, wireless communication channels have increased sharply, and communication bandwidth has been further improved, creating favorable technical conditions for the development of digital libraries services. With the help of advanced wireless communication technology, digital library services are ubiquitous, and readers can access the digital information resources and services provided by the family library anytime and anywhere through any terminal application. XML can be used to express semistructured data. It is not obvious to the model that missing information and data with complex structure can be expressed very well. Especially in many web systems, XML is already a standard form of data exchange and expression. Therefore, the efficient management of XML data has important applications in internet systems. Libraries should continuously change traditional service concepts and service models. By introducing new technical conditions and constructing new service models, it provides readers with richer collections of digital resources, so that readers can obtain information and conduct learning activities anytime and anywhere, and the public has equal access to information [16, 17]. The digital library is a new library concept and an important form of searching for future libraries. The nature of the digital library form and the inequality of services have expanded the scope and scope of services and expanded the content of the project [18].

Wireless network application is the current research hotspot in the computer field, and it is also a convenient way for the library to achieve technological leapfrogging and improve the level of reader service. Based on wireless network technology, it expounds the research of library development with the help of advanced technology for database development [19, 20]. Digital technology converts traditional historical materials of paper texts into compact discs through digital compression and encoding. The digitization of historical materials and fundamental changes in their external form have expanded the scope of propaganda and application of historical materials. The application itself of the digital library has relatively simple requirements, but the requirements for storage and retrieval are relatively high. "Cloud" has a considerable scale, supporting users to obtain application services at any location and using various terminals. Measures such as fault tolerance of multiple copies of data and interchangeability of computing nodes are used to ensure high reliability of services. Cloud computing is better than using a reliable local computer. The cloud platform uses massive data storage and cloud search technology to build a general data storage and retrieval platform. On this basis, in independent development of digital data management platform, each unit can use this platform to build its own digital library, and the built system runs directly on the cloud platform. For such applications that require large amounts of stored data and relatively high computing resources, the cloud platform is used to achieve their advantages are particularly obvious. With the development and popularization of digital data management technology,

research on digital archives oriented to specific fields has gradually received attention, and a large number of enterprise-oriented digital data management systems and electronic archives management systems have also emerged [21]. When the library's service system fails, cloud computing technology can clone these massive data in a very short time and start a new server to provide services, which achieves the goal of uninterrupted service.

2.2. Principles of Dynamic and Static Encoding of XML Data. Various parameters are obtained directly and accurately, and the returned data information can supplement the monitoring points that cannot be realized by SNMP monitoring method, greatly expanding the scope and types of monitoring. Data management should be based on the XML data model, starting from the basic characteristics of XML data, designing the physical storage concept of XML database, fully realizing the XML data model, supporting multiple index features, being able to support multiple XML auxiliary indexes, and effectively creating data queries. In order to find an effective management solution that supports your unique tree system, a native XML data management solution is the key to solving the problem [22]. Since the underlying storage expression has an important influence on the query and optimization of the upper layer, it is also an effective method to improve the response time of processing XML data. The previous link record was added to keep the virtual address of the mobile record unchanged, and then the correctness of various index structures was maintained; the relocation record maintains the document order storage of the actual node record and solves the problem between the efficient update of the XML database and the index maintenance, a contradiction. Adding positioning records maintains the document order storage of the actual node records, which solves the contradiction between the efficient update of data and the difficulty of index maintenance (simulation is shown in Figure 1). Regarding the internal data of the Wushu Historical Archives as a node according to the sequence of book names, the driving force for this category can be the initials or the type of historical materials, such as literature, history, or astronomy. A large number of historical data points such as heuristic direction information were used. In such a group of current visiting departments, multiple pages similar to the current page are repeatedly selected as the meeting point for the next visit. Inside the database, the files are divided into each node and stored [23]. The principle and working method of XML database storage at the granularity of nodes regard each category as a different level, and each level can be expressed as

$$HC_i = ec_{;ceil}(\log pto_j) \sqrt{\alpha}. \quad (1)$$

HC_i represents each node in the space, and the utilization rate of the layer mark pto_j is restricted by the level within a fixed length. The encoding between two lines

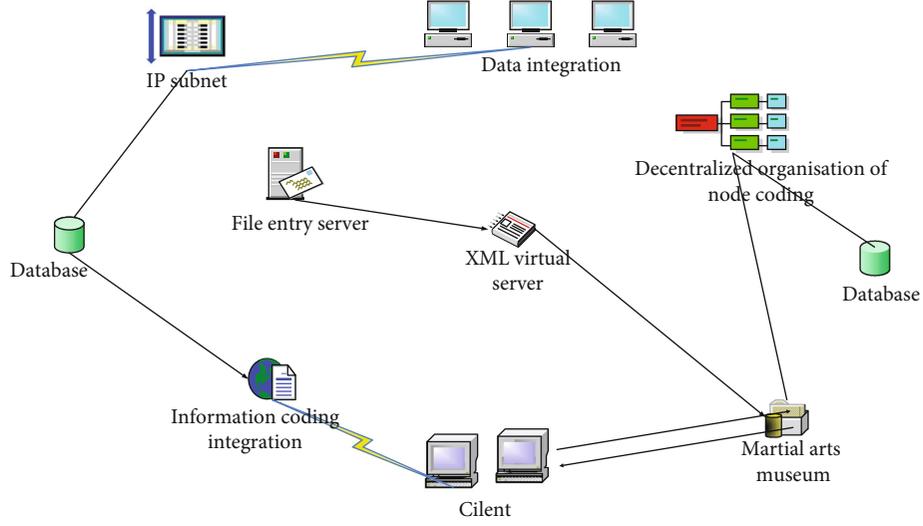


FIGURE 1: Document order storage.

can be expressed as

$$2^{li+1} - 1 = \sum_{i=1}^{L_f} 2^{i-1} \leq L_{pto} \leq 1, \quad (2)$$

$$Hl_i = l_p(L_{COSe} + i) + L_c(I^e \times \beta).$$

According to the characteristics of the documentation, serialization is performed in combination with the fixed-length field of the smaller layer markcos e , which can greatly save the space utilization of communication technology data [24]. The best matching space is as follows:

$$Q_{pto} = \frac{1}{0.375} \times \sqrt{\frac{2}{ay}} \times \sqrt{\frac{\beta ef}{\beta il}}, \quad (3)$$

$$W_t = \frac{w_{pto}}{(1 + r \times (\chi - r_0 \times (\lambda - \mu)))}.$$

w_{pto} is the weight probability of the best matching node in the space. In the case of heterogeneous nodes, through energy consumption r_0 , there is the overall threshold of the space

$$A(v_i) = \frac{x_i}{1 - \rho_i \times (t \bmod (1/\rho_i))}, \quad (4)$$

$$T_{ip} = \sqrt{(a_{ei} - m_{\sin i})^2 + (b_{ei} - n_{\sin e})^2 (1/v_{e=0}^{i-1})}.$$

The energy consumption of node deployment within the overall threshold is positively correlated with the density a_i^e and remaining energy b_i^e of each node. The weighting coefficient is used to search for stored files in the XML database or search for a certain node [25], the process of emphasizing the importance of different characteristic elements of historical materials in the historical archives in the entire element system and assigning a certain charac-

teristic value to this element. The characteristic value is generally represented by a numerical value, which can be expressed as

$$D_{ei} = T_i / \sin m_e^i + T_{i+1} (\cos \mu_{i-1}^{e=0}), \quad (5)$$

$$d_{\min}(w_t) = \sum_{t \notin W} (a, b) + \text{long}_{\min}(w, \cos_t).$$

long_{\min} represents the minimum weight value of the data aggregation of the node code in a fixed length. The data encoding nodes are randomly arranged in a decentralized cluster, and the storage nodes are expanded:

$${}_{i+1 \rightarrow \phi}^G l \left\{ \left| \frac{t_l}{a_{i+1}/(a_i + b_{i+1})} - \frac{a - b_i}{a + b} \right| < \theta \right\} = 1, \quad (6)$$

$$\Delta = (X_{i=1} \times a_{i=0}) / (a + b) \int \rho_i.$$

$a_i + b_{i+1}$ means any two types of historical materials in the random arrangement of scattered clusters.

2.3. XML Data Bidirectional Path Constraint Mode. The indexing technology of XML database, that is, the abstract path of XML data structure, is one of the most important technologies of XML database [26]. The indexing technologies used by mainstream XML databases vary, and all nodes at the same level are stored. In the same layer of the tree, because the elements of the same layer are all of the same type or the same name, when there is no path index, if we want to find the lower elements of a certain layer, we need to traverse all the above levels backtracking. Document order is an important concept of the structural characteristics of XML documents. Data-centric XML documents focus on the data in the document rather than the document format. The data in the document is the node with moderate data strength, and it is planned in sequence according to

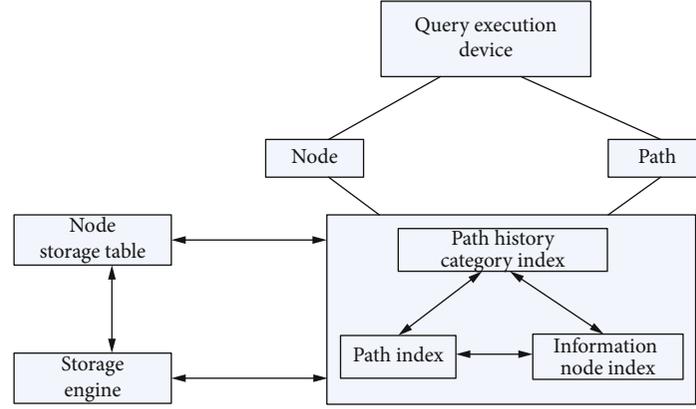


FIGURE 2: XML database index.

the string value of the attribute node. Nodes for relative and absolute paths are combined. The data storage and analysis capabilities of XML databases simplify the process of indexing, searching, and sorting (Figure 2). This study uses a two-way path method to simplify this process and applies the shortest path optimization technology and optimization algorithm in a large-scale network to convert an absolute path into a node before converting the path into the full path DTD specification syntax analysis expansion name defined by us. For the sequence form [27], the content of XML is read into the DataTable of the DataSet, and then the rows and columns are determined in the DataTable to obtain the corresponding data; the purpose of XML data classification is also to classify XML documents into specified categories, so as to facilitate the classification management and optimization of XML documents and effectively detect the order; the new and old relationships between nodes are as follows:

$$R_{ij}^a = \min R_{ij}^a, \min_1 \leq \theta \leq bR_{ik}^{a-1} + c_{kj},$$

$$\omega_{i,j} = R_{ij}^{a-1} + b_{ij} \int \theta(a, b, c). \quad (7)$$

The minimum weight of more than $b-c$ paths is recorded as $\omega_{i,j}$, where (a, b, c) is a sequence text format node [28], which is calculated as a matrix sorting to improve the efficiency of detecting nodes:

$$T_{ij} = \sum_{c=1}^b \theta_{i=1}^{j-1} \times (a+b)_{ij}, \quad (8)$$

$$E_{ij}^{a-1} = a_{ic}/b_{ic} \sqrt{r^{2 \log(a-1)} - 1}.$$

Use the PFloyd algorithm to filter the node information number of the matrix, and update it x times:

$$X_{\min} = (a \longrightarrow b \longrightarrow c) \sum_{i \leq a, b, c}^{\log i} \alpha_{m,n} \times \beta_{m,n}. \quad (9)$$

Let $(a \longrightarrow b \longrightarrow c) = 0$, and the absolute ratio of the element with the same name in the bidirectional path index constraint mode is

$$\frac{2\alpha + i}{\max [3\alpha + a, 3\beta + b]} \leq \sum_{i,j}^{a=1}, \quad (10)$$

$$Average^\sigma \leq \lim_{i,j \rightarrow \sigma} \frac{(2a+b)(i+1)}{(4a+b)(j-1)}.$$

$Average^\sigma$ is the progressive absolute value of the data node storage in the space, and the path choice under the large-scale decentralized problem of multinode data storage is [29]

$$\sum_{i=1}^{\pi} t_{ij} \leq \sum_{j=1}^{|r|-a} j_{b,c}, \quad (11)$$

$$\frac{|r|m}{m+1} \leq a \leq \frac{1}{R-r}.$$

After the data node completes the data fusion, the minimum remaining energy $|r|$ is maximized.

3. Metadata Storage Test

3.1. Experimental Content

- (1) Research on existing distributed storage-related technologies and related technologies in retrieval
- (2) Research and implement the XML database of the storage and organization mechanism of massive resources in digital libraries suitable for large-scale data sets
- (3) Select SGML, HTML, and XML databases for comparative analysis, and perform data statistics on storing, loading, and querying XML documents in the database

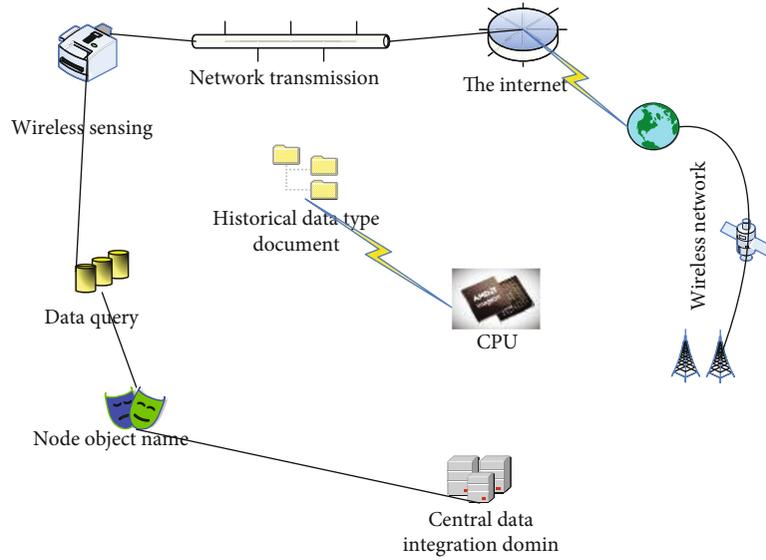


FIGURE 3: Simplified model of XML database.

TABLE 1: General case of document data storage time-consuming.

File name	Document 1	Document 2	Document 3	Document 4	Document 5
Document size	19.2 kb	71.3 kb	123.8 kb	289.4 kb	555.9 kb
General storage time	0.24 s	0.42 s	0.62ss	0.88 s	0.97 s

- (4) Use the XML database data processing structure model
- (5) Apply the digital library system architecture.

3.2. Experimental Arrangement

- (1) Adopt a two-tier architecture consisting of a transport layer that achieves load balancing, caching, and prefetching and a storage layer that combines distributed file systems and reliable storage. The basic logical unit and entity of information storage in the digital library system structure is a method of expressing information memory in a certain structured digital expression form. The entire system is managed by relational data
- (2) Propose a comprehensive distributed index and query mechanism for massive data in digital libraries
- (3) Study the feasibility of constructing a new type of digital library system architecture and mass data storage organization mechanism by XML database.

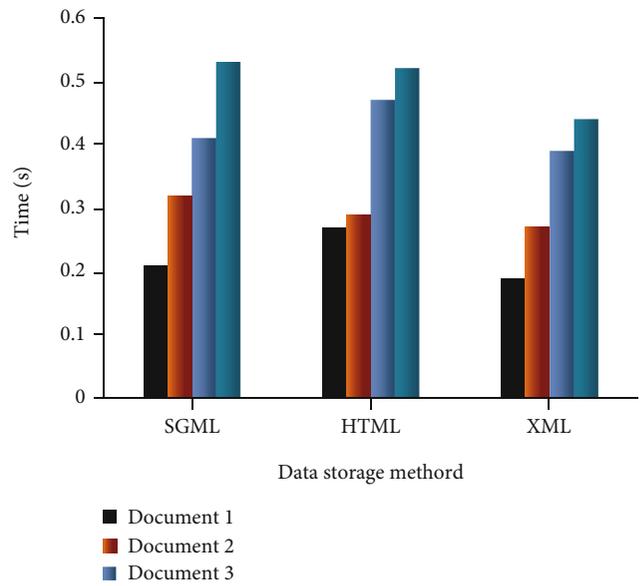


FIGURE 4: Database document storage time.

3.3. Characteristics of Wireless Communication Network and XML Metadata Storage. In addition to ensuring the credibility of the cloud storage platform, the data access control mechanism ensures that the confidentiality, availability, integrity, and consistency of data are the basic security mechanism that cloud storage services need to provide. The network communication connection protocol module includes two parts: the server and the client. The client

encapsulates the user’s request into a network data packet in a predetermined format and then sends it to the server; the server unzips the data packet according to the predetermined format and instantiates the request object, the request to the server framework. The diversification of methods and ways for users to check data requires that users should not be inclined in the method, degree of professionalism, and sequence of document retrieval when organizing digital

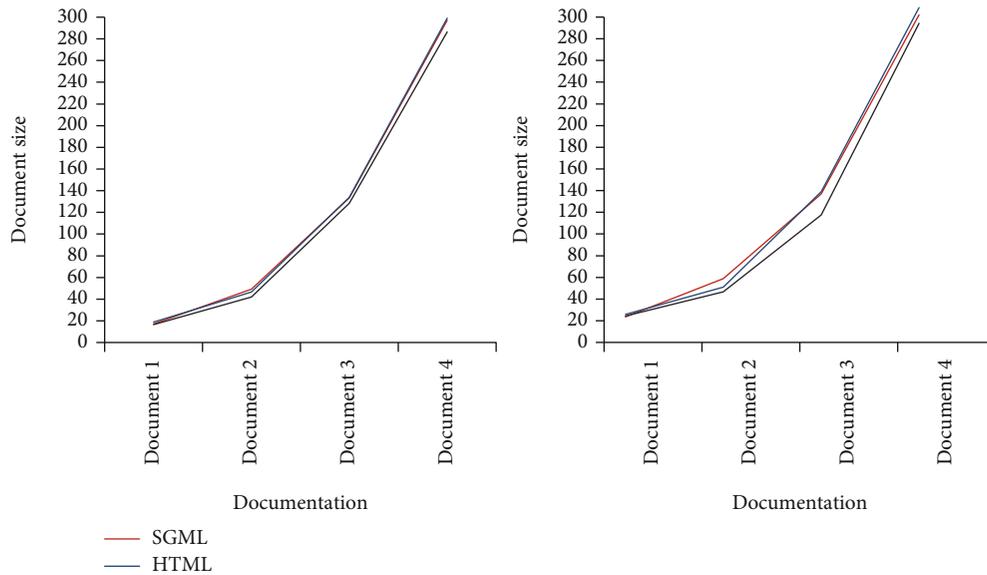


FIGURE 5: Three database storage methods of closed path and open path index and also the size relationship before and after document storage.

information. Indexing service provides a mechanism for discovering digital objects and receives retrieval requests from users. With the rapid development of computer technology, the degree of informatization of society is getting deeper and deeper, and a large amount of data is rapidly generated along with this process. The database management system plays a very important role in the management of these data. As the foundation of various functions of memory database, storage management has a very important position. Storage management not only directly affects the performance of in-memory databases but also affects various upper-level functions, so the research on in-memory database storage management is of great significance. There are a large number of XML documents emerging in the mobile internet. How to manage XML data effectively and in a timely manner has become a hotspot for the entire mobile internet industry. In order to explore the internal connections and manifestations of data, this research carried out the storage and query of XML documents in the database analysis. First, a simple simulation of the model is carried out on the basic processing structure of XML data, as shown in Figure 3.

Because, in the column storage database, the relational table is stored as multiple separate files, and each file corresponds to a data column, the traditional row storage database method is not applicable. For column store databases, a conventional method is to first read the first data column file and find out the location set of tuples that meet the corresponding query conditions. The overall memory database should include the query client, query engine, storage engine, and other components. The client is the bridge between the user and the in-memory database system. The query engine is responsible for generating logical plans, physical plans, and executing physical plans. The technical point in the query engine is how to optimize the generation and execution of these plans. The storage engine is responsible for storing data. The technical points are to improve data

compression rate, increase data reading speed, and ensure storage reliability, load balance, and other issues. According to relevant literature statistics, the average time of storing files in various databases in general is shown in Table 1.

The calculation and processing of network data need to consume more time and storage space, making parallel computing more and more important in high-performance computing algorithms. We compare the storage time of SGML\HTML\XML three data storage technologies for documents of different sizes. The data is shown in Figure 4.

XML database inherits the rich markup function of SGML and the less vigilant features of HTML and defines the data structure in an open and self-describing way. XML can describe data content and structural characteristics at the same time. Through structural characteristics, you can understand the semantic relationship between data. SGML is too complex and requires a lot of expensive software to operate, so it is too simple and lacks flexibility if it cannot be effectively promoted and used in practice. We conduct research and statistics on the differences in document size changes, document query, document loading, and one-way and two-way path indexes under the two different situations of SGML, HTML, and XML storage and open and closed path indexing. The first is the study of the size changes of the three database documents before and after storage of the status of the different path indexes, as shown in Figure 5.

Query technology is one of the key technologies of the database system, through which users can query the data of interest from the database. The current research on query technology in column store databases mainly includes the execution of query operations and the execution of query statements. The cost of data transmission in the network is much more than that of processing data. It is necessary to compress the data in the process of data transmission. There is a certain connection between the loading and storage time of the document and the size of the document in terms of

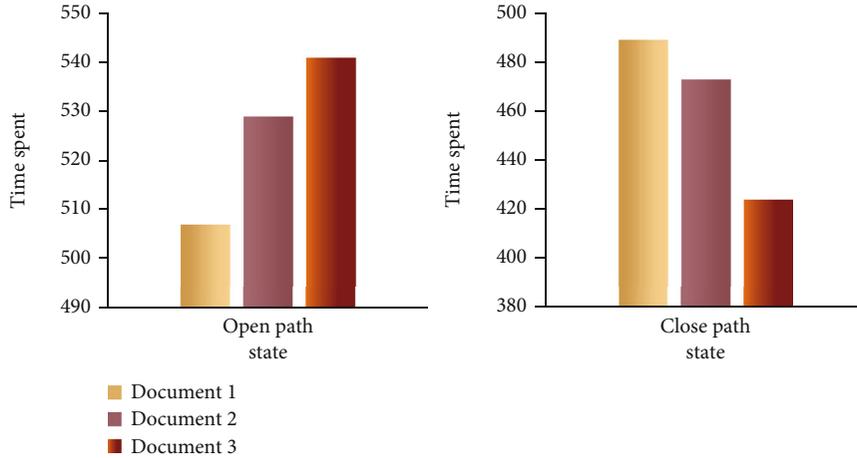


FIGURE 6: Close the path and open the path index document query XML database test.

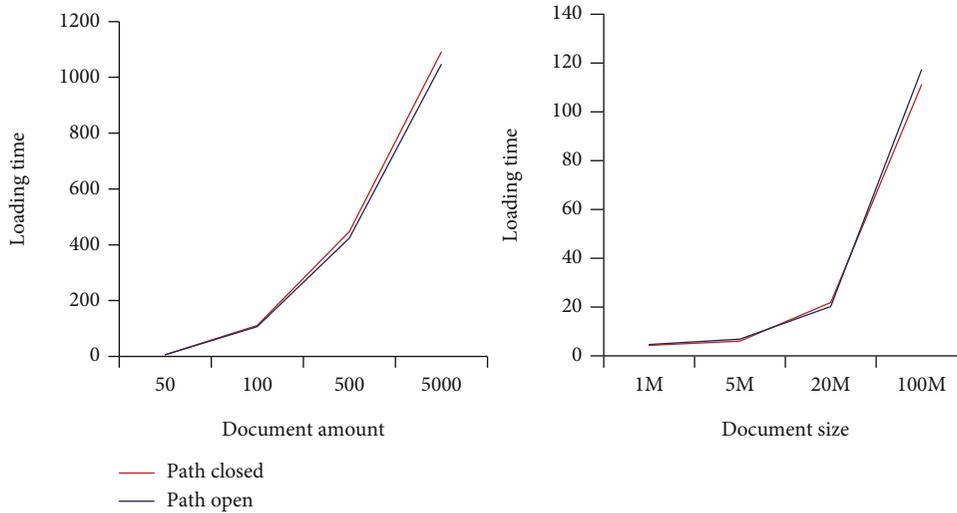


FIGURE 7: Load document test with closed path and bidirectional path index.

experimental data. The experimental data shows that the larger the document, the longer the overall loading time and the longer the storage time. The storage time of documents and the size of documents increase linearly and greatly facilitate data exchange between related technologies and systems such as data mining. The design of XML database is more reliable than the SGML and HTML format, which helps to reduce the risk of data loss caused by file damage. The damage of some components will not affect the normal reading of other components. The time spent on document query was tested in different path index states, as shown in Figure 6. Under the premise of the two variables of random document and fixed document, the time spent on document loading under different path index states is counted, as shown in Figure 7.

According to Figures 6 and 7, a comparative analysis of the time it takes to load and query documents in the XML database shows that when the path index is closed for document query, the time taken is shorter than when the path index is opened, and the query time difference between doc-

ument 1 and document 2 is greater than 50. The query time difference of document 3 is greater than 100. In addition, a simulation experiment for random documents and fixed-size documents found that the loading time of random documents with closed path indexing is longer than that of fixed-size documents; the loading time of fixed-size documents with open path indexing is more than that of random documents' time to load. In order to highlight the advantages of path indexing in XML databases, we conduct data research on the reconstruction of document indexes of XML, HTML, and SGML databases for bidirectional path indexing and unidirectional path indexing, as shown in Figure 8.

By reducing the data size, the storage space occupied by the data is reduced, and the data density in the cache is increased during the query process, thereby improving the storage performance and query performance of column storage data. The data in Figure 8 shows that whether it is a two-way path index or a one-way path index, the reconstruction of XML databases has obvious advantages over HTML and

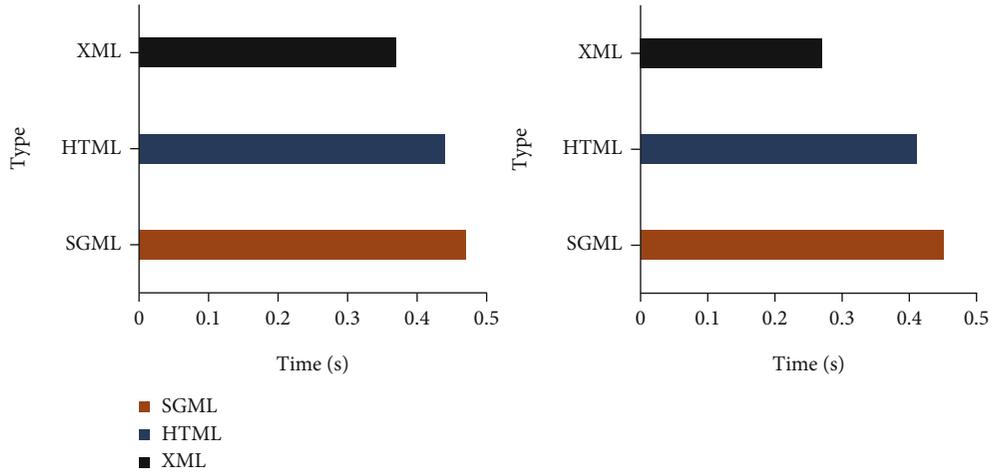


FIGURE 8: XML, HTML, and SGML two-way path index and one-way path index.

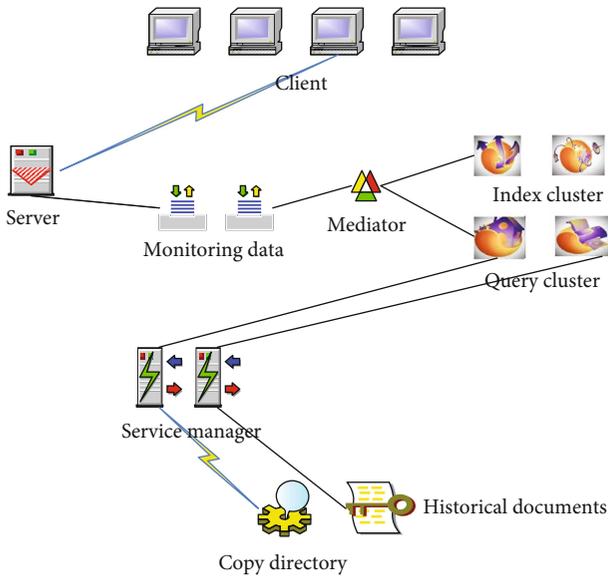


FIGURE 9: Information retrieval framework diagram.

SGML databases. The time consumed by the XML database in the case of bidirectional path indexing is 0.37 s, and the time consumed in the case of unidirectional path indexing is 0.27 s, which is much faster than HTML and SGML.

3.4. XML Database and User Analysis of Wushu Historical Archives. The distributed system architecture and storage index mechanism of the new digital library is a highly reliable, highly available, scalable, and high-performance underlying module in the million-volume digital library service platform, ensuring the efficient management of high-level resources and supporting multiple levels in the digital library. The rapid resource positioning of book reading, personalized recommendation, social network, and other services is an important part of the distributed service system of digital libraries (with information retrieval framework diagram; Figure 9). For the storage and management of resource files, traditional commercial

TABLE 2: Basic user situation table.

Questionnaire category	Group A (15 people)	Group B (15 people)
Average age	32.12	29.78
Average daily reading time of historical materials (month)	8.51 ± 1.12	4.79 ± 0.98
Average time since first contact with online historical archives	≤1 year	≤1/4 year
Average preference of XML database(1-10)	8.99	9.01

TABLE 3: Normal reading behavior of XML database users.

User behavior	Percentage calculated value
Page refresh or resize	77.134%
Read next page	17.058%
Browse other pages	5.808%

network-attached storage technologies are mostly used. The XML database studied in this experiment, which plays a huge role in data storage and management, is one of them. Based on the user’s reading log, this part studies the feasibility of constructing a new type of digital library system architecture and mass data storage organization mechanism based on XML database. Firstly, the reading behaviors of 30 users in groups A and B divided into males and females in a random questionnaire are counted in the online martial arts historical archives. The basic information and normal reading behaviors of the XML database are recorded in Tables 2 and 3, respectively.

The random survey users we selected all have a certain understanding of the XML database, and they are all enthusiasts of martial arts historical materials. They conduct investigations and analyses from different age levels, different educational backgrounds, and time of contact with online historical materials to get more information detailed and more comprehensive data. The data presented in Table 3 shows that in the normal reading behaviors or

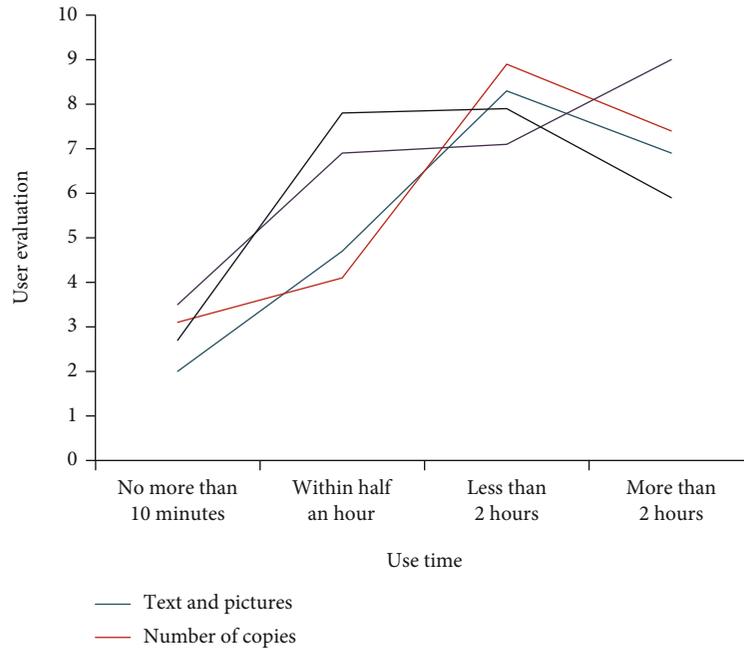


FIGURE 10: Analysis of the difference between user usage time and evaluation.

actions of the 30 users' online historical archives database, the frequency of page refreshing and page resizing is as high as 77.134%, and most people just focus on it. On the current page on-demand browsing instead of jumping reading, the reason for this experiment is to allow users to evaluate the four aspects of text and pictures, the number of copies, the design of the operation page, and the overall sense of use (1-10 points). According to the user's use time, user age, and user education classification, three data statistics are carried out. The first is to investigate the influence of the user's use time on different aspects of evaluation, as shown in Figure 10.

Relevant data shows that in the evaluation of XML database, within 10 minutes of users' browsing time, there is a positive growth trend between the ratings of text and pictures, the number of copies, the design of the operation page, and the overall sense of use and the user's use time; the browsing time is half within hours to two hours, the number of copies and the scores of text and pictures are increasing rapidly, while the design of the operation page and the overall sense of use are in a flat trend, and the user's attitude towards this change is relatively smooth; and when the browsing time exceeds two hours, the overall score is in a downward trend; especially, the overall user experience declines most obviously, and the user score of the operation page design is still on the rise. This survey has great pertinence and reference for the design of XML database. The following is the scoring statistics based on the user's age and the user's educational background, as shown in Figures 11 and 12.

Users of different age groups pay different attention to different aspects. Users under the age of 10 have the most stringent requirements for text and pictures and therefore give the lowest score of about 3 points in this survey; users

aged 11-18 think that the number of copies is too small to meet their needs; users between the ages of 19 and 25 gave the highest score to the number of copies, which was also the highest score in the audience. Users in the 26-35 age group show great interest in database page operation design. Among the 30 random users selected this time, users with a high school degree or above have a high evaluation score for data processing in the XML database, while those with a junior high school degree have a poor evaluation. In this regard, the difficulty of processing data in the XML database can be adjusted to achieve the purpose of being suitable for more people.

4. Discussion

With the rapid growth of big data and the widespread use of stored data in the field of big data, academia at home and abroad has carried out extensive research on several key technologies of storage databases and has achieved a series of results. In order to manage the large-scale data and obtain valuable information from it, the column storage database has shown a powerful advantage and has also attracted widespread attention from the academic and industrial circles. In the era of big data, the value range and scale of data columns are increasing rapidly, making the defects and deficiencies of these technologies increasingly obvious, and they need to be improved and optimized. In this experiment, through the above-mentioned research on data compression technology, it is found that data compression can reduce redundant data and improve storage space utilization under appropriate circumstances. When the actual data volume of the database is greater than the upper limit set by the user or the upper limit of the physical memory available in the system, the in-

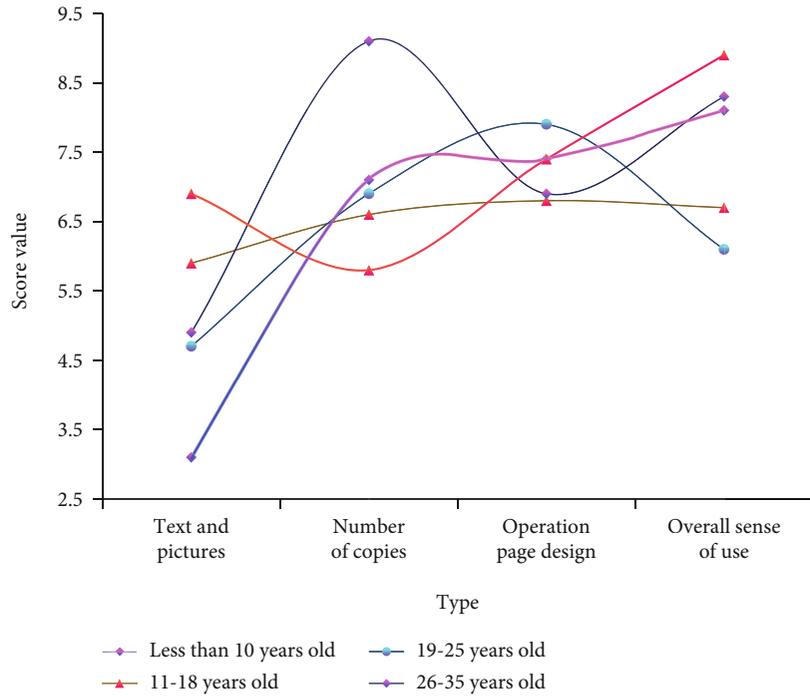


FIGURE 11: Analysis of the difference between user age and evaluation.

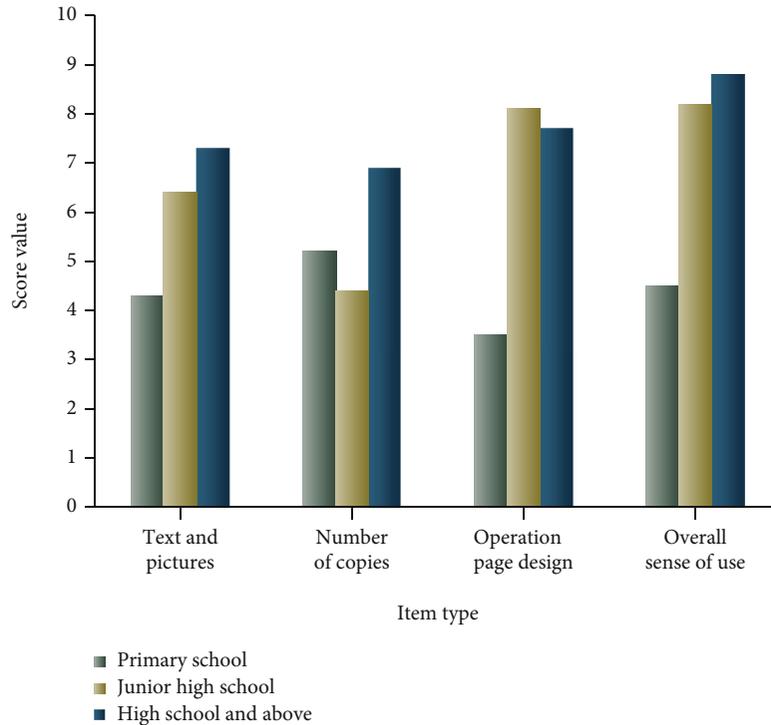


FIGURE 12: Analysis of differences between user's educational background and evaluation.

memory database will inevitably not be able to load all the data into the memory when it is started. The parallel independent selection strategy of XML database can significantly improve the accuracy and execution speed of the compression algorithm selection strategy under large-scale data sets.

5. Conclusion

Database servers often have completed functions, high stability and reliability, and can manage massive amounts of data. They are suitable for complex large-data application scenarios. Such applications often have a high degree of

concurrency, and application crashes will not occur, causing the database server to crash. On the contrary, some applications are relatively simple, with low concurrency, and the amount of data is not too large. It is often unnecessary to use those advanced functions in the database. The embedded mode database is more suitable for this type of application. In the era of continuous technological development, this storage system integrates popular technologies, draws on the representative research results of the latest columnar memory database and improves it, and realizes data storage, compression, fault-tolerant recovery, and query. Optimizing these aspects was continued to achieve a fast, low memory consumption data snapshot method suitable for query-intensive situations. XML database can provide protection at the hardware and system level and build a credible computing platform.

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

The author declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

- [1] Y. Hagiwara, "Wireless communication apparatus, wireless communication method, and storage medium," *Urbanism & Architecture*, vol. 40, no. 6, pp. 90–96, 2016.
- [2] M. Chi, J. Liu, and J. Yang, "ColdStore: a storage system for archival data," *Wireless Personal Communications*, vol. 111, no. 4, pp. 2325–2351, 2020.
- [3] W. Jun, Y. Qiong, W. Yue, and J. Yingying, "Distributed data storage solution under sink failures in wireless sensor networks," *The Journal of China Universities of Posts and Telecommunications*, vol. 24, no. 2, pp. 72–102, 2017.
- [4] H. Ding, C. Zhang, Y. Cai, and Y. Fang, "Smart cities on wheels: a newly emerging vehicular cognitive capability harvesting network for data transportation," *IEEE Wireless Communications*, vol. 25, no. 2, pp. 160–169, 2018.
- [5] W. Liu, T. Deng, Yang et al., "Towards robust surface skeleton extraction and its applications in 3D wireless sensor networks," *IEEE/ACM Transactions on Networking*, vol. 24, no. 6, pp. 3300–3313, 2016.
- [6] I. H. Jawhar, N. Mohamed, Z. Trabelsi, and J. al-Jaroodi, "Architectures and strategies for efficient communication in wireless sensor networks using unmanned aerial vehicles," *Unmanned Systems*, vol. 4, no. 4, pp. 289–305, 2016.
- [7] P. Gafurova, A. Elizarov, and E. K. Lipachev, "Basic services of factory metadata digital mathematical library Lobachevskii-Dml," *Russian Digital Libraries Journal*, vol. 23, no. 3, pp. 336–381, 2020.
- [8] T. Sui, D. Marelli, X. Sun, and M. Fu, "Multi-sensor state estimation over lossy channels using coded measurements," *Automatica*, vol. 111, 2020.
- [9] T. M. Adams, "Coding with XML for efficiencies in cataloging and metadata," *Journal of Electronic Resources Librarianship*, vol. 31, no. 1, pp. 49–50, 2019.
- [10] P. R. Olmedo, "Technical-legal management standards for digital legislative information services," *Revista Chilena de Derecho y Tecnología*, vol. 6, no. 2, pp. 57–95, 2017.
- [11] C. Guccio, M. Martorana, I. Mazza, G. Pignataro, and I. Rizzo, "An assessment of the performance of Italian public historical archives: preservation vs utilisation," *Journal of Policy Modeling*, vol. 42, no. 6, pp. 1270–1286, 2020.
- [12] R. Michalak, "Building community, fostering collaboration, and engaging bridge program students with a college's historical archives," *Journal of Western Archives*, vol. 10, no. 2, pp. 4–4, 2019.
- [13] J. Silovsky, J. Nouza, and M. Kucharova, "Search for speaker identity in historical oral archives," *Multimedia Tools & Applications*, vol. 75, no. 7, pp. 3767–3786, 2016.
- [14] B. Li, G. Xiao, R. Lu, R. Deng, and H. Bao, "On feasibility and limitations of detecting false data injection attacks on power grid state estimation using d-facts devices," *IEEE Transactions on Industrial Informatics*, vol. 16, 2020.
- [15] H. Yang, Z. Yi, X. A. Wang, Y. Su, Z. Tu, and X. Yang, "Improved lightweight cloud storage auditing protocol for shared medical data," *Wireless Communications and Mobile Computing*, vol. 2021, Article ID 8886763, 13 pages, 2021.
- [16] J. Zhong and X. Xiong, "Data security storage method for power distribution internet of things in cyber-physical energy systems," *Wireless Communications and Mobile Computing*, vol. 2021, Article ID 6694729, 15 pages, 2021.
- [17] Y. He, L. Dai, and H. Zhang, "Multi-branch deep residual learning for clustering and beamforming in user-centric network," *IEEE Communications Letters*, vol. 24, 2020.
- [18] N. T. Rao, D. Bhattacharyya, and H. J. Kim, "Analyzing the increase of network life time for wireless sensor network with smart local moving algorithm using NS2 simulator," *International Journal of Future Generation Communication and Networking*, vol. 10, no. 9, pp. 75–84, 2017.
- [19] A. M. Rajabalinej, "A framework for requirement elicitation, analysis, documentation and prioritisation under uncertainty," *Chinese Journal of Biotechnology*, vol. 9, no. 4, pp. 247–255, 2017.
- [20] Z. Xu and Z. Cai, "Privacy-preserved data sharing towards multiple parties in industrial IoTs," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 5, pp. 968–979, 2020.
- [21] F. Wang, W. Zhang, X. Nian et al., "Refining the late-Holocene coastline and delta development of the northern Yangtze River delta: combining historical archives and OSL dating," *The Holocene*, vol. 29, no. 9, pp. 1439–1449, 2019.
- [22] A. Valls, F. Garcia, M. Ramirez, and J. Benlloch, "A combined use of GPR data with historical archives for identifying pavement construction periods of Valencian Silos (16th Century, Spain)," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 9, no. 1, pp. 98–107, 2016.
- [23] N. T. T. Linh, Y. Kumeda, M. Matsushita, T. Hironiwa, and M. Furuta, "Application of gamma radiation for disinfection of fungi in a large volume of historical archives damaged by flood following typhoon Hagibis 2019, Japan: a case report," *Biocontrol Science*, vol. 26, no. 1, pp. 55–59, 2021.
- [24] S. Lenninger, "Culture in the layers of contemporary discourses and historical archives: a review of Anna Maria Lorusso's Cultural Semiotics," *Public Journal of Semiotics*, vol. 8, no. 1, pp. 67–77, 2017.

- [25] K. Pačliková, A. Weinfurter, M. Vohnoutová et al., “Geoinformatics and crowdsourcing in cultural heritage: a tool for managing historical archives,” *Agris On-line Papers in Economics and Informatics*, vol. 10, no. 2, pp. 73–83, 2018.
- [26] K. Mackinnon, “Ethical approaches to youth data in historical web archives (dispatch),” *Studies in Social Justice*, vol. 15, no. 3, pp. 442–449, 2021.
- [27] S. Pejović and J. Katelan, “Historical archives and born digital records,” *Atlanti*, vol. 27, no. 1, pp. 63–78, 2017.
- [28] D. Schlenker, “The historical archives of the European Union in Florence: research in a multilingual and transnational archives,” *Atlanti*, vol. 26, no. 2, pp. 59–64, 2016.
- [29] I. V. Tunkina, “Academy archives as repositories of Russia’s historical memory,” *Herald of the Russian Academy of Sciences*, vol. 89, no. 6, pp. 550–557, 2019.