Standardized Storage of Sports Data Based on XML

Qiufen Wang[1,*] and Yongxin Wang[2]

[1]*Nanyang Institute of Technology/College of Computer and Information Engineering, Nanyang, China;* [2]*Nanyang Institute of Technology/Department of Physical Education, Nanyang, China*

Abstract: Sports data feature strong timeliness, multiple types, many specifications, large data amount and very complex storage. The information technology has been more widely applied in sports data storage, and also achieved a wealth of research findings. In terms of sports data, the paper proposes a standardized data storage solution based on XML, makes a brief description of the methods and principles of various data concerning sports training programs, and then puts forward the logical structure of data storage. The XML data file system and database system grounded on XML storage sports industry is simply discussed and a high efficient, standardized storage solution is raised for sports data.

1. INTRODUCTION

With the rapid development and application of computer technology and network technology, it has penetrated into all aspects of social life and also has a wide range of applications in the field of sports [1-4]. In regard to sports data, computer technology and network technology have achieved full range of penetration and intervention [5-7]. On the one hand, it ensures that the convenience and security of data storage; on the one hand, it also builds a scientific and reasonable information system for the development of sports. For example, understand the development trend of sports can thus promote sports training to be scientific and systematic. As information technology develops and international exchanges enhance, information technology has been gradually increasingly applied in the sports industry, and has made outstanding achievements.

In terms of sports data storage, in the past, the traditional paper media were used to store data and manage traditional sports industry records. However, traditional paper media featured very limited amount of memory and very complicated inquiry. After introducing IT to the sports industry, all kinds of sports industry data have begun to be stored on the hard disk, tape drives, CD-ROM and other equipment in digital form, such that the sports data storage capacity can increase dramatically, and the retrieval and accessibility of data can become very easy [8-11]. Visibly, the use of information technology can enormously improve the storage capacity of sports data, while facilitating the development and utilization of data, thereby organizing the data together quickly and shapely and making contributions to the development of the sports industry.

2. SPORTS INDUSTRY DATA AND INFORMATION TECHNOLOGY

2.1. An Overview of Sports Industry Data

The sports industry is a very complex industry and can be divided into many branches. Each of the branches contains a lot of complex data. How to scientifically and efficiently acquire, store, classify, retrieve and update these data is the key to ensuring the healthy development of the sports industry. For sports training, as an example, there are mainly the following types of sports data: (1) sports training content, training performance, the completion of technical movements, training videos and other data. These data are different in formats and there are many data types, and the total amount of data is very large. (2) Individual physiological and biochemical indexes of athletes. These are data regarding athletes' body function conditions available from various instruments in sports training. (3) The basic personal data of athletes, such as date of birth, gender, place of origin, technical features and sports injuries. There is very slight variation of these data, so they are relatively stable. (4) Data on sports training programs. Below these types of data are illustrated to study the standardized storage methods of these non-normalized data by utilizing XML storage technology.

2.2. XML Technical Overview

The most commonly seen storage technology in the existing research in the field of data storage is XML based on relational databases. From the perspective of data management, the XML data can be deemed as hierarchical structure or semi-structured data.

Since XML data are a meta-markup language, it is characterized by extremely strong flexibility, scalability and openness. Users can mark flexibly according to the needs and clearly describes the defined marks, so that these marks can be used to store all kinds of data and data of different specifications can be saved together. The sports industry data are diverse and of many specifications, and there are a large

number of advantages of XML storage. For example, first understand the characteristics of various sports training data, and then define a common mark collection, on which coaches can mark individually according to their own definition of training, thus facilitating coaches to use data. In addition, due to the timeliness characteristic of sports data, an appropriate XML document can be devised to resolve the timeliness problem, better store sports data and promote the sustainable development of the sports industry while employing the XML storage technology.

2.3. An Overview of Sports Industry Data and Information Technology Research

With reference to the sports training data, since sports training is a very complex system, there are enormous complex data. The key foundation of scientifically conducting sports training is well acquiring, storing, classifying, retrieving and updating data. In recent years, significant attention has been paid to researching mechatronics, multimodal user interface, wired and wireless data output, multimedia databases, data warehouses, data mining technology and training data management, among many others. Further, practical results have been achieved in this respect.

When studying weight training data management and evaluation systems, Zhong Yaping, *et al.* proposed a framework structure and mode of entry for training data management domain measurement system based on multi-modal user interface under different application environments, because training data acquisition features diversity and space-time uncertainty, while the training data management and evaluation system based on a single graphical interface has failed to meet the needs of management personnel and is impossible for high efficient data training data management and evaluation.

Due to temporal characteristic in sports training data, a lot of data are obtained in sports training scenarios and target tracking. In that way, how can we use computer technology or these data? Xiong Wenping presented a hybrid tracking method, while Shen Yuejun, *et al.* put forward a real-time tracking method for athletes during training of ball games, and suggested the feature selection, state transition and implementation methods for target tracking in ball games. The computer technology is utilized to track every athlete in the field, obtain their space-time information, and automatically calculate various corresponding types of motion parameters, such as instantaneous velocity and trajectory, so as to provide valuable data for the late management as well as promoting the development of sports.

In sports target tracking process, the position information of the target can be acquired by matching the image, where the correlated matching algorithm is the most classic. Below is the schematic diagram for the correlated matching algorithm. Now suppose the template is denoted with M (U, V); in the current frame image, the region to be searched is denoted with S (X, Y); the current frame image is I (M, N), then

$$R\left(x, y\right) = \sum_{M=0}^{U} \sum_{V=0}^{V} \left| M\left(u, v\right) - S(s + u, y + v) \right| \tag{1}$$

where $0 \le x \le X-U+1$ and $0 \le y \le Y-V+1$. The template searches correlation coefficients of the pixels one by one in the search region. The location where the correlation coefficient is the minimum value is exactly the search area that the target is located in. Beyond that, there is still normalized cross correlation evaluation function, mean square error evaluation function and other commonly used criteria. The normalized cross correlation evaluation function is:

3. XML STANDARDIZED STORAGE ANALYSIS OF SPORTS DATA

3.1. Develop XML File Specifications, Describe Different Types of Data

First, the sports training projects incorporates both individual projects and collective projects. There are relatively less types of data in individual projects, making them relatively simple, whereas the data in collective projects are very complex. For example, in volleyball games, the data of every athlete's various performances should be recorded clearly, as well as the performance of the whole team, including the time and place for sports training, individual performance and position of the team, the overall performance and the coordination condition of the whole team. From the perspective of confrontation angle of sports training programs, in regard to extremely antagonistic sports like table tennis and fencing, the recorded data should also highlight the characteristics of this sport, together with the opponent's technique and tactical characteristics, *etc*.

Then it is the description of different sources of data. We can divide the abovementioned first data description into several parts: the time and place of sports, sports content, sports performance and units, technical movements and evaluation, tactics and evaluation. The second athlete physiological and biochemical indicators are generally the body state data obtained from the detection equipment after a certain kind of sport, and the data play a great role to the late sports data analysis. The data have strong timeliness but are very fixed, and therefore we can describe such data in data acquisition time and object, the type of data indicators, specific data values and units. The other two types of data are mostly stable data with little changes, so it is even simpler to describe. It is fine with a simple description directly based on the content.

For example, here is a simple XML file concerning the specific time and performance and more of a particular sports item of an athlete. It is very simple and demonstrates the advantages of XML in sports data storage.

```
<List of training projects>
< Training projects >
< Player > Li </ Player >
< Time >2014-12-10 10:30:00</ Time >
< Training content >
< Item >50M</ Item >
<Performance>6~20</ Performance >
< Unit > second </ Unit >
```
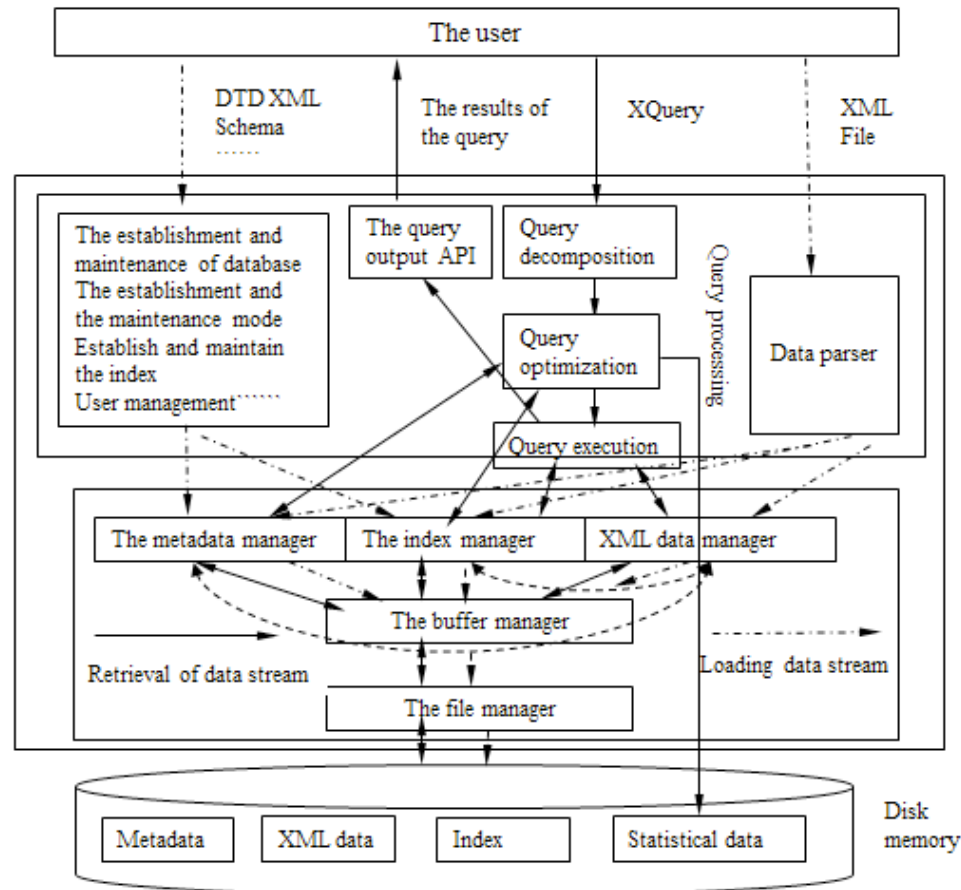
Fig. (1). Pure XML database.

3.2. Logical Structure of Data Storage

Sports data storage is not a person's data, but those from multiple coaches and athletes. The collection of these data will be very complicated. If a unified specification or data format is lacking to organize and manage data, these data will become so complicated that cannot achieve effective storage and management. Hence, we have to construct a logical structure that can store data, device a reasonable form of data that helps determine what data in which XML files and also what kind of correlation among different XML files. In the construction of the logical structure, we not only need to consider the storage and retrieval efficiency of data, but also take into account the modifications and updates of the data. Given the mobility of athletes, the data need to change with the changes of athletes, so as to ensure that the data of each athlete and coach are relatively independent.

For example, (1) create a separate XML file for the subordinate relationship between the coach and the athlete; (2) create a separate XML file for the basic data of the coach; (3) create a separate XML file for the basic data of the athlete; (4) create a XML file for the training results of a {certain} athlete; (5) create a XML file for the physiological and biochemical indicators of a {certain} athlete; (6) create a XML file for the training activates of a {certain} athlete; (7) create a XML file for the training plans of a {certain} athlete; (8) create a XML file for the training data of the entire team. In this logic structure, when each XML file is related

to the description of specific sports data, the description methods of sports data in Section 3.1 can be adopted to maximize the openness, scalability, flexibility characteristics of XML files, ensure that the logical and rational storage of data and facilitate administrators to use them. In addition, if the athlete's data are directly assigned to be self-managed by their coaches, then the above XML files from (3) to (8) can be integrated. The above content portrays the logical structure of the sports data storage. It is clear that the logical structure of the sports industry data storage based on XML is relatively simple, as long as various types of data generated are described and classified correctly and a logical structure is built according to the characteristics of sports items.

3.3. XML Database

The physical structure and the management mode should be determined as per the amount of data. For instance, a large sports team would produce many sports data and involve a large number of management staff. Then a XML database can be created to ensure data storage and retrieval, to guarantee real-time dynamic updates and management of data, to make full use of sports data for scientific training of sports items, to allow the administrator to master the dynamics of various coaches and athletes, so as to devise personalized, targeted approaches and measures and enhance the efficiency and quality of sports. Here, XML databases have expanded XML supportive module and completed the database
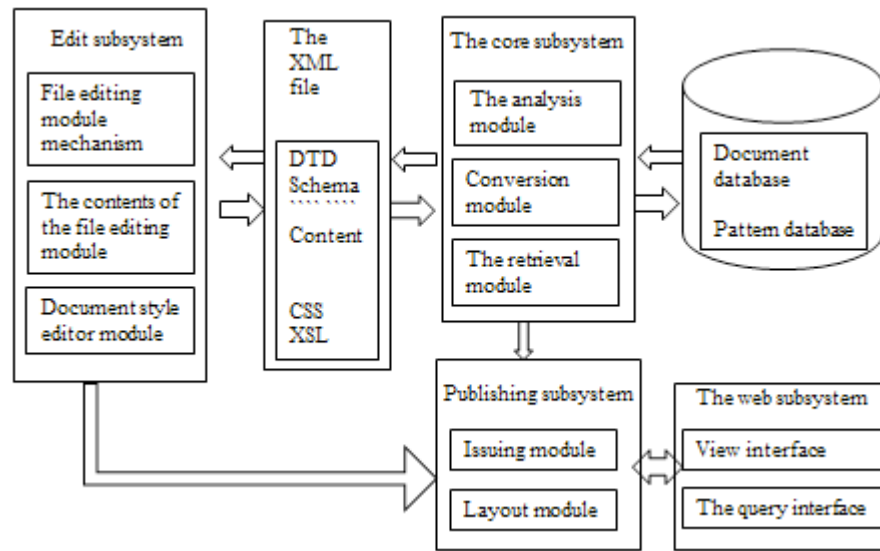
Fig. (2). Structure of an integrated XML document management system.

that satisfies format conversion and transfer between XML data type and database type, on the basis of relational databases and object-oriented databases. The pure XML database that stores and manages sports data is constructed as shown in the Fig. (**1**). This database is the same as with traditional relational databases and can be classified into four modules, which are user interface, query processing, storage management and physical storage.

3.4. XML File System

There is no need to create a database with only sports data of a coach or several coaches and their athletes. In this case, some simple XML file systems are enough to guarantee the storage, retrieval and use needs of the sports data. The Fig. (**2**) shows the structure of an integrated XML document management system.

XML file system is the simplest way for storage. It directly stores XML as text files for data storage, organization and management by way of documents. It is very easy to implement, with no need of databases and object storage management. Further, if storing the entire XML documents in a file, then these data can be accessed directly in multiple text editors and XML tools, without having to convert the stored data again or reconstruct the query results.

However, when there are so much data in the XML file system that reach a critical point, there will be some confusion, because the XML file system does not have a good memory management mechanism. So, when there is a small amount of uncomplicated data, the XML file system can be applied.

CONCLUSION

With the rapid development of the network information age, computer technology has penetrated deeply into the sports field. Extensive and profound investigations have been done concerning data acquisition, storage, management systems, technique analysis and simulation, and training ac-

cessory systems of sports, which end up a wealth of valuable research findings. However, there have not been many outcomes that can be applied to specific sports practices. In term of data storage, more research should be done to strengthen the accuracy and integration of data and develop more data storage technologies and products suitable for different scenarios, thus making data storage and management easier, time-sensitive and highly efficient, realizing the standardized storage of sports industry data, and making them better assist in sports training. This paper presents a standardized storage technology of XML-based sports data, introduces the description methods and logical structure of sports data, simply points out the types of XML file structures applicable for sports data storage under different scenarios, and makes sure to meet the basic needs of data storage and data retrieval.

CONFLICT OF INTEREST

The authors confirm that this article content has no conflict of interest.

ACKNOWLEDGEMENTS

REFERENCES

[1]   Y. Zhong, "Application and prospects of information technology in sports training," *Journal of Wuhan Institute of Physical Education*, vol. 42, no. 6, pp. 52-55, 2008.

[2]   J. Lin, "Application and research of IT at home and abroad in modern sports," *China Science & Technology Panorama Magazine*, no. 23, pp. 37-38, 2010.

[3]   X. Lan, and J. Ren, "Research on XML storage technology of Oracle XML," *Journal of Geodesy and Geodynamics*, vol. 29, no. 6, pp. 139-143,156, 2009.

[4]   C. Feng, "Research of XML mapping mechanism based on relation database system," *Coal Technology*, vol. 29, no. 12, pp. 136-137, 2010.

[5]   F. D. Davis, P. B. Richard, and R. W. Paul, "User acceptance of computer technology: a comparison of two theoretical models," *Management Science,* vol. 35, no. 8, pp. 982-1003, 1989.

[6]   M. Igbaria, I. Juhani, and M. Hazem, "Why do individuals use computer technology? A Finnish case study," *Information & Management,* vol. 29, no. 5, pp. 227-238, 1995.

[7]   K. A. Frank, Z. Yong, and B. Kathryn, "Social capital and the diffusion of innovations within organizations: The case of computer technology in schools," *Sociology of Education,* vol. 77, no. 2, pp. 148-171, 2004.

[8]   D.G. Liebermann, L. Katz, M.D. Hughes, R.M. Bartlett, J. McClements, and I.M. Franks, "Advances in the application of information technology to sport performance," *Journal of Sports Sciences,* vol. 20, no. 10, pp. 755-769, 2002.

[9]   Y. Yoo, H. Ola, and L. Kalle, "Research commentary-The new organizing logic of digital innovation: An agenda for information systems research," *Information Systems Research*, vol. 21, no. 4, pp. 724-735, 2010.

[10]   S.P. Deshpande, and V.M. Thakare, "Data mining system and applications: A review." *International Journal of Distributed and Parallel Systems (IJDPS)*, vol. 1, no. 1, pp. 32-44, 2010.

[11]   J. D. James, and D. R. Stephen, "Comparing sport consumer motivations across multiple sports," *Sport Marketing Quarterly*, vol. 13, no. 1, pp. 17-25, 2004.