Application of Data Mining Pipeline in Analysis of Technique and Tactics of Volleyball Matches

Kunrong Zhu[1], Baokui Wang[2]
[1]Sport Department, Jiangxi Agricultural University, Nanchang, 330045, China
[2]Physical Culture Institute, Jinggangshan University, Ji'an, 343009, China

Abstract: With the rapid development of computer technology and network technology, people ushered in a new era of network with information and knowledge as the main features. People's exposure to digital information has grown exponentially. The continuous and strong growth in the amount of information has brought many challenges to the people. Sports competition has now become an important symbol of the progress of national and national civilization. As the most powerful tool for modern human development, computer technology will inevitably play an important role in the development of sports competition. Therefore, this paper studies the application of data mining algorithm in the analysis of the technique and tactics of volleyball game. We integrate the data mining and the pattern finding models to solve the data analytic tasks. The results are validated through experiment.

Keywords: Data mining algorithm; Association rules; Data analysis; Volleyball match; Tactical analysis

1. Introduction

At present, the analysis of sports data information usually adopts conventional statistical methods: sampling theory, hypothesis testing, decision theory, estimation theory, time series, etc. Most of statistical analysis techniques require perfect mathematical theory and rigorous use skills as high, and most of them belong to sampling studies, and there are different levels of information loss. In addition, it is difficult to discover the underlying laws of data by traditional statistical methods. With the increase of data in the sports field in recent years, physical fitness monitoring data for various indicators, various kinds of

sports competition data, and so on, more and more, the formation of the massive data. At the same time, computer application technology has obvious potential for development in storage and simulation analysis of data. Therefore, how to find the interrelationships among various factors in the massive data and find out the changing rules between them, and further in-depth analysis of the data has become one of the important directions of computer development. Therefore, the research and application of data mining technology has become an indispensable part of sports science research.
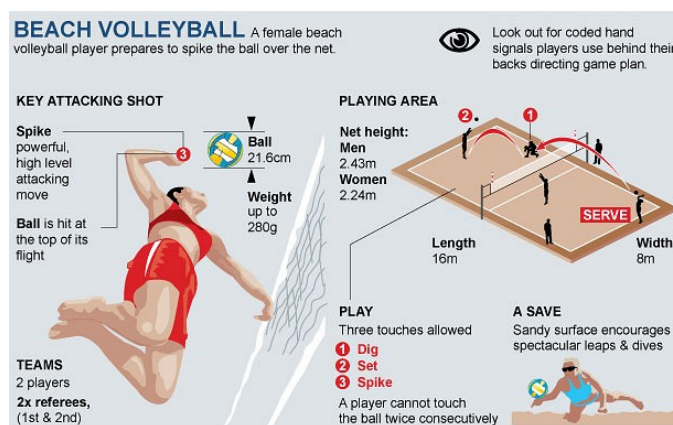

Figure 1. The volleyball principles

In the field of sports, there will be a great deal of data and information in athletics training, actual competitions, school sports management, and national fitness testing. How to use these data effectively and give full play to its potential value is exactly the problem that the scientific and technical personnel in sports and related fields urgently need to solve with a cluster-based data transmission method to ensure the security of data transmission through the identity digital signature scheme and to speed up the data transmission speed by reducing the computational complexity and security verification time. Therefore, the principles can be summarized as the following aspects.

For the volleyball players to wear the inertial sensor, the initializing calibration while through the sensor to the volleyball throwing ball and the swing stroke movement data collection. The basic idea is as follows, on the basis of each inertial sensor for volleyball players in the volleyball and stroke gesture joint profile data, the real-time coordinate conversion processing is completed all the joints in the human body model for the volleyball ball and stroke action corresponding tracking, obtain the corresponding number.

The attitude data of the inertial sensor worn on the body of the volleyball player is the coordinate of the inertial sensor itself, so when collecting the data of the volleyball throwing and the swing of the ball, and it is necessary to convert the action attitude data obtained above to the skeleton mark system in the human skeleton model.

Volleyball is a collective project, the coaches also have reliable and effective information about the team and the opponents, whether according to the information to make effective and targeted training plan that can to master in the competition of the initiative and win the game at the end of the play a vital role. The game statistics and processing of the data is scientific training, the most direct information, information can be divided into three types, namely, the use of information before the game, game information and game information. Before the game, trainers and policy makers could be based on statistical data to understand the players and the team's basic information about competitors, individually and collectively, basic conditions for development of targeted training programs to develop the team's technical and tactical play, do know thy enemy and know thyself, and targeted. In the game, the trainers can be timely feedback game statistics with the command and do contingency. After the game, coaches, athletes, researchers and policy makers, etc. can be used in combination with the statistical data for after the game.

2. The Data Assisted Volleyball Match Research Pattern

2.1. Data assisted volleyball match research pattern

2.1.1. Components and the details

Establishing a 3D solid model for athletes. The athlete's modeling uses Pro E modeling software to model each part of the body. Then the data of the Pro E model is converted into the data needed in the program to complete the athlete's human body modeling.

Coach interface is divided into two parts: volleyball movement dialog and player activity dialog. The player activity dialog box is used to enter the movement of each player in an attack. Here, we can choose whether to enter the six players, whether it is one or two, whether it is an attack or a cover. In the lower dialog box, we can further detail the parameters of the player's activities.

The motion process is calculated according to the given parameters. This is the core of the whole program design. First, the basic posture of human is established, and then the motion process is combined by using basic attitude. People's basic attitudes include standing upright, squatting, walking posture, jumping posture and other special features in volleyball. The combined actions include some necessary actions such as moving, jumping, one pass, second pass, with the attack and so on.

In volleyball competitions, the results of the analysis of on-the-spot technical and tactical data are the important basis for strategic deployment of coaches. Therefore, the collection, statistics, and analysis of on-the-spot technical and tactical data have also become a key factor affecting coaches' decision-making. However, it is difficult to rely on manual work to complete statistical work on data. Due to the speed constraints, it is impossible to record large amounts of comprehensive and accurate data, and it is even more impossible to perform scientific systems on the collected data within a limited time. The analysis cannot meet the data requirements for scientific guidance. Coaches can only rely on experience and less data to guide the use of computers for statistical analysis is a way to solve this problem. However, there are also several problems:

Data collection. Since volleyball is a turn to change the score, in the course of each round, there are two teams of 12 athletes on the field. The volleyball exercise requires high speed, which also makes the recorder must be in a very short time. Recording the situation on the field, including the number of players, the ball's landing area, and a series of data on tactics, etc., requires high real-time performance.

Analysis of data. Since there are many tactical changes and accidents in volleyball, it is necessary to prevent the possibility of making erroneous analysis results based on a small amount of data. There is a high requirement for the ability of the system to analyze data. It is more and more necessary to apply the data mining algorithm to the technical and tactical analysis of volleyball competitions.
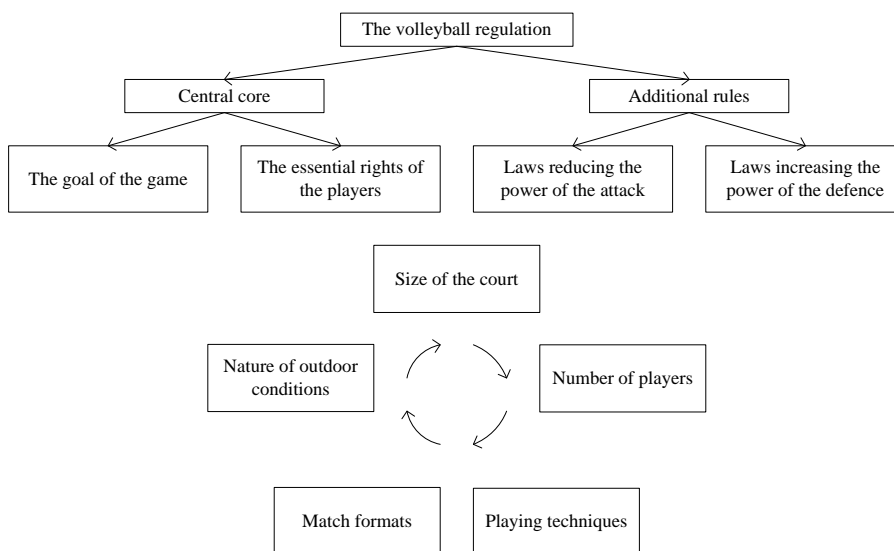
Figure 2. The volleyball regulations

3. Application Analysis of Data Mining Algorithm in Analysis of Techniques and Tactics of Volleyball Matches

3.1. Data mining algorithm

Data mining emerged in the late 1980s and is a new field of application value in database research. It is a cross-disciplinary discipline that integrates artificial intelligence, database technology, pattern recognition, machine learning, statistics, and data visualization. And other fields of theory and technology. At the KDD96 international conference, according to the elucidation of well-known scholars in the field of knowledge discovery, a uniformly accepted definition of knowledge discovery (KDD) is expressed as: "The implications for the database are unknown, potentially valuable, non-trivial patterns. extract". It is generally considered that data mining is a major process of knowledge discovery and does not make a strict distinction between the two. According to the definition proposed by frawley and shapir et al., data mining (Data Mining) refers to the extraction of hidden data from a large number of incomplete, noisy, fuzzy, and random data that people do not know in advance. But is potentially a useful process of information and knowledge. Its processing of data is not limited to queries and access, but it can also find potential connections between data. This definition includes several layers of meaning: Data is the "source" of data mining. They come from databases, data warehouses, certain non-database systems, and network data. Data sources must be real and numerous; the result of data mining is the user experience. Interested, conceptual knowledge: Discovered knowledge

needs to be acceptable, understandable, applicable, and support specific problem discovery.

$$\mu A : U \in [0,1], u \in \mu A \quad (1)$$

$$\begin{bmatrix} C1n \\ C2n \\ ... \\ Cnn \end{bmatrix} = \begin{bmatrix} Q11, Q12,...,Q1n \\ Q21, Q22,...,Q2n \\ ... \\ Qn1, Qn2,...,Qnn \end{bmatrix} \begin{bmatrix} C1n \\ C2n \\ ... \\ Cnn \end{bmatrix} \quad (2)$$

The system reliability difference calculation method is: After the system reliability is calculated, a small increment is added to each item in the state transition probability matrix, and then the system is recalculated by using the state transition probability matrix added to the increment. The difference between the reliability and the front and rear system reliability is the system reliability difference. The larger the difference, the greater the influence of the state transition process corresponding to the value that is currently changed in the state transition probability on the system reliability.

3.2. Application of data mining in sports field

Association rules are rules that describe different associations between different data items in a database. The discovery process is generally divided into two steps: frequent item set acquisition and rule generation. Association rule mining is to find the hidden association rules among the itemsets in the database through association analysis, and combine the existing information to speculate on the unknown issues.

Rough set theory can be used to find inaccurate data or structural relationships within noise data. It was proposed by Polish scholar Z. Pawlak et al. in the 1970s. The advantage of this method is that it does not need to provide any information other than the analysis data. This method

is widely used for the classification and acquisition of uncertain or incomplete information, and the description of the analyzed problems is also more objective.

In competitive sports, the technical and tactical analysis of various items will generate a large amount of data. The traditional methods of competition analysis are generally based on artificial observation of competitions, mathematical statistics combined with the experience of researchers, and it is difficult to draw conclusions on the science. Therefore, it has certain limitations. With the development of computer technology, the reduction of manual input and the refinement of technical and tactical analysis have driven the development of technical and tactical collection and analysis systems, such as table tennis, tennis, football and fencing, jointly developed by Tongji University and Shanghai Institute of Physical Education. The technical and tactical collection and analysis system has helped coaches accumulate a large amount of technical and tactical data for further analysis of technical and tactical mining. At present, the application of data mining in technical and tactical diagnostic analysis mainly focuses on the analysis of the athlete's own technical and tactical features, analysis of the advantages and disadvantages of different opponents, the analysis of the technical and tactical characteristics of the opponents, and the maximum win-loss technique for consecutive shots the tactical sequence, the influence of the various technical and tactical indicators on the winning weight of the game, the mathematical analog diagnosis of ball games, etc.

3.3. Application of data mining in volleyball competition

In order to collect the data needed to implement the data mining algorithm, it is necessary to record the execution process of each technical action in the volleyball game. Due to the fact that the tactical movements of the athletes in the game change more and the various actions are completed in an instant, the process of recording the game is challenging. Mentioned earlier in the software Data Volleyball designed by the Italian Data Project company, in order to solve this problem, a process-based script description language was designed. The language uses the mnemonic method to encode basic technical moves in volleyball matches. However, this script description language design method requires the recorder to record a series of information including the player's number, technical action, type of technology, starting area, and ending area in a situation where the situation on the game site changes rapidly. This makes the recorder's workload a lot.

In order to collect game data more quickly and accurately, it is necessary to design a pattern-based script description language. The so-called pattern-based script description language means that: Since the tactics used by both sides of the game in high-level volleyball matches have a certain number of patterns, these patterns can be used as some fixed formulas in the script description language design process. Only a small amount of information on this mode is required during the user's recording of the game, and other information can be inferred based on the tactical rules of the volleyball game.

In the aspect of acquisition mode, the main advantage of the keyboard acquisition mode is that the acquisition speed is faster. The disadvantage is that the accuracy is slightly lower than the mouse acquisition. The mouse acquisition has advantages in the accuracy of acquisition, but because of the limited acquisition interface, it cannot cover all the scripts, so it has certain limitations. In this scenario, the combination of mouse acquisition and keyboard acquisition is used to improve the collection efficiency.

Since the pattern-based script description language is designed for technical and tactical information, it is necessary to classify and integrate the collected data and integrate the data existing in different data tables into a new, small-scale data warehouse. The separation of data warehouse and operational database is due to the different structure, content and usage of data in these two systems. Decision support requires historical data, while operational databases generally do not maintain historical data. In this case, although the data in the operation database is very rich, it is often not enough for decision-making. Different data types, such as offensive and defensive processes, will generate different data warehouses.

The volleyball technical and tactical analysis system realizes an on-the-spot technical and tactical statistical analysis system for volleyball matches that integratethe data acquisition, data statistics, data analysis, data management, and video management functions. Through this software, coaches can use the computer to record and analyze the purpose of the game data to provide technical support for coaches' strategy formulation. Its main contents include: the application of data collection and association rules analysis in volleyball competitions, and the application of data mining algorithms based on Markov process in the analysis of tactics and tactics of activities.The application of the "dual-division" training method in the training of lesser men's volleyball requires that the training concept of "teaching according to individual aptitude" should be established first, and training activities should be carried out in accordance with the actual abilities of the athletes. The training of sportsmanship should be emphasized and the volleyball skills of the younger men's volleyball players should be promoted considered overall improvement of mental quality. Before training, first of all, through certain tests, we can understand the current actual ability of each athlete, must not be arbitrarily divided, should pay attention to the es-

tablishment of "in accordance with their aptitude" training concept, the volleyball skill level uneven athletes, according to actual ability Uniformly, establish different training goals, training methods and training require-

ments. In addition, this "grouping" and "stratification" are not immutable, and they need to be adjusted and changed in accordance with the actual performance and learning of the younger men's volleyball players.
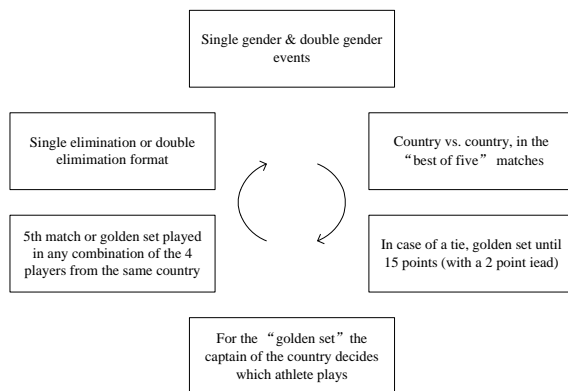

Figure 3. The training tips for reference

Therefore, for the training, the training patterns are summarized into some aspects. In the process of volleyball training for younger men, the coach can establish different teaching goals and develop different teaching contents according to the actual abilities of the athletes. For example, in the process of instructing students on the A-level to perform volleyball training, the A-level athletes may be required to be proficient in learning various volleyball skills while maintaining the stability of the skills and the coordination of the movements. While continuously improving the overall strength of the A-layer athletes, more potential athletes will be tapped.In guiding B layer of volleyball skills training in the process of the athletes, coaches can ask the B level athletes focus to strengthen the mastery level of volleyball skills and flexible application level, to carry out a lot of the training activities to improve B layer of the comprehensive ability of athletes. In the course of coaching c-layer athletes in the training of volleyball skills, the coach can combine the game activities with the volleyball training by carrying out some interesting and the competitive game activities.

A technical and tactical algorithm development kit was introduced into the technical and tactical analysis system, and a comprehensive analysis module was added to the original system. Dig up the collected game data and extract useful information to guide the game according to the laboratory's existing table tennis technical and tactical systems and volleyball tactics analysis system. In the table tennis tactics analysis system, mining algorithms such as association rule analysis, cluster analysis and data mining method based on Markov process are mainly used. In the volleyball analysis system, mining algorithms such as association rules analysis and data mining methods based on Markov process are used.

Data mining is an information technology with rich practicality and powerful vitality. Its application in the sports field has a certain extent, but it also shows some problems in practical applications. The R&D team in the sports field is usually composed of sports science and technology personnel who master the characteristics of sports items and technical and tactical characteristics, and information technology personnel who master data mining technologies. In the process of completing research projects, it takes a lot of time and effort to communicate and coordinate. There is a shortage of composite sports science and technology personnel who are familiar with sports project knowledge and are familiar with data mining technology.

4. Conclusion

This paper studies the difficulties of data mining algorithms applied to sports computing, and proposes a program to analyze the technical skills of volleyball matches based on Markov process data mining algorithms. The program improves the collection efficiency by searching for frequent scripts during the data collection process. In the process of data preprocessing, a method of setting thresholds is proposed to make the data satisfy the algorithm application conditions. With the continuous development of data mining technology and continuous research by sports science and technology personnel, both the theoretical research of data mining in sports and the development of practical data mining tools can bring great convenience to sports management decision-making and scientific research and considerable benefits. Therefore, only by successfully solving the above problems that can data mining technology play a greater role in the scientific development of the sports field, and it

can have a broader development prospect in the sports field.

References

[1] Wang Z. An optimized data mining algorithm application in volleyball match technique and competition tactics analysis. BoletínTécnico. 2017, 55(17), 109.

[2] Giannopoulos N., Vagenas G., Noutsos K., Barzouka K., Bergeles N. Somatotype, level of competition, and performance in attack in elite male volleyball. Journal of Human Kinetics. 2017, 58(1), 131-140.

[3] Sato K., Watanabe K., Mizuno S., Manabe M., Yano H., Iwata H. Development and assessment of a block machine for volleyball attack training. Advanced Robotics. 2017, 31(21), 1144-1156.

[4] Zhang S., Wang H., Huang W., You Z. Plant diseased leaf segmentation and recognition by fusion of superpixel, K-means and PHOG. Optik-International Journal for Light and Electron Optics. 2018, 157, 866-872.

[5] Benson A.J., Eys M. Understanding the consequences of newcomer integration processes: the sport team socialization tactics questionnaire. Journal of Sport and Exercise Psychology. 2017, 39(1), 13-28.

[6] García-Ordóñez E., Touriño-González C., Iglesias-Pérez M.D.C. Effects of situational variables on offensive performance indicators in elite water polo. International Journal of Performance Analysis in Sport. 2017, 17(1-2), 164-176.

[7] Yang Y., Li D. Robust player detection and tracking in broadcast soccer video based on enhanced particle filter. Journal of Visual Communication and Image Representation. 2017, 46, 81-94.

[8] Schrapf N., Alsaied S., Tilp M. Tactical interaction of offensive and defensive teams in team handball analysed by artificial neural networks. Mathematical and Computer Modelling of Dynamical Systems. 2017, 23(4), 363-371.

[9] Hastie P.A., Ward J.K., Brock S.J. Effect of graded competition on student opportunities for participation and success rates during a season of Sport Education. Physical Education and Sport Pedagogy. 2017, 22(3), 316-327.

[10] Leite W.S., Almeida C.H. Competitive-level and mid-term effects on the magnitude of home advantage in portuguese futsal. International Journal of Performance Analysis in Sport. 2018, 18(1), 184-194.

[11] Huang J., Yan E., Cheung G., Nagappan N., Zimmermann T. Master maker: Understanding gaming skill through practice and habit from gameplay behavior. Topics in Cognitive Science. 2018, 9(2), 437-466.

[12] Sotomayor A. Caribbean soccer: hispanoamericanismo and the identity politics of fútbol in puerto rico 1898–1920s. The Latin Americanist. 2017, 61(2), 193-224.