Hindawi
Mobile Information Systems
Volume 2023, Article ID 9802708, 1 page
https://doi.org/10.1155/2023/9802708

Retraction

Retracted: Application of Data Mining in Offensive Tactical Research of Tennis Match

Mobile Information Systems

Received 29 August 2023; Accepted 29 August 2023; Published 30 August 2023

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] H. Yang and X. Li, "Application of Data Mining in Offensive Tactical Research of Tennis Match," *Mobile Information Systems*, vol. 2022, Article ID 5435027, 11 pages, 2022.

Hindawi
Mobile Information Systems
Volume 2022, Article ID 5435027, 11 pages
https://doi.org/10.1155/2022/5435027

Research Article

Application of Data Mining in Offensive Tactical Research of Tennis Match

Huahua Yang[1] **and Xia Li**[2]

[1]*Department of Physical Education, Guangzhou City University of Technology, Guangzhou 510000, Guangdong, China*
[2]*School of Sports Training, Guangzhou Sport University, Guangzhou 510800, Guangdong, China*

Correspondence should be addressed to Huahua Yang; yanghh@gcu.edu.cn

Received 1 June 2022; Revised 15 July 2022; Accepted 25 July 2022; Published 22 August 2022

Academic Editor: Imran Shafique Ansari

With the continuous development and growth of information technology and data scale, data mining technology has gradually matured and applied in various fields. With the continuous improvement in sports information level, the sports data are also increasing. It is the general trend to use data mining technology in the field of sports data analysis. In this article, data mining technology is used in the research of offensive tactical data of tennis, which aims to provide scientific data support for tactical decision-making in tennis training and competition and improve the winning rate of tennis competition. Firstly, based on the requirement analysis, the goal and structure framework of the tennis tactical analysis system are determined. The main functions of the system include tactical data collection, tactical statistics, and tactical comprehensive analysis. Then, the data collection ability, speed, and tactical statistic function of the tennis tactical analysis system are tested. The accuracy rate of data collection for five common offensive tactics in 10 matches of the tennis tactical analysis system is 94.4%, 97.1%, 100%, 97.4%, and 96.3%, respectively. This shows that the tennis tactical analysis system developed in this article can accurately record the tactical data in the game in real time. After the test, the system is used to analyze the tactics used by a tennis player in 2015 and analyze the characteristics of various tactics of the serve round and receive round. The results show that W-C-C is the main tactic of the serve round in the singles match, with the utilization rate of 100% and the scoring rate of W-C-C-Q tactics of 100%; the most important landing area is the inner corner of the right area, with the distribution rate of 30% and the scoring rate of 73.33%. The tactics of receiving round are mainly C-C-C-C tactics, with the utilization rate of 66.67% and the scoring rates of C-C-C-C tactics and C-Q-C-C tactics are 100%; the most important landing area is the middle road, with the distribution rate of 41.67% and the scoring rate of 46.67%. To sum up, the application of data mining in the research of offensive tactics in tennis matches has high analysis efficiency and accuracy, which is worthy of promotion.

1. Introduction

1.1. Background Significance. It is an important way to improve the winning probability and competitive level to scientifically analyze the relevant data of competitive sports and provide scientific and technological support for the tactical decision-making of competition. The popularity of tennis in the world is high, and the major international events are also in full swing. The competitive ability of athletes in the competition is mainly affected by their own physical fitness, skill level, tactical ability, and psychological ability [1]. The success of tactics directly determines the outcome of the game, so it is necessary to study tactics and provide scientific data for tactical decision-making. The application of data mining technology in various fields has brought remarkable economic benefits, and its application in the field of sports is still in its infancy. Using data mining technology to build a tactical analysis system to find effective information in the massive game data is of great significance to improve the tactical level.

1.2. Related Works. Data mining technology is widely concerned by the information industry and society because of its huge amount of data and the urgent need to

transform these data into useful information and knowledge. [2] Recent studies using this technique have successfully calibrated the results to estimate several parameters in different fields. However, the effective use of data in some areas is still developing, such as sports, which has shown moderate growth. In view of the shortcomings of the current sports stationery industry consumption data system, Zhang combines the K-means spatial clustering, fusion decision tree, naive Bayes, and other data mining algorithms with data warehouse technology and applies them to the sports stationery industry [3]. Taking the consumption data system as the research object, he made the geographical spatial feature clustering, customer segmentation, and consumption preference prediction analysis on the consumption of sports stationery industry. He constructed the data fusion system model of sports cultural products industry based on data mining and expounded the architecture, technical route, and function realization of the model. Their research results provide a theoretical basis for the application of data mining technology in the research of offensive tactics in tennis matches. The Markov process model has been widely used in system reliability evaluation. The key factor analysis method based on the Markov model is discussed, and the selection conditions of each increment are analyzed in detail. On the basis of the above discussion, the data mining algorithm for the key factors of the system is given.

Before serving in tennis doubles, the serving team and receiving team can use several tactical variants. The purpose is to analyze the frequency and efficiency of tactical changes in the elite doubles service team and receiving team. Carboch et al. analyze 1067 points of 18 doubles matches in the professional men's top tennis championship and summarizes the strategies and specific tactical variables in the performance of the current top-level men's doubles matches, which is helpful to the preset setting of training amount in these matches [4]. According to the tactical needs of professional tennis, coaches can use this information to adjust training hours. Alexander Raschke and Lames applied the video-based tactical training method that was successfully established in other sports to 10- to 14-year-old Tennis Championship athletes and evaluated its effect through field experiments [5]. In order to maximize the practical relevance, he conducted a field experiment with 12 members in each group. The experimental group took part in the video tactical training for 12 weeks, while the control group continued the conventional training without video. He uses a specially developed video-based tactical test VITT to test the players' understanding of tactical behavior and uses a specially developed competition tactical test MATT to analyze the players' tactical behavior in the formal competition. New test VITT and MATT provide complex information about tactical behavior as performance analysis tools. His method is very novel, but there are some shortcomings in the process of data analysis. Relevant scholars have studied the algorithms, application fields, and tactics of sports events in data mining as well as the actual correlation, but they have not explained the specific application of event tactics in data mining. This article discusses the application of data mining in the study of offensive tactics in tennis matches.

1.3. Innovative Points in This Article. In order to improve the scientific decision-making of tennis game tactics and improve the winning probability and the tactical level in competition, this article studies the tennis offensive tactics based on data mining technology. The innovation of this article is as follows: (1) the demand of the tennis tactical analysis system is analyzed, and the main function of the system is to achieve tactical data collection, tactical statistics, and tactical comprehensive analysis. (2) Using association rules and clustering analysis algorithm of data mining technology to build a tennis tactical analysis system and test its data collection ability, speed, and tactical statistics function. (3) After the test, the tennis tactical analysis system is used to analyze the tactical application in specific matches, and the utilization rate, gain and loss rate, and other detailed information of each tactic are obtained. In this article, the apriori algorithm is used for analysis, which has its advantages in tactical analysis, and data mining technology is used to conduct comprehensive calculation research to obtain data.

2. Data Mining Technology and Offensive Tactics in Tennis Match

2.1. Data Mining Technology

2.1.1. Classification Algorithm. In the field of data mining, classification algorithm can get an objective function through learning and map the data to a predefined class [6]. Classification is widely studied as a method of data mining.

These algorithms all have their own advantages and disadvantages in classification. This article selects which method by explaining its advantages and disadvantages. The advantages and disadvantages have been added to the text. As a commonly used classification method, there are neural network, decision tree classification, support vector machine, and naive Bayes classification. Classification technology is suitable for the prediction of data sets, but it is not very effective for sorting classification. In terms of sorting and classification, clustering algorithms and decision tree models are not many choices. This article uses these two algorithms to explain sorting and classification.

Step 1: the calculation of the output value of each unit in the hidden layer of the BP neural network algorithm is shown in the following formula:

$$y_j = f\left(\sum_{i=1}^{n} e_{ij}x_i - d_j\right). \quad (1)$$

Step 2: the threshold value in the neuron model is written into the joined weight, so that $e_{0j} = -d_j$, $x_0 = -1$. Formula (1) can be transformed as follows:

$$y_j = f\left(\sum_{i=0}^{n} e_{ij}x_i\right). \tag{2}$$

The third calculation formula is as follows:

$$O_k = f\left(\sum_{j=1}^{m} z_{ij}y_j - d_k\right), \tag{3}$$

Step 3: similarly, if $z_{0k} = -d_k, y_0 = -1$, formula (3) becomes

$$O_k = f\left(\sum_{j=0}^{m} z_{jk}y_j\right). \tag{4}$$

Step 4: The difference is calculated between the actual output value and the expected output value and judged whether the difference is less than the previously set error range. If it is less, go back to the Step 2, select the next sample to continue training; if it is greater, carry out error backpropagation and adjust the weight to reduce the difference [7]. The weight correction of the output layer is expressed as follows:

$$\Delta z_{ij} = -\mu\frac{\partial E}{\partial z_{ij}} = -\mu\frac{\partial E}{\partial \text{net}_k}\cdot\frac{\partial \text{net}_k}{\partial z_{ij}}. \tag{5}$$

Step 5: it is to skip to the next sample pair in order and continue training until all samples have completed the training.

Step 6: new samples are imported into the neural network for testing after the training of all samples.

The decision tree model provides a fast method to judge the conditions for obtaining numerical values. The entropy required by the sample is shown as follows:

$$E(x_1, x_2, \ldots, x_m) = -\sum_{i=1}^{m} p_i\log_2(p_i), \tag{6}$$

where x is the number of samples. The entropy of non-category attribute F dividing T data sample set is shown in the following formula:

$$E(F) = -\sum_{j=1}^{v}\frac{x_{1j} + \cdots + x_{mj}}{X}E\left(x_{ij}, \ldots, x_{mj}\right). \tag{7}$$

The information gain ratio of the data sample set is expressed as follows:

$$\begin{aligned} GR(F, X) &= \frac{G(F, X)}{SI(F, X)}, \\ G(F, X) &= E(x_1, x_2, \ldots, x_m) - E(F), \\ SI(F, X) &= -\sum_{i=1}^{c}\frac{|X_i|}{|X|}\log_2\frac{|X_i|}{|X|}. \end{aligned} \tag{8}$$

2.1.2. Correlation Analysis. Association analysis can find meaningful relations hidden in massive data, and the association rules between two disjoint itemsets can be expressed by $X \longrightarrow Y$ [8]. The accuracy and applicability of association rule analysis are generally expressed by confidence and support, which are expressed as follows:

$$\frac{\text{confidence} = |\{R:\ X \cup Y \subseteq R, R \in Q\}|}{|\{R:\ X \subseteq R, R \in Q\}| \times 100\%}. \tag{9}$$

$$\frac{\text{support} = |\{R:\ X \cup Y \subseteq R, R \in Q\}|}{|Q| \times 100\%}. \tag{10}$$

Support is used to filter and delete meaningless rules, and trust is used to infer the reliability of rules. The original method of correlation analysis is to calculate the support and reliability of possible rules, but the number of rules that can be extracted from the data set increases exponentially according to the amount of data, so the amount of calculation is large. Because of its large computing capacity and storage, it is relatively large in actual data. Correlation analysis is suitable, but after correlation analysis, data information needs to be classified by clustering to facilitate calculation.

The apriori algorithm is a classical algorithm for mining frequent itemsets in association rules [9]. Based on the important property that the non-empty set of a group of frequent items must also be frequent, it uses the iterative method of successive search to find a group of frequent items to generate relevant rules. The general step is to first access the processed data, calculate the minimum support and minimum confidence, then connect and prune, generate all frequent itemsets, and finally output strong association rules. The apriori algorithm can analyze the data information in the experiment to form a set of proprietary analysis rules to deal with a large amount of data. This algorithm is used in the offensive tactics in the text.

2.1.3. Cluster Analysis. By transmitting the data of tactical information to the computer, according to the different association rules of its tactics, the training selection is carried out, and it is inappropriate to filter the selected data and the data in the actual tactics. Cluster analysis divides the sample data into different groups according to specific metrics, and the data in the same group have similar characteristics. The clustering algorithm has advantages and disadvantages for different application schemes and characteristics of different data samples, and it is applied to data sets of different algorithms, and the results are compared to obtain the possible global optimal solution [10]. The commonly used distance measurement methods in cluster analysis include Mahalanobis distance, and Euclidean distance [11, 12], and their calculations are shown in formulas (11)–(13), respectively:

$$d_{ij} = \left[\left(x_i - x_j\right)^t \left(x_i - x_j\right)\right]^{1/2}. \quad (11)$$

$$d^2 = \left(x_i - x_j\right)^T S^{-1} \left(x_i - x_j\right). \quad (12)$$

$$d(i,j) = \sqrt{\left(x_{i1} - x_{j1}\right)^2 + \left(x_{i2} - x_{j2}\right)^2 + \cdots + \left(x_{in} - x_{jn}\right)^2}. \quad (13)$$

Clustering analysis is inductive, which can analyze the data without predetermined classification criteria, and has nothing to do with the labels of known clusters. According to the principle of maximizing the similarity within the class and minimizing the similarity between the classes, the cluster and grouping are carried out. Through calculation and reasonable classification, the observations of the same class are relatively close, and the observations of different classes are quite different. Each cluster generated can be treated as an object class, and then rules can be exported. Clustering can enhance the understanding of objective reality, which is the precondition of concept description and bias analysis.

The K-means algorithm is a typical clustering algorithm, which randomly selects K objects in the cluster as the initial clustering center. Then, all the objects are assigned to the center closest to the object, and the n objects are divided into k clusters. The average value of each cluster is recalculated and repeated until the maximum number of iterations is reached [13].

BP neural network algorithm: this algorithm has nonlinear mapping, self-learning, and adaptive abilities. The generalization algorithm has strong fault tolerance, but the convergence speed is slow, and its structure selection is different, which is suitable for the contradiction between prediction ability and training ability, as well as the problem of sample dependence. Apriori algorithm: Data are organized horizontally, which are suitable for mining association rules of transactional databases, but scanning the database multiple times generates a large I/O load and significantly increases the computing time. K-means algorithm: it is also a kind of clustering algorithm, which is simple and fast, and has high efficiency for processing large data sets. Algorithms often end up in local optima, but are sensitive to noise, and a small amount of this type of data can have a large impact on the average.

2.2. Tennis Competition Information Collection Technology

2.2.1. Wireless sensor network. The wireless sensor network is composed of multiple wireless sensor nodes with the same or different functions. These sensor nodes are interconnected with the wireless network to sense, collect, and process the information of the sensing object in the geographical area covered by the network and send it to the observer [14]. The development of a microcomputer system provides the possibility for the miniaturization of sensors. The development of microprocessing technology promotes the intellectualization of sensors. The combination of microelectromechanical systematization technology and radio frequency communication technology promotes the birth of wireless sensors and their networks.

The hardware resource of a wireless sensor network is limited, so the protocol level should not be too complex. The capacity of the power supply battery used by the node is limited, and the depletion of battery energy means that the node is invalid. Therefore, the premise of energy saving must be considered in the design of wireless sensor networks. All nodes in wireless sensor networks are equal and decentralized. The failure of a single node will not affect the operation of the network. Nodes organize their own network through layered protocols and distributed algorithms and do not need to rely on preset network facilities. All nodes in the wireless sensor network have their own transmission methods, and each layer has information transmission, through the description of the task, signal processing, time synchronization, and coordinated operation.

The key technology of a wireless sensor network is closely related to its performance. The first is the network protocol, which is similar to the protocol framework of the Internet. It can guarantee the efficient communication of the network and improve the utilization rate of energy and the quality of service [15]. The second is network security, which is also the premise of system availability, generally from the two aspects of routing security and security protocol. The third is energy management technology, which must extend the life of the network as far as possible. The fourth is data fusion technology, which can reduce the amount of data transmission and improve the accuracy and credibility of information.

2.2.2. Short-Range Wireless Communication Protocol. Short-range wireless communication system is generally composed of a wireless transmitter and a wireless receiver, and they are powered by batteries. Its wireless transmission power is very low, generally less than 100 microwatts [16]. Therefore, the communication range is relatively small, within a few hundred meters. Short-range wireless communication uses an omnidirectional antenna and circuit board antenna, and there is no need to apply for a wireless channel.

ZigBee wireless network protocol is between wireless marking technology and Bluetooth, which is mainly used for short-range wireless connection [17]. ZigBee protocol sensor power consumption is very low, to relay the way to transfer data, so the communication efficiency is very high. The speed of data transmission is low and the cost is greatly reduced. It has a short delay, complete data checking, and weighting function, which ensures security. Each ZigBee protocol network device can be connected with 254 devices, with large capacity and excellent network topology capability.

The standard handshake process is added to the MiWi wireless network protocol, which simplifies the operation of link disconnection and channel jump [18]. After the device is powered on, the first task is to shake hands with other devices. The device will send a beacon request first, and the device that can connect with other devices will respond to its

request. After the initiating device collects the beacon, it determines a beacon and uses it to establish a handshake to send association request and data request. The handshake protocol is completed when the device receiving the connection sends an association response. MiWi wireless network protocol has the characteristics of small program space. It can enter the sleep state after the end of communication and turn off the RF when the device is idle.

Wi-Fi protocol is widely used in daily life, and its radio wave coverage is large, and the radius can reach about 100 meters [19]. Only in the coverage area, the Internet can be accessed, without the cost of network wiring access, which can save costs. The transmission speed is fast, which can meet the needs of informatization, but the communication quality is general, and the security performance is poor.

2.2.3. Tennis Match Information Collection System. The information collection system of tennis match is responsible for collecting relevant tactical data. The general information collection system includes a video monitoring window and edit bar, technical and tactical indicators, and a simulation field. In the video monitoring window, we can analyze the video of tennis match and adjust the video progress at will. As shown in Figure 1, while watching the video, according to the actual situation of tactical application, you can input relevant information from the tactical indicators on the right side in order. Once the tactical information input is completed, the system will automatically jump to the blank input state and input the next technical information.

Since the system automatically defaults to singles, it is more complex to collect data in doubles. In doubles, when inputting tactical information in each match, we must reselect the basic information such as the formation and batting order of both players. The information collection system completes the comprehensive collection of tennis game tactical information and provides the basis for the exploration and analysis of the tactical data in the game [20].

2.3. Tennis Competition Tactics

2.3.1. Analysis of Site Factors. The Australian Open is a hard-medium speed court, its material structure determines that the rebound angle and speed of the ball are maintained at a medium level, and the rebound stability of the ball is high [21]. That is to say, the irregular rebound ball is less and the rebound speed is fast. The smoother the court is, the higher the hardness is and the smaller the impact of the ball rebound is. It can reflect all the spinning balls. On the hard court, the speed of the ball is medium, the inside and outside corners of the serve angle keep balance, and the scoring probability is equal. Because the surface of the hard court is relatively balanced, there are few chances for the ball to bounce irregularly, and the obstacles to players' judgment can almost be ignored.

The French Open is a laterite slow court, which is not as flat and hard as the hard ground made of concrete and other mixed materials. Granular structure makes the site rougher and the friction coefficient larger. The rotation of the ball becomes stronger, the reaction time of players increases, and

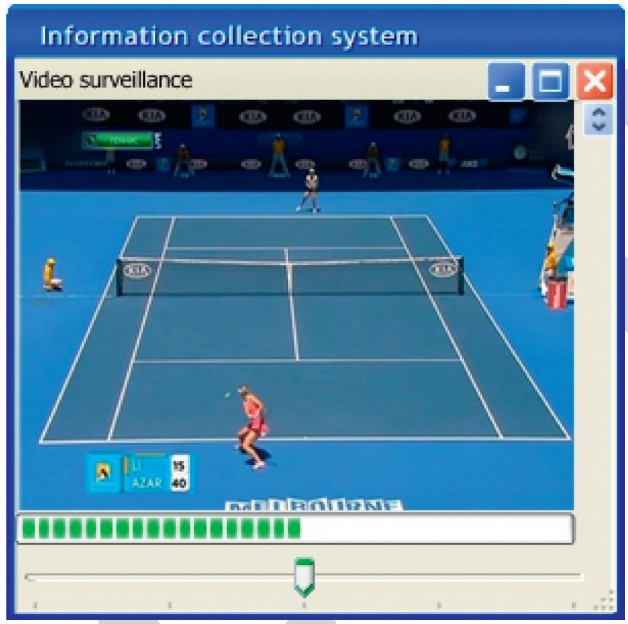

Figure 1: Video surveillance window of tennis match.

the preparation time is also more abundant. The characteristics of high friction provide the maximum sliding space for athletes, which can further test the psychological quality and strong physical strength of athletes. The clay field is easily affected by the weather and air humidity, the viscosity of the field becomes higher, the weight of the ball increases, the speed of the ball decreases, and the irregularity of the ball bounce increases [22]. The occurrence rate of the middle line becomes lower, the occurrence rate of the drop zone in the serve stage becomes lower, and the utilization rate of the inner corner and the outer corner becomes higher.

Wimbledon is a grass fast track, affected by the high cost of soil, grass quality, climate, and maintenance; it is difficult to popularize in the world. In order to reduce the condition and meet the performance of turf, artificial turf was made. The turf has less friction and the ball rebounds quickly after falling. In the serve stage, the middle of the inner corner and the outer corner is the choice of the landing point [23]. The emergence rate of the middle is rising, and the fast-middle serve is easy to produce body tracking effect, which has an advantage. In the stage of receiving service, the utilization rate of cutting ball increases, while the utilization rate of forehand and backhand decreases. In the stalemate stage, it will obviously reduce the number of shots.

2.3.2. Structure of Tennis Game. In tennis training, it is assumed that the players choose to serve the inside corner and the outside corner. If the inside corner serves the topspin, and the outside corner serves the side spin and topspin, they will achieve their goals; if they fail to serve, they will not achieve their goals. This is a single-player game problem, and its structure includes players, strategies, and benefits. If the goal is achieved, the benefit is positive; if the goal is not achieved, the benefit is negative. If the information owned by the players increases, such as the basic information of the other

party, game image, tactical characteristics, physical condition, and psychological pressure resistance ability [24], we can choose a more reasonable and effective game strategy based on the information. The higher the correctness of the decision, the greater the income.

A tennis game involves two people, strategic planning, serve and receive tactics, rhythm control, and every moment in the fight for wisdom and courage. The tactical thinking of athletes in the process of competition is particularly important in the doubles game. In the actual game, not every player has the same strategy choice. In many games, the choice of strategy is not only different, but the number of choices is also different.

The limited number of opponents in each game is a "limited strategic game"; at least a few players have unlimited strategies, which is an "unlimited strategic game." In tennis competitions, only a few strategies can be selected at most, and the options of infinite game basically do not exist. In tennis tactics, there are only two or three types of games that are prepared, studied, and used most flexibly.

2.3.3. Types of Tennis Tactics. In order to defeat the opponent in the tennis match, or to show their expected competitive ability, the tactics and actions are tennis tactics. In the Australian Open, top athletes use the tactics of attack and defense, exchange attack and defense, and attack and defense stalemate [25]. On the court, players can make use of the characteristics of service to form tactics in the service stage; the ability to receive service is enhanced, the tactics of defense turning to attack are obvious, and scoring opportunities are actively created. Australian Open serve tactics, whether one serve or two serves, each path is chosen equally. The serve stage is also the first step of the attack. If the attack is successful, you can score immediately, or press the other side to make mistakes to score, creating an opportunity to attack. The receiving tactics of the Australian Open are usually passive, and they usually choose to return the ball in the middle line. In the stalemate stage of the Australian Open, the opportunities of both sides were balanced, and no one had an absolute advantage.

In the French Open, the tactics of serving, receiving, and bottom line attack are obviously switched. Compared with the hard court, the threat of serve tactics is weaker. The application of receiving tactics is more prominent. Technology is the means to achieve tactics, tactics are relatively stable, but the performance of the technology is relatively single. In the serve tactics of the French Open, when one serve and two serves, you can take the initiative to choose the inside corner and the outside corner. French Open receiving is passive in tactics, but because of its irregular slow bounce, the oblique line rate and straight-line rate of return ball increased obviously. In the French stalemate stage, the chances of both players are balanced, and the use rate of forehand and backhand is high and low. Generally speaking, the use rate of the backhand is higher than that of the forehand.

In Wimbledon, with the change of physical environment on the field, the speed is accelerated, and the tactical utilization rate of the boundary line is reduced. In Wimbledon serve tactics, inside and outside corners are the target of active selection. Serve stage is an important means of scoring. In Wimbledon's receiving tactics, passivity becomes more obvious. With the rapid service ability of the players and the speed of the lawn court, the occurrence rate of springback using cutting technology has increased significantly. In the stalemate stage of Wimbledon, the exchange of attack and defense becomes less, and the players choose to attack and defend each other obliquely and straightly. Some common terms used in tennis are double fault, net, out of bounds, high ball, serve, game point, inventory, match point, lob, winning point, chipping, and netting.

3. Experiments on Offensive Tactics of Tennis Match Based on Data Mining

3.1. Construction of the Tennis Tactical Analysis System

3.1.1. System Requirement Analysis. The tennis tactical analysis system is a ball analysis system that studies the collection and analysis of tennis match information. Traditional statistical methods can only count the frequency of technical movements and the score of athletes, but cannot analyze the potential practice and tactical cooperation between tactics. The application of data mining technology can solve these problems.

Based on the characteristics of tennis form, the tennis tactical analysis system must be able to get the following information: the relationship between tactical and technical actions in the game, the relationship between two tactics that users are interested in, the situation of winning and losing points, and cooperation of tactics.

Based on the above analysis, the goal of the tennis tactical analysis system should include: to find the relationship between different tactics in the game; to find the technical action matching with the designated tactics; to find the key tactics that affect the outcome of the game; simple interface, simple operation, short human-computer interaction time; the conclusion can be displayed in the form of visualization.

3.1.2. System Design. The structural framework of the tennis tactical analysis system is shown in Figure 2. On the basis of the original system, the mining algorithm of tactical analysis is abstracted, and the middleware of developing class library is created, so that the tennis tactical analysis system can be developed by using class library, and the mining results are displayed.

The overall function of the system is divided into three modules, which are tactical data collection, tactical statistics, and tactical comprehensive analysis. The tactical data collection part is the basis of data mining, which inputs the relevant game data into the database [26].

3.2. Application of Data Mining Technology in Tactical Analysis

3.2.1. Application of Association Rules in the Tennis Tactical Analysis System. The process of data mining includes the collection, extraction, processing, and mining analysis of the tennis tactical information data.

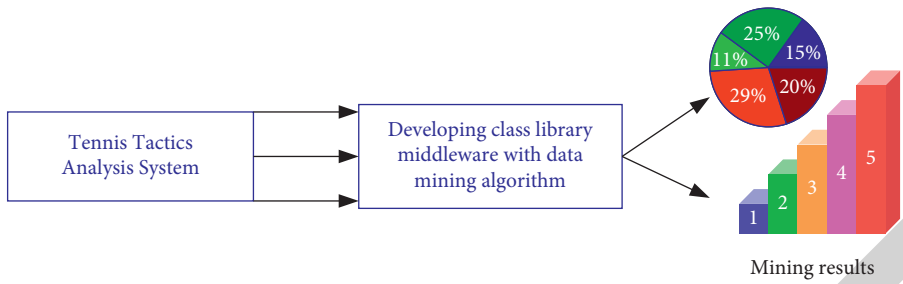

Figure 2: System architecture.

In the process of data collection, the tactical data that need to be collected mainly include the hit area, hit position, and technology. The two halves of the tennis court are divided into six areas, which are coded for the convenience of recording, as shown in Table 1.

Then, according to the needs of the decision content, the data that meet the requirements from the database are extracted, and the abnormal data is checked and processed. The original data format is transformed to meet the requirements of mining tools. Finally, the association rule algorithm is used for pattern mining.

3.2.2. Application of Cluster Analysis in the Tennis Tactical Analysis System. The application of cluster analysis in the tennis tactical analysis should first preprocess the original data to improve the quality of data. Data cleaning is used to delete the records whose attribute values are empty in serve and receive tactics, then the data are normalized, and the K-means algorithm is used to cluster the data.

3.3. Test and Application of the Tennis Tactical Analysis System. This article mainly tests the data collection ability, speed, and tactical statistics function of the tennis tactical analysis system. This article collects data from the training samples of the game and analyzes whether the collected data can be applied to the current tactics by recording the movements in the tennis game and whether the tactical statistics function can accurately and quickly calculate the success rate and failure rate of a tactic and complete the statistics of tennis landing points and sports routes. Ten well-known tennis matches are selected as the test data, and five kinds of offensive tactics commonly used by tennis players are selected as the collection objects. The times of the tactics in each match are recorded and compared with the actual times, so as to judge the data collection ability and speed of the system.

After testing, the system is used to analyze the tactical application of a tennis player in recent years and to analyze the tactical characteristics of serving and receiving, including usage rate, points gained and lost, serving distribution rate, and service point.

4. Discussion on the Results of Offensive Tactics in Tennis Match Based on Data Mining

4.1. Test Results. The five common offensive tactics in tennis match are numbered as A, B, C, D, and E. The tennis tactical analysis system was used to record the times of these five tactics in each game and compared them with the actual times. The statistical results are shown in Table 2.

Table 1: Comparison table of field landing area code.

Code	Landing area (upper half)	Code	Landing area (lower half)
U1	Middle and backcourt right location	L1	Middle and back right location
U2	Middle and backcourt middle road	L2	Middle and backcourt middle road
U3	Middle and backcourt left location	L3	Middle and backcourt left location
U4	Frontcourt right location	L4	Frontcourt right location
U5	Frontcourt middle road	L5	Frontcourt middle road
U6	Frontcourt left location	L6	Frontcourt left location

As shown in Table 2, in the statistics of the tennis tactical analysis system, A tactics appear 34 times in 10 games, B tactics appear 34 times, C tactics appear 31 times, D tactics appear 37 times, and E tactics appear 26 times. The actual data of five tactics in 10 games are shown in Figure 3.

As shown in Figure 3, in the actual data, A tactics appear 36 times in 10 games, B tactics appear 35 times, C tactics appear 31 times, D tactics appear 38 times, and E tactics appear 27 times. By comparing the two sets of data, a′ indicates the actual number of times a tactic appears in the game.

As shown in Figure 4, in the 10 games, there are five games of data statistical errors, respectively, the first game of A tactics missed once, the third game of A tactics missed once, the seventh game of E tactics missed once, the ninth game of B tactics missed once, and the tenth game of D tactics missed once. This shows that there is no misjudgment in the system, only a few missing records.

The accuracy of each tactical data collection is calculated, and the results are shown in Table 3.

As shown in Table 3, the accuracy of data collection of the tennis tactical analysis system for five common offensive tactics in 10 games is 94.4%, 97.1%, 100%, 97.4%, and 96.3% respectively. This shows that the tennis tactical analysis system developed in this article can accurately record the tactical data in the game in real time.

Table 2: The results of statistics of attack tactics.

Matches played	A	B	C	D	E
1	3	4	3	5	2
2	4	3	5	4	3
3	4	2	3	5	1
4	2	5	4	3	2
5	5	2	4	3	3
6	3	4	2	5	1
7	5	4	2	6	1
8	1	4	3	2	4
9	2	2	4	3	3
10	5	4	1	1	6
Total	34	34	31	37	26

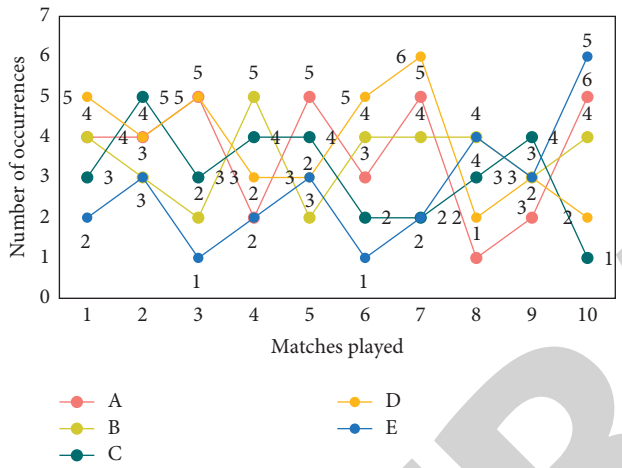

Figure 3: Actual times of attack tactics.

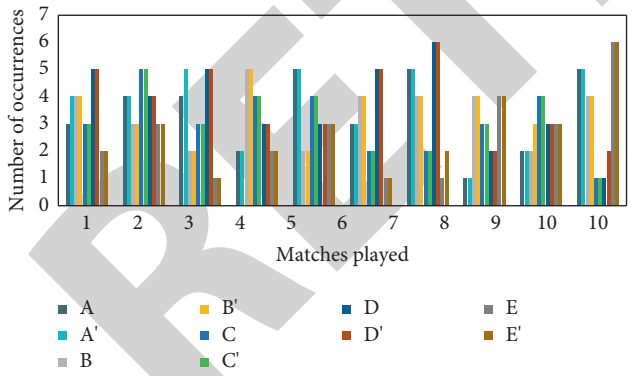

Figure 4: Comparison of system statistical data and actual data.

4.2. Application Results of Tactical Analysis

4.2.1. Analysis of Serve Round Tactics. This article analyzes the tactical utilization rate of a tennis player in a single serve round in 2015, mainly including the first three strokes and the first four strokes. The single technique is coded, the no-hit technique is W, the draw technique is C, and the slice technique is Q. The tactics used in the first three shots are W-C-C, and the tactics used in the first four shots are W-C-C-C and W-C-C-Q. The utilization rate is shown in Figure 5.

Table 3: Accuracy of data acquisition.

Tactics	A	B	C	D	E
Statistical times	34	34	31	37	26
Actual times	36	35	31	38	27
Accuracy	94.4%	97.1%	100%	97.4%	96.3%

As shown in Figure 5, the first three shots of the player's serve round tactics are mainly W-C-C tactics, with a utilization rate of 100%; the first four shots are mainly W-C-C-C tactics, with a utilization rate of 90%, and the utilization rate of W-C-C-Q tactics is 10%. Therefore, in the single serve round, the main tactics used by the player are W-C-C and W-C-C-C.

Then the scoring rate of each tactic in the serve round is calculated and the results are shown in Figure 6.

As shown in Figure 6, the scoring rate of W-C-C tactics in the first three shots, the scoring rate of W-C-C-C tactics in the first four shots, and the scoring rate of W-C-C-Q tactics are 15.34%, 95.8%, and 100%, respectively. This shows that the main scoring tactics of the serve round are W-C-C-C and W-C-C-Q.

This article analyzes the distribution rate and the rate of gain and loss of points of the player in the service round. The area of service point is divided into four areas, namely LI, LO, RI, and RO, and the results are given in Table 4.

As shown in Table 4, the distribution rates of the left inner and outer corners are 21.67% and 23.33%, respectively, and the right inner and outer corners are 30% and 25%, respectively. It shows that the most important area of the player's serve wheel is the inner corner of the right area.

As shown in Figure 7, the scores of the left inner and outer corners were 46.15% and 50%, respectively, and the scores of the right inner and outer corners were 55.56% and 73.33%, respectively. It shows that the main scoring area of the serve round is the outer corner of the right area.

4.2.2. Tactical Analysis of Receiving and Serving Round. This article analyzes the tactical utilization rate of the player's receiving and serving round in a single match in 2015, mainly including the first three strokes and the first four strokes. The tactics used in the first three shots are C-C-C and Q-C-Q, and the tactics used in the first four shots are C-C-C-C and C-Q-C-C; the utilization rate is shown in Figure 8.

As shown in Figure 8, the first three shots of the player's return round hitting tactics are mainly C-C-C tactics, the utilization rate is 60%, and the utilization rate of Q-C-Q tactics is 40%; the first four shots are mainly C-C-C-C tactics, the utilization rate is 66.67%, and the utilization rate of C-Q-C-C tactics is 33.33%. Therefore, in the single serve round, the player's main tactics are C-C-C and C-C-C-C.

The scoring rate of each tactic of receiving and serving round is shown in Figure 9.

As shown in Figure 9, the scoring rate of C-C-C tactics in the first three shots is 98.8%, and that of Q-C-Q tactics is 95%; the scoring rates of C-C-C-C tactics and C-Q-C-C tactics in the first four shots are 100%. This shows that the main scoring tactics of the serve round are C-C-C-C tactics and C-Q-C-C tactics.

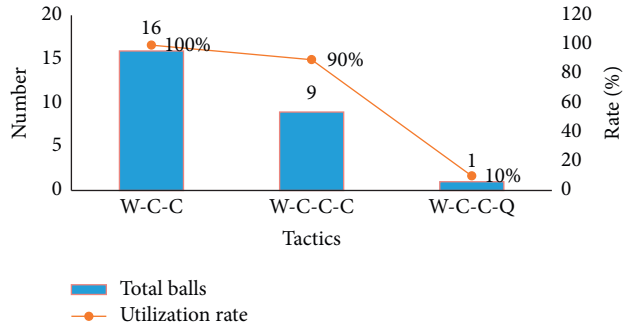

Figure 5: Utilization rate of single serve round hitting tactics.

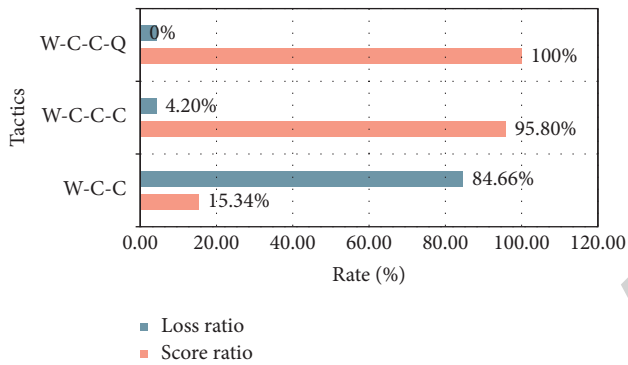

Figure 6: The ratio of winning and losing points in single serve round.

Table 4: Distribution of service points and winning and losing points of service round.

Region	Total balls	Utilization ratio (%)	Win	Lose	Score ratio (%)	Loss ratio (%)
RI	18	30	10	8	55.56	44.44
RO	15	25	11	4	73.33	26.67
LI	13	21.67	6	7	46.15	53.85
LO	14	23.33	7	7	50	50

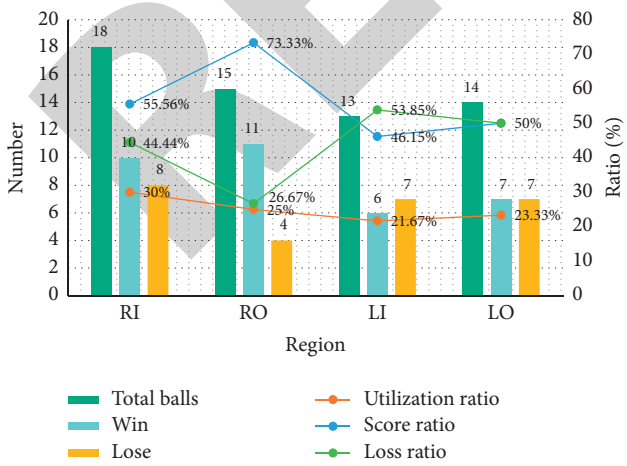

Figure 7: Distribution rate of service point and gain and loss rate of service round.

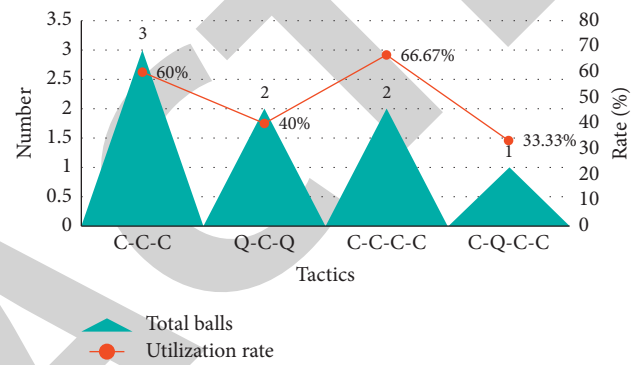

Figure 8: Tactical utilization rate of receiving round.

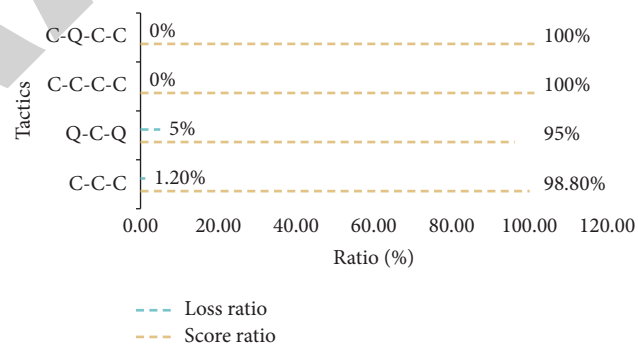

Figure 9: The percentage of points gained and lost in single game receiving and serving round tactics.

This article analyzes the distribution rate of the drop point and the gain and loss rate of the player in the receiving round. The area of the receiving point is divided into middle, right, and left, and the results are tabulated in Table 5.

As shown in Table 5, the distribution rates of the drop points in the middle, right, and left areas are 41.67%, 27.78%, and 30.55%, respectively, which indicates that the most important drop point area of the receiving and serving wheel is the middle.

As shown in Figure 10, the scoring rates of the middle, right, and left areas are 46.67%, 30%, and 27.27%, respectively, indicating that the main scoring area of the receiving and serving round is the middle.

Table 5: Distribution of service point and score of receiving round.

Region	Total balls	Utilization ratio (%)	Win	Lose	Score ratio (%)	Loss ratio (%)
Middle road	15	41.67	7	8	46.67	53.33
Right area	10	27.78	3	7	30	70
Left area	11	30.55	3	8	27.27	72.73

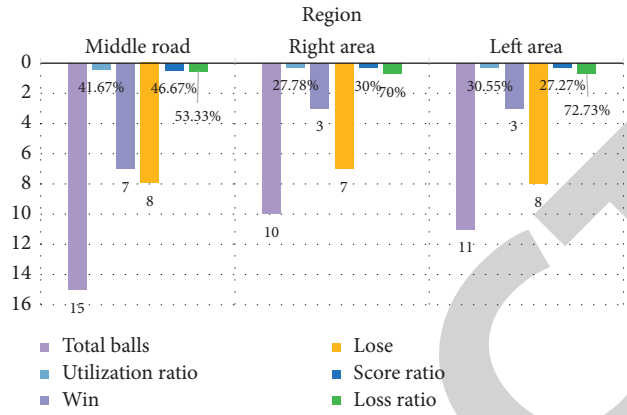

Figure 10: Distribution rate of service point and gain and loss rate of receiving round.

5. Conclusions

It is an important way to improve the winning probability and competitive level to scientifically analyze the relevant data of tennis match and provide scientific and technological support for tactical decision-making. Data mining technology can classify and encode techniques and tactics, convert them into scripts to describe the basic sentence patterns of the language, and input the technical actions in sports into the computer. According to different technical routines, different tactics are changed, and tactics are the winning ones. Critical, tactical analysis plays an important role. The application of data mining technology in various fields has brought remarkable economic benefits, but its application in the field of sports is still in its infancy. Using data mining technology to build a tactical analysis system to find effective information in the massive game data is of great significance to improve the tactical level. After recognizing the function of data mining, it has been applied in many sports fields abroad, and a lot of important information has been mined from the historical statistical data of competitive sports, and a series of research ideas, methods, and formulas have been proposed accordingly. With professional sports data mining tools, many sports organizations have hired specialized data mining personnel to predict the future prospects of sports teams, and data mining will certainly be the mainstream method of sports data analysis in the future.

In this article, data mining technology is used in the research of offensive tactical data of tennis, which aims to provide scientific data support for tactical decision-making in tennis training and competition and improve the winning rate of tennis competition. This article constructs the tennis tactical analysis system. Firstly, based on the demand analysis, the goal and structure framework of the tennis tactical analysis system are determined. The main functions of the system include tactical data collection, tactical statistics, and tactical comprehensive analysis. Then the data collection ability, speed, and tactical statistics function of the tennis tactical analysis system are tested.

After the test, the system is used to analyze the tactics used by a tennis player in 2015 and analyze the characteristics of various tactics of the serve round and receive round. Whether it is test results or practical application, the system has a strong and fast data acquisition ability and speed, can accurately count the use of tactics, and accurately analyze the gain and loss rate of tactics. This provides data support for the tactical decision-making in the follow-up competition, which is conducive to the improvement in the tactical application level and is worthy of promotion and use.

Data Availability

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

Conflicts of Interest

The authors state that this article has no conflicts of interest.

References

[1] B. A. Gennadievich, "Machine learning and data mining activity results when using projectiles in different sports," *International Journal of Advanced Trends in Computer Science and Engineering*, vol. 9, no. 3, pp. 3157–3160, 2020.

[2] R. Parmezan Bonidia, J. Duilio Brancher, and R. Marques Busto, "Data mining in sports: a systematic review," *IEEE Latin America Transactions*, vol. 16, no. 1, pp. 232–239, 2018.

[3] L. Zhang, "Design of a sports culture data fusion system based on a data mining algorithm," *Personal and Ubiquitous Computing*, vol. 24, no. 1, pp. 75–86, 2020.

[4] J. Carboch, J. Carboch, M. Cabela, and K. J, "Tactics in tennis doubles: analysis of the formations used by the serving and receiving teams," *International Journal of Physical Education, Fitness and Sports*, vol. 9, no. 2, pp. 45–50, 2020.

[5] A. Raschke and M. Lames, "Video-based tactic training in tennis," *German Journal of Exercise and Sport Research*, vol. 49, no. 3, pp. 345–350, 2019.

[6] I. Kavakiotis, O. Tsave, A. Salifoglou, N. Maglaveras, I. Vlahavas, and I. Chouvarda, "Machine learning and data mining methods in diabetes research," *Computational and Structural Biotechnology Journal*, vol. 15, no. C, pp. 104–116, 2017.

[7] A. L. Buczak and E. Guven, "A survey of data mining and machine learning methods for cyber security intrusion detection," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 2, pp. 1153–1176, 2016.

[8] L. Xu, C. Jiang, J. Wang, Y Jian, and R Yong, "Information security in big data: privacy and data mining," *IEEE Access*, vol. 2, no. 2, pp. 1149–1176, 2014.

[9] C. Helma, T. Cramer, S. Kramer, and L. De Raedt, "Data mining and machine learning techniques for the identification of mutagenicity inducing substructures and structure activity relationships of noncongeneric compounds," *Journal of Chemical Information and Computer Sciences*, vol. 44, no. 4, pp. 1402–1411, 2004.

[10] X. Zhao and J. Tang, "Crime in urban areas: a data mining perspective," *Acm Sigkdd Explorations Newsletter*, vol. 20, no. 1, pp. 1–12, 2018.

[11] M. Nematollahi, R. Akbari, S. Nikeghbalian, and C. Salehnasab, "Classification models to predict survival of kidney transplant recipients using two intelligent techniques of data mining and logistic regression," *International Journal of Organ Transplantation Medicine*, vol. 8, no. 2, pp. 119–122, 2017.

[12] J. . Keller, "DARPA adds two companies to OFFSET swarm reconnaissance drone research project[J]," *Military & Aerospace Electronics*, vol. 29, no. 4, pp. 30-31, 2018.

[13] H. Y. Jeong and T. S. Yoon, "Analysis of hepatitis C virus using data mining algorithm -Apriori, decision tree," *International Journal on Bioinformatics & Biosciences*, vol. 7, no. 3, pp. 01–11, 2017.

[14] M. P. Couper, "The future of modes of data collection," *Public Opinion Quarterly*, vol. 75, no. 5, pp. 889–908, 2011.

[15] M. Dong, K. Ota, and A. Liu, "RMER: reliable and energy-efficient data collection for large-scale wireless sensor networks," *IEEE Internet of Things Journal*, vol. 3, no. 4, pp. 511–519, 2016.

[16] A. Misra, A. Gooze, K. Watkins, M. Asad, and C. A. Le Dantec, "Crowdsourcing and its application to transportation data collection and management," *Transportation Research Record*, vol. 2414, no. 1, pp. 1–8, 2014.

[17] N. Tamaki, T. Mukai, Y. Ishii et al., "Comparative study of thallium emission myocardial tomography with 180° and 360° data collection," *Journal of Nuclear Medicine*, vol. 23, no. 8, pp. 661–666, 2017.

[18] A. Tzikas, D. R. Holmes, S. Gafoor et al., "Percutaneous left atrial appendage occlusion: the Munich consensus document on definitions, endpoints, and data collection requirements for clinical studies," *Eurointervention*, vol. 12, no. 1, pp. 103–111, 2016.

[19] C. C. Chang and T. L. Lee, "Fencing tactics analysis in broadcast video: a point-by-point Analytical system," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 29, no. 7, pp. 2162–2175, 2019.

[20] G. Kadiravan, P. Sujatha, T. Asvany et al., "Metaheuristic clustering protocol for healthcare data collection in mobile wireless multimedia sensor networks," *Computers, Materials & Continua*, vol. 66, no. 3, pp. 3215–3231, 2021.

[21] A. Mavvidis, G. Dallas, A. Theodosiou, and I. Mavvidis, "Improvement of theoretical tennis knowledge in student coaches at the sports university of Greece," *OALib*, vol. 05, no. 10, pp. 1–8, 2018.

[22] D. Schranner, L. Scherer, G. P. Lynch et al., "In-play cooling interventions for simulated match-play tennis in hot/humid conditions," *Medicine & Science in Sports & Exercise*, vol. 49, no. 5, pp. 991–998, 2017.

[23] Galé-Ansodi, Carlos, J. Castellano, and O. Usabiaga, "Physical profile of young tennis players in the tennis match-play using global positioning systems," *Journal of Physical Education and Sport*, vol. 17, no. 2, pp. 826–832, 2017.

[24] P. P. Dharma, H. Setijono, and E. Mintarto, "Impact of anthropometric factor on serve velocities evaluation tennis match at the 2019 US Open Grandslam tournament," *JOSSAE Journal of Sport Science and Education*, vol. 5, no. 1, p. 48, 2020.

[25] M. T. Smith, M. Reid, S. Kovalchik, T. O. Woods, and R. Duffield, "Heat stress incident prevalence and tennis matchplay performance at the Australian Open," *Journal of Science and Medicine in Sport*, vol. 21, no. 5, pp. 467–472, 2018.

[26] O. I. Khalaf and G. M. Abdulsahib, "Optimized Dynamic Storage of Data (ODSD) in IoT Based on Blockchain for Wireless Sensor Networks," *Peer-to-Peer Netw*, vol. 14, pp. 2858–2873, 2021.