Table tennis video data mining based on performance optimization of artificial fish swarm algorithm

Hai Wang [1,*], Wenbo Qu [2], Qunli Shen[3]

[1] *Department of Physical Education, Shanghai Business School, Shanghai, 201400, China*

[2] *School of Humanities and Law, Shanghai Business School, Shanghai, 201400, China*

[3] *College of Computer and Information , Shanghai Business School, Shanghai, 201400, China*

Received 1 June 2014, www.cmnt.lv

Abstract

In view of the traditional AFSA still has problems such as the low optimal performance and poor efficiency in data mining for table tennis video, a video data mining model for table tennis match is proposed based on improved AFSA. First, the traverse range of chaotic motion enlarges to the value range of AFSA optimization variables by leading in chaotic Logistic mapping. And then it increases the optimal artificial fish state information on the basis of original artificial fish behaviour, so guides the artificial fish to quickly close to the global optimal, improves the speed of the algorithm. Finally, adaptive optimize the search strategy of traditional AFSA, and apply the improved algorithm to the video data mining for table tennis match. The simulation experiments show that the video data mining model for table tennis match based on improved AFSA has better ability of optimization, and can dig out more attributes and types in the table tennis game video.

Keywords: Improved AFSA, Table Tennis Match, Data Mining, Chaotic Logistic mapping, Behavioral Optimization, Adaptive Strategy

1 Introduction

With gradually in-depth knowledge and understanding on the data mining technology and the significant economic benefits of data mining produced in the practical application, the foreign countries have the application of data mining technology in competitive sports [1]. And with progressively realized the informatization of our sports enterprise, a large amount of data accumulated in the field of competitive sports, how to make the fullest use of these data and find the valuable important information by ignored, has become an important task for the researchers [2]. Table tennis is a confrontational competitive sport, it not only requires the athletes with high technical level, and also has higher demand to the use of tactics, because the use of tactics play a decisive role on the outcome in most of competition. Therefore, it is necessary in analysis applications of data mining and knowledge discovery for table tennis match, its purpose is to provide a large number of complicated information related to the database table tennis for real time and deep-seated analysis, find a real valuable information and knowledge, so as to provide reference and basis for the tactics using by coaches decision [3].

It has a lot of common methods to sports video data mining, M.S.Lew extracts the lens of goal by billiards, the color of the table and ball bag, tracing the motion curve of chromosphere in snooker competition [4]. A.Ekin makes the match video clips of games in play and suspension by counting distribution ratio of the field main color in video image of basketball game [5]. Ohno Y judge the competition process by tracking both the players and ball's trajectories according to the contrasting color features of soccer field and football players jerseys. Video motion characteristics refer to the global motion features of sports video [6]. G.S. Pingali make use of "MPEG-7 movement behaviour as descriptor" describing the motion characteristics of golf [7]. Leonardi et al. describe the video motion by tennis players and tennis video motion trajectory description [8]. S. Intille et al. montage soccer video by analysing the trajectory and the mutual movement relations of video object [9]. Prandini. M et al. test goal events in football match by sound descriptor characterization of the intensity of the sound, and combined with the camera movement information [10]. D.Q.Zhang put forward that determined the scores for the baseball game by the points of display video images. Australian research and development of DV Coach, the system can be carried out statistics on the ball tactics, and plays a very important role in the 2000 Sydney Olympic Games [11]. Italian Data Project company develop a series of sports statistics, analysis software, adopted by several national sports teams and sports clubs, it helps the volleyball, basketball and tennis ball on-the-spot technical statistics and analysis process of the Project, provided objective data of field technology and tactics in competition for coaches [12]. NBA coaches take the data mining tools which called

* Corresponding author's E-mail: iamwanghai2005@163.com

Advanced Scout from IBM auxiliary make decision about substitutions on-the-spot, achieved a good result. Basketball is representative analysis software technique, it is different to collect and analysis methods for volleyball tactics, it doesn't give each a basketball technique action coding, also does not use video search video file binding technology, but focus on collecting all kinds of basketball special technology, so as to analyze each tactics of a basketball game. The analysis statistical software of basketball technique exist problems such as weak real-time motion control, small database capacity, poor video extraction ability, low efficiency of numerous statistics in competition etc[13]. Like basketball on-the-spot technical statistics of the microcomputer application system is widely used in the current domestic CBA and CUBA leagues and for the NBA and NCAA system is done with VBA kernel abroad, does not support do complex logic with a lot of basketball game data [14].

Because of the problems such as weak optimization ability and low the efficiency in data mining of AFSA, the improved AFSA is proposed, it optimizes the initial population of traditional algorithm, behavior of fish school and search strategy.

2 The convergence analysis of AFSA

AFSA is an artificial intelligence algorithm based on fish school feeding behavior in nature, makes use of a top-down optimization mechanism, each of the artificial fish is as a packaging body. Its basic idea is to construct the coordinate the organization behavior of fish school by a simple and effective way, to realize the random behavior, foraging behavior, group behavior and train behavior of fish school, finally, to find the global optimal solution by evaluating a variety of behaviors.

Define the parameters of the artificial fish algorithm: *step* sets as the artificial fish moving step length, *visual* as visual field, *try_number* as the number of attempts by artificial fish, δ as artificial fish crowded degree factor, *n* as the total number of artificial fish. Similar to the reproduction, crossover and mutation operators of genetic algorithm, the fish algorithm including four standard behaviors:

(1) AF-Prey

See Equation (1).

$$X_j = X_i + visual \cdot rand \tag{1}$$

In equation, X_i as the current state of the artificial fish, X_j as the random position within the scope of vision, Y_i and Y_j as artificial fish food concentration (fitness function value) in X_i and X_j respectively. If $Y_i < Y_j$, in the current time *t* status X_i update for as follow. See Equation (2).

$$X_i^{t+1} = X_i^t + \frac{X_j - X_i^t}{\left\| X_j - X_i^t \right\|} \cdot step \cdot rand \tag{2}$$

Otherwise update for Equation (3).

$$X_i^{t+1} = X_i^t + visual \cdot rand \tag{3}$$

(2) AF-Swarm

X_i as the current state of the artificial fish, X_c as the centre position of artificial fish, n_f as the number of artificial fish in vision scope, *n* as the total number of artificial fish. If it meets the condition $Y_c > Y_i, \frac{n_f}{n} < \delta$, the position of the next moment artificial fish is update as follow. See Equation (4).

$$X_i^{t+1} = X_i^t + \frac{X_c - X_i^t}{\left\| X_c - X_i^t \right\|} \cdot step \cdot rand \tag{4}$$

Otherwise, carries out AF-Prey.

(3) AF-Follow

X_i as the current state of the artificial fish, X_j as the highest location of the food concentrations in visual field of artificial fish. If it meets the condition $Y_j > Y_i, \frac{n_f}{n} < \delta$, the current position of artificial fish updates as follow. See Equation (5).

$$X_i^{t+1} = X_i^t + \frac{X_j - X_i^t}{\left\| X_j - X_i^t \right\|} \cdot step \cdot rand \tag{5}$$

Otherwise, carries out AF-Prey.

(4) AF-Random

This behavior is a kind of default behavior of artificial fish, artificial fish random walk in the visual field. See Equation (6).

$$X_i^{t+1} = X_i^t \cdot visual \cdot rand \tag{6}$$

After executing the basic behavior of fish, evaluate artificial fish behavior according to the value size of the fitness function, and coordinate the next behavior of artificial fish until converge to global optimal solution finally.

The convergent algorithm is an important indicator for reflection of the algorithm performance. Each artificial fish instinctively swam to higher food concentration in AF-Prey, makes the artificial fish state to a better state, so that provides a basis to the algorithm convergence. Each artificial fish closes to the nearest the optimal state in AF-Follow, speed up the artificial fish close to the better condition of velocity, that makes artificial fish into local extremum escape from local minima in process of following, incline to global extreme value, increases the speed of convergence, and also enhances the global convergence. Each artificial fish closes to the optimal condition of partner centre position in AF-Swarm, which makes artificial fish into local extremum escape from local minima in process of following, incline to global extreme value, increases the speed of convergence, and also enhances the performance of stability and global convergence of the algorithm. Parameters is the essential factor of the effect on performed behavior of fish swarm

and then effect on the convergence performance when performing fish swarm behavior.

3 The improved AFSA

3.1 THE INITIALIZATION BASED ON THE CHAOS FISH SWARM

Logistic map is a kind of chaotic system which very simple but has been widely used, its definition is shown below. See Equation (7).

$$Z_{k+1} = \mu Z_k (1 - Z_k), Z_k \in (0,1) \quad (7)$$

In Equation, Z_k sets as real number sequence, μ sets as control parameters of system.

The idea of chaotic search introduced to the process of initialized fish swarm, use the characteristics of ergodicity of chaos movement, namely the chaotic motion not repeatedly traverse all state within a certain range according to its own "rules", closes to the optimal value areas after the chaotic movement, it can be used to improve the basic fish algorithm in the convergence speed when on a large fish swarm. The basic idea is made the chaotic map introduce to the state of chaos optimization variables, and enlarge traverse range of the chaotic motion to the scope of optimization variables, and then using chaotic variables to search optimization.

With solving the maximum of complex functions for an example, the mathematical model of optimization problems described as follow. See Equation (8).

$$\max f(x) = f(x_1, x_2, ..., x_n), x_i \in [a_i, b_i], i = 1, 2, ..., n \quad (8)$$

The steps of the application of chaos initialization in the AFSA are as follows:

(1) The known biggest number of iterations chaotic variables as M , we can get the corresponding initial chaotic sequence $Z^0 = (Z_1^0, Z_2^0, ..., Z_n^0)$, the scope of the sequence Z^0 is $(0,1)$, but need to skip the fixed points of Logistic map.

(2) We can obtain $x^0 = (x_1^0, x_2^0, ..., x_n^0)$ according to $x_i^0 = a_i + (b_i - a_i) \times Z_i^0$, and calculate $f^0 = f(x^0)$, at the same time make the number of iterations as $k = 0$.

(3) According to Logistic map $Z_{k+1} = \mu Z_k (1 - Z_k)$, the chaotic variables sequence Z_i^{k+1} obtained by constant iterative method. Then according to equation $x_i^{k+1} = a_i + (b_i - a_i) \times Z_i^{k+1}$, we calculate the function variable sequence x_i^{k+1} and the adaptive value of function $f(x_i^{k+1})$.

(4) When $k < M$, if $f(x^{k+1}) \geq f^0$, order $x^0 = x^{k+1}$ and $f^0 = f(x^{k+1})$. When $f(x^{k+1}) < f^0$, x^0 and f^0 stay the same, order $k = k + 1$ and $Z^k = Z^{k+1}$, go to step (2), chaotic search to stop until reach the preset maximum cycle times .

(5) Get a new chaotic sequence x^0 when the search is completed, it is a new initial state of artificial fish obtained by the initialization of chaos; the artificial fish's position has been close to the optimal area.

Making use of the position of the fish swarm by initializing idea chaos, have not changed essential randomness in the initialization of AFSA, and enhance the performance of search ergodicity, improve the convergence speed of basic fish algorithm in large fish scale. After the initialization of chaos fish, make the artificial fish swarm optimize preliminary in a large scope of optimization, improve the convergence performance of the basic algorithm.

3.2 THE OPTIMIZING BEHAVIOR OF THE ARTIFICIAL FISH SWARM

In the basic of AFSA, N sets as the number of artificial fish in search domain, No. i artificial fish as a feasible solution vector X_i , $X_i = (x_1^i, x_2^i, ..., x_D^i)$. Then $X_i(t)$ sets as the current state of artificial fish, $X_i(t+1)$ as the next state of artificial fish, each iteration artificial fish according to the existing state choose from four kinds of behavior to perform, then update the ego state, add a vector $\Delta X_i(t+1)$ based on the original state $X_i(t)$, so the update formula of location of artificial fish in the AFSA can describe following formula. See Equation (9) and (10).

$$\Delta X_i(t+1) = rand() \cdot step \cdot [X_{next}(t+1) - X_i(t)] \quad (9)$$

$$X_i(t+1) = X_i(t) + \Delta X_i(t+1) \quad (10)$$

In equation, $rand()$ sets as a random number between 0 and 1, the value difference of the function value is smaller than a preset value eps , and then randomly selected some artificial fish, these random set the parameters of the artificial fish, expression as shown below. See Equation (16).

$$X_i(t+1) = X_i(t) + \beta \cdot visual \cdot rand() \quad (16)$$

In equation, β can be a parameter or a mutation function.

3.3 THE SEARCH ADJUSTMENT BASED ON THE ADAPTIVE STRATEGY

In the standard AFSA, adjust the relationship between the global search and local search ability is the key to solve the problem, in order to make the adjustment more flexible and effective, the adaptive strategy introduced for the optimization. First, define the optimal rate of change K and the variance changes σ as follows. See Equation (17) and (18).

$$k = \frac{f(t) - f(t-n)}{f(t-n)} \quad (17)$$

$$\sigma = D(f(t), f(t-n), f(t-2n)) \quad (18)$$

$f(t)$ sets as the optimal adaptive value when populations in t generation, $f(t-n)$ as the optimal adaptive value in $(t-n)$ generation, $f(t-2n))$ as the optimal adaptive value in $(t-2n)$ generation, and then K as the rate of relative change of the optimal adaptive value in n generation, D as change variance in $2n$ generation.

Then use the game theory as measure of whether the parameters change, make the algorithm adaptive adjustment of parameters in the process of operation, specific implementation is as follows. See Equation (19).

$$\begin{cases} step = f(step), \delta = f(\delta) & K \le \theta, \sigma \le \phi \\ step = step, \delta = \delta & K > \theta, \sigma > \phi \end{cases} \tag{19}$$

$f(step)$ sets as step length adjusting according to fish algorithm performance, generally the principle is took little to big, because according to the above section optimization theory, the preliminary big step in favour of finding the optimal value, reduce late step can improve search precision, prevent local convergence; $f(\delta)$ sets as crowded degree factor by adjusted accordingly, θ, ϕ as evaluation coefficient, adjust according to the different problems.

4 The performance simulation of algorithm

In order to verify the effectiveness of the improved algorithm proposed in this paper, simulation experiments on it. First test the improved AFSA performance by the following typical function, test function is as follows. See Equation (20).

$$F = \frac{\sin^2 \sqrt{x^2 + y^2} - 0.5}{(1 + 0.001(x^2 + y^2))} - 0.5 \tag{20}$$

The maximum number of iterations is 100 in the algorithm, $N = 10$, $T = 60$, $\delta = 0.618$, $visual = 8$, $step = 0.6$, the optimization curve of AFSA in the test function as follows. Please see Figure 1, Figure 2 and Figure 3.

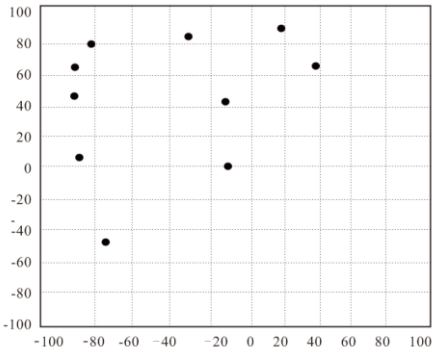

FIGURE 1. The initial state distribution of artificial fish swarm

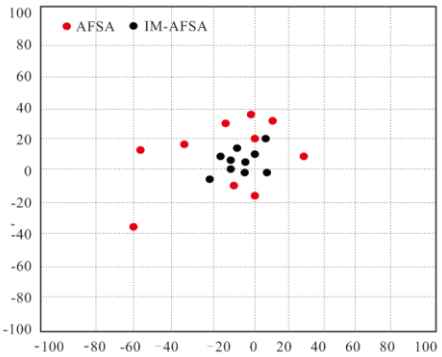

FIGURE 2. After 50 iterations of state distribution of artificial fish

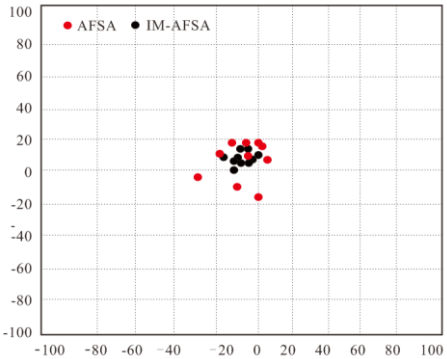

FIGURE 3. After 100 iterations of state distribution of artificial fish

Then the improved AFSA apply to video data mining in the table tennis competition, the result please See Table 1.

TABLE 1. The results of data mining algorithm based on improved artificial fish

Iteration time	Number of attributes		Type Number	
	AFSA	IM-AFSA	AFSA	IM-AFSA
100	22	25	3	5
120	34	58	5	8
140	41	74	8	11
160	46	92	9	17
180	52	114	12	21
200	66	128	13	25

Seen from the above simulation results, the searching capability of the improved AFSA is greatly better than the traditional algorithm, and has well application in the video data mining of the table tennis match, it can dig more game data.

5 Conclusions

With the continuous development of sports information, sports field has accumulated a amount of data, data mining is applied in sports field analysis especially antagonistic competitive sports analysis will become a trend. A video data mining model of table tennis match is proposed based on the improved AFSA, seen from the simulation results, the improved algorithm has better optimization ability, and can dig out more rules of classification.

With the advancement in networking and multimedia technologies enables the distribution and sharing of multimedia content widely. In the meantime, piracy becomes increasingly rampant as the customers can easily

duplicate and redistribute the received multimedia content to a large audience.

References

[1] Xu Yun. (2012) Application of Data Mining in Sports. *Journal of Wuhan Institute of Physical Education*, **46**(11), 27-30.

[2] Zhao HuiQun. (2013) A StudY of Data Mining Technology Applied on Technique and Tactic Analysis of Sport ComPetitions. *Journal of Beijing University of Physical Education*, **31**(5), 712-715.

[3] LiHuiling.(2013) The Application of Data Warehouse and Data Mining Technology in College Gym Data Analysis. *Journal of Guangzhou Physical Education Institute*, **25**(5), 126-128.

[4] Mao Jie. (2013) Application of data warehouse in competitive sports. *Journal of Wuhan Institute of Physical Education*, **37**(3), 175-176.

[5] Mao Jie. (2012) Basketball Technology Statistical Software Development. *Journal of Wuhan Institute of Physical Education*, **46**(2), 70-73.

[6] Ma Donglai. (2014) Improved method of apriori algorithm based on data attributes. *Chinese Agricul Tural Mechanization*, **35**(4), 253-256.

[7] Yu Lijuan.(2013)Systematic Study and Application of Analysis of Technique and Tactics in Table Tennis Matches. *Journal of Shanghai Physical Education Institute*, **32**(6), 39-43.

[8] Shan Donghong. (2014) Application Research of Data Mining Technology in the Internet Information Retrieval. *Bulletin of Science and Technology*, **30**(3), 161-164.

[9] NING Guangfang. (2014) Quality Parameters of Mass Movement Data Mining Method. *Computer Simulation*, **31**(6), 398-401.

[10] Zhao Yangqing. (2011) Application of Sequential Model Digging-Out to Analysis of Table Tennis Matches Techniques and Tactics. *Journal of Shanghai Physical Education Institute*, **32**(2), 83-85.

[11]Zhao Zhigang. (2014) A Probability-weighted Association Rules Mining Algorithm Based on Vector. *Computer Engineering & Science*, **36**(2),354-358.

[12] DING Yazhi. (2014) Improved SPRINT Algorithm and Its Application in the Physical Data Analysis. *China Sport Science*, **34**(6), 90-96.

[13] Ji Gremlin. (2014) A Survey of Spatiotemporal Data Mining for Big Data. *Journal of Nanjing Normal University（Natural Science Edition)*,**37**(1),1-7.

[14]Meng Haidong. (2014) Research And Implementation of Quantitative Association Rules Mining Based on Data field. *Computer Applications and Software*, **31**(7),40-42.

Authors

Hai Wang, 04.04,1977, China

Current position, grades:lecturer at Department of Physical Education, Shanghai Business School, Shanghai,
University studies: a master degree education. graduated from the physical education and sports training specialty of Shanghai University of Sport， in 2008
Scientific interest: sports teaching and training.

Wenbo Qu, 16.08.1965, China

Current position, grades: Professor, School of Humanities and Law, Shanghai Business School, Shanghai
University studies: master's degree, doctor of Science degree from Harbin Institute of Technology China in 2002
Scientific interest: Infinite matrix theory, data mining, quantum information

Qunli Shen, 13.05.1973, China

Current position, grades: associate Professor, College of information and computer, Shanghai Business School, Shanghai
University studies: Dr. Tongji University
Scientific interest: Data mining, information management and information system of computer.