The Prediction Model Study of Basketball Simulation Competition Result Based on Homogeneous Markov

Yonggan Wang

Institute of physical education, Langfang Teacher's College, Langfang, China

w_angyonggan@126.com

Keywords: basketball; homogeneous markov; prediction; simulation

Abstract. Basketball game is one of the most popular world sports in which the largest number of people participate, especially NBA has spread all over the world. The people enjoy the results of the basketball competition process, but they will pay more attention to the competition results. Therefore, it is extremely important to simulate the competition process and even predict the competition results. In the paper, the study was carried on by two aspects. On the one hand, the current situation of a basketball team is analyzed and the development of the basketball team is predicted by data mining technology. Then competition result is predicted on the base of the analyzation and prediction; On the other hand, the homogeneous Markov chain is introduced into the basketball game to obtain the data of the main factors by establishing the state transition probability matrix. Meanwhile, the future development of the competition is predicted by the homogeneous Markov model and the reasonable conclusion is made, that the state of a good basketball is stable.

Introduction

Basketball was invented by American James Naismith in 1891. The basketball game is simple in the beginning, and there is no limitation in the size of the venue and the number of the people participating in the game. In 1892, the basketball game was introduced into Mexico and then spread, popularized and developed all over the world. In 1932, the men's basketball entered the Olympic Games officially. In 1946, the U.S. professional basketball league appeared and became the NBA now. The basketball could not only improve the vitality, but also promote the personality development. At present, the basketball game has become one part of the people's life. People can enjoy themselves from the game, but they will pay more attention to the competition result. Thus, it is important to make the prediction on the basketball competition result scientifically and objectively.

In the paper, the scientific prediction and relative research are made in terms of data mining and homogeneous Markov model.

Data Mining

In the era of information, data mining technology has been used widely. It could not only obtain data from the practice but also deal with these data. Thus, the important information could be found. In this paper, the competition result is regarded as the research object, the competition process is analyzed by using the technology and the competition result is predicted. Meanwhile, the factors which affected the competition result are obtained. Thus, we can improve the competition process to attain the aims of improving the performance.

As for the basketball game, the routine training and the preparation before the competition will affect the final result. The basketball game is a team sport, and the actual performance of each player will also affect the final result. A certain factor may affect this player, but may not affect that player. Each player should have his own differences. Thus, we could not predict the result with only one model. According to different situation in the competition, each player will have the different performances. Finally, we can acquire many factors affecting the results of the competition, and their relationship is complex. Therefore, the data we required are obscure and subjective. The principle of relevance was adopted in the paper.

Because each match has a lot of data and the data is timeliness, the data in the beginning could not be adopted. The main factors are adopted in the paper when establishing the prediction model.

$$v = b_1 v_1 + b_2 v_2,\ \ b_1 + b_2 = 1$$

Among them, $v_1 = E(I)$ is team participating in the contest I 's mathematical expectation, v_2 is the predicted result. b_1 and b_2 should be adjusted according to the prediction results, as follows:

Let V be the prediction result, the actual result of the team participating in the contest I is T. When $|V-T| > \sigma$, b_1 and b_2 should re-value according to the equations $\begin{cases} b_1 + b_2 = 1 \\ b_1 v_1 + b_2 v_2 = T \end{cases}$.Making use of the rank in the augmented matrix,

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & v_1 - v_2 & T - v_1 \end{pmatrix}$$

When $v_1 - v_2 \neq 0$, b_1 and b_2 is obtained by solving the equations. When $v_1 - v_2 = 0$ and $T - v_1 = 0$, the equations have infinitely many solutions and b_1 and b_2 need not to be adjusted; When $v_1 - v_2 = 0$ and $T - v_1 \neq 0$, the equations have not the solutions. At this time, $\varepsilon = \frac{2}{5}(T - v_1)$. The neural network model was utilized to deal with the data. When L is in middle level, $v_2 > T$, which indicates that v_2 could be used and the error of v_1 is relatively great. b_1 is magnified and b_2 is reduced, let $v_1' = T + \varepsilon L$ instead v_1, the equations $\begin{cases} b_1 + b_2 = 1 \\ b_1 v_1' + b_2 v_2 = T \end{cases}$ is obtained, b_1 and b_2 can be solved.

When L is in low level, $v_2 < T$, which indicates that v_1 could be used and the error of v_2 is relatively great. b_2 is magnified and b_1 is reduced, let $v_2' = T + \varepsilon L$ instead v_2, the equations $\begin{cases} b_1 + b_2 = 1 \\ b_1 v_1 + b_2 v_2' = T \end{cases}$ is obtained, b_1 and b_2 can be solved.

If a value in the b_1 and b_2 keep at the minimum level, it indicates that the prediction results are inaccurate. b_1 also indicates that the state of the basketball player is at the peak or in the low ebb.

To sum up this article, the prediction result of a team in the next game has little difference with the recent result, which indicates that $E(I)$ is stable. So it is scientific to regard $E(I)$ as the factor of prediction results.

Homogeneous Markov Model

Markov is a famous Russian mathematician, and Markov process is a special method which describes the development process of things.

As we all know, the development of the things will change as time changes. As for the development of the things, the state of the past and the present need to be investigated comprehensively to predict the future. However, as for the development of the things, if you know the current state, you can predict the future state. For example, when we play the Chinese chess, how to move the chessman next step is related to the current position, but there is no need to know where it was. It is unrelated to the value, which was called the ineffectiveness theory. The development process of the things of ineffectiveness theory is called the Markov process.

Assume that the number of state of the variable is n, P_{ij} is one-step transition probability which transfers from i to j by one step transfer. The matrix which is arranged and constituted is called the transition probability matrix:

$$P(n) = \begin{bmatrix} P_{11}(n) & P_{12}(n) & \cdots & P_{1n}(n) \\ P_{21}(n) & P_{22}(n) & \cdots & P_{2n}(n) \\ \cdots & \cdots & \cdots & \cdots \\ P_{n1}(n) & P_{n2}(n) & \cdots & P_{nn}(n) \end{bmatrix}$$

Multi-step transition probability is spread on the base of the one-step transition probability. If the system is in the state i at time t_0, the system is in the state j at time t_n after transfer of n step. The quantitative index of transfer probability is called the n-step transition probability and marked $P(x_n = j \mid x_0 = i) = P_{ij}(n)$. n-step transition probability matrix:

$$P(n) = \begin{bmatrix} P_{11}(n) & P_{12}(n) & \cdots & P_{1n}(n) \\ P_{21}(n) & P_{22}(n) & \cdots & P_{2n}(n) \\ \cdots & \cdots & \cdots & \cdots \\ P_{n1}(n) & P_{n2}(n) & \cdots & P_{nn}(n) \end{bmatrix}$$

Markov chain is an important stochastic process, and its state space is limited and infinite. After some time, the system will transform from one state to another state. This process only relies on the current state, but not relate to the history.

$\{X(n),\ n = 0,1,2,\cdots\}$ is a stochastic process, $E = \{0,1,2,\cdots\}$ is a state space. As for an arbitrary group of integer time $0 \le n_1 < n_2 < \cdots < n_k$ and arbitrary state $i_1, i_2, \cdots, i_k \in E$, conditional probability exists.

$$P\{X(n_k) = i_k \mid X(n_1) = i_1, X(n_2) = i_2, \cdots, X(n_{k-1}) = i_{k-1}\} = P\{X(n_k) = i_k \mid X(n_{k-1}) = i_{k-1}\}$$

Namely, the process $\{X(n),\ n = 0,1,2,\cdots\}$ future state is related to the current state, and is unrelated to the past state, $\{X(n),\ n = 0,1,2,\cdots\}$ is a Markov chain of discrete time parameter.

$p_{ij}(m,k) = P\{X(m+k) = j \mid X(m) = i\},\ i,j \in E$ is called the k-step transitional probability of Markov chain. If the k-step transitional probability is only related to k, is unrelated to time start. $\{X(n)\}$ is called the homogeneous Markov chain of discrete time. k-step transition probability matrix is:

$$P(m,k) = \begin{bmatrix} P_{00}(m,k) & P_{01}(m,k & \cdots & P_{0n}(m,k) & \cdots \\ P_{10}(m,k) & P_{11}(m,k) & \cdots & P_{1n}(m,k) & \cdots \\ \vdots & \vdots & & \vdots & \\ P_{j0}(m,k) & P_{j1}(m,k) & \cdots & P_{jn}(m,k) & \\ \vdots & \vdots & & \vdots & \end{bmatrix}$$

$\{X(t), t \ge 0\}$ is a continuous time parameter stochastic process, $E = \{0,1,2,\cdots\}$ is a state space. If this arbitrary n, $0 < t_1 < t_2 < \cdots < t_n < t_{n+1}$ and $i_1, i_2, \cdots, i_n, i_{n+1} \in E$,

$$P\{X(t_{n+1}) = i_{n+1} \mid X(t_k) = i_k,\ k = 1,2,\cdots,n\}$$
$$= P\{X(t_{n+1}) = i_{n+1} \mid X(t_n) = i_n\}$$

Then, $\{X(t), t \ge 0\}$ is called the Markov chain of continuous time parameter.

$p_{ij}(s,t) = P\{X(s+t) = j \mid X(s) = i\},\ i,j \in E$ is called the transition probability function. If the transition probability function is only related to the time interval t, and is unrelated to the starting point of the time s, then $\{X(t), t \ge 0\}$ is called the homogeneous Markov chain of continuous time parameter.

In general, we require transition probability function to satisfy the following continuity condition:

$$\lim_{t \to 0^+} p_{ij}(0,t) = \delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \ne j \end{cases}$$

It is well known that Markov theory has found wide applications in the prediction field. In this paper, the homogeneous Markov chain is introduced into the basketball game, and the data of main factors are obtained by establishing the state transition probability matrix. Meanwhile, the homogeneous Markov model is used to predict the future trends and get the competition results.

Let B_0 be the final competition result of a basketball team in the preseason, $B_i (i = 1,\cdots,k)$ is the competition result in the regular season. P is one-step transition probability matrix, $P(k)$ is a k-step transition probability matrix, then $P(k) = P^k$, that is,

$$P(k) = P^k = \begin{bmatrix} P_{11} & P_{12} & \cdots & P_{1n} \\ P_{21} & P_{22} & \cdots & P_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ P_{n1} & P_{n2} & \cdots & P_{nn} \end{bmatrix}^k$$

$P(k) = P^k = P(k-1)P$ is obtained from the formula. Because B_0 is known, the transition probability p_{ij} from B_0 to B_1 is obtained, thus the prediction results of the first formal race $B_1 = B_0 P$ is obtained. Similarly, the prediction results of $B_i (i = 2, \cdots, k)$ is obtained, that is

$$B_2 = B_1 P = B_0 P^2 , \quad B_3 = B_2 P = B_0 P^3$$

$$\cdots \cdots \quad B_k = B_{k-1} P = B_0 P^k$$

From the homogeneous Markov Model, P is non-diagonal matrix, then the system which is stable is marked B, that is, $B = \lim_{k \to \infty} B_k$. Then, $B_k = B_0 P^k$ is substitute to B,

$$B = \lim_{k \to \infty} B_k = \lim_{k \to \infty} B_0 P^k = B_0 P^\infty$$

Let $P^\infty = \begin{bmatrix} b & 1-b \\ b & 1-b \end{bmatrix}$, $0 \le b \le 1$, $b = 0$ means losing the game, and $b = 1$ means winning the game,

$$B = B_0 P^\infty = (y_1 \quad y_2) \begin{bmatrix} b & 1-b \\ b & 1-b \end{bmatrix}$$

$$= (y_1 b + y_2 b \quad y_1(1-b) + y_2(1-b))$$

$$= (b(y_1 + y_2) \quad (1-b)(y_1 + y_2))$$

$$= (b \quad 1-b)$$

The formula shows that as long as P is non-diagonal matrix, the final result of the regular season is not affected whether win or lose the final competition in the preseason.

Conclusions

Basketball is one of the most popular sports in the world, and people pay more attention to the win or loss of the competition. Thus, predicting the competition results is very important, which is helpful to improve the record of the basketball team. In this paper, the current situation of a team sport is analyzed and the development tendency of the team is predicted by data mining and the homogeneous Markov model. On that basis, the competition process is simulated, the competition result is predicted and the scientific and reasonable result is obtained.

References

[1] Yang Weiguo, Ji Lingjun, Sun Jie. Mixing property of mth order hidden non homogeneous Markov models[J]. Journal of Jiang Su University, 2009 (30) 6: 645-648.

[2] Han Le. Research on the application of the modified non-homogeneous Markov model. [J] Bioinformatics, 2004,2:27-28

[3] Wu Ji. On the Strategy of Simulation Training in College Basketball Training. [J] Journal of Sports Science2011 (30) 2: 209-210.

[4] Wu Shuangwen. Influences of basketball rule transformations on technique and tactics. [J] Journal of Wu Han Institute of physical Education, 2011 (45) 1: 72-75.

[5] Zhao Weihua, Shu Jian. Application of Matrix Rank in Judging the Ergodicity of Homogeneous Markov Chains. [J] Journal of Nan Tong University2009 (8) 1: 80-82.

[6] Yu Shaoyong, Han Jian, Xie Weiping. Measurement of Athlete on Teamer's Trust and the Relation with Relevant Variables for Basketball Sports. [J] Sports Science2009 (29) 2: 30-38.

[7] Zhou Rui, Xu Renzuo, Yang Xiaoqing. A Non-Homogeneous Markov Model With Fault Repair. [J]Mini-macro Systems, 2000 (21) 3: 242-245.

[8] Yu Shaoyong. Relation among Team Trust, Athlete Satisfaction and Team Cohesion of Basketball Players. [J]Sport Science, 2010 (30) 6: 85-91.

[9] Ran Jie, Liu Limei. The Application of Homogeneous Markov Chain and Its Realization Based on Matlab. [J]Computer Learning, 2009, 2: 19-21.