Acta Technica 62 (2017), No. 6B, 547–558

Superiority selection method based on nonparametric Bayesian algorithm

Peng Dai[1]

Abstract. This paper proposes a kind of selection method of basketball skill management advantages for players based on non-parametric Bayesian algorithm to improve the accuracy of basketball skill management algorithm for players. Firstly, according to the unknown model number problem in discovery process of basketball skill management data partitioning and hiding model, the data partitioning and model discovery are conducted by using hidden Markov algorithm in Dirichlet process and non-parametric Bayesian factorial analysis; secondly, the basketball skill management model is predicted and analyzed based on non-parametric Bayesian algorithm, achieving effective promotion of the basketball skill management level.

Key words. Non-parametric, Bayesian algorithm, Basketball skill, Management.

1. Introduction

In March, 2013, Yao Ming, newly elected as member of CPPCC, submitted a proposal, named Reforming the Management System, Opening the Talents Selection and Promoting Sustainable Development of Chinese Basketball. Such proposal has same goal with this paper. In proposal, Yao Ming sharply points out that three problems exist in reserve force cultivation in Chinese basketball: single and closed talent cultivation channel-the excessively short talent cultivation period violates objective law of basketball talent cultivation; the professional basketball, namely occupational basketball, isolates from school basketball and social basketball, the basketball talents are faced with exhaustion. After respectively analyzing these problems, Yao Ming puts forwards some suggestions: reforming Chinese basketball management system, establishing open talents selection platform, combining sports with education and holding students basketball leagues with new concepts. In fact, the most important and critical part in suggestions is "Chinese basketball management system reform", which actually suggests that Chinese basketball overcomes the defects of "the whole nation system" for sports development dominated by government, develops towards

[1]College of physical education, Huainan Normal University, Anhui, Huainan, 232038

essence of competitive sports and becomes professional and market-oriented, making the Chinese basketball at forefront of professionalism and marketization faced with the comprehensive market-oriented reform impacts. These suggests the current requirements of Chinese reform and opening-up and spanning development and conform to the strategic conception "Chinese Dream" proposed by General Secretary Xi Jinping.

There is no doubt that the leading force of sports development in China is the government. Lu Yuanzhen, the professor in South China Normal University, points out that the competitive sports system in China can be summarized as "government-hold", "Olympic strategy", "professional team" and "the national games", the government-hold sports are the basic property of this system. The government is not only the decision maker and organizer, but also the investor and operator. Therefore, the government is also beneficiary and risk bearer. The essence of competitive sports is game. But the competitive sports are greatly politicized in China, which causes special phenomenon of Chinese sports management system in world sports development. According to the analysis by professor Lu, such system is beneficial to unified resource allocation. But at the same time, it is not beneficial to market initiative playing, social organization growth and development and not beneficial for society to share competitive sports responsibilities, thus restricting the diversity of competitive sports resource source and sustainable development and limiting continuous development scale expansion. According to such system design, many senior managers from sports circles propose different development ideas. Some hold that market-oriented system and the whole nation system can be effectively integrated. For example, Li Lingwei, member of Committee of the National People's Congress and Deputy Director of Network Management Center, said in NPC and CPPCC, "the promotion of "Li Na mode" (on the basis of international marketization operation mode) "the whole nation system" is feasible in China".

According to the basketball skill management advantages selection in non physical education major recruitment and the unknown model number problem in discovery process of basketball skill management data partitioning and hiding model, this paper conducts the data partitioning and model discovery by using hidden Markov algorithm in Dirichlet process and non-parametric Bayesian factorial analysis, predicts and analyzes the basketball skill management model based on non-parametric Bayesian algorithm.

2. Basketball skill management model

The experimental data sets from the students in the physical major of Class 3, Grade 2003 in an American university are selected as test objects. Firstly, the stationary shot with single hand over the shoulder is selected as example and the key parameter acquisition process for skill actions are clearly indicated. Firstly, set 7 links for the stationary shot with single hand over the shoulder based on action accomplishment process, including ready stance, take-off leg action, force use in waist and abdomen, elbow raise, loading arm, wrist bending and finger stirring. Then ask the students wearing force sensor to take shooting exercise, collect image information by using three CCD cameras, use MATTAB programming environment to process data based on image data to obtain relevant information, such as position, range, strength, bending and rotation degree of arms, legs and trunks etc. Then compare the actions of different exercisers with standard actions, determine the merit ranking index for 7 links mentioned above, give corresponding scores with 5, 4, 3, 2, 1 based on the merit ranking of actions (excellent, good, fair, poor, bad) and obtain the skilled action parameters for 20 exercisers.

According to the establishment principle of Bayesian reasoning rules, select the key parameter values of skilled action from different exercisers as the sample vectors. The stationary shot with single hand over the shoulder includes 7 links. Therefore, the nerve cell in input layer is set as 7. The number of nerve cell in competitive layer is established based on type of error action in skilled action exercises or evaluation level. The following error actions may exist in stationary shot exercise with single hand over the shoulder: A. unbalanced force use in waist and abdomen, stiff action; B. incorrect finger stirring, non-rotation or lateral rotation of basketball and incorrect force use speed; C. incorrect wrist bending and finger stirring, non-rotation or lateral rotation of basketball; D. unreasonable body stretching in shooting, insufficient wrist bending and finger stirring, non-rotation of basketball; E. no apparent errors; F. excessively outward stretching of elbow, no elbow raising forwards and insufficient angle of put; G. wrong ready stance or holding. In view of the speed and stability of Bayesian rule training, the learning rate is set as 0.1.

After establishing the Bayesian rules, invoke function to initialize rule weight and threshold, then take exercises and tests. Bayesian rules put forwards normalization requirements to sample data. Therefore, 20 students from basketball major are selected as research objects. The skilled action parameters are normalized and its formula is shown as below:

$$X = X_i - X_{\min}/ X_{\max} - X_{\min},. \qquad (1)$$

Where, Xrefers to the normalization coefficient of "i" index data; $X_{\max}$ and $X_{\min}$ respectively refer to the maximum value and minimum value in X_i index group. Then, the normalized sample data should be input in Bayesian rules for simulation training. When the training times reach at 800, Bayesian rules will classify different exercisers and obtain the results (as shown in Table 1). The skilled action level evaluation results (as shown in Table 2) for different exercisers can be obtained at the same time.

Table 1. Testing results of students' shooting skill error actions through bayesian rules

Number of exerciser	Type of error
1,6,12,16	Unbalanced force use in waist and abdomen, stiff action
11,17,19	Incorrect finger stirring, non-rotation or lateral rotation of basketball and incorrect force use speed
3,7,9,13	Incorrect wrist bending and finger stirring, non-rotation or lateral rotation of basketball
5,10	Unreasonable body stretching in shooting, insufficient wrist bending and finger stirring, non-rotation of basketball
14,18,20	No apparent errors
4,8	Excessively outward stretching of elbow, no elbow raising forwards and insufficient angle of shot
2,5	Wrong ready stance or holding

Table 2. Comparison of bayesian rules evaluation results with teacher evaluation results for different exercisers' action level

Type	Excellent	Good	Fair	Poor	Bad
Bayesian rules evaluation results	8 4 18 20	3 7 9 13	1 6 12 16	4 5 10 15	2 11 17 19
Teacher evaluation results	14 18 20	3 8 7 9 13	1 6 12 16	4 10 15	2 5 11 17 19

3. Algorithm theory description

3.1. Non-parametric Bayesian algorithm

HMM requires pre-assigning state number and estimate parameters by EM algorithm. For this purpose, the hidden Markov algorithm based on hierarchy Dirichlet process (HDP-HMM) is proposed, achieving automatically deducing state number from data. Such model can be interpreted as grouping number dynamic upgrading process using hierarchy Dirichlet process based on state assignment. The infinite Markov model partition can be made by combining Gaussian state with active level model. Once the active level of partition activity time is obtained, the separate analyzer can be operated for each model.

Dirichlet process $DP(\gamma, H)$ is distribution of discrete random probability measure G on $(\mathcal{Q}, \mathcal{B})$. The alternative definition is given here [13∼13]:

$$G = \sum_{k=1}^{\infty} \beta_k \delta_{\varphi_k} \,. \tag{2}$$

$$\varphi_k \overset{iid}{\sim} H, k = 1, \cdots, \infty \,. \tag{3}$$

$$\beta = (\beta_k)_{k=1}^{\infty} . \tag{4}$$

Where, β refers to weight constructed through "connection and fracture" process, simply denoted as$\beta \sim GEM(\gamma)$. Dirichlet process has been widely used in prior distribution mixing measure of mixed Bayesian model, typically the mixed model of Dirichlet process [14].

3.2. HDP-HMM data stream partition

During HDP process, the method replacing state number with Dirichlet prior number is proposed. Therefore, the unknown state number in HMM [15] can be defined by reference to HDP method. If HDP is used as non-parametric building block, HDP-HMM random process can be described as:

$$\begin{cases} G_0 \sim DP(\gamma, H \times S), \theta_t \overset{iid}{\sim} G_k \, , \\ G_k \overset{iid}{\sim} DP(\alpha, G_0), y_t \sim F(\theta_{t-1}) \, . \end{cases} \tag{5}$$

Where, $k = 1, 2, \cdots \infty$, $t = 1, 2, \cdots, T$. There are Ttime steps in total. Tcan be defined as data hours collected in a day. HDP-HMM is shown in Fig. 1.

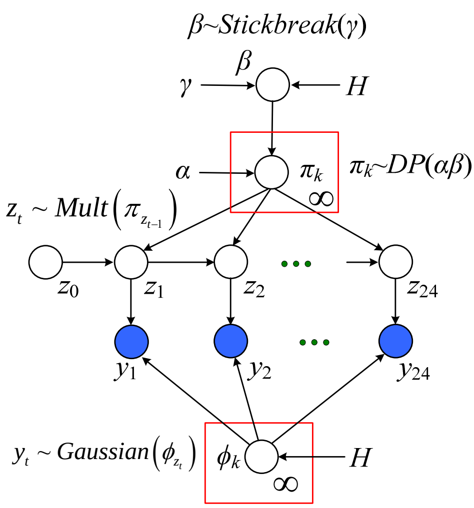


Fig. 1. HDP-HMM model representation

The relevant parameters in Fig. 1 should meet the following distribution requirements:

$$\begin{cases} \beta \sim GEM(\gamma), \pi_k \sim DP(\alpha, \beta) \, , \\ \varphi_k \sim H, Z_t \sim \pi_{Z_{t-1}}, y_t \sim F(\varphi_{Z_t}) \, . \end{cases} \tag{6}$$

3.3. HDP-HMM process reasoning

HMM shall be utilized firstly to classify data as coherent slice at the first stage, and time dynamicity of data can be considered sufficiently. Number of coherent slice of video is unknown, so it shall be estimated. The primary goal is to implement rough data partitioning with hour as interval, and 34 input data points can be gained, and variable $\{y_t\}$ and $\{z_t\}$ observed correspondingly play a potential state variable role in standard HMM. H is basic measure being able to sample parameter$\{\varphi_k\}$. In model proposed, y_t is modeled as univariate Gaussian process, and φ_k is a tuple $\left\{\mu_k, \sigma_k^2\right\}$, of which μ_k and σ_k^2 are unknown and considered as random variables. H is used as conjugate prior, and H will meet Gaussian invGamma distribution in the case.

Gibbs reasoning is adopted in HMM, and potential state z_t and weight β_k are sampled sequentially, and emission probability and transition probability are gained respectively through explicit integration of parameter $\{\varphi_k\}$ and $\{\pi_k\}$. For example, given $z_{t-1} = i$ and $z_{t+1} = j$ according to previous iteration, conditional gibbs's distribution form of z_t sampling is:

(1) In z_t sampling, and conditional probability of z_t is considered as:

$$\begin{aligned} &p(z_t = k|z_{-t}, y, \beta, H) \propto \\ &p(y_t|z_t = k, z_{-t}, y_{-t}, H) \times p(z_t = k|z_{-t}, \alpha, \beta) \end{aligned} \tag{7}$$

In the formula, $p(y_t|z_t = k, z_{-t}, y_{-t}, H)$ is likelihood value of observed value y_t, and the value can be expressed through other forms:

$$\int_{\varphi_k} p(y_t|z_t = k, \varphi_k)p(\varphi_k|y_{-t}, z_{-t}, H)d\varphi_k, . \tag{8}$$

Formula (7) can be based on conjugation property analysis. In the formula (7), order term $p(z_t = k|z_{-t}, \alpha, \beta)$ is transient process. Assumed that n_{ij} is number of transition from state i to$j$ and n_{*j} is number of all states transited to state j. Similarly, it can be defined that n_{i*} is process number of transition leaving from state i, and then Markov feature of transient process can be described as:

$$\begin{aligned} &p(z_t = k|z_{-t}, \alpha, \beta) \propto \\ &p(z_t = k|z_{t-1}, \alpha, \beta) \times p(z_t = k|z_{t+1}, \alpha, \beta) \end{aligned} \tag{9}$$

Then above probability can be calculated on the basis of 4 kinds of condition:

$$\begin{aligned} &p(z_t = k|z_{-t}, \alpha, \beta) \propto \\ &\begin{cases} (n_{z_{t-1},k} + \alpha\beta_k)\dfrac{n_{k,z_{t+1}} + \alpha\beta_{z_{t+1}}}{n_{k*} + \alpha}, k \le K, k \ne z_{t-1} \\ (n_{z_{t-1},k} + \alpha\beta_k)\dfrac{n_{k,z_{t+1}} + 1 + \alpha\beta_{z_{t+1}}}{n_{k*} + 1 + \alpha}, z_{t-1} = k = z_{t+1} \\ (n_{z_{t-1},k} + \alpha\beta_k)\dfrac{n_{k,z_{t+1}} + \alpha\beta_{z_{t+1}}}{n_{k*} + 1 + \alpha}, z_{t-1} = k \ne z_{t+1} \\ \alpha\beta_{new}\beta_{Z_t+1}, k = K + 1 \end{cases} \end{aligned} \tag{10}$$

(3) Computational process of weight parameterβ and hyper-parameter α and γ is the same with that of HDP algorithm, and see literature [16] for details. To improve robustness of algorithm, assumed that hyper-parameter α and γ follow Gamma distribution, and are sampled again in Gibbs iteration of every time.

Analytic algorithm can be described as: assumed that $X \in \mathcal{R}^{d\times n}$ has n feature vectors of d-dimensional center; C is covariance matrix and SVD is decomposed into:

$$C = U \sum U^T . \tag{11}$$

Feature vector sourced from U is divided into 2 groups:

$$C = [U_1, U_2] \begin{bmatrix} \sum_1 & 0 \\ 0 & \sum_2 \end{bmatrix} U^T . \tag{12}$$

Where, selection of $\sum_1$ and $\sum_2$ shall follow:

$$tr\left(\sum_1\right) \bigg/ \left(tr\left(\sum_1\right) + tr\left(\sum_2\right)\right) = 0.9 . \tag{13}$$

That is to say that the most important feature vector shall be chosen to make total energy coverage rate be 90%. U_1 is called as principal subspace while U_2 is called as residual subspace. Analytic algorithm compares test vector and detection threshold value (λ) by projecting test vector into residual subspace U_2, and it is also function of non-principal characteristic value in residual subspace simultaneously.

4. Empirical analysis

This paper chooses coaching performance data of basketball player over the past century, including numerical values of G, W L, Y and H. Specially, H here is just contest champion time; taking basketball for example, it includes number of entrance into 4 top, time of entrance into post season and number of grand champion etc. Because most players have not participated in regular season or other famous contest, the data is not complete and therefore data shall be preprocessed to replace missing value with zero. To guarantee stability and reliability of data, original data and processed data are analyzed for reliability in this paper.

4.1. Evaluation result

This paper defines 4 decisive indexes to measure coaching performance of player, respectively being Y, G, WL and H. Comprehensive score of player shall be calculated according to historical data of 4 decisive indexes and top 5 players shall be determined through ranking according to comprehensive score.

(1)Standard data setting. For different ball game, this paper divides ranks on the basis of different game data because reasonable result cannot be gained by calculating different levels through single standard: firstly, handle historical data of decisive index of each game and calculate median of each index respectively; then data greater than median shall be trisected, and data lower than median shall be halved. Quantile gained after processing shall be used as essential data for rank division and rank divided will increase with increase of numerical value of index.

(2)System evaluation result. This paper generates training sample by embodying even-distributed interpolation method firstly and training sample is based on essential data described above. Then train original Bayes rule including training sample repeatedly for 100 times to gain training Bayes rule; finally, comprehensive score of player shall be calculated according to historical data of 4 decisive indexes with training Bayes rule, and comprehensive score shall be ranked to gain top 5 players in each ball game (take basketball as example):

Table 3. Top 5 basketball players

Rank	Name	Y	G	WL	H	Comprehensive Score
1	Mike	39	1276	0.763	68	7.254
2	Adolph	41	1065	0.821	70	7.048
3	Dean	36	1132	0.775	70	6.958
4	Jim	40	1258	0.685	56	6.943
5	Bob	42	1271	0.704	46	6.752

(3)Result of rapid clustering method. This paper uses 4 indexes chosen as characteristic values to perform rapid clustering. Result shows that 88 individuals are used as basketball samples and ranking results of fuzzy Bayes rule belong to the optimum group gained through rapid clustering, which shows correctness of index choice and rank diversion in Bayes rule. Experimental result is reasonable.

(4)Evaluation result of analytic hierarchy process (AHP). For all judgment matrixes passing consistency check, weight vector shall be calculated further. After weight vector is gained, comprehensive score of each player shall be calculated respectively and it shall be compared with ranking result of fuzzy Bayes rule, and judgment matrixes shall be modified repeatedly to make ranking condition of 2 models be conforming. A judgment matrix applicable to all sport games will be gained finally, and its CR value is 0.021, passing consistency check. Feature vector shall be normalized to gain $a_1 = 0.4694$, $a_2 = 0.3146$,$a_3 = 0.1371$, and $a_4 = 0.0789$. After experiment, comparative result of AHP rank and fuzzy Bayes rule rank is as shown in Table 4 (take basketball as example).

Table 4. Comparison of top 5 basketball players

Rank	T-S Fuzzy Bayes Rule	AHP	Algorithm of This Paper
1	Mike	Mike	Mike
2	Dean	Adolph	Adolph
3	Adolph	Jim	Dean
4	Jim	Dean	Jim
5	Bob	Bob	Bob

Top 20 players in each game shall be compared for conformance simultaneously to inspect whether AHP is applicable to player performance evaluation of different games further. Result of conformance comparison analysis is that in sport game, basketball, if specific rank is not considered, conformance of output result of top 5 is 100% while conformance of output result of top 20 is 75%, 70% and 70% respectively.

4.2. Performance comparison of algorithm

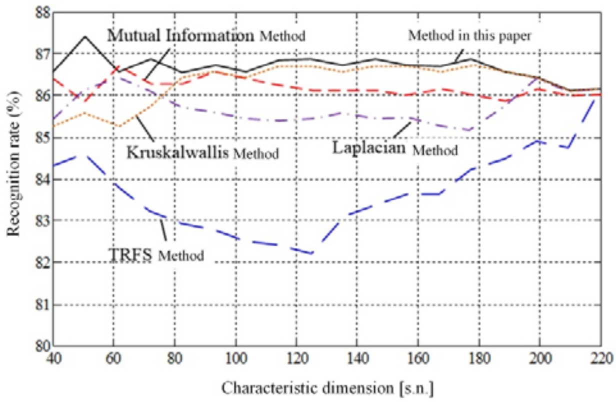
In comparison experiment, testing method proposed in this paper is compared with Mutual Information method, TRFS method, Kruskalwallis method and Laplacian method. data04, datal4 and data24 testing objects in basketball skill management dataset are chosen as experimental objects. See Fig.2a to Fig.2c for comparison condition of basketball skill management data with different characteristic dimensions.

According to comparison result of testing accuracy of basketball skill management in Fig.2, in 3 groups of scene chosen, variation tendency of testing accuracy of basketball skill management generated with change of characteristic dimension is different. In data04 test set, variation tendency of recognition rate is moderate with change of characteristic dimension, but in data14 test set, recognition rate increases firstly and turns to saturation later with increase of characteristic dimension; in data24 test set, recognition rate is in smooth and steady condition firstly and decreases rapidly with increase of characteristic dimension. Reason of above difference lies in difference of scene complexity. data04 test set is relatively simple, and for data14 and data24, characteristic redundancy is relatively great because of complex scene, thus affecting testing accuracy of basketball skill management and variation tendency.

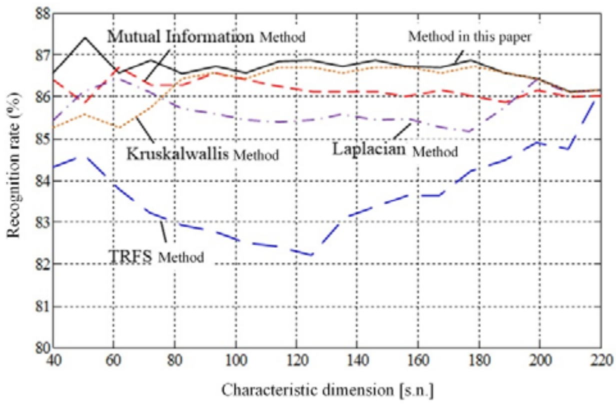
Through result of comparison with 4 kinds of different basketball skill management recognition algorithms chosen, algorithm proposed has better structure and dependency keeping performance in feature extraction, so better basketball skill management testing result can be gained. In the case of the same characteristic dimension setting, in terms of recognition rate, algorithm of this paper is superior to comparison algorithm chosen.

5. Conclusion

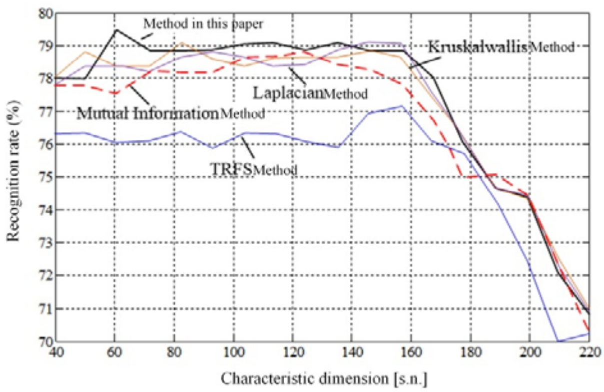
This paper proposes a kind of selection method of basketball skill player management advantages based on non-parametric Bayesian algorithm. Aimed at the unknown model number problem in discovery process of basketball skill player management data partitioning and hiding model, the basketball skill player management model is predicted and analyzed based on non-parametric Bayesian algorithm, which has achieved effective promotion of the basketball skill player management level. In next step, how algorithm shall come into play in actual basketball player training will be researched mainly.



(a) Test result of data04



(b) Test result of data14



(c) Test result of data24

Fig. 2. Comparison of testing accuracy of different dimensions

References

[1] Blakemore C L, Others A: (1992) *Comparison of Students Taught Basketball Skills Using Mastery and Nonmastery Learning Methods.*[J]. Journal of Teaching in Physical Education, 11(3):235-247.

[2] Girotra K, Terwiesch C, Ulrich K T: (2010) *Idea Generation and the Quality of the Best Idea*[J]. Management Science, 56(4):591-605.

[3] Wright P M, Mcmahan G C: (2011) *Exploring human capital: putting 'human' back into strategic human resource management*[J]. Human Resource Management Journal, 21(2):93–104.

[4] Beaudry P, Green D A, Sand B M: (2016) *The Great Reversal in the Demand for Skill and Cognitive Tasks*[J]. Journal of Labor Economics, 34(Volume 34, Number S1):S199-S247.

[5] Pan C Y: (2010) *Effects of water exercise swimming program on aquatic skills and social behaviors in children with autism spectrum disorders.*[J]. Autism: The International Journal of Research and Practice, 14(1):9-28.

[6] Salleh N, Mendes E, Grundy J: (2010) *Empirical Studies of Pair Programming for CS/SE Teaching in Higher Education: A Systematic Literature Review*[J]. IEEE Transactions on Software Engineering, 37(4):509-525.

[7] Bayer P J, Ross S L: (2010) *Identifying Individual and Group Effects in the Presence of Sorting: A Neighborhood Effects Application*[J]. Ssrn Electronic Journal, 2010(10-50).

[8] Angelika Hauke, Julia Flintrop, Emmanuelle Brun, et al.: (2011) *The impact of work-related psychosocial stressors on the onset of musculoskeletal disorders in specific body regions: A review and meta-analysis of 54 longitudinal studies*[J]. Work & Stress, 25(25):243-256.

[9] Mj C E S, Moreira C H, Gonçalves C E, et al.: (2010) *Growth, maturation, functional capacities and sport-specific skills in 12-13 year-old- basketball players.*[J]. Journal of Sports Medicine & Physical Fitness, 50(2):174.

[10] Kampfe J, Sedlmeier P, Renkewitz F: (2011) *The Impact of Background Music on Adult Listeners: A Meta-Analysis.*[J]. Psychology of Music, 39(4):424.

[11] Longo U G, Loppini M, Berton A, et al.: (2012) *The FIFA 11+ program is effective in preventing injuries in elite male basketball players: a cluster randomized controlled trial.*[J]. American Journal of Sports Medicine, 40(5):996-1005.

[12] Schlee R P, Harich K R: (2010) *Knowledge and Skill Requirements for Marketing Jobs in the 21st Century.*[J]. Journal of Marketing Education, 32(3):341-352.

Received May 7, 2017