IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received February 19, 2021, accepted March 1, 2021, date of publication March 8, 2021, date of current version March 17, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3064261

Blockchain-Enabled HMM Model for Sports Performance Prediction

PING CAO[1], GUOQING ZHU[2], QINGGUO ZHANG[1,3], FAN WANG[4], (Graduate Student Member, IEEE), YUWEN LIU[4], AND RUICHAO MO[5]

[1]Teaching Department of Sports, Qufu Normal University, Rizhao 276826, China
[2]School of Physical Education and Health, Linyi University, Linyi 276007, China
[3]China Basketball College, Beijing Sport University, Beijing 100084, China
[4]School of Computer Science, Qufu Normal University, Rizhao 276826, China
[5]School of Computer and Software, Nanjing University of Information Science and Technology, Nanjing 210044, China

Corresponding author: Guoqing Zhu (zhqg2005@126.com)

This work was supported by the Financial and Science Technology Plan Project of Xinjiang Production and Construction Corps under Grant 2020DB005.

ABSTRACT The historical training or exam data of an athlete produced in the past sport exercise or test activities have provided a promising way to objectively and accurately evaluate the real-time sport performance of the athlete. However, the continuous generation of sport training or exam data has placed a heavy transmission and processing burden on the traditional centralized data processing paradigm (e.g., cloud platform). Considering this drawback, a decentralized blockchain-based athlete sport data transmission and utilization solution is proposed in this research work. Moreover, the available athlete sport data produced in past sport exercise or test activities is often sparse and time-related, which call for a robust and time-aware data fusion and processing solution. In this situation, HMM model is employed in this article to cope with the data sparsity and dynamics and further make accurate sports performance prediction for athletes accordingly. Finally, we design a set of experiments on a real-world dataset to validate the feasibility of our proposal in terms of effectiveness and efficiency.

INDEX TERMS Sports performance, HMM, blockchain, prediction.

I. INTRODUCTION

The rapid development of IoT (Internet of Things) and AI (Artificial Intelligence) has brought people a more convenient and comfortable lifestyle [1], [2] [3], [4]. People only need to move their fingers instead of physical work to finish complex tasks in some cases, which has led to ever-increasing sub-health phenomena. In order to improve the physical fitness of the people, more and more countries or governments attach great importance to sports activities. Specifically, there are lots of sports competitions held to answer the national call and inspire citizens' desire for sports.

Sports performance is an important aspect that reflects the physical health and competitive level of athletes. Predicting sports performance can help sports coach to clarify the future development trend of athletes and formulate reasonable training plans for them. Furthermore, the prediction results play a crucial role in tracking and monitoring the development characteristics of athletes' physical fitness. In this situation, it is becoming a necessity to implement accurate sports prediction to provide significant guidance for efficient physical education.

Generally, sports performance prediction of athletes is based on large-scale score data, which usually requires uploading the scores of different athletes in different sports events during a certain period of time. However, large-scale data storage often consumes a lot of management resources. The data center must not only complete the heavy data uploading work, but also ensure the correctness and security of the data. Due to the complexity of social relations [5], [6] [7], data errors and data tampering often occur, which may render athletic contests infeasible in proceeding normally. In worse cases, an athlete's career may be directly affected. In addition, the physical and mental state of an athlete is one of the potential determinants of athletic performance. However, the complex distribution of sports data makes it challenging to analyze data characteristics accurately. Furthermore, numerical information often fails to reflect the

The associate editor coordinating the review of this manuscript and approving it for publication was Zhipeng Cai.

physical and mental state of athletes. The immeasurability of athlete states has become a major obstacle to improve the accuracy of sports performance prediction.

In light of the above-mentioned challenges, we utilize blockchain technology for data storage and management. Blockchain technology ensures the openness, transparency and unforgeability of data [8]. Both competition organizers and referees are entitled to upload athletes' sports performance. In fact, blockchain technology can not only simplify data management, but also alleviate the pressure of data uploading and proofreading for data center administrators. Then we construct an hidden markov model (HMM) to explore the changes in athletes' physical and mental states over time.

In conclusion, our contributions in this article are threefold.

- We introduce blockchain technology into sports data storage and management, in order to reduce data maintenance costs and ensure data security.
- We construct an HMM model to explore the physical and mental changes of athletes in multiple competitions, and perform the prediction of athletes' next sports performance. To the best of our knowledge, this is the first work that applies HMM to sports performance prediction.
- We conduct a wide range of experiments on the WS-DREAM dataset to demonstrate the feasibility and effectiveness of our proposal. Experimental results indicate that our method performs well in terms of prediction accuracy.

The rest of article is organized as follows. In the following section, we review related work for sports performance prediction. In Section III, we discuss the motivation of our proposal. Following the motivation section, we detail our presented sports performance prediction model. Finally, we conclude this article.

II. RELATED WORK

At present, sports competitions are widely popular all over the world. For athletes, it is a matter of concern for athletes to perform well and win the competition. Nowadays,the topic of prediction attracts peoples attention [9], and the research on the sports performance prediction is also a challenging research. Due to the result of sports competition data is time series data, the first method to predict the result of sports competition is the time series data analysis method. Yong *et al.* [10] collected the marathon results of Boston over the years, and then through statistical analysis, analyzed the athletes' performance from multiple perspectives such as age and gender, and then took the mean value and established the autoregressive moving average model, so as to predict the athletes' next competition performance. However, the competition is affected by many uncertain factors, which makes it impossible to make a reliable prediction only through data analysis. Later, multiple linear regression prediction models began to emerge. Krzyszkowski *et al.* [11] proposed a canonical regression model to predict javelin performance. Specifically, it analyzed the biomechanical data of athletes, and then used the canonical regression model to predict the throwing distance of athletes. And Li *et al.* [12] proposed a prediction approach for sports team performance based on data envelopment analysis (DEA) methodology and data-driven technique. Firstly, it used multiple logistic regression to predict the relationship between the winning probability and the result of the game, and then used the team's historical data to predict the result of the game, which can be used to guide the team's future training. The result of sports competition is often affected by many factors, such as the state of athletes, the environment of competition and so on. Wang *et al.* [13] analyzed the competitive sports regression prediction model based on support vector machine and artificial intelligence, and predicted the individual skill level of athletes, which is helpful to improve the performance of the competition.

With the continuous development of fuzzy theory [14], [15] and grey theory [16], there are many nonlinear sports performance prediction models. The reliability of the model is improved. Liu *et al.* [17] constructed a theoretical framework of grey prediction of competitive sports, which laid a good foundation for the application of grey prediction theory in competitive sports. Huang *et al.* [18] put forward a prediction model of track and field performance in the Olympic Games, which is combined with grey modeling method to predict the development trend of track and field. Liu *et al.* [19] combined the fuzzy method with SVM technology to predict the basketball game results. A Hybrid Fuzzy SVM model (HFSVM) is proposed. This model can achieve better accuracy than SVM in basketball performance prediction.

Therefore, many researchers have carried out studies related to neural networks [20], [21]. Maszczyk *et al.* [22] used a 4-3-1 perceptron network to predict the sports performance of 116 javelin athletes. This study compared the prediction performance of regression model and neural model, and it proved that the prediction ability of neural model is significantly better than regression model. Horvat *et al.* [23] also compared many methods based on machine learning. Comparative analysis found that most articles use machine learning algorithm before feature selection and feature extraction, and data segmentation of time series data. Neural network is the most commonly used method of data segmentation. In order to grasp the development trend of sports events in time, Li [24] improved the traditional back-propagation algorithm and proposed an improved adaptive back-propagation neural network (BPNN) to predict the sports performance. In order to improve the accuracy of sports performance prediction and alleviate the slow speed of sports performance prediction, Zhang *et al.* [25] proposed a prediction model based on firefly optimized neural network. Experiments showed that the accuracy of this method is high, but it didn't consider some other factors that affect the game, such as athletes' physical state [26], psychological state [27] and so on. Moreover, there are many advanced prediction applied in various fields, such as stock prediction [28], software aging prediction [29] and the total number of reported COVID-19 cases and deaths prediction [30]. The application of deep neural network in prediction is also developing rapidly [31], [32]. For example, Liu *et al.* [33] proposed a deep memory neural network for video prediction and Wan *et al.* used Long Short-Term Memory (LSTM) [34] for time series prediction. Therefore, there is still room for improvement in the method of sports performance prediction.

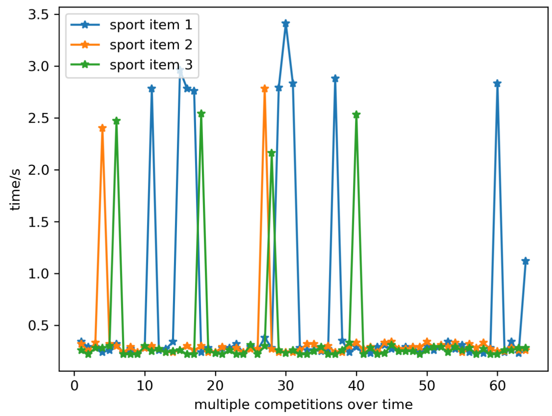
The time series analysis method used for sports performance prediction did not consider the uncertain factors that affect the stadium, only for the analysis of historical data rules, can not achieve reliable accuracy. However, the training of machine learning model often needs a lot of data, and the training time of the model is long and the interpretability is poor. In summary, in order to consider the uncertain factors of the competition field, such as athletes' psychological state and physical state. Moreover, to achieve better results of sports performance prediction, we use HMM to predict sports performance. This method not only considers the athlete's physical state and the influence of historical competition, but also achieves good accuracy, and the training speed of the model is fast and easy to realize.

III. MOTIVATION

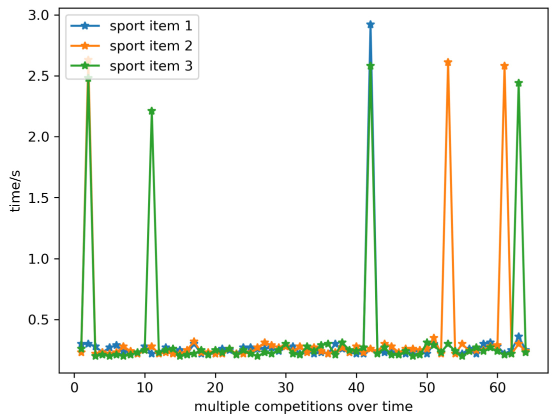
In this section, we employ Fig. 1 to discuss the motivation of this article, where Fig. 1(a) and Fig. 1(b) depict the sports performances of two athletes in multiple sport competitions over time, respectively.

As shown in Fig 1, we can observe that the performances of the two athletes in all three sports events fluctuate with the change of time. Occasionally, an athlete's performance may be excellent or poor. This is because athletes have different physical and mental states at different times, resulting in different levels of their own abilities. In this situation, we can infer that the physical and mental state of an athlete is a significant factor in determining sports performance, and exploring the state information of an athlete can potentially improve prediction accuracy of sports performance. However, the physical and mental state of an athlete often fails to be observed with eyes and measured with numerical values, which makes it challenging for us to regard state information as reference information for sports performance prediction.

Considering the above-mentioned challenges, we argue that a promising way for objectively and accurately predicting the real time sport performances of an athlete is to make full use of all the valuable athlete information stored by different data stakeholders. In this situation, as the precondition of objective and accurate athlete performance evaluation, how to guarantee the reliability and authenticity of each piece of involved athlete training or exam data is becoming a crucial but challenging task that requires intensive study. To achieve this goal, in this article, we use blockchain technology to secure the reliability and authenticity of athlete data that is probably shared, integrated and transmitted among different parties, and HMM model to accurately predict athlete performances based on the time-aware athlete training or exam data produced in the past and protected by the previously mentioned blockchain technology. The details of our proposal will be specified in the next section.



(a) Sports performance of $user_1$.



(b) Sports performance of $user_2$.

FIGURE 1. Sports performances of two users w.r.t. time.

IV. SPORTS PERFORMANCE PREDICTION

In this section, we study the detailed process of forecasting athletes' sports performance algorithm based on blockchain and HMM, named $STPP_{HMM}$. Concretely, we first introduce the problem definitions in Section IV-A. In Section IV-B, we discuss the details of our proposal.

A. DEFINITIONS

Consider the sports performance sequences $G = \{g_1, g_2, \ldots, g_n\}$ of a group of athletes, where each sequence g_i represents a set of sport grades of the athlete u_i, i.e., $g_i = \{r_1, r_2, \ldots, r_{m_i}\}$. Each sports performance record r refers to the sports score of an athlete in m competitions held in chronological order, e.g., 11.2 seconds in 100-meter sprint, 1.79 meters in high jump and so on. According to the sequences G, we aim to predict the next sports performance for the athlete u_{target}. Concretely, given sequence $g_{target} = \{r_1, r_2, \ldots, r_t\}$, our goal is to explore the athlete's state pattern over time and forecast his/her next sports performance r_{t+1}.

B. BLOCKCHAIN-ENABLED SPORTS PERFORMANCE PREDICTION BASED ON HMM: $stpp_{HMM}$

In this section, we propose a novel sports performance prediction method based on blockchain technology and HMM, named $STPP_{HMM}$. Specifically, we first employ blockchain technology to implement data storage and management. With the aid of blockchain technology, referees of different competitions can independently complete the score uploading task. Open and transparent blockchain data enables mutual trust between athletes and competition managers. The data in blockchain is prohibited from being tampered with, which guarantees the security of official sport data. Afterwards, we employ HMM to model the target athlete's next sports performance prediction problem, and conduct parameter estimation in HMM. Finally, we utilize the learnt HMM to predict sports performance.

Hidden markov model (HMM) is a generative model of dynamic bayesian network with simple structure, which is employed for processing time series data [35]. In this article, we utilize HMM to model an athlete's performance variation pattern and the dependency between his/her physical and mental states over time. In fact, the state transition sequences in HMM can comprehensively simulate sports performance data. Concretely, the visible state sequences can model the athlete's performance variation pattern, and the hidden state sequences can indicate the athlete's physical and mental conditions, e.g., muscle, skeleton and mood. To the best of our knowledge, this work is the first to employ HMM for sports performance prediction. The details of our proposal consist of the following three steps.

1) MODEL THE PROBLEM

In this step, we employ HMM to model the next sports performance prediction problem. In our problem, the underlying states (i.e., physical and mental condition) of an athlete is represented by a set of hidden states $\mathcal{S} = \{s_1, s_2, \ldots, s_M\}$; the athlete's performance levels are represented by a set of observation states $\mathcal{O} = \{o_1, o_2, \ldots, o_N\}$. In general, HMM consists of the following three main parameters that comprehensively interpret the characteristics of an HMM model: (1) initial probability distribution $\Pi = (\pi_{s_1}, \pi_{s_2}, \ldots, \pi_{s_M})$, where $s_i \in \mathcal{S}$ and $\sum_{i=1}^{M} \pi_{s_i} = 1$; (2) state transition matrix $A = (a_{ij})_{M \times M}$, where a_{ij} is the probability of transition from the hidden state s_i to s_j, and $\sum_{j=1}^{M} a_{ij} = 1$; (3) observation probability matrix $B = (b_{ij})_{M \times N}$, where b_{ij} denotes the probability of emission from hidden state $s_i \in \mathcal{S}$ to observation state $o_j \in \mathcal{O}$, and $\sum_{j=1}^{N} b_{ij} = 1$. Intuitively, we employ Fig. 2 to depict HMM model with probabilistic parameters.

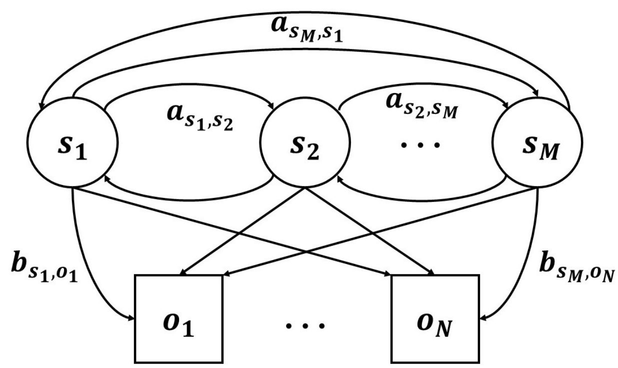


FIGURE 2. HMM with probabilistic parameters.

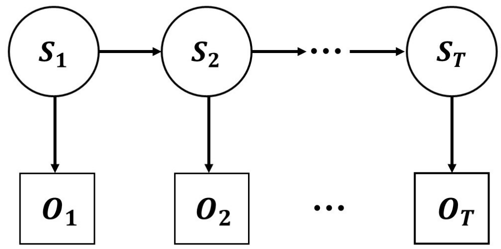


FIGURE 3. State transition graph of HMM.

We regard an athlete's sports performances in a period of time as a sequence of performance levels whose length is T, i.e., $g = O_1 O_2 \ldots O_T$ (abbreviated as $g = O_{1:T}$), which will be used to train the HMM in the next step of our algorithm. In this observation sequence, $O_t \in \mathcal{O}$ is the observed performance level at time t, where $1 \leq t \leq T$. Each O_t is corresponding to a hidden state $S_t \in \mathcal{S}$, denoting the unobservable state of the athlete at time t. Intuitively, we employ Fig. 3 to depict the potential dependence of the observation sequence $g = O_{1:T}$ and the corresponding hidden state sequence. In the following, we will train the HMM model based on the observation sequence $g = O_{1:T}$ and conduct sports performance prediction for an athlete.

2) PARAMETER ESTIMATION FOR HMM

There are lots of researchers devoting themselves into HMM parameter estimation methods and receiving great achievements. One of the most classic methods is the Baum-Welch algorithm [36], which trains parameters by maximizing the likelihood of the observation sequence. In Baum-Welch algorithm, the overall likelihood of each observation sequence is the sum of the likelihood of all possible routes through the hidden states, as specified in Eq.(1). It is worth noting that the state transition matrix and the observation probability matrix do not change over time.

$$P(O_{1:T}) = \sum_{S_1=s_1}^{s_M} \cdots \sum_{S_T=s_1}^{s_M} \pi S_1 \prod_{t=2}^{T} a_{S_{t-1},S_t} \prod_{t=1}^{T} b_{S_t,O_t} \quad (1)$$

Eq.(1) adopts a simple but violent way that exhausts all possible routes. However, it is a time-consuming and labor-intensive way when there is a long observation sequence to be estimated. In order to address this problem, the literature [37] optimizes Eq.(1) and proposes Eq.(2) to reduce the time cost through matrix multiplication. In Eq.(2), π is the initial probability distribution, A is the state transition matrix. Furthermore, we transform the each column of the observation probability matrix B into an $M \times M$ diagonal matrix $\mathcal{B}_{O_t}$, where each diagonal element is b_{s_i,O_t} $(1 \leq i \leq M)$, and other elements are 0. Afterwards, we apply Baum-Welch algorithm to the observation sequence to estimate the initial probability distribution, the state transition matrix and the observation probability matrix.

$$L_T = P(O_{1:T}) = \pi \mathcal{B}_{O_1} A \mathcal{B}_{O_2} \ldots A \mathcal{B}_{O_T} \mathbf{1}^T \tag{2}$$

In order to obtain the most suitable number of hidden states (i.e., M), we adjust the parameters of M in Section V to optimize the prediction accuracy.

3) SPORTS PERFORMANCE PREDICTION

Given the performance $g_{target} = O_{1:t}$ of the target athlete u_{target} as well as the learnt HMM, we can predict his/her athletic performance at time $t+1$, including the following two steps:

$$S_{t+1} = arg\ max_{s_i \in \mathcal{S}} P(S_{t+1} = s_i | O_{1:t}) \tag{3}$$
$$O_{t+1} = arg\ max_{o_j \in \mathcal{O}} b_{S_{t+1},o_j} \tag{4}$$

where we first explore the hidden state S_{t+1} with the greatest possibility, and then predict the most likely sports performance O_{t+1} at time $t+1$ based on the given hidden sate S_{t+1}.

In this situation, another problem is how to explore the next hidden state S_{t+1} according to the given observation sequence $O_{1:t}$. Due to the homogeneous markov hypothesis and the observation independence hypothesis, $P(S_{t+1} = s_j | O_{1:t})$ can be solved recursively by the state transition probability $a_{S_t,S_{t+1}}$ and the possibility of the previous state $P(S_t | O_{1:t})$, as specified as follows:

$$P(S_{t+1} = s_j | O_{1:t}) = \sum_{i=1}^{M} a_{s_i,s_j} P(S_t = s_i | O_{1:t}) \tag{5}$$

where the calculation strategy of $P(S_t = s_i | O_{1:t})$ is the famous "filtering" algorithm, as shown in Eq.(6).

$$P(S_t = s_i | O_{1:t}) = (\pi \mathcal{B}_{O_1} A \mathcal{B}_{O_2} A \ldots \mathcal{B}_{O_{t-1}} A_i b_{S_t,O_t}) / L_t \tag{6}$$

where A_i represents the i_{th} column of the state transition matrix A, and L_t can be calculated by Eq.(2).

In summary, we first employ Eq.(5) to solve all the probabilities of $P(S_{t+1} = s_j)$, $1 \leq j \leq M$, and select the most likely hidden state S_{t+1} by Eq.(3). Next, we employ Eq.(4) to find the most likely operation O_{t+1} according to the given S_{t+1}.

V. EXPERIMENTS

In this section, we conduct a wide range of experiments on the WS-DREAM dataset to demonstrate the feasibility and effectiveness of our proposal. Specifically, Section V-A introduces the background and evaluation indicators of our experiments, and Section V-B shows the experimental performance.

A. EXPERIMENT DATASET AND EVALUATION CRITERIA

Our experiment adopts the real QoS dataset WS-DREAM [38] to simulate the athlete's sports performance. The WS-DREAM dataset includes 4,532 web services invoked by 142 users at 64 time slots. Thus, we regard 142 users as 142 athletes, 4,532 web services as 4,532 sport items, 64 time slots as 64 competitions held in chronological order, and quality of service (e.g., response time) as sports performance. Under the simulation of the dataset, we compare our $STPP_{HMM}$ with average and random methods to test the performance of our prediction.

In addition, we test the performance of the three methods in terms of the following three indicators:

- MAE (Mean Absolute Error, the smaller the better): measure the deviation between the predictions and the actual results.
- RMSE (Root Mean Square Error, the smaller the better): similar to MAE, measure the standard deviation of the differences between the predictions and the observations.
- Time cost: measure the prediction efficiency.

The experiment settings are as follows: (1) hardware: 2.70GHz CPU, 8.0GB RAM; (2) software: Windows 10 and Python 3.6.

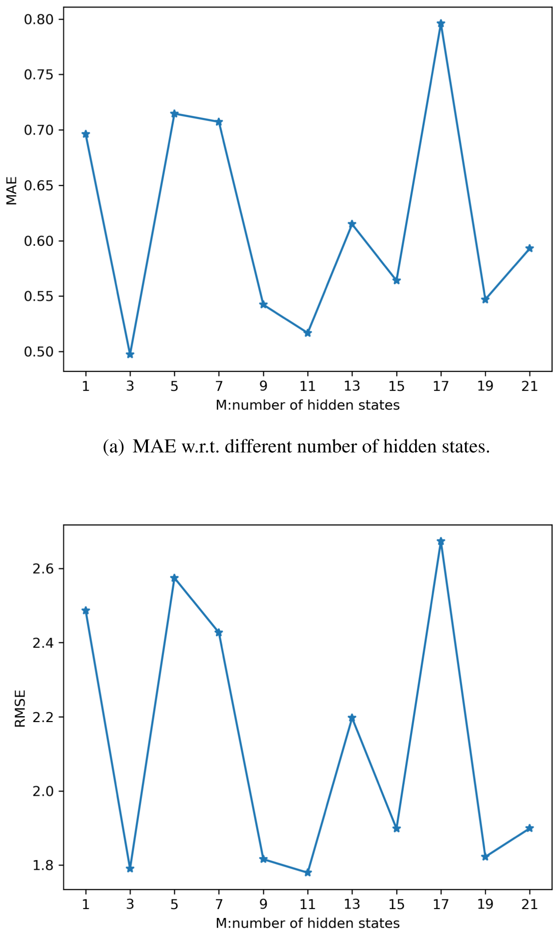
B. EXPERIMENT RESULTS

We design three profiles in this section to analyze the experimental results of the three methods. Concretely, profile 1 adjusts the parameter M to ensure the optimal experimental performance of our proposal; profile 2 compares the three methods in terms of MAE and RMSE; profile 3 evaluates the time cost of our $STPP_{HMM}$.

1) IDENTIFY OPTIMAL NUMBER OF HIDDEN STATES

The number of hidden states play a significant role in HMM model, thus it is of great necessity to select the most suitable M to achieve the best performance of our proposal. Specifically, we randomly select a web service as a sport item to test the prediction accuracy of sports performance. The experimental results are shown in Fig. 4.

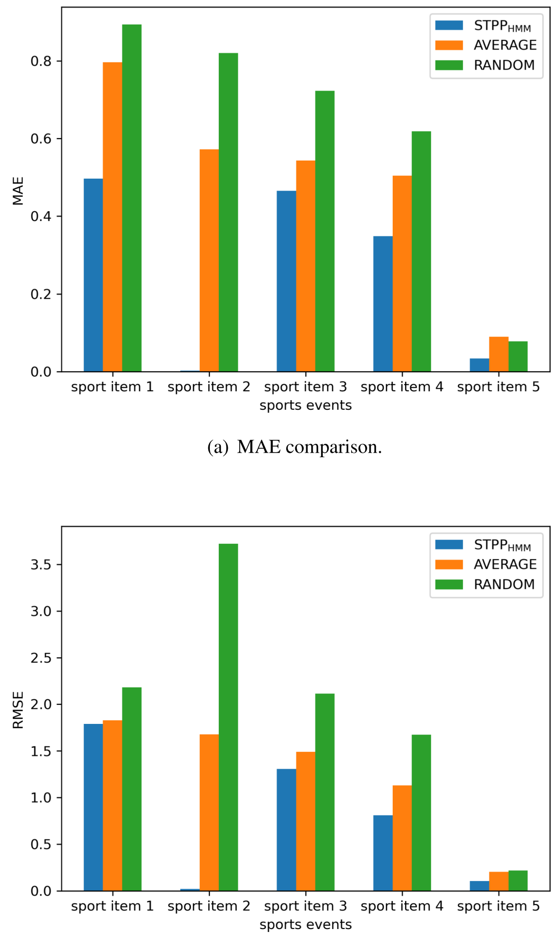
As shown in Fig. 4, the accuracy of our $STPP_{HMM}$ (including MAE and RMSE) is affected dramatically as the number of hidden states changes. We can observe that our method performs well in terms of both MAE and RMSE when $M = 3$. According to the above analysis, we set M as 3 to complete the subsequent parameter training and prediction of our model.



(a) MAE w.r.t. different number of hidden states.

(b) RMSE w.r.t. different number of hidden states.

FIGURE 4. Prediction accuracy w.r.t. the number of hidden states.



(a) MAE comparison.

(b) RMSE comparison.

FIGURE 5. Prediction accuracy comparison of three approaches.

2) ACCURACY COMPARISON OF THREE SPORTS PERFORMANCE PREDICTION METHODS

In this profile, we compare the accuracy of the three sports performance prediction approaches through MAE and RMSE. Specifically, we randomly select five web services as five sport items to evaluate the prediction accuracy. The experimental results are shown in Fig. 5.

In Fig. 5, we can observe that our $STPP_{HMM}$ performs best in the accuracy of sports performance prediction (i.e., both MAE and RMSE are the smallest); this is because our model not only considers the data distribution of athlete performances, but also the athletes' states in each competitions. On the contrary, average and random algorithms fail to explore the physical and mental states of athletes, resulting in poor prediction accuracy. Therefore, our proposal outperforms the other two algorithms in terms of MAE and RMSE.

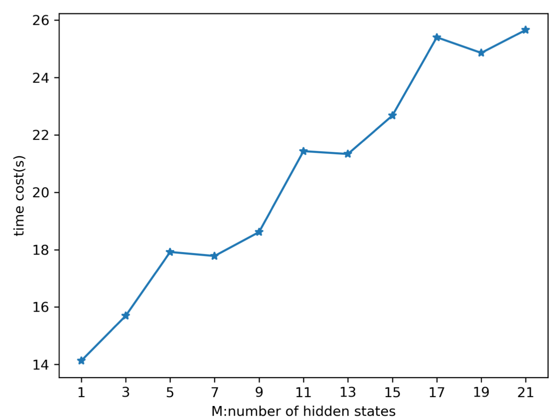


FIGURE 6. Time cost w.r.t. different number of hidden states.

3) TIME COST EVALUATION OF $STPP_{HMM}$

In this profile, we test the efficiency of our sports performance prediction method. Specifically, we change the value of M from 1 to 21 and randomly select a web service as a sport item to measure the predictive efficiency of our proposal. The experimental results are shown in Fig. 6.

As shown in Fig. 6, the time cost continues to increase with the growth of M (the number of hidden states); this is because more training and computing time is required when the number of underlying states increase. In summary, the time cost of our proposal is often positively correlated with the number of hidden states. Due to that we set M as 3 in our $STPP_{HMM}$, our model has a relatively high efficiency.

VI. CONCLUSION

The large scale of athletes' training data and competition performances provide a valuable reference for accurate sports performance prediction. Reasonable and accurate prediction of sports performance is conductive to formulating future training plans or development directions for athletes. However, the increasing volume of data makes it challenging for cloud platform to conduct data storage and management. In addition, the existing researches related to sports performance prediction often fail to take athletes' physical and mental state into consideration, which usually leads to inaccurate prediction. Considering these issues, this article propose a robust and accurate sports prediction approach based on decentralized blockchain and HMM. In this situation, our proposal can not only relieve the management burden of the data processing center, but also ensure the accuracy of sports performance prediction. Finally, we conduct extensive experiments on the WS-DREAM dataset. The experimental results demonstrate the effectiveness and accuracy of our proposal.

In the future, we will work on more sports-related research. On the one hand, we will perform recommendations [39], [40] associated with sports competitions or sports events for athletes based on their physical fitness indicators. On the other hand, we will consider the privacy issues of athletes and protect the privacy of athletes during training [41], [42] [43].

REFERENCES

[1] Z. Cai and Z. He, "Trading private range counting over big IoT data," in *Proc. IEEE 39th Int. Conf. Distrib. Comput. Syst. (ICDCS)*, Jul. 2019, pp. 144–153.

[2] Y. Wang, Y. Gao, Y. Li, and X. Tong, "A worker-selection incentive mechanism for optimizing platform-centric mobile crowdsourcing systems," *Comput. Netw.*, vol. 171, Apr. 2020, Art. no. 107144.

[3] T. Liu, Y. Wang, Y. Li, X. Tong, L. Qi, and N. Jiang, "Privacy protection based on stream cipher for spatiotemporal data in IoT," *IEEE Internet Things J.*, vol. 7, no. 9, pp. 7928–7940, Sep. 2020.

[4] S. Zhu, Z. Cai, H. Hu, Y. Li, and W. Li, "ZkCrowd: A hybrid blockchain-based crowdsourcing platform," *IEEE Trans. Ind. Informat.*, vol. 16, no. 6, pp. 4196–4205, Jun. 2020.

[5] Z. Cai, Z. He, X. Guan, and Y. Li, "Collective data-sanitization for preventing sensitive information inference attacks in social networks," *IEEE Trans. Depend. Sec. Comput.*, vol. 15, no. 4, pp. 577–590, Jul./Aug. 2016.

[6] H. Kou, H. Liu, Y. Duan, W. Gong, Y. Xu, X. Xu, and L. Qi, "Building trust/distrust relationships on signed social service network through privacy-aware link prediction process," *Appl. Soft Comput.*, vol. 100, Mar. 2021, Art. no. 106942.

[7] X. Zhou, W. Liang, K. I.-K. Wang, R. Huang, and Q. Jin, "Academic influence aware and multidimensional network analysis for research collaboration navigation based on scholarly big data," *IEEE Trans. Emerg. Topics Comput.*, early access, Jul. 26, 2019, doi: 10.1109/TETC.2018.2860051.

[8] Z. Sun, Y. Wang, Z. Cai, T. Liu, X. Tong, and N. Jiang, "A two-stage privacy protection mechanism based on blockchain in mobile crowdsourcing," *Int. J. Intell. Syst.*, 2021, doi: 10.1002/int.22371.

[9] L. Qi, C. Hu, X. Zhang, M. R. Khosravi, S. Sharma, S. Pang, and T. Wang, "Privacy-aware data fusion and prediction with spatial-temporal context for smart city industrial environment," *IEEE Trans. Ind. Informat.*, early access, Jul. 28, 2020, doi: 10.1109/TII.2020.3012157.

[10] W. Yong, P. Lingyun, and W. Jia, "Statistical analysis and ARMA modeling for the big data of marathon score," *Sci. Sports*, vol. 35, no. 6, pp. 375–385, Dec. 2020.

[11] J. Krzyszkowski and K. Kipp, "Prediction of throwing distance in the men's and women's javelin final at the 2017 IAAF world championships," *J. Sports Sci.*, pp. 1–5, Jan. 2021, doi: 10.1080/02640414.2020.1856462.

[12] Y. Li, L. Wang, and F. Li, "A data-driven prediction approach for sports team performance and its application to national basketball association," *Omega*, vol. 98, Jan. 2021, Art. no. 102123.

[13] J. Wang and H. Qu, "Analysis of regression prediction model of competitive sports based on SVM and artificial intelligence," *J. Intell. Fuzzy Syst.*, vol. 39, no. 4, pp. 5859–5869, Oct. 2020, doi: 10.3233/JIFS-189061.

[14] Y. Jin, W. Cao, M. Wu, and Y. Yuan, "Accurate fuzzy predictive models through complexity reduction based on decision of needed fuzzy rules," *Neurocomputing*, vol. 323, pp. 344–351, Jan. 2019.

[15] O. Cagcag Yolcu, E. Bas, E. Egrioglu, and U. Yolcu, "A new intuitionistic fuzzy functions approach based on hesitation margin for time-series prediction," *Soft Comput.*, vol. 24, no. 11, pp. 8211–8222, Jun. 2020.

[16] S.-F. Liu and J. Forrest, "Advances in grey systems theory and its applications," in *Proc. IEEE Int. Conf. Grey Syst. Intell. Services*, Nov. 2007, pp. 1–6.

[17] J. Liu, J.-G. Shen, Y. Li, and Y.-F. Wang, "The construction of the framework of grey prediction theory for competitive sports," *J. Grey Syst.*, vol. 26, no. 4, p. 89, 2014.

[18] H. Chang Mei, S. Wei Hua, and X. Xiao Cong, "Research on construction and application of the GM(1,1) forecast model of olympics track and field achievements," *Grey Syst., Theory Appl.*, vol. 2, no. 2, pp. 178–196, Aug. 2012, doi: 10.1108/20439371211260153.

[19] S. Jain and H. Kaur, "Machine learning approaches to predict basketball game outcome," in *Proc. 3rd Int. Conf. Adv. Comput., Commun. Autom. (ICACCA) (Fall)*, Sep. 2017, pp. 1–7.

[20] X. Zhou, Y. Hu, W. Liang, J. Ma, and Q. Jin, "Variational LSTM enhanced anomaly detection for industrial big data," *IEEE Trans. Ind. Informat.*, vol. 17, no. 5, pp. 3469–3477, May 2021, doi: 10.1109/TII.2020.3022432.

[21] X. Zhou, W. Liang, K. I.-K. Wang, H. Wang, L. T. Yang, and Q. Jin, "Deep-learning-enhanced human activity recognition for Internet of healthcare things," *IEEE Internet Things J.*, vol. 7, no. 7, pp. 6429–6438, Jul. 2020.

[22] A. Maszczyk, A. Gołaś, P. Pietraszewski, R. Roczniok, A. Zając, and A. Stanula, "Application of neural and regression models in sports results prediction," *Procedia Social Behav. Sci.*, vol. 117, pp. 482–487, Mar. 2014.

[23] T. Horvat and J. Job, "The use of machine learning in sport outcome prediction: A review," *WIREs Data Mining Knowl. Discovery*, vol. 10, no. 5, Sep. 2020, Art. no. e1380.

[24] H. Li, "Analysis on the construction of sports match prediction model using neural network," *Soft Comput.*, vol. 24, no. 11, pp. 8343–8353, Jun. 2020.

[25] F. Zhang, "Research on improving prediction accuracy of sports performance by using glowworm algorithm to optimize neural network," *Int. J. Inf. Educ. Technol.*, vol. 9, no. 4, pp. 1–4, 2019, doi: 10.18178/IJIET.2019.

[26] M. A. Dos Santos, F. C. Moreno, and I. Valantine, "Editorial: The management of emotions in sports organizations," *Frontiers Psychol.*, vol. 11, Oct. 2020, doi: 10.3389/fpsyg.2020.607041.2020.

[27] R. Anderson, S. J. Hanrahan, and C. J. Mallett, "Investigating the optimal psychological state for peak performance in australian elite athletes," *J. Appl. Sport Psychol.*, vol. 26, no. 3, pp. 318–333, Jul. 2014.

[28] Y. Wang, L. Wang, F. Yang, W. Di, and Q. Chang, "Advantages of direct input-to-output connections in neural networks: The elman network for stock index forecasting," *Inf. Sci.*, vol. 547, pp. 1066–1079, Feb. 2021.

[29] J. Liu and L. Meng, "Integrating artificial bee colony algorithm and BP neural network for software aging prediction in IoT environment," *IEEE Access*, vol. 7, pp. 32941–32948, 2019.

[30] T. Reddy, Z. Shkedy, C. Janse van Rensburg, H. Mwambi, P. Debba, K. Zuma, and S. Manda, "Short-term real-time prediction of total number of reported COVID-19 cases and deaths in South Africa: A data driven approach," *BMC Med. Res. Methodol.*, vol. 21, no. 1, pp. 1–11, Dec. 2021.

[31] X. Zhou, Y. Li, and W. Liang, "CNN-RNN based intelligent recommendation for online medical pre-diagnosis support," *IEEE/ACM Trans. Comput. Biol. Bioinf.*, early access, May 14, 2020, doi: 10.1109/TCBB.2020.2994780.
[32] X. Zhou, W. Liang, S. Shimizu, J. Ma, and Q. Jin, "Siamese neural network based few-shot learning for anomaly detection in industrial cyber-physical systems," *IEEE Trans. Ind. Informat.*, early access, Dec. 31, 2021, doi: 10.1109/TII.2020.3047675.
[33] Z. Liu, X. Chai, and X. Chen, "Deep memory and prediction neural network for video prediction," *Neurocomputing*, vol. 331, pp. 235–241, Feb. 2019.
[34] H. Wan, S. Guo, K. Yin, X. Liang, and Y. Lin, "CTS-LSTM: LSTM-based neural networks for correlatedtime series prediction," *Knowl.-Based Syst.*, vol. 191, Mar. 2020, Art. no. 105239.
[35] H. Cheng, J. Ye, and Z. Zhu, "What's your next move: User activity prediction in location-based social networks," in *Proc. SDM*, 2013, pp. 171–179.
[36] L. E. Baum, T. Petrie, G. Soules, and N. Weiss, "A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains," *Ann. Math. Statist.*, vol. 41, no. 1, pp. 164–171, Feb. 1970.
[37] W. Zucchini, I. L. MacDonald, and R. Langrock, *Hidden Markov Models for Time Series an Introduction Using R*. London, U.K.: Chapman & Hall, 2017, doi: 10.1007/s13253-016-0273-2.2009.
[38] Z. Zheng, Y. Zhang, and M. R. Lyu, "Investigating QoS of real-world Web services," *IEEE Trans. Services Comput.*, vol. 7, no. 1, pp. 32–39, Jan. 2014.
[39] L. Wang, X. Zhang, R. Wang, C. Yan, H. Kou, and L. Qi, "Diversified service recommendation with high accuracy and efficiency," *Knowl.-Based Syst.*, vol. 204, Sep. 2020, Art. no. 106196.
[40] L. Qi, Q. He, F. Chen, X. Zhang, W. Dou, and Q. Ni, "Data-driven Web APIs recommendation for building Web applications," *IEEE Trans. Big Data*, early access, Feb. 24, 2020, doi: 10.1109/TBDATA.2020.2975587.
[41] Z. Cai and X. Zheng, "A private and efficient mechanism for data uploading in smart cyber-physical systems," *IEEE Trans. Netw. Sci. Eng.*, vol. 7, no. 2, pp. 766–775, Apr. 2020.
[42] L. Qi, X. Wang, X. Xu, W. Dou, and S. Li, "Privacy-aware cross-platform service recommendation based on enhanced locality-sensitive hashing," *IEEE Trans. Netw. Sci. Eng.*, early access, Jan. 27, 2020, doi: 10.1109/TNSE.2020.2969489.
[43] Z. Cai, X. Zheng, and J. Yu, "A differential-private framework for urban traffic flows estimation via taxi companies," *IEEE Trans. Ind. Informat.*, vol. 15, no. 12, pp. 6492–6499, Dec. 2019.

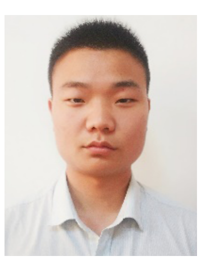
PING CAO received the master's degree from the Department of Physical Education, Qufu Normal University, China, in 2003. She is currently a Lecturer with the Department of Physical Education, Qufu Normal University. Her research interests include sports culture and school sports.

GUOQING ZHU was born in Pingyi, Shandong, China, in 1972. He received the master's degree from Beijing Sport University, China, in 1998. He is currently with the College of Physical Education and Health, Linyi University. His research interests include physical education and training, humanities, and the sociology of sport.

QINGGUO ZHANG received the master's degree from Beijing Sport University, China, in 2006, where he is currently pursuing the Ph.D. degree. He is currently a Lecturer with the Department of Physical Education, Qufu Normal University. His research interests include sport education and exercise training.

FAN WANG (Graduate Student Member, IEEE) received the bachelor's degree from the Department of Mathematics, Qufu Normal University, China, in 2019, where she is currently pursuing the master's degree with the School of Computer Science. Her research interests include big data analyses and recommender systems.

YUWEN LIU received the bachelor's degree from the School of Information Science and Engineering, Qufu Normal University, China, in 2019, where she is currently pursuing the master's degree with the School of Computer Science. Her research interests include big data analyses and recommender systems.



RUICHAO MO received the B.S. degree in software engineering from the Binjiang College, Nanjing University of Information Science and Technology, in 2019, where he is currently pursuing the master's degree in computer science and technology. His research interests include edge computing, service computing, and deep reinforcement learning.