Categorical Learning-based Line-up Prediction in the Drafting Process of MOBA Games

Run-Lu Zhu, Jie Wang, Sheng-Jie Zhao*

School of Software Engineering, Tongji University, China
runluzhu@tongji.edu.cn, wangjie_tongji@tongji.edu.cn, shengjiezhao@tongji.edu.cn

Abstract

In recent years, the e-sports industry has witnessed a rapid development, setting a wave of social and economic influence all over the world. Among these, multiplayer online battle arena (MOBA) games, represented by Dota2 account for the largest share in the number of games, audience, and prizes. As the beginning of the MOBA games, the *drafting phase* is of great importance due to its impact on the outcome of the game. Therefore, line-up prediction in the drafting process is an important research field in e-sports. Nonetheless, different MOBA games have different hero pools, different rules for drafting, and different non-player factors, while the available game data are sparse, hindering the prediction of line-up in the drafting phase. To solve this problem, we i) identify the key steps in lineup prediction through statistical analysis, ii) based on which we propose an orderless prediction method that achieves 7% higher accuracy than existing models, and iii) further improve the accuracy by 1.6% with an attention layer in line-up predictions. Our finding of orderless key steps in MOBA game lineup drafting may also be used to expand the lineup dataset, which is crutial to learning-based automatic game analysis.

Keywords: MOBA game, Line-up prediction, LSTM

1 Introduction

Multiplayer online battle arena (MOBA) games have become one of the most popular game types in recent years, generating a revenue of 358 billion per year. Due to its social and economic influence among young audience, MOBA-related e-sports have been officially admitted into the Asian Games in 2022. For instance, Tencent Games, who specializes in MOBA gaming platforms, yielded approximately $20 billion in 2018 alone. On the other hand, in line with professional gaming data collected by eSportsflag, Dota2 is the most popular Chinese e-sports, where China alone is represented by over one hundred professional teams [1]. Dota2 focuses on tactics, team coordination and skills, which attracts the largest participants and audiences. Before a game begins, both teams need to go through a hero selection process, which is called *drafting*. In this process, the coach of each team needs to ban or pick heroes alternatively, to form a powerful *line-up* of heroes, which will directly affect the outcome of a game.

Due to the complexity induced by the large hero pool and sophisticated rules, professional Dota2 games require automation to assist the analysis. Especially in the drafting process, given a hero pool of just moderate size, e.g., 40 heroes, the computation complexity using a complete tree search method will be intractable, let alone human analysis [2].

In this line of research, existing literature on drafting selection can be broadly categorized into three types, namely, feature selection, line-up recommendation and drafting process prediction. At present, the features of MOBA games have not been fully explored in the research on the prediction of the drafting, resulting in a low (around 13%) prediction accuracy on key steps, which refers to the action steps that has the most impacts on the line-up.

We focus on the drafting process, thoroughly analyze the stats of line-up drafting process and apply machine learning techniques to predict the heroes to be selected. The overall structure of this paper can be divided into three parts. In the first part, we propose key steps for the use of boxplots for lineup draft, and use CNN, DNN, Bi-LSTM and LSTM to construct the corresponding disorder prediction model, and the prediction accuracy is improved by 7%. In the second part, K-W Test in rank sum test is used to propose the property of identical distribution of adjacent steps and expand the dataset. In the third part, we introduce attention mechanism to improve LSTM prediction accuracy by 1.6%. As shown in Figure 1. The proposed model outperforms existing methods shedding lights on predicting roster drafts and game designs.

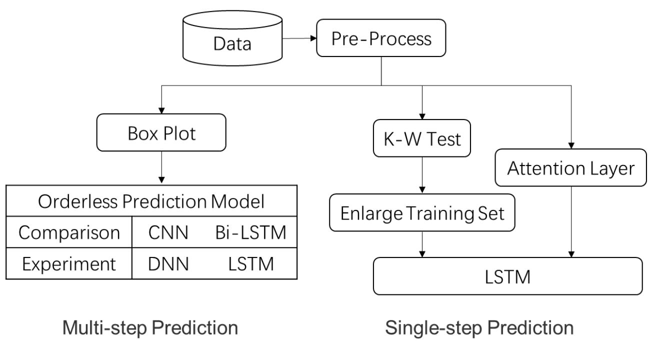

Figure 1. Structure of this paper

*Corresponding Author: Run-Lu Zhu; E-mail: runluzhu@tongji.edu.cn
DOI: 10.53106/160792642023032402019

2 Related Work

Research on the drafting process of MOBA games, mainly includes directions on feature selection, line-up recommendation and prediction of the drafting process.

Considering the complexity of a match in Dota2, most of the current feature selection work focus on representing heroes. Wang, Li, Xiao et al. chose 17 features to represent heroes in Dota2, such as hero damage, etc [3]. Zhang, et al. used word2vec to generate vectors to represent heroes in Dota2, and proposed an improved Bi-LSTM model for drafting recommendation [4]. These efforts show that a hero character is game-specific and hence not the focus of our study.

The line-up recommendation is similarly to the prediction of the drafting process we studied, but ignores the heroes who were banned during the drafting process. On this direction, Hong, Lee, Yang used neural networks to study the best team formation in MOBA games [5]. Similarly, Tanuar et al. proposed a champion recommendation system for League of Legends [6]. Chen et al. designed a long-term value estimation mechanism but also ignored the banned heroes [7].

On the direction of draft prediction, literature [8] and [9] are closest to our research, which predict the heroes selected in the drafting process in Dota2 or other MOBA games. However, these works are all single-step predictions and the accuracy of rolling prediction for the key steps is quite low (around 13%, presented in Sec. 3).

Mortelier and Rioult [10] use the LSTM to predict the number of surviving players in two teams within the next five seconds based on the player trajectories and hero's trajectories in the competition, as well as the characteristics of all the players on the team.

Yang and Pan et al. [11] focused their attention on predicting the outcome of MOBA games and proposed a Two-Stage Spatial-Temporal Network. Although the model could predict the outcome of MOBA games in real time, it ignored the link of determining the lineup before the start of the games.

Zhang et al. [12] proposed a framework to accurately assess player level based on deep learning, which will help us to expand the dataset with match data from non-professional players in the future.

3 An Orderless Prediction Model for Key Steps

Draft is the first thing in MOBA games, which can be played in ordinary and professional game modes. This paper studies the professional game data of Dota2, whose drafting rules are shown in the Table 1, to minimize the influence brought by different player levels. In the drafting phase, Team-0 is the first team to act in draft, while Team-1 is the second; B is short for *ban*, P for *pick*. The draft proceeds from left to right. Note that once a hero is banned or picked by one team, the other team cannot select this hero in its draft any more.

For most MOBA games, there usually exists a small group of heroes that are stronger than other in a particular meta, which are then called meta strong heroes. In the early stages of the line-up draft, coaches tend to ban or pick these powerful heroes. Consequently, the heroes that are selected.

"ban" or "pick" in the first few steps of the roster draft tend to concentrated on a small group of heroes. This concentration phenomenon can be observed throughout the dataset, e.g., the professional game data of version 6.88 shown in Figure 2. The vertical axis represents the number of times a hero was selected (including "ban" and "pick") in this version, and sorted from small to large.

Table 1. The order of picks and bans in professional model

Step	1	2	3	4	5	**6**	**7**	8	9	10	11	12	13	14	15	16	17	18	19	20
Operation	B	B	B	B	P	**P**	**P**	P	B	B	B	B	P	P	P	P	B	B	P	P
Team-0	●		●		●			●		●		●		●		●		●	●	
Team-1		●		●		●	●		●		●		●		●		●			●

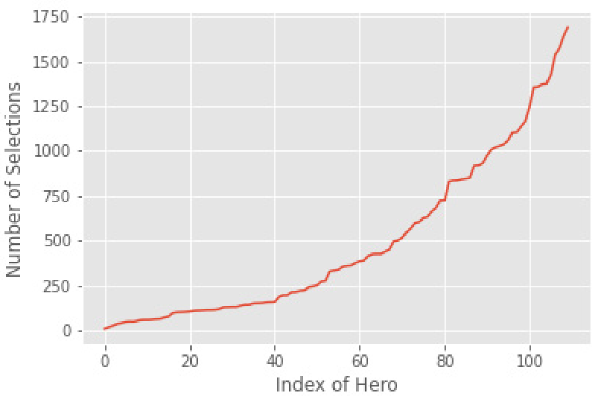

Figure 2. Histogram of hero selection

From Figure 2, the number of selections varies greatly, in which some heroes were selected much more frequently than other heroes, namely, the meta strong heroes. However, the distribution of these strong heroes selected in the line-up drafting process is not revealed by this histogram, so further statistics are needed to understand the relationship between strong heroes and the line-up drafting process.

3.1 Statistical Analysis of Line-up Drafts

In this section, we use descriptive statistics to analyze the distribution and statistical characteristics of the data, which is all professional game data of Dota2 under version 6.88, with a total of 2668 pieces obtained from the OPENDOTA (http://www.opendota.com) platform.

3.1.1 Statistical Metrics

First, we group the data by roster draft steps, with each group containing the number of times each hero was selected on that group's corresponding step. The sorting form is shown in Table 2, where $Num_{i,j}$ represents the number of times the hero i is selected in step j [13].

Table 2. Line-up draft data format

	Step 1	···	Step 20
Hero 1	$Num_{1,1}$	···	$Num_{1,20}$
···	···	···	···
Hero 110	$Num_{110,1}$	···	$Num_{110,20}$

The line-up draft data is then presented in the form of a boxplot, also known as Box-whisker Plot, which can reflect the discrete distribution of one or more sets of data.

In Figure 3, the horizontal axis represents the steps in the draft of the line-up, N_i on the vertical axis represents the number of the i-th hero selected in the N-th step. In order to clearly see the distribution of heroes on the entire line-up draft, this section takes the logarithm of N_i. The inverted triangle represents the mean value; the five-pointed star represents the outliers, the upper bottom, dashed line, and lower bottom of the box represent the upper, middle, and lower quartiles of the data, respectively.

As can be observed from Figure 3, individual outliers appeared on the outside of the box towards the end of the roster draft session. This suggests that some teams may have chosen uncommon heroes as their "secret weapons". Given that these choices are highly correlated with the team's preferences, we exclude these outliers in this paper.

3.1.2 Identifying Key Steps

A key observation from Figure 3 is that the first six steps are concentrated on a small group of heroes, which can be considered as strong heroes in the current meta, but it is not the case for later steps. In the first six steps, the mean is significantly larger than the median, which indicates that outliers concentrate on the larger side. Moreover, the boxes in the first seven steps are wide, indicating that the data in the previous steps fluctuates greatly, while the opposite is true in the later steps. From the seventh step, the hero selection becomes scattered and hence more difficult to predict.

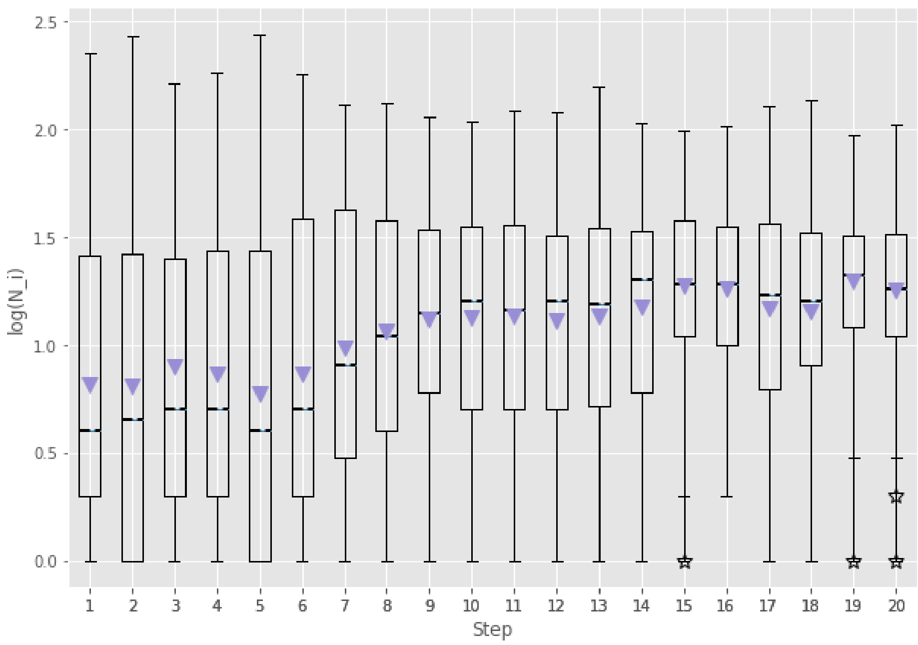

Figure 3. Boxplot of hero selection

Among the first six steps, the first four steps are all "ban" operations, and in these steps, the heroes selected by the coaches do not play. Therefore, the coaches tacitly select a few heroes, that is, the strong meta heroes, so as not to reveal their plans at the beginning.

Excluding these "ban" operations, the distribution of the fifth, sixth and seventh steps are also very concentrated. Among them, the fifth step is the first time that the two teams can choose heroes to play. From the perspective of the coach, it is bound to choose the version strong hero, which is also the reason that the lower quartile line and lower boundary are almost coincidental of the fifth step. The sixth and seventh steps are the other teams' selection operation twice in a row, which we refer to as the key steps. Compared with other steps, the key steps select two heroes required by the line-up together, rather than one ban or pick action. As they construct a continuous selection that composes a large portion (2/5) of the line-up, it is reasonable to believe that coaches make later selections based on the two heroes selected in the key steps, which is in fact un-order. Therefore, this paper proposes to leverage this orderless characteristic in the prediction.

Key steps are important for both teams in the match. For teams that make choices in later steps, predictions about these steps can guide coaches and players to make rational choices. For the team that chooses first, it needs to choose a different hero in the fifth step, which is the last step before the key step that hinders the selection of the team later, which makes the prediction in the key step slightly different. Thus, the predictions made based on their different choices can in turn guide the role selection of the first-choice team in the fifth step. Although we cannot determine the multiple choices of the first team in the fifth step, the prediction methods are consistent, so this paper omits the experiment of this scheme.

3.2 Neural Network Sequence Prediction Model

When choosing heroes at the key steps, the coach, who takes the entire line-up into consideration, does not choose the heroes one by one, but considers them together before picking. The key steps can be predicted by orderless methods. To emulate this human drafting, we consider four categorical machine learning algorithms.

3.2.1 Methods

A. LSTM

LSTM is a neural network topology structure, which can effectively use input data and memorize long-term data in the past, and hence achieved success in many sequence prediction scenarios, such as natural language processing and computer vision [14].

Since drafting is a multi-steps process, we use LSTM rolling prediction, which takes the predicted value of the sixth step as a feature and then predicting the seventh step, as shown in Figure 4. We regard the result by this method as the baseline.

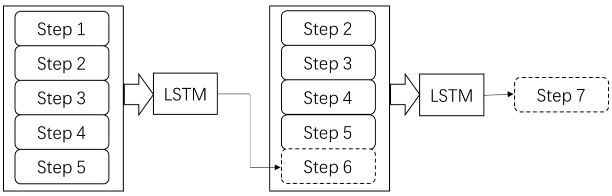

Figure 4. LSTM rolling prediction process

Observing the orderless property of key steps, we design a new prediction model. Specifically, the first five step sequence is used as the input data, and the output of our LSTM is fully connected to two nodes respectively which represent the predicted value of the key steps as illustrated in Figure 5. In addition to building models based on LSTM, we also adopted other methods for comparison.

B. DNN

DNN, the fully connected neural network, also known as the multi-layer perceptron (MLP), is the simplest neural network. The DNN network structure used here includes an input layer, a dense layer, and an output layer [15].

C. CNN

CNN, the convolutional neural network, is also one of the commonly used methods in sequence prediction. Compared with recurrent neural networks, CNN have faster training speed and less model complexity [16]. We construct a basic network based on CNN including an input layer, a convolutional layer, a maximum pooling layer and a global average pooling layer, and the output layer for the key steps that are fully connected [17].

D. Bi-LSTM

Bi-LSTM is a variant of LSTM, which is composed of a forward and a backward LSTM [18]. The output of Bi-LSTM is spliced by the output of forward LSTM and the output of backward LSTM. In this paper, the spliced vectors are fully connected to the sixth and seventh steps.

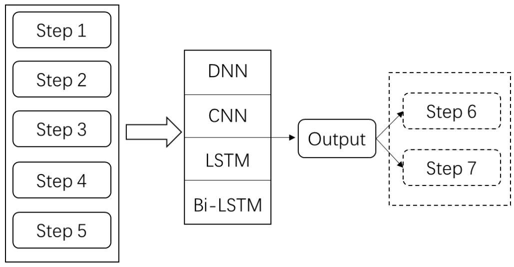

Figure 5. Orderless prediction model

3.2.2 Hyperparameter Settings

The hyperparameters of the Dota2 line-up draft prediction model are listed in Table 3. Values of the hyperparameters are obtained by experimenting on the validation set to achieve the lowest loss.

Table 3. Hyperparameters for prediction models

Hyperparameters	Settings
Dimensions of the hero vector	110
Hidden unit dimension of LSTM	256
Epochs	2
Learning rate	0.01
Loss function	Categorical_crossentropy
Optimization method	Stochastic Gradient Descent
Batch_size	1
Output layer dimensions	110

3.3 Experiment and Result Analysis

This section introduces the experimental results and analysis of the recursive prediction model and the disordered prediction model.

3.3.1 Experimental Evaluation Methods

In this experiment, one-hot coding is used to represent heroes, and the output of the prediction model is also a 110-dimensional vector. Using one-hot coding to represent the game character can avoid the unique related attributes of the game character itself and the particularity of the game itself, which makes the method in this paper generic. Each dimension in the vector corresponds to the probability of the hero being selected. Therefore, in this experiment, the value of each dimension of the output vector is sorted by the probability of the corresponding hero, and the accuracy of the model prediction is calculated using hit rate. Since neural networks generally do not have the function of checking duplicates, we exclude the probability values corresponding to the heroes that have appeared in the sequence when sorting the probability values of the output vector.

3.3.2 Comparison of Results

In this paper, the recursive prediction LSTM model is used as the baseline, and a single-layer LSTM is selected as the network layer. In practice, we find that increasing the number of LSTM layers does not improve prediction accuracy, but it increases the number of parameters and training time exponentially, as shown in Table 4.

It can be seen from Table 4, increasing the number of layers of LSTM reduces the performance of the model, which may be due to the over-fitting caused by the deep number of layers in the network. Therefore, if not specified, the LSTM models adopted in the remaining of this paper are all configured with one single layer.

The prediction results of different neural network models are shown in Table 5. Among them, Rolling-LSTM means that the traditional recursive prediction model which is based on a single-layer LSTM (baseline), Orderless-LSTM refers to the un-order prediction model based on the single-layer LSTM. DNN, CNN, and Bi-LSTM all use the un-order prediction model.

Bi-LSTM has the highest accuracy rate, but requires the most training time. Bi-LSTM has fewer parameters compared to single-layer LSTM, because the dimension of the output vector generated by the forward LSTM and the backward LSTM are both half the length of the vector generated by the single-layer LSTM. CNN has the lowest complexity, but also the lowest accuracy. DNN performs well in both regards, i.e., the complexity of DNN is low and the accuracy is almost to the best, which can be useful in quick or real-time predictions.

Table 4. Recursive LSTM model prediction results

Number of layers	Prediction accuracy	Train-Time/s	Parameters
1	13.26%	55	404,078
2	12.18%	170	929,390
3	12.06%	480	1,454,702

Table 5. Different neural network prediction results

Architecture	Rolling-LSTM	Orderless-LSTM	DNN	CNN	Bi-LSTM
Prediction accuracy	13.26%	20.38%	20.12%	15.36%	21.17%
Train_time/s	55	53	10	9	58
Parameters	404,078	404,078	169,326	113,006	273,006

4 Inter-step Selection Analysis

This section analyzes the relationship between steps on line-up draft. First, the line-up drafting process should be described in statistical language. In all professional games under a version, the number of times the j-th hero selected on the i-th step in the line-up draft is a discrete random variable, denoted by $X_{j,i}$, whose possible values are {1, 2, ⋯, K}, where K represents the number of professional games. The line-up draft is set to a multidimensional random variable as $\left(X_1, X_2, \cdots, X_{20}\right)$ composed of $X = X_i$. Therefore, in a professional game under a specific version, the number of times a hero is selected in each step of the draft is an observation of a multidimensional random variable, denoted as $\left(x_{j,1}, x_{j,2}, \cdots, x_{j,20}\right)$, with $\sum_{j=1}^{110} x_{j,i} = K$.

4.1 Rank Sum Test for Line-up Drafts

Through the study of key steps, this paper concludes that there is disorder property in hero selection of six or seven steps and using the corresponding disorder prediction model can improve the prediction accuracy. A natural thought is whether this property also exists in other steps. From the boxplot that the selection of game characters exhibits a skewed distribution, for which we use nonparametric test to analyze the steps instead of parametric tests. Specifically, the K-W (Kruskal-Wallis) test method, which is one kind of rank sum test in the nonparametric test, is adopted to analyze the relationship between steps [19].

For m independent simple random samples $(X_1, ..., X_{n_i})(i = 1, ..., m)$, all observations for the samples are merged and sorted from small to large, in which the rank is the serial number; $R_i\left(i = 1, \cdots, m\right)$ represents the rank sum of the n_i observations of the i-th sample; the rank sum statistics amount is calculated as:

$$H = \frac{12}{N\left(N+1\right)} \sum_{i=1}^{S} \frac{R_i^2}{n_i} - 3(N+1). \quad (1)$$

If r data in each sample are the same, let $t_i\left(1, \cdots, r\right)$ be the number of occurrences of the i-th observation in each sample in all observations, and the rank sum becomes:

$$H' = \frac{N\left(N^2-1\right)}{\sum_{i=1}\left(t_i^3 - t_i\right)} H. \quad (2)$$

Then we have two hypotheses for line-up drafting, H0: the distribution of heroes being selected is the same; H1: the distribution of heroes being selected differs among steps. We select the P value method to test this hypothesis, in which threshold α is set to 0.05 [20].

The K-W test is performed on all steps in the line-up draft, and the result is H = 146.8122, P = 3.0889e-22 ≪ α, so the null hypothesis can be rejected. That is, there are at least two steps in the line-up draft, and the distribution of heroes being selected is different.

Though not all steps on the line-up draft follow the same distribution, the steps can be grouped to be tested. The K-W test is performed on the line-up drafts after they are grouped from front to back, whose results are shown in Figure 6, where the columns represent the start steps in the group, and the rows represent the end steps, and the value in each cell represents the P value after K-W test for the group which contains steps from start to end.

In the lower triangle area of the Figure 6, the cell with P value < 0.05 is represented by 0, which means that the distribution of the corresponding grouping of this cell is different in at least two steps.

With the sixth and seventh steps as the boundary, the line-up draft can be divided into two groups with different distributions. This also verifies the rationality of this paper to identify the sixth and seventh steps as the key steps. Moreover, the P value between adjacent steps is greater than 0.05, indicating that the hero selection between adjacent steps is identically distributed.

Start step

End step

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
2	0.943	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0.369	0.229	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	0.534	0.463	0.62	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	0.4	0.282	0.175	0.184	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6	0.526	0.423	0.331	0.355	0.246	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7	0.144	0.12	0.112	0.072	0	0.147	0	0	0	0	0	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0.068	0.333	0	0	0	0	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0.253	0.451	0	0	0	0	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0.299	0.699	0.995	0	0	0	0	0	0	0	0	0	0
11	0	0	0	0	0	0	0.339	0.823	0.989	0.88	0	0	0	0	0	0	0	0	0
12	0	0	0	0	0	0	0.446	0.923	0.998	0.983	0.87	0	0	0	0	0	0	0	0
13	0	0	0	0	0	0	0.508	0.958	0.999	0.995	0.974	0.825	0	0	0	0	0	0	0
14	0	0	0	0	0	0	0.428	0.903	0.972	0.932	0.862	0.702	0.574	0	0	0	0	0	0
15	0	0	0	0	0	0	0.079	0.342	0.451	0.351	0.29	0.223	0.231	0.304	0	0	0	0	0
16	0	0	0	0	0	0	0	0.172	0.28	0.227	0.217	0.201	0.288	0.516	0.702	0	0	0	0
17	0	0	0	0	0	0	0	0.235	0.373	0.32	0.313	0.294	0.377	0.489	0.371	0.297	0	0	0
18	0	0	0	0	0	0	0.056	0.304	0.457	0.403	0.39	0.364	0.421	0.457	0.316	0.394	0.822	0	0
19	0	0	0	0	0	0	0	0.109	0.209	0.188	0.2	0.209	0.289	0.375	0.284	0.248	0.172	0.076	0
20	0	0	0	0	0	0	0	0.096	0.205	0.195	0.221	0.249	0.359	0.482	0.398	0.363	0.271	0.191	0.426

Figure 6. Rank sum test results

4.2 Expansion of the Training Set

Based on this property of same distribution among adjacent steps, this paper designs a scheme to expand the training dataset, the small size of which hinders many studies on learning-based automated analysis. Specifically, for the second to nineteenth steps, the hero selection of each step is exchanged with the hero selection of the adjacent steps. In this way, we obtain four different training sets from the original training set.

Then we compare the prediction effects of the corresponding models on the same validation set, and the results are shown in Figure 5. The horizontal axis "Step Number" represents the serial number of the step, the vertical axis "Accuracy" represents the average accuracy rate of the prediction model under ten-fold cross-validation, "base_acc" represents the result of the original training data set, and "bef_acc" represents the result of the augmented dataset after the prediction target step is swapped with the previous step. Similarly, "aft_acc" represents the swapping with the next step, and "all_acc" represents the result of swapping with both the previous and next steps.

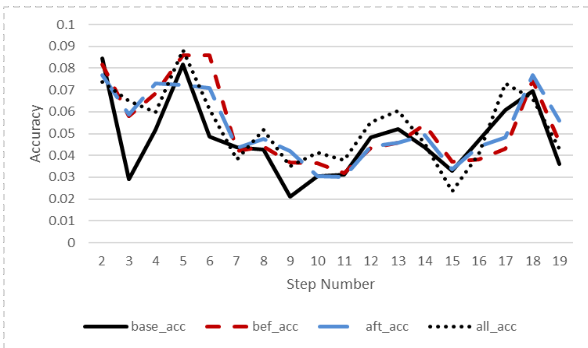

Figure 7. Predictions results

The average value of the prediction accuracy of the four models is shown in Table 6.

Table 6. Four models' average prediction accuracy

Training set	Average prediction accuracy
Original training set	4.8%
the previous step	5.3%
the next step	5.2%
the previous and next steps	5.3%

Combining Figure 7 and Table 6, after exchanging the hero selection on adjacent steps to expand the training set, we see a small improvement (about 0.5%) of the model's prediction ability. Due to the small size of the original dataset, improvement from the expansion is limited. However, we believe that there is room for further improvement when more game data are available.

5 Attention-enhanced Prediction

Attention Mechanism [21], which enables neural networks to select specific inputs, has achieved success in the field of image processing [22], and natural language processing [23]. Inspired by its advantage in sequence prediction problems, we propose to further improve the prediction accuracy by introducing an extra attention layer [24].

5.1 Attention Layer

In this section, we add a self-attention layer into model to further improve the prediction accuracy, as shown in Figure 8.

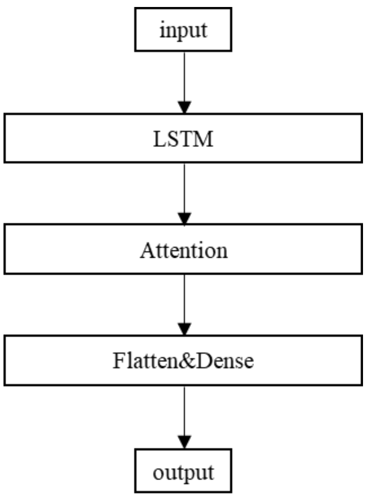

Figure 8. Prediction model with an attention layer

5.2 Result and Analysis

Prediction results of the attention-enhanced prediction model (labeled "attention_acc") are compared with the previous model (labeled "base_acc") in Section 3, as shown in Figure 8. The abscissa represents the target step which contains the second to twentieth steps for the convenience of comparison, and the ordinate represents the average accuracy of prediction on the validation set.

As can be seen from Figure 9, the prediction accuracy of the attention-enhanced model (black line) is higher throughout the steps with an average increase of 1.6%.

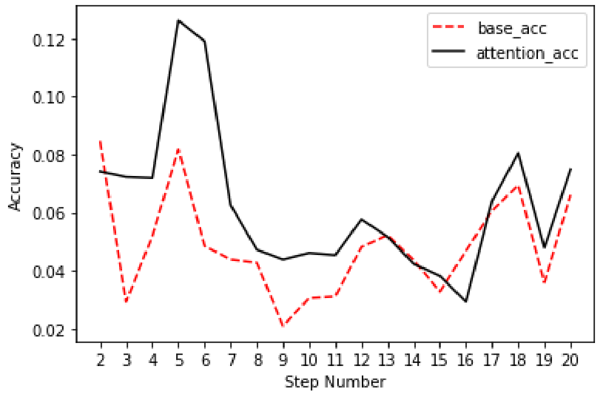

Figure 9. Prediction results w & w/o attention layer

Moreover, we find that the hero selection at most of the second to twentieth steps has a larger distribution of attention values on the first three steps. Two detailed analyses are shown in Figure 10 and Figure 11.

Figure 10 shows one distribution, in which the target prediction step has great attention value on the first and third steps; Figure 11 shows the other distribution, in which the target prediction step has great attention value on the second step.

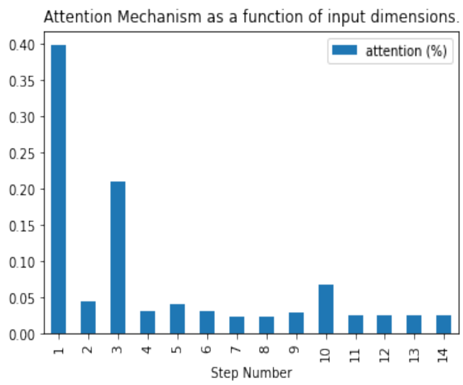

Figure 10. Attention distribution of thirteenth step

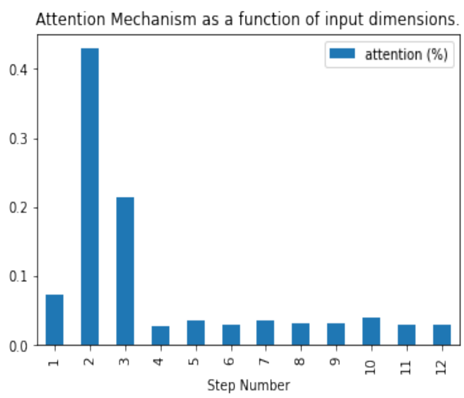

Figure 11. Attention distribution of fifteenth step

We highlight two observations from this result: i) The heroes banned by the two teams in the first three steps of the line-up draft may have established the “frame” of the line-up draft in this game, which can be observed from the attention distribution map of all steps. ii) If one team bans a hero that restraints their line-up, this team may also pick it, preventing the other team from picking that hero. These two observations may seem obvious in the roster draft, but the attention weight shows that the roles that teams ban in the early line-up draft are an important factor in predicting the hero selection.

6 Conclusion

This paper studies the line-up prediction problem for MOBA games. Specifically, we identify the non-game-specific, rule-based factors in the drafting process to tackle the low prediction accuracy problem. Contributions of this paper can be summarized as three parts: First, we identify the key steps in the drafting process through descriptive statistics, by which we propose an orderless prediction model that outperforms existing methods. Second, we find that hero selections at adjacent steps exhibit the same distribution characteristics, leveraging which dataset can be expanded for learning-based analysis. Finally, we use attention mechanism to strengthen the prediction accuracy, whose values also reveal the trend and tactics in the drafting process.

Other popular MOBA games, like *League of Legends* and *Honour of Kings*, differ slightly from *Dota2* in terms of the rules of lineup selection, often due to different game designers’ understanding of first-hand selection. Moreover, these games are similar in terms of how the game is played, how the characters are portrayed, and how matches are won. In addition, the method adopted in this paper only uses line-up draft sequence data and does not involve the attribute values of game characters, so it is a general method and can be easily extended to other MOBA games, which facilitates auto-mated analysis of games, providing useful insights to MOBA gamers, coaches, and game designers.

References

[1] M. Maciej, W. Kempińska, Characteristics of the Chinese Gaming and Esports Market. Applications for Polish Game Manufacturers, *Przegląd Politologiczny*, No. 3, pp. 87-108, September, 2020.

[2] D. Ye, G. Chen, W. Zhang, S. Chen, B. Yuan, B. Liu, J. Chen, Z. Liu, F. H. Qiu, H. S. Yu, Y. Y. T. Yin, B. Shi, L. Wang, T. F. Shi, Q. Fu, W. Yang, L. X. Huang, W. Liu, Towards playing full moba games with deep reinforcement learning, *NIPS’20: Proceedings of the 34th International Conference on Neural Information Processing Systems*, Vancouver, BC, Canada, 2020, pp. 621-632.

[3] N. Wang, N. Li, L. Xiao, G. Yang, Y. Zhou, Outcome prediction of dota2 using machine learning methods, *Proceedings of 2018 International Conference on Mathematics and Artificial Intelligence*, Chengdu, China, 2018, pp. 61-67.

[4] L. Zhang, C. Xu, Y. Gao, Y. Han, X. Du, Z. Tian, Improved Dota2 lineup recommendation model based on a bidirectional LSTM, *Tsinghua Science and Technology*, Vol. 25, No. 6, pp. 712-720, December, 2020.

[5] S. J. Hong, S. K. Lee, S. I. Yang, Champion Recommendation System of League of Legends, *2020 International Conference on Information and Communication Technology Convergence*, Jeju, Korea (South), 2020, pp. 1252-1254.

[6] E. Tanuar, B. S. Abbas, A. Trisetyarso, C. Kang, F. L. Gaol, W. Suparta, Back propagation neural network experiment on team matchmaking MOBA game, *2018 International Conference on Information and Communications Technology*, Yogyakarta, Indonesia, 2018, pp. 240-243.

[7] S. Chen, M. Zhu, D. Ye, W. Zhang, Q. Fu, W. Yang, Which Heroes to Pick? Learning to Draft in MOBA Games With Neural Networks and Tree Search, *IEEE Transactions on Games*, Vol. 13, No. 4, pp. 410-421, December, 2021.

[8] A. Summerville, M. Cook, B. Steenhuisen, Draft-analysis of the ancients: predicting draft picks in dota 2 using machine learning, *Twelfth Artificial Intelligence and Interactive Digital Entertainment Conference*, Vol. 12, No. 2, pp. 100-106, June, 2016.

[9] D. Gourdeau, L. Archambault, Discriminative Neural Network for Hero Selection in Professional Heroes of the Storm and DOTA 2, *IEEE Transactions on Games*, Vol. 13, No. 4, pp. 380-387, December, 2021.

[10] A. Mortelier, F. Rioult, Addressing the Complexity of MOBA Games through Simple Geometric Clues, *Conference on Computational Collective Intelligence Technologies and Applications*, Springer, Cham, pp. 643-649, September, 2022.

[11] Z. Yang, Z. Pan, Y. Wang, F. Cai, S. Shi, S. Huang, W. Bi, X. Liu, Interpretable Real-Time Win Prediction for Honor of Kings--a Popular Mobile MOBA Esport, *IEEE Transactions on Games*, February, 2022.

[12] C. Zhang, K. Wang, H. Chen, G. Fan, Y. Li, B. Zheng, QuickSkill: Novice Skill Estimation in Online Multiplayer Games, *Proceedings of the 31st ACM International Conference on Information & Knowledge Management*, pp. 3644-3653, October, 2022.

[13] B. S. Everitt, D. C. Howell, *Encyclopedia of Statistics in Behavioral Science–Volume 2*, John Wiley & Sons, Ltd, 2021.

[14] M. Cornia, L. Baraldi, G. Serra, R. Cucchiara, Predicting Human Eye Fixations via an LSTM-Based Saliency Attentive Model, *IEEE Transactions on Image Processing*, Vol. 27, No. 10, pp. 5142-5154, October, 2018.

[15] T. N. Sainath, O. Vinyals, A. Senior, H. Sak, Convolutional, Long Short-Term Memory, fully connected Deep Neural Networks, *2015 IEEE International Conference on Acoustics, Speech and Signal Processing*, South Brisbane, QLD, Australia, 2015, pp. 4580-4584.

[16] A. Kroshchanka, V. Golovko, The Reduction of Fully Connected Neural Network Parameters Using the Pre-training Technique, *2021 11th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and*

Applications, Cracow, Poland, 2021, pp. 937-941.

[17] J. Hu, L. Shen, G. Sun, Squeeze-and-excitation networks, *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 7132-7141, June, 2018.

[18] M. A. Istiake Sunny, M. M. S. Maswood, A. G. Alharbi, Deep Learning-Based Stock Price Prediction Using LSTM and Bi-Directional LSTM Model, *2020 2nd Novel Intelligent and Leading Emerging Sciences Conference*, Giza, Egypt, 2020, pp. 87-92.

[19] S. S. Shapiro, M. B. Wilk, An analysis of variance test for normality (complete samples), *Biometrika*, Vol. 52, No. 3/4, pp. 591-611, December, 1965.

[20] ALGORITHM AS R94 APPL, *STATIST*, Vol. 44, No. 4, 1995.

[21] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, I. Polosukhin, Attention is all you need, *Advances in neural information processing systems*, Vol. 30, 2017.

[22] M. H. Guo, T. X. Xu, J. J. Liu, Z. N. Liu, P. T. Jiang, T. J. Mu, S. H. Zhang, R. R. Martin, M. M. Cheng, S. M. Hu, Attention mechanisms in computer vision: A survey, *Computational Visual Media*, Vol. 8, pp. 331-368, September, 2022.

[23] Q. Cheng, H. Li, Q. Wu, F. Meng, L. Xu, K. N. Ngan, Learn to Pay Attention Via Switchable Attention for Image Recognition, *2020 IEEE Conference on Multimedia Information Processing and Retrieval*, Shenzhen, China, 2020, pp. 291-296.

[24] T. Ueda, M. Okada, N. Mori, K. Hashimoto, A Method to Estimate Request Sentences using LSTM with Self-Attention Mechanism, 2019 8th International *Congress on Advanced Applied Informatics*, Toyama, Japan, 2019, pp. 7-10.

Biographies

Run-Lu Zhu is currently a graduate student at the School of Software Engineering, Tongji University, Shanghai, China. His research interests include sequence prediction and artificial intelligence.

Jie Wang is an Assistant Professor with the School of Software Engineering, Tongji University, Shanghai, China. She received her Ph.D. in Computer Engineering from North Carolina State University at Raleigh, NC. Her research interests include mobile/edge computing, mobile big data, performance evaluation of networked systems (including wireless networks, IoT, cyber-physical systems).

Sheng-Jie Zhao (Senior Member, IEEE) received the B.S. degree in electrical engineering from the University of Science and Technology of China, Hefei, China, the M.S. degree in electrical and computer engineering from the China Aerospace Institute, Beijing, China, and the Ph.D. degree in electrical and computer engineering from Texas A&M University, College Station, TX, USA. His research interests include artificial intelligence, big data, wireless communications, image processing, and signal processing.