REINFORCEMENT LEARNING FRAMEWORK FOR EVALUATING FOOTBALL ACTIONS AND PLAYERS

RIKITO ENDO[1]**, KAZUSHI IKEDA**[1]

[1]Graduate School of Science and Technology, Nara Institute of Science and Technology, Ikoma 6300192, Japan
E-MAIL: kazushi@is.naist.jp

Abstract:

Association football (a.k.a. soccer) is one of the most popular sports in the world and thus has become a big business. A football club is keen on signing and contracting good players at a reasonable price depending on the players' contribution. The previous work by Liu and Schulte formulated the problem of evaluating the contribution of a player's play in the framework of reinforcement learning, however, their metrics have not completely succeeded in capturing the overall influence of players on team performance. To improve the performance of evaluation, the present paper proposes some modifications to their method, that is, feature optimization and normalization in the assessment. The effectiveness of our method was confirmed using a real dataset, the 2022–2023 English Premier League season (380 games).

Keywords:

Football player rating; Deep reinforcement learning; State-action value function

1. Introduction

Football is one of the most popular sports in the world. In fact, more than 1.4 billion people watched the final game of the 2022 FIFA World Cup in Qatar, and more than 262 billion people watched the entire tournament through various platforms [1]. The global popularity makes football a large marketplace, and football clubs spend a lot of money to sign players. For example, 7.36 billion dollars were spent on football players' transfers, and 2.0 billion euros were the total prize money for the UEFA Champions League (the European clubs' tournament held in 2022–2023). Thus, club managers and scouts are interested in signing and retaining good players with a reasonable value since a player's quality affects their success both on the pitch and in the field of business. However, it is difficult to evaluate players, and players with high salaries do not perform well in many cases since they are subjectively evaluated based on the experience and intuition of the managers and scouts. Therefore, objective methods to evaluate players are desired.

One approach to an objective metric of players is to focus on a specific type of actions in a game. Power et al. [3] evaluated the probability of the success of a pass and the probability that the pass ends in a goal within 10 seconds. These models make it possible to evaluate the difficulty and the creativity of the passes in the games. As the shot evaluation, Green [2] proposed an xG model that is defined as the probability of the shot's success. Although they work to some extent, these metrics cannot evaluate different types of plays or different position players due to the specificity.

A solution to this problem is to introduce a state-action value function, which is popular in the framework of reinforcement learning. Liu and Schulte took a deep reinforcement learning approach for player evaluation, which can introduce context-aware factors. They first showed its effectiveness in ice hockey [4] and then extended it for football by considering the so-called home advantage [5].

This paper proposes to modify their method to improve performance by introducing the time factor, which is supposed to affect the actions of players. To confirm its effectiveness, we calculated the correlation between the metrics and the salaries of the players.

2. Materials and Methods

2.1. Dataset

A play-by-play dataset was used in this study, which consisted of the logs of actions with the location and the time in all Premier League games in the 2022–2023 season (Table 1). The features used in our study are listed in Table 2, where the duration is the difference between the remaining time of the last

TABLE 1. Play-by-Play Data

minute	second	Manpower	side	team_name	name	action	x	y	outcome	goal
64	49	0	away	Arsenal	A. Sambi Lokonga	pass	62	51	1	0
64	50	0	away	Arsenal	O. Zinchenko	pass	65	40	1	0
64	54	0	away	Arsenal	A. Sambi Lokonga	pass	59	15	0	0
64	58	0	home	Manchester United	Diogo Dalot	pass	24	86	1	0
64	59	0	home	Manchester United	C. Eriksen	pass	27	57	1	0
65	1	0	home	Manchester United	Bruno Fernandes	pass	43	52	1	0
65	5	0	home	Manchester United	M. Rashford	shot	82	52	1	1

action and that of the current action.

2.2. Model Archtecture

Our model was the same as the deep reinforcement learning proposed by Liu and Schulte [5] (Fig. 1), which was named the Two-Tower Long Short Term Memory (LSTM) model. This is an extension of the LSTM model used in [4] so that it can capture the home advantage, that is, the home team and the away team have different models. Each tower used a Stacked LSTM (a.k.a. a multi-layer LSTM) to abstract input data layer-by-layer and to extract advanced features since an LSTM can treat complex situations of a game by summarizing the match history and actions into a final hidden state.

The model is given State (features) and Action in the current Play sequence as inputs and outputs the Q-value, that is, the triplet of $Q_{\text{home}}, Q_{\text{away}}, Q_{\text{neither}}$. The triplet can be regarded as the probabilities of the current episode ending with a home goal, an away goal, and a game-ending without the goal, respectively. The Q value was determined by normalizing the final hidden state mentioned above.

2.3. Model Training

The two-tower neural network was trained with a Temporal Difference (TD) prediction method called SARSA in the same way as Liu and Schulte [5]. More concretely, the training processes are as follows: A dynamic-possession LSTM was used to control the trace length during the training of a function for $Q_{\text{team}}(s, a)$ for the play features observed in the dataset. The output of the model at time-step t was sent from the tower of the team holding the ball to the hidden layer and the hidden layer estimates Q-values for the two consecutive actions and states based on the TD loss function,

$$\mathcal{L}(\theta) = \sum_{team \in T} E\Big[\big(r_{team,t+1} + \hat{Q}_{team}(s_{t+1}, a_{t+1}) \quad (1)$$

$$-\hat{Q}_{team}(s_t, a_t)\big)^2\Big]. \quad (2)$$

2.4. Features

The original model in [5] did not use the end location although the end location has a significant impact on upcoming chances than the starting location in soccer games. Thus, the features of the proposed model were the ones in [5] and the end location defined as the starting location of the following action.

2.5. Evaluation of Players

Each player was evaluated in a similar way to Liu and Schulte [5]. They proposed a metric for the contribution of an action to a goal, named the Goal Impact Metric (GIM), defined as the difference between the Q values before and after the action. For example, the GIM of Action Carry in Table 3) is calculated as

$$\text{GIM} = (\text{Current } Q_{\text{home}}) - (\text{Last } Q_{\text{home}}) \quad (3)$$

$$= 0.334 - 0.288 = 0.046. \quad (4)$$

Each player was evaluated by aggregating the GIMs of the players through a season as was done in [5]. The value is called the rating. The difference of our rating from that in [5] was that ours normalized the aggregated value by the playing time while the original did not.

2.6. Validation

In the previous work [5], the validity of the ratings was confirmed in terms of the actual numbers of goals and assists, as well as the general subjective evaluation. More concretely, the correlation between the GIMs earned by a team and its actual goals was calculated since the GIMs quantify the contribution to the goals.

TABLE 2. Feature Description

Feature	Description
Side	Home, Away
Location	Position on the pitch, scaled from [0, 100]
Action	Type of Action Performed
Duration	Time Spent on the Current Action
Manpower	Difference in the Number of Players on the Field
Goal Impact	Influence of the Action on the Probability of Scoring

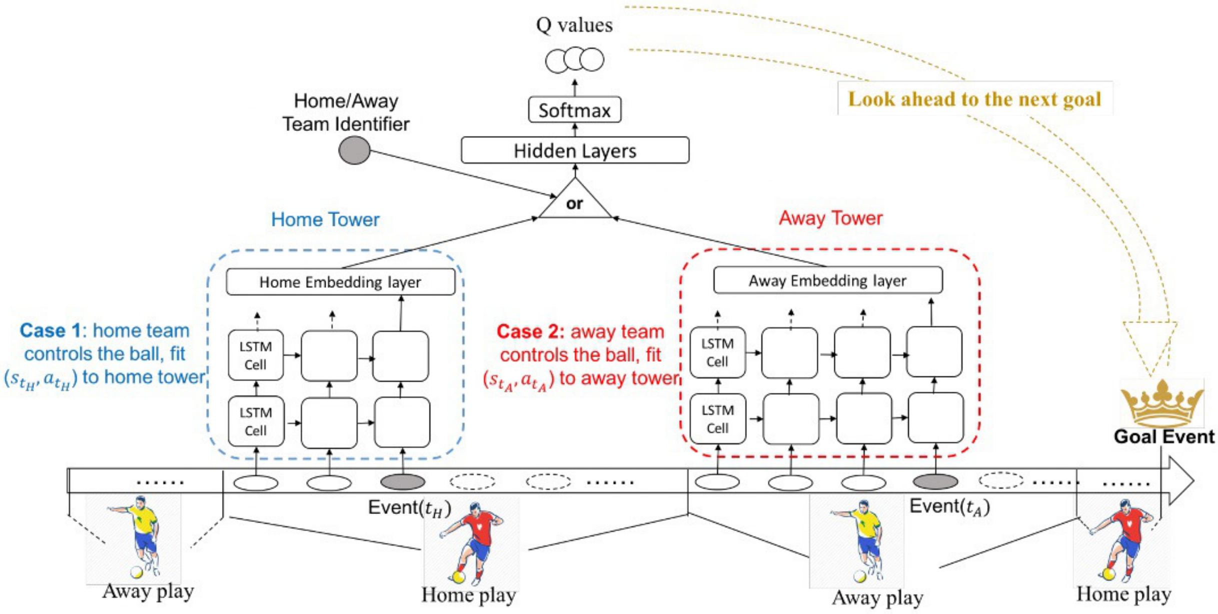


FIGURE 1. Two Tower Deep Reinforcement Learning Model Architecture for Football [5]

TABLE 3. Example of GIM Calculation

Side	Team	Name	Action	Q_home	Q_away	GIM
Home	Manchester United	A. Garnacho	Pass	0.289	0.095	−0.001
Home	Manchester United	Bruno Fernandes	Pass	0.288	0.093	−0.001
Home	Manchester United	A. Wan-Bissaka	Carry	0.334	0.081	0.046
Home	Manchester United	A. Wan-Bissaka	Offensive Duel	0.299	0.092	−0.035

Although the goals are important for the evaluation of players, they are not enough since they are microscopic values. In this study, therefore, players' salaries were also paid attention to as an objective index, that is, the correlation between players' salaries and the GIMs was calculated. This is because the main purpose of the general manager of a football club is to sign good players at a low cost. At the same time, a player's salary must reflect the player's ability to some extent.

3. Result

3.1. Q-Value Transition

To see how the Q-values were affected by the addition of the new feature (end location), we compared the Q-value transitions of the previous model and ours over a game time (Fig. 2). As a result, the transitions had no significant difference in time. The Q-values for both home and away went down in time and they took a high value at the moment of scoring.

3.2. Correlations between GIMs and Goals

To see how much the index reflects the goal possibility, we calculated the correlation between the GIMs per action and the goal numbers (Fig. 3). Our proposed model had a statistically significant positive correlation ($R = 0.520$, $p = 0.0022$), while the previous model did a negative one ($R = -0.344$) although it was not statistically significant ($p = 0.174$) under the condition of $\alpha = 0.05$. The same tendency was found even when we calculated the correlation between the goal numbers and the total GIMs through the season ($R = 0.482$, $p = 0.032$, for our model; $R = -0.245$, $p = 0.290$, for the previous model).

3.3. Correlations between GIMs and Salaries

To see how much the index reflects players' evaluation from the economical viewpoint, we calculated the correlation between the GIMs per action and the players' salaries (Fig. 4). Our proposed model had a higher correlation ($R = 0.230$) than the previous model ($R = -0.057$), which is statistically significant ($p < 10^{-6}$). The same tendency was found even when we calculated the correlation between the GIMs per 90 mins and the players' salaries ($R = 0.150$ for our model, $R = -0.031$ for the previous model; $p < 10^{-4}$).

4. Discussion

The goal of our study is to develop an objective evaluation index for football players using a statistical model for estimating the value of each action under a given status. Thus, we employed the deep reinforcement learning framework and proposed to improve the previous model in [5], that is, adding the end location of the action to the feature set and normalizing the evaluation by the playing time. Our proposed method showed significantly higher correlations with both the numbers of goals and players' salaries.

Although our model seems more suitable to evaluate players than the previous model, there are some limitations. One is that the dataset in this study contains actions related to the ball but does not contain information about the positions of the other players, which is important for tactics. To overcome this problem, more informative datasets such as the tracking data of players are necessary.

In addition, our model assigns high Q-values only to the actions immediately prior to the decisive chance of the match. That is, the actions unlikely to directly lead to a goal is almost neglected. For example, even if a defender nullifies the defense of an opposing forward player and passes to a midfielder, it is not evaluated as a great action. A possible solution to this problem is to introduce other rewards for events that more frequently occur than goals, such as shots for the goal mouth or entering the penalty area. However, soccer games are not so simple as such an idea works so well. That is why this problem is interesting and to be studied in the future.

References

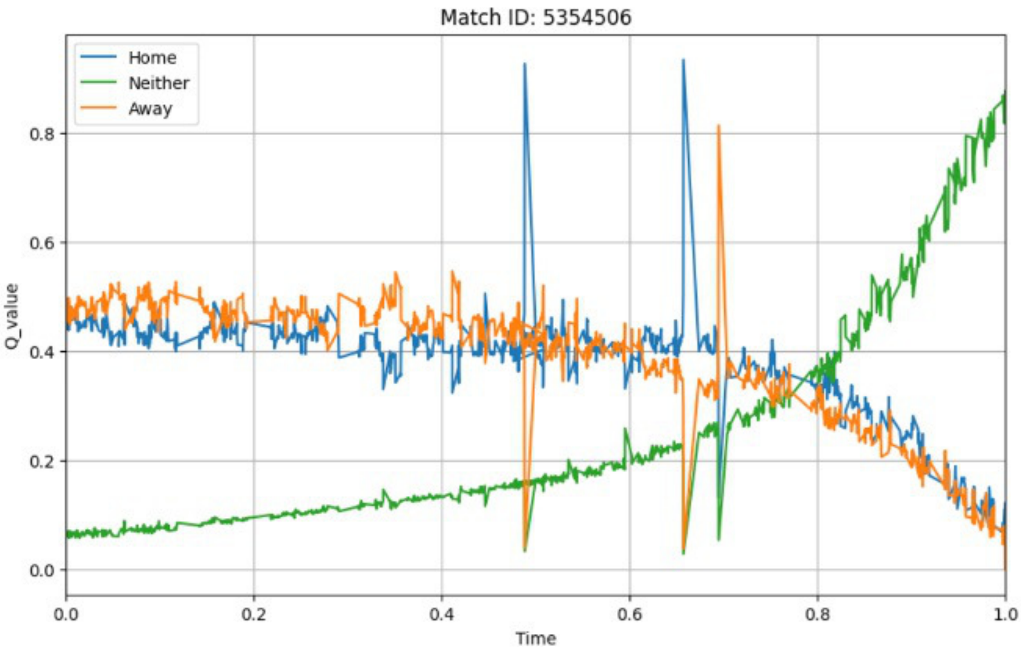
[1] FIFA World Cup Qatar 2022 in Numbers, `https://publications.fifa.com/en/annual-report`

[2] S. Green, "Assessing the performance of Premier League goalscorers," OptaPro Blog, 2012.

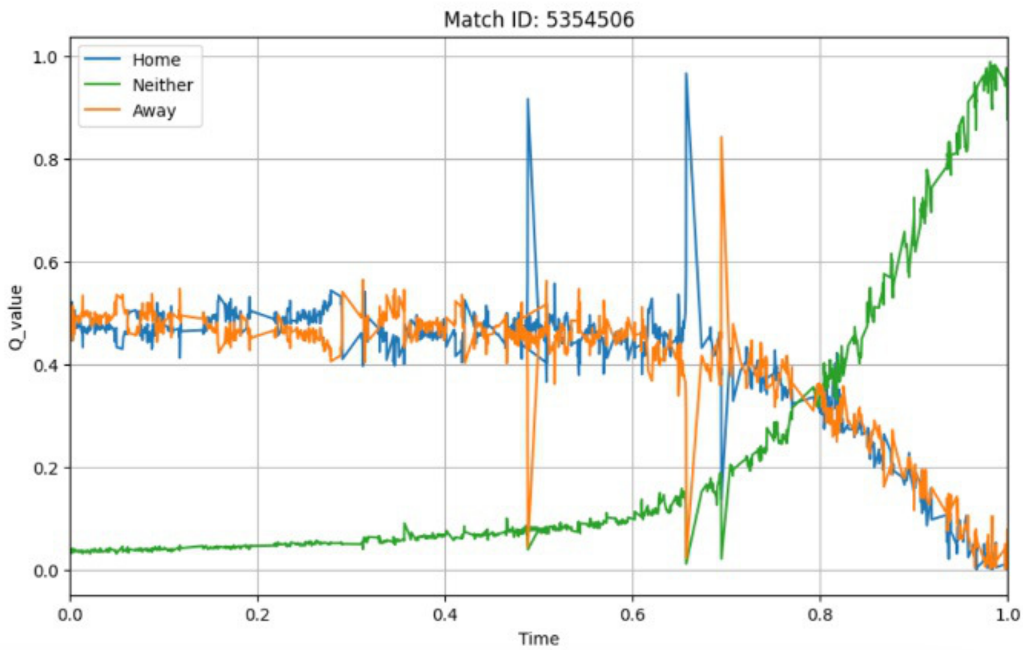
[3] P. Paul, et al., "Not all passes are created equal: Objectively measuring the risk and reward of passes in soccer from tracking data," Proc. 23rd ACM SIGKDD, 2017.

[4] G. Liu and O. Schulte, "Deep reinforcement learning in ice hockey for context-aware player evaluation," arXiv:1805.11088, 2018.

[5] G. Liu, et al., "Deep soccer analytics: learning an action-value function for evaluating soccer players," Data Mining and Knowledge Discovery, vol. 34, pp. 1531–1559, 2020.

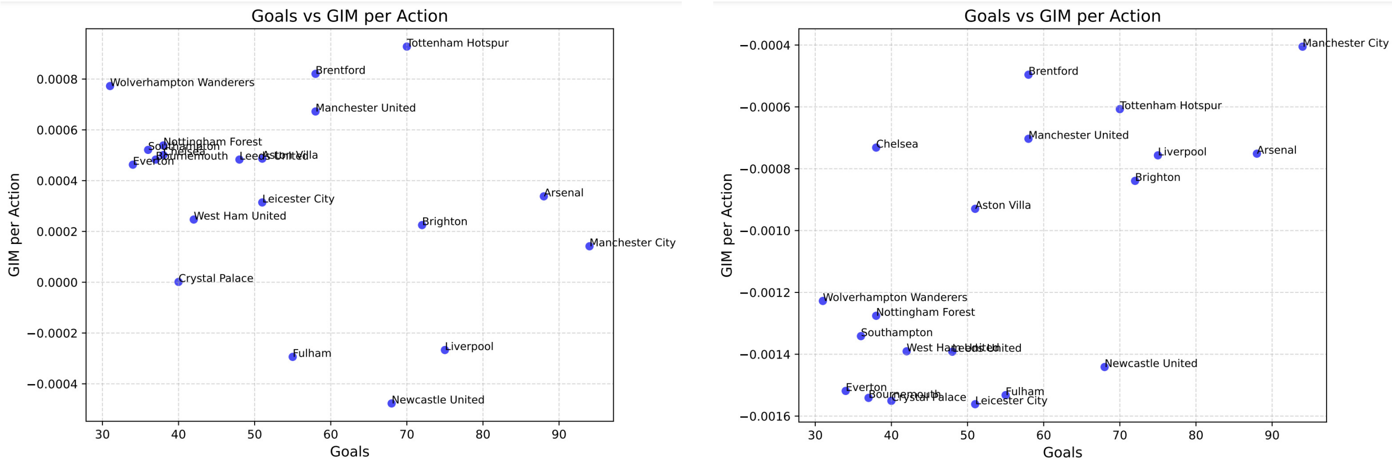


Previous model [5]



Our model

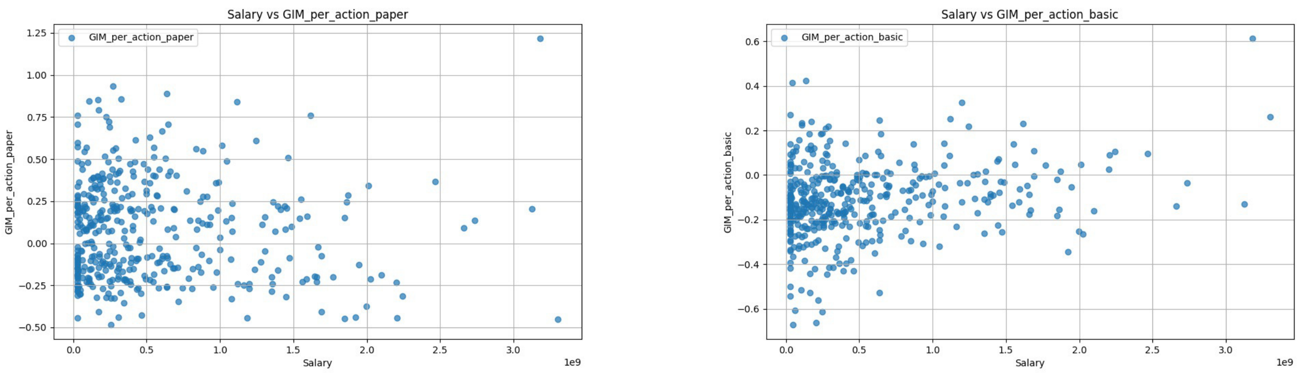
FIGURE 2. Comparison of Q-value changes



Previous model [5] ($R = -0.344$)

Our model ($R = 0.520$)

FIGURE 3. GIMs per action and the numbers of goals



Previous model [5] ($R = -0.057$)

Our model ($R = 0.230$)

FIGURE 4. GIMs per action and the salaries