Team cooperation research based on complex network method

Junxiao Wang1, Long Fan1,2 and Zhaoyang Wang1

1Communication University of China, No.1 Dingfuzhuang East Street, Chaoyang District, Beijing, China
2Email: fanlong@cuc.edu.cn

Abstract. Football is one of the sports that can reflect the importance of teamwork. The passing on the football field can be regarded as a closely connected network. This paper takes Everton as the research object, and a passing network is set up with players as nodes. Based on the method of complex network, this paper summarizes the tactical style of the team from the perspective of the whole season. At the same time, this paper selects two representative games and analyzes them dynamically according to minutes, and finds out the problem that the team does not implement the passing tactics in place. In addition, this paper constructs node performance evaluation indicator system, from three aspects of attack, transmission and defense, using AHP to determine the weight, in order to evaluate player's personal ability. According to the network model and performance evaluation system, we give some suggestions from the formation, tactics and players. The 4-2-3-1 formation get the highest evaluation. The tactics of passing and defending counterattack should be designed. In addition, the Everton is necessary to introduce some midfields and left defenders.

1. Introduction
Our conceptual understanding of team success has advanced significantly over the past 50 years. One of the most useful settings for exploring the team process is team sports. Basketball, volleyball and other ball sports are an important part of it. Duch J (2010) and others constructed indicators through quantitative analysis of player pass success rate and other factors to quantify the performance of individual players in teamwork [1]. Bin Liao (2017) and others studied the quality of the pass and constructed a pass network [2]. In the current field of ball team sports research, football-related research has been favored by many scholars. CintiaP (2015) and others established H indicators by measuring factors such as the average player passing volume of the team, and performed performance analysis on the team as a whole [3]. Bekkers J (2017) and others studied the mobility of players and teams [4]. Gursakal N(2018) and others used complex network knowledge to study the network pattern in football [5]. Buldu JM (2019) and others built a passing network, and analyzed the passing network statically and dynamically by calculating the team's average shortest distance, clustering coefficient and other indicators [6]. Pappalardo L (2019) and others studied the changes in space and time in football games, evaluated team performance, and further found tactical strategies [7].

Team success is much more than the sum of the abilities of individual players. Rather, it is based on many other factors that involve how well the teammates play together. This paper takes the football team cooperation as an example to analyze the cooperation strategy in the game in order to contribute to the cooperative research of social groups.
Firstly, this paper uses multiple-scale variables to consider players passing quality, and use entropy method to calculate the weight of each variable to find out the best players in each position. Then, this paper introduces the network attribute parameters such as average centroid and average clustering coefficient to summarize the team's tactical style from the perspective of the whole season. In addition, this paper selects two representative games and makes dynamic analysis according to minutes. Afterwards, $H_W$ is regarded as the performance index of personal cooperation in this paper. $H_W$ is divided into three aspects: attack, defense and transmission, which are calculated by the weighting of its constituent factors. The weight of each factor is determined by AHP. Finally, we analyzed and summarized the tactical style and common strategies of the team.

2. Methods

2.1. Passing network model

2.1.1. Passing quality. The passing quality reflects the importance of the player in the team. By analyzing and evaluating the factors of a player’s passing quality, it can be concluded that the player's passing quality can be reflected by the number of successful passes, shots, long passes, crosses and smart-passes.

**Entropy Weight Method** Therefore, we introduce an evaluation system related to passing quality, using the entropy weight method to calculate the weight value occupied by each factor, so as to obtain the importance of each player. The entropy weight method determines the weight of the indicator by the information entropy, that is, the value of the information. Normalize the original data to obtain the matrix:

$$y_{ij} = \frac{x_{ij}}{\sum_{i=1}^{n} x_{ij}}; \ 0 \leq y_{ij} \leq 1 \quad (1)$$

We can get the information entropy of the indicators:

$$e_j = -k \sum_{i=1}^{n} y_{ij} \ln y_{ij} \quad (2)$$

Where the constant $k$ is related to $m$, the number of system samples: $k = (\ln m)^{-1}$

Since the information entropy $e_j$ can be used to measure the utility value of the indicators, when it is completely disordered, $e_j = 1$, and the utility value of the information on the comprehensive evaluation of $e_j$ is 0. So the information utility value of the indicators $h_j$ depends on the difference between the index's information entropy $e_j$ and 1.

$$h_j = 1 - e_j \quad (3)$$

The entropy weight method is used to estimate the weight of each indicator, and its essence is to calculate with the value coefficient of the indicator information. The higher the value coefficient, the greater the importance to the evaluation, so does the weight of the indicators.

$$w_j = \frac{h_j}{\sum_{i=1}^{n} h_j} \quad (4)$$

2.1.2. Team position. The average centroid of the team correspond to the average position of all passes of the network. The closer the center of mass is to the frontcourt, the more opportunities the team has to goal and the more advantage they have on the field. The center of mass can be obtained by averaging the coordinates of the ball each time the player handles the ball.

**The Average Centroid** The average centroid of the team correspond to the average position of all passes of the network. The closer the center of mass is to the frontcourt, the more opportunities the team has to goal and the more advantage they have on the field. The center of mass can be obtained by averaging the coordinates of the ball each time the player handles the ball.

$$C_x = \frac{\sum_{i=1}^{n} x_i}{n} \quad (5)$$
\[ C_y = \frac{\sum_{i=1}^{n} y_i}{n} \] (6)

x, y represents the coordinates of each contact with the football, \( C_x, C_y \) represents the average centroid of the team after handling the ball for n times.

**Average Shortest Path** We define the average shortest path of the whole team as:

\[ d = \frac{1}{N(N-1)} \sum_{i<j} D_{ij} \] (7)

N refers to the number of players on the court. It shows how the players pass like a team on the court.

**Average Clustering Coefficient** Since the network is weighted, we not only account for the number of nodes connected between them but how the link weights are distributed. This is the case of passing networks, where the number of passes between pairs of players is not constant. In this way, we use the weighted clustering coefficient \( C_w(i) \) to measure the likelihood that neighbours of a given player \( i \) will also be connected between them.

\[ C_w(i) = \frac{\sum_{j,k} w_{ij} w_{jk} w_{ik}}{\sum_{j,k} w_{ij} w_{jk}} \] (8)

Where \( j \) and \( k \) are any two players of the team and \( w_{ij} \) and \( w_{jk} \) the number of passes between a third player \( i \) and both them. So we get the clustering coefficient of the whole network by averaging \( C_w(i) \) over all players:

\[ C(i) = \frac{1}{N} \sum_{i=1}^{N} C_w(i) \] (9)

It is a measure of local robustness.

2.1.3. **Node performance evaluation indicator system.** In the previous passing network model, the Everton corresponds to the pass network while the nodes correspond to each player. We analyze the main factors that affect the change of the passing network and nodes in a match and use these factors to establish our performance evaluation system from the node levels. Node performance evaluation indicator system consists of offense, defense, and transmission ability. As shown in Figure 1.

**AHP** This paper uses the AHP to calculate personal \( H \) index. The system for evaluating individual offensive capabilities is based on the internal logical relationship, and the evaluation index is used as a representative to form an orderly hierarchical structure. The indicators of each layer are analyzed in pairs for comparison. We construct a judgment matrix \( A = \{a_{ij}\}_{m \times n} \), calculate the maximum characteristic root \( m \times n \) of the comparison judgment matrix \( A \), and use \( m \times n \) to solve the characteristic equation [8].

\[ AX = \lambda_{m \times n} X \] (10)

And obtain the corresponding eigenvector \( X \). After \( A \) passes the consistency check, we normalize the eigenvector and obtain the weight of each index.

\[ w = \left\{ \frac{x_1}{\sum_{i=1}^{n} x_i}, \frac{x_2}{\sum_{i=1}^{n} x_i}, \ldots, \frac{x_n}{\sum_{i=1}^{n} x_i} \right\} = \{w_1, w_2, \ldots, w_l\} \] (11)

**Figure 1.** Node performance evaluation indicator system.
3. Empirical analysis

3.1. Introduction of data

The data in this article comes from the Soccer Data Challenge initiative (https://sobigdata-soccerchallenge.it/) and, to the best of our knowledge, it is the largest collection of soccer-logs ever released to the public. The collection covers European seven prominent male soccer competitions and records in detail the events that occur in each football match. The data we choose in this article consists of the detailed data of the Everton team in the Premier League of the 17-18 season, including all 38 games they played against their 19 opponents (they played each opposing team twice). Overall, the data covers 23,429 passes between 366 players (30 Everton players, and 336 players from opposing teams), and 59,271 game events.

Using the player's position on the field and the order of appearance to label each player, forwards are recorded as $F_1, F_2$, midfielders are recorded as $M_1, M_2$, and defenders are recorded as $D_1, D_2$, etc.

3.2. Model building and analysis

3.2.1. Passing quality. According to the principle of entropy method to determine the weight, the weight of each element can be obtained as shown in Table 1.

| Indicator         | Weight |
|-------------------|--------|
| Successful Pass   | 0.2107 |
| Smart-pass        | 0.1885 |
| Shot              | 0.1937 |
| Cross             | 0.2065 |
| Long Pass         | 0.2047 |

Table 1. Each indicator's weight.

| Player | Index   | Player | Index   | Player | Index |
|--------|---------|--------|---------|--------|-------|
| M1     | 359.950 | D1     | 220.322 | F2     | 278.451|
| M3     | 255.003 | D5     | 205.384 | F1     | 90.636 |
| M6     | 178.132 | D4     | 204.419 | F6     | 72.570 |
| M4     | 149.466 | D3     | 188.537 | F5     | 60.477 |
| M12    | 50.354  | D2     | 150.571 | F4     | 52.861 |

Table 2. Important index of top 5 players on each position.

From this, we get the important index of each player, which can be reflected in the passing network. (See Table 2 to have the important index of top 5 players on each position) The larger the important index, the larger the corresponding graph.

It can be seen that the defender's passing ability is relatively average and generally maintains a high level. There is a passing master in both middle-field and frontcourt (M1, F2), but each has its own hidden dangers: Everton has 13 midfielders, but most of them do not have the ability to pass the football successfully all the time, or not getting enough playing time, and no more forwards have great passing ability except $F_2$.

3.2.2. Static analysis of passing network. Putting the required data into the passing quality and average shortest distance model, based on the analysis and calculation above, we have a static passing network.

Model in Figure 2:

The point represents the player. The position is the average position of the player on the field this season. It is divided into goalkeeper (green), defender (yellow), midfielder (blue) and forward (red). The size of the point and player label scaling strictly according to the player importance index.
The connection represents the pass between players, the connection thickness is strictly scaled according to the number of passes. As a directed graph, the arrow points to the player who receive the ball.

Figure 2. Static passing network model in the season.

Based on Figure 2, we summarize the tactical style of the Everton:
Relying on the center axis transmission from center defender to midfielder to center forward.
The offense is sorted through a certain midfield core, but the connection between the midfield and the forward is weak, even weaker than the connection between the backcourt and the forward.
After passing to the midfield core, the ball is passed to the right side, and pass cross to start a side attack. The left side is weaker than the right.
The goalkeeper often forms a direct connection with the striker, creating opportunities by long-range-pass punches, which is very aggressive.

3.2.3. Dynamic analysis of passing network. We selected two games (22nd and 30th games) with different results under the same coach, formation (4-2-3-1) and Home-away conditions. (See Figure 3, Figure 4 to get the dynamic indicators of the two games) We construct the nodes every 15 touches. From the perspective of the passing network, we consider the changes of some time-related variables and compare the similarities and differences between the two games in detail. Here is our dynamic analysis of the network model:
The same point: The Everton’s clustering coefficient is always at a disadvantage in most cases, regardless of winning or losing, which means that the Everton is not good at passing the football. The players are unable to execute the coaches’ tactics, resulting in bad cooperation between team members.
Differences:
- In the case of 2-0, the team's average centroid changes a lot, and at the same time, the change rate is faster than the other match, indicating that the area covered by the team's offense is larger.
- In the case of 2-0, the team's average center of mass is biased to the forward field, while the opponent's average center of mass is at the centerline. At 0-2, the team's average center of mass and the opponent's center are near the centerline.
- In the case of 0-2, the average clustering coefficient of the team continued to be at a low level for many times, indicating the Everton's passing ball has been stagnated for a long time.
3.2.4. Node performance evaluation. According to the principle of determining the indicator weight by analytic hierarchy process, it is calculated that the offense judgment matrix $A_1 = \begin{bmatrix} 1 & 0.33 & 0.2 \\ 3 & 1 & 0.2 \\ 5 & 5 & 1 \end{bmatrix}$, and the $H_w$ matrix $A_2 = \begin{bmatrix} 1 & 7 & 9 \\ 0.14 & 1 & 3 \\ 0.11 & 0.33 & 1 \end{bmatrix}$.

Table 3 and Table 4 are the indicator weights that constitute the offensive ability and the indicator weights that constitute personal performance $H_w$.

**Table 3.** Weight of offensive ability.

| Indicator                    | Weight |
|------------------------------|--------|
| Accelerating ability         | 0.2107 |
| Place kick ability           | 0.1885 |
| Frontcourt attack ability     | 0.1937 |
Table 4. Weight of $H_w$.

| Indicator | Weight |
|-----------|--------|
| Transmission | 0.068 |
| Offensive | 0.776 |
| Defence | 0.154 |

According to the weight of each indicator, we obtain the attributes of the 30 players in Everton, analyze the influence of nodes on the network, summarize their respective characteristics and disadvantages, and calculate their comprehensive capabilities, that is, what they do to contribute the team. Table 5, Table 6, Table 7 are the top 3 players with $H_w$ values in defenders, midfielders and forwards respectively.

Table 5. Top 3 and average $H_w$ of defender.

|       | Transmission | Offensive | Defence | $H_w$ |
|-------|--------------|-----------|---------|-------|
| D1    | 851          | 583       | 51.93   | 188.93|
| D5    | 625          | 410       | 89.62   | 175.92|
| D4    | 568          | 252       | 98.63   | 154.54|
| Average | 619.4       | 548.4     | 89.45   | 196.84|

Table 6. Top 3 and average $H_w$ of midfielder.

|       | Transmission | Offensive | Defence | $H_w$ |
|-------|--------------|-----------|---------|-------|
| M1    | 1255         | 729       | 149.24  | 314.79|
| M6    | 508          | 512       | 134.23  | 218.35|
| M3    | 887          | 487       | 87.4    | 204.07|
| Average | 297.15      | 245.92    | 47.3    | 95.191|

Table 7. Top 3 and average $H_w$ of forward.

|       | Transmission | Offensive | Defence | $H_w$ |
|-------|--------------|-----------|---------|-------|
| F2    | 859          | 491       | 157.38  | 257.12|
| F1    | 238          | 897       | 65.9    | 206.42|
| F4    | 125          | 435       | 44.05   | 110.15|
| Average | 394.83      | 374.16    | 71.86   | 140.81|

3.3. Tactical analysis

We can know that our team is most accustomed to using 4-2-3-1 and 4-3-3 formations. The most commonly used offensive methods are long pass and wing cross. To check whether these strategies are generally effective, we take Tottenham Hotspur as an example to analyse the passing network. Everton lost to this opponent twice this season. The comparison of the passing networks of the two teams is shown in Figure 5 and Figure 6.

The Everton used the 4-2-3-1 formation in this game, while the opponent used a formation similar to 3-5-1-1. Therefore, the Everton put $D_5$, which is often on the right, to the left to strengthen the offense. The Everton still used the long-range-pass punches between $G_1$ and $F_3$ and wing crosses as the basic offensive methods, but the opponents defensive coverage area is larger (determined by the area of the passing network), these two conventional attack methods had little effect.

The Everton's $H_w$ is similar to its opponent. Reflected in the number of shots and corners, the Everton is even higher than the opponent, which shows that the Everton performance in this game is acceptable.

Analyse: Both teams used the 4-2-3-1 formation in this game, but the tactical execution results are very different:
The opponents center of mass is very forward, causing the Everton to be tired to defend. The long-range-pass punches and wing crosses are not well executed.

For it is away combat, in addition to the core midfielder M1, the Everton midfield and frontcourt players performed significantly weaker than the home field. Relying too much on core performance caused the liquidity and rhythm of the passing network in this game to be disrupted, and the $H_w$ is significantly lower than the opponent.

To sum up, 4-2-3-1 is the most frequently used and most effective formation of the Everton. However, it is easy for opponents to make targeted arrangements when the Everton relies only on long-range-pass punches and wing crosses.

\begin{figure}[h]
\centering
\includegraphics[width=0.8\textwidth]{figures/everton.png}
\caption{(a) Everton's passing network in match 4. (b) the opponent's passing network in match 4.}
\end{figure}

\begin{figure}[h]
\centering
\includegraphics[width=0.8\textwidth]{figures/everton.png}
\caption{(a) Everton's passing network in match 23. (b) the opponent's passing network in match 23.}
\end{figure}

4. Conclusions
This article takes the Everton of the 17-18 season in the Premier League as the research object, and uses players as nodes to establish a pass network. First, we use the average shortest path in the complex network, the average clustering coefficient and other parameters to analyze the overall tactical style of this team to explore the defects and improvements in cooperation. Then we establish the player's personal performance indicator model to evaluate the player's personal ability.

We found that the team's 4-2-3-1 formation has the highest evaluation. The long-pass punches is commonly used, but it is easily targeted by opponents; at the same time, this team has a high level of defense, but the offensive method is relatively simple; midfielder M1 is the core, but overly dependent on this player; and the player's personal ability is weak on the left side. We compared these conclusions with the team's real performance in the 17-18 season and the professional evaluation, and the results are basically consistent. According to the analysis above, we suggest that the team should...
further improve the passing ability; at the same time, they should consider introducing excellent midfielders and left-wing players to share midfield pressure.

Although we have analyzed separately from the perspective of team members' capabilities and the overall strength of the team, from the perspective of team dynamics and social network analysis, team atmosphere, leadership style, and personal psychological status should be considered.

Reference

[1] Duch J, Waitzman J S and Amaral L 2010 Quantifying the performance of individual players in a team activity[J] *PloS one* 5(6)

[2] Liao B, Zhang T, Guo B L, Yu J, Biu Y F, Zhang X G and Liu Y 2017 Research on passing quality quantification based on graphX passing network [J]. *Computer Science* 44(12) 175-182 + 201

[3] Cintia P, Giannotti F, Pappalardo L et al. 2015 The harsh rule of the goals: Data-driven performance indicators for football teams[C]/2015 IEEE International Conference on Data Science and Advanced Analytics (DSAA). IEEE 1-10

[4] Bekkers J and Dabadghao S 2017 Flow Motifs in Soccer: What can passing behavior tell us?[J]. *Journal of Sports Analytics* (Preprint) 1-13

[5] Gursakal N, Yilmaz F M, Cobanoglu H O et al. 2018 Network motifs in football[J]. *Turkish Journal of Sport and Exercise* 20(3) 263-272

[6] Buldu J M, Busquets J and Echegoyen I 2019 Defining a historic football team: Using Network Science to analyze Guardiolas FC Barcelona[J]. *Scientific reports* 9(1) 1-14

[7] Pappalardo L, Cintia P, Rossi A et al. 2019 A public data set of spatio-temporal match events in soccer competitions[J]. *Scientific data* 6(1) 1-15

[8] Wang J and Zhang J S 2001 Comparing several methods of assuring weight vector in synthetical evaluation [J]. *Journal of Hebei University of Technology* (02) 52-57