ISSN 2618-1576 Vol. 3, Issue 3: 10-16, DOI: 10.25236/FSR.2021.030303

Football team cooperation strategy based on directional weighted network

Xinjue Li

School of Civil and Environmental Engineering, Harbin Institute of Technology, Shenzhen, Shenzhen, Guangdong, 518000, China

Abstract: *Football is a typical team sport. A successful football game often requires tacit cooperation between players. In this paper, the method of block modeling was adopted to discuss how to evaluate teamwork. After analyzing the impact of teamwork competition results, we give specific strategies that can improve teamwork and put forward our own views on the general model of teamwork.*

Keywords: *directed weighting network, LM--BP neural network, structure strategy, team cooperation model*

1. Introduction

Facing the increasingly complex social environment, people are faced with more and more challenges, but it is impossible to effectively deal with these challenges only by relying on individuals. Teamwork is an important way to solve challenging problems at present. How can we judge the success of teamwork? What kind of collaboration can increase the chances of success? Many researchers describe measures of teamwork. One of the most effective ways to explore the team process is team sports. Under certain rules, players from both sides of the competition fight for victory through various actions, cooperation, strategies and a series of other ways. However, the success of a team is not simply determined by the superposition of individual abilities, but also depends on the connections between players, which may have a great influence on the final result of the whole game. In football, players are closely related to each other, so it is necessary to explore the ultimate impact of various interactions between players on the outcome of the game. Based on the conclusion, adjust the team's strategy to maximize the victory in the fierce competition.

To help the Huskies study their players on the field for better results later in the season, we need to analyze 38 games against 19 other teams from the previous season. Through the analysis of all kinds of game events of the team participants in each game, including the players on the field and the substitutes on the field, this paper discusses the influence of the interaction between the players on the result of the whole game.

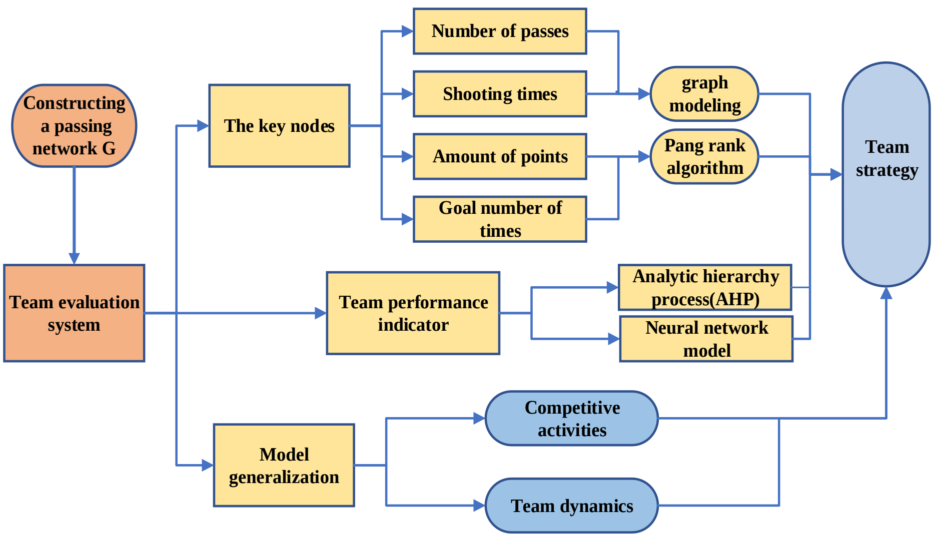

Figure 1

ISSN 2618-1576 Vol. 3, Issue 3: 10-16, DOI: 10.25236/FSR.2021.030303

2. A passing network model based on key nodes

In order to better describe the position relationship of players on the field, we first set up a network to represent the real complex field environment with the network. It is assumed that each player is a node, and the passing between players constitutes their connection. Then quantify how tight the nodes are by how many connections they have, and pick out the key points, the key players in the team.

2.1 Establishment of passing network

2.1.1 Basic structure parameters of the figure

A graph *G* refers to a binary group *(V(G),E(G))*, in which the non-empty finite set is called the vertex set. *E(G)* is the set of disordered or ordered element pairs in the vertex set *V(G)*, which is called the edge set. The set of points *N* and the set of edges *E* constitute a graph, represented as *G=(N,E)*. Where, *N* is the number of nodes in the graph, that is, the number of players on the field. *E* is the number of edges in the graph, that is, the number of passes between each player. If each edge of the graph is given a real number *w(e)*, call *w(e)* the weight of edge *e*, *G* with the weight of the edge. It's called a weighting graph [1].

Because the passing of each player in the match has the directivity, so the sides have the directivity. Meanwhile, the number of passing between different players is different, so the weight of each side is not the same. The graph theory algorithm is used to set up the passing network quickly and intuitively.

You can use adjacency matrices to visualize graphs. The adjacency matrix of a directed weighting graph, and

$$a_{ij}=\begin{cases} w_{ij}, & \text{if}(v_i,v_j)\in E, \text{and } w_{ij} \text{ is the weight,} \\ 0, & i=j, \\ \infty, & \text{if}(v_i,v_j)\notin E. \end{cases} \tag{1}$$

In a directed graph, the number of edges derived from vertex *v* is called the exit degree of vertex *v*, denoted *d+(v)*, and the number of edges introduced from vertex *v* is denoted as the entry degree of *v*, denoted as *d -(v)*. *D (v)= d+(v)+d -(v)* is denoted as the degree or degree of vertex *v*. It can be concluded that the degree of vertices can reflect the importance of nodes in the network.

2.2 Determination of key team members

2.2.1 Visualization of diagrams

The data given in the title is huge. We extracted all the passing data of 30 players in 38 games, and realized the visual processing of all the passing data of 30 players in all the games of the whole season with the help of Gephi software. As shown in figure 2(a,b), the average clustering coefficient of the network C=0.844.

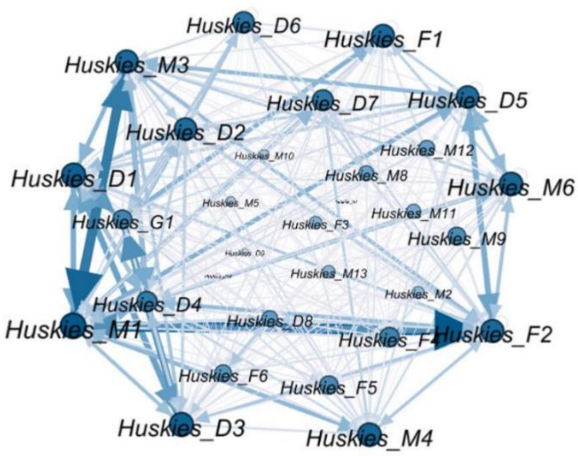

Figure 2a: Schematic diagram of passing network generated by Gephi software.

ISSN 2618-1576 Vol. 3, Issue 3: 10-16, DOI: 10.25236/FSR.2021.030303

In Gephi, the size of nodes depends on the center degree of feature vectors, which can intuitively represent the importance of nodes. The thickness of edges between different nodes depends on the weight of edges. The average clustering coefficient C=0.844.

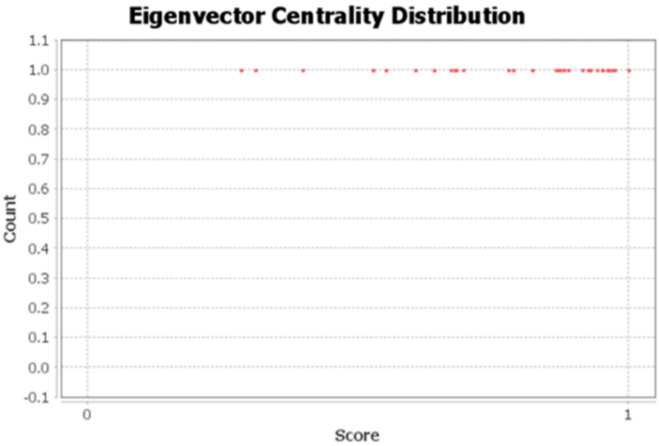

Figure 2b: Feature vector centrality distribution of nodes in the network

Based on this theory, we extracted 6 key nodes from the network diagram, which correspond to 6 key players, namely M1, M3, D1, D3, F1 and F2. Among the six players, there are two strikers, two midfielders and two defenders. Screening key players by the number of passes between nodes is only a preliminary judgment and needs further verification.

2.2.2 Verification of key nodes

PageRank algorithm is widely used in Google search engine page ranking, it is mainly in the relevant pages between the random walk to identify the importance of the page.

The local community detection method based on Page Rank algorithm USES the low-dimensional embedding of graph based on random walk of seed nodes for community detection. This algorithm can detect multiple possible overlapping areas. The following is the verification process [2].

In a directed graph *G (V,E)*, consider the original random walk, assume its initial distribution vector is p, and select some nodes in this graph *G* as seed nodes, denoted as S. Elements in p can be determined by seed nodes, like (2)

$$P_0(v) = \frac{l_s(v)}{|S|} \tag{2}$$

Where, if *v* is a member of s, then ls of *v* is equal to 1; Otherwise the ls (v) = 0. So the random walk of step t can be expressed as equation (3).

$$P_{t+1}(v) = \sum_{u \in V} \frac{A_{u,v}}{d(u)} P_t(u) \tag{3}$$

Here, *Au,v* is the member of the adjacency matrix, and d(u) is the degree of the node *u*. According to equation (3), we define a simple graph embedding, as shown in equation (4):

$$v \in V \to P(v) = (P_1(v), P_2(v), \cdots P_T(v)) \in [0,1]^T \tag{4}$$

The *T* here is the number of steps in the random walk, and in general, the recommended value for T is 2 or 3. The PageRank used here is the personalized random walk, the transfer matrix used is as follows (5), we let $r_t(v)$ represent the probability of node v at time t, then the probability at time *t+1* can be obtained through the transfer matrix formula (6), where $r_0(v)=ls(v)/s$, and a represents the damping coefficient of personalized PageRank, which is generally 0.85.

$$A = cP + (1-c)ve \tag{5}$$

ISSN 2618-1576 Vol. 3, Issue 3: 10-16, DOI: 10.25236/FSR.2021.030303

$$r_{t+1}(v) = (1-a)\,\mathrm{r}_0(v) + a\sum_{u\in V}\frac{A_{u,v}}{d(u)} r_t(u) \tag{6}$$

For a given T value, PageRank algorithm will get the rank of each node at step *T*, we define the top k nodes in a set, namely s*k*={v*1*,... ,v*k*}, and these *k* nodes satisfy Rt(v*k*)>, 0. In the set S*k*, we find the optimal *k* through the fractional function f, making the set *S* union S*k* become an optimal association. So we define an optimal *k*, as shown in equation (7).

$$k^* = \arg\max_k \ f(S \cup S_k) \tag{7}$$

The optimal community obtained corresponds to the key node, which represents the node with the largest overlap region.

2.3 Determination of binary configuration

After identifying the key players and the key pairs, we began to narrow down the size of the analysis, from all players to the key players and the key pairs, from the whole season to every game. The selected fields were analyzed.

As shown in figure 3, in the network diagram of the court, we regard the whole court as a plane coordinate system. The coordinates of the four top angles are (0,0), (0,100), (100,100), and (100,100) respectively. The Numbers in these coordinates are proportional relations rather than actual distances. Based on the data provided in the appendix, we tracked the number of passes made between the Huskies' 14 players in the second match, analyzing the characteristics and effects of key nodes and key pairs.

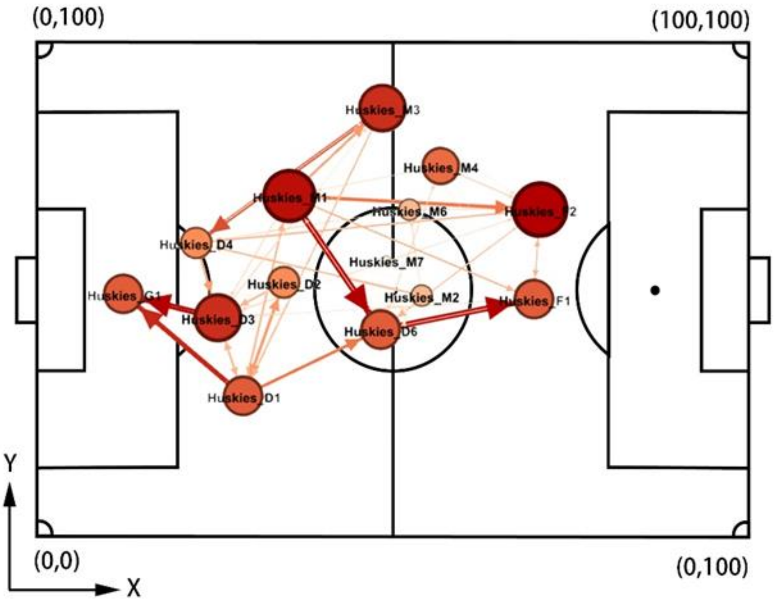

Figure 3: Network diagram

After clarifying our thoughts, we started to select the matches and sorted them in order from large to small according to the passing times between players in the whole season as mentioned above, as shown in figure 4, in which several key pairs with large contributions were selected and sorted into table 1

Table 1

Node1	Node2	number
Huskies_M1	Huskies_F2	182
Huskies_M3	Huskies_M1	168
Huskies_M1	Huskies_M3	143
Huskies_D3	Huskies_G1	120
Huskies_F2	Huskies_M1	117
Huskies_D1	Huskies_G1	107
Huskies_D1	Huskies_D3	105

ISSN 2618-1576 Vol. 3, Issue 3: 10-16, DOI: 10.25236/FSR.2021.030303

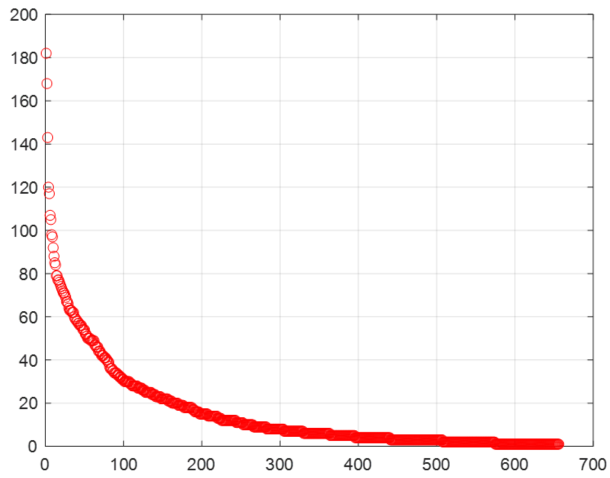

Figure 4: Scatter chart of the number of passes between players

It can be found that there are five groups of nodes that all pass more than 100 times, and M1 and M3, M1 and F2 are two-way, D1 and D3, D3 and G1, and D1 and G1 are one-way. We selected the pair of nodes M1 and F2 with the largest number of passes, selected the games they participated in and won, and further selected the data of the three games with the largest number of passes for visualization processing.

Before the visualization, the wavelet maximum method is used to extract the abnormal data, clean the data and remove the wrong data. Improve the result accuracy. The average clustering coefficients of the passing network were 0.689,0.649 and 0.697, respectively.

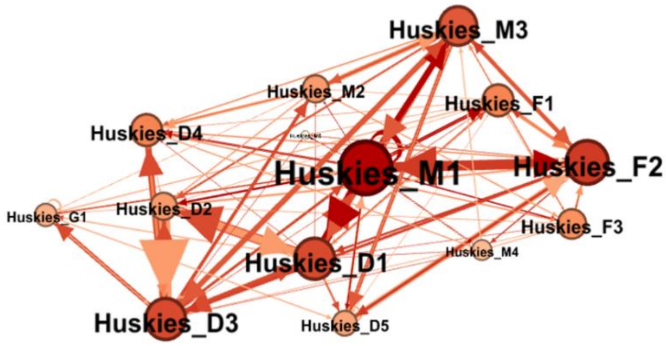

Figure 5(a): Match 1st

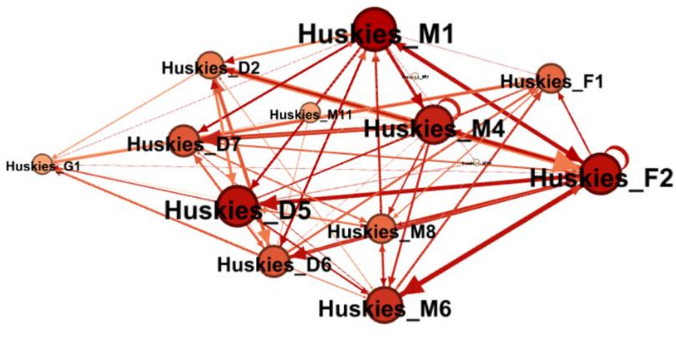

Figure 5(b): Match 14th

ISSN 2618-1576 Vol. 3, Issue 3: 10-16, DOI: 10.25236/FSR.2021.030303

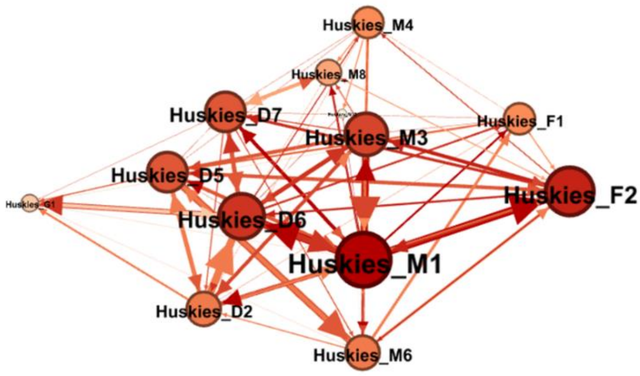

Figure 5(c): Match 18th

Through observation, we can find that the center of feature vector of M1 and F2 in these three games has obvious advantages over other players in the same game, and the weight between them is also at the leading level. Therefore, we can approximate that these two nodes can be used as a binary configuration. According to the same analysis method, similar rules can be obtained for the other four groups of nodes, which can be used to further measure teamwork and analyze and optimize the structural configuration of the game.

2.4 Competitor data analysis

By comparing the data of the season we can find there are still some gaps when compared with rival team, as shown in figure by comparing the ball, score and win, can be found that the data are below average index, but this does not necessarily indicate team member's cooperation is not successful, the assessment team, we should go through more index analysis.

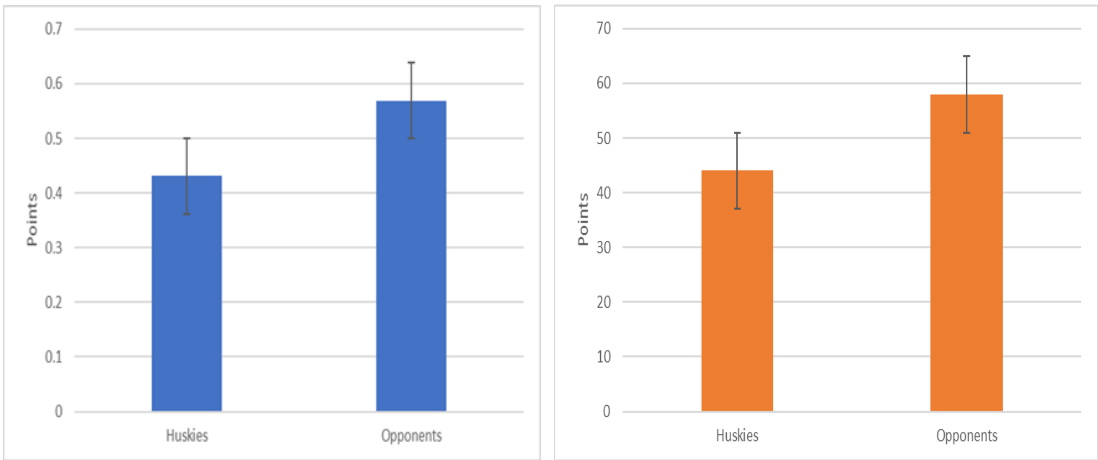

Figure 6: Data comparison diagrams

3. Model Strengths, Weaknesses and improvements

3.1 Strengths and Weaknesses

Strengths

The network model is simple and clear, and the visualization is easy.

The hierarchical analysis method combines the neural network algorithm to innovate and make scientific and reasonable on the basis of the basic model.

Based on the improved neural network algorithm of lm algorithm, the LM algorithm is added to the conventional neural network algorithm, which improves the accuracy of the results.

Quantifying qualitative indicators for analysis, such as flexibility.

ISSN 2618-1576 Vol. 3, Issue 3: 10-16, DOI: 10.25236/FSR.2021.030303

Weaknesses

The subjective method of hierarchical analysis is strong.

Neural network algorithm requires a lot of training, the training in the topic is limited, which will cause the deviation of the result.

3.2 Model Improvements

When selecting key nodes, take the type of team members into account and select them with different indicators.

The hierarchical analysis method increases the number of criteria when selecting indicators.

Replace hierarchical analysis with more objective and accurate algorithms.

References

[1] S.H. Mousavi, M. Khansari, R. Rahmani. A fully scalable big data framework for Botnet detection based on network traffic analysis [J]. Information Sciences, 2020, 512.
[2] Amin Hashemi, Mohammad Bagher Dowlatshahi, Hossein Nezamabadi-pour. MGFS: A multi-label graph-based feature selection algorithm via PageRank centrality [J]. Expert Systems with Applications, 2020, 142.
[3] Seda Turk, Gokham Sahin. Corrigendum to "Multi-criteria decision-making in the location selection for a solar PV power plant using AHP" [Measurement 129 (2018) 218–226] [J]. Measurement, 2020, 153.