

Team Evaluation Model Based on Principal Component Analysis

Yanan Duan

Institute of mechanical engineering, North China Electric Power University, Hebei 071000, China.

1518013645@qq.com

Abstract

First, starting from the individual in the team, a quantitative analysis of the performance of the 30 Huskies players in these 38 games. Calculate the individual game conditions of these 30 players and calculate the personal characteristics reflected in the game to build a player evaluation model. Then we use PCA to reduce the dimension to get the comprehensive ranking of these 30 players. After that, we use hierarchical clustering to classify 30 players into four categories, and analyze the four types of players according to the results of the principal component scoring and the index evaluation model. Then from the team itself, the two teams that beat the Husky team were counted, and the correlation analysis was performed. It was found that two teams showed very different characteristics, and the results were analyzed. Finally, according to the situation of Huskies' own players and the situation of other teams, make specific recommendations to improve the team's winning rate.

Keywords

Principal Component Analysis, Hierarchical Cluster Analysis, Correlational Analyses.

1. Introduction

Social connections are becoming closer and closer, and mathematical analysis can also be used for football. We analyze player data in 38 games and analyze to make useful suggestions for the team.

2. Huskies player characteristics analysis

2.1 Centrality

Degree Centrality: It refers to the total number of neighboring nodes directly connected to the node, reflecting the ability of the node to directly access the network's mobile content and its status and influence in the network. In a social network, if there is a direct connection between one actor and many other actors, then the actor is at the center. The larger the value, the higher his status in the network.

Betweenness centrality: It refers to the number of nodes that pass through the shortest path of all node pairs in the network. It is a kind of control index, which mainly studies how much an actor can control the cooperative activities between other actors in the network. The betweenness centrality indicator can measure a player's ability as a medium. If an actor is on multiple cross-network paths, this person also occupies an important position, plays an important intermediary role, and has a high degree of intermediary centrality.

Closeness Centrality: If a point is connected to many other points through a relatively short path, we say that the point has a higher proximity to centrality. Closeness Centrality is the sum of the shortest paths from a node to all other nodes in the network. This indicator reflects the degree to which nodes are centered in the network, it reflects the extent to which the actor's information is not controlled by others. It uses the concept of distance to calculate the central degree of a node. The closer it is to other nodes, the stronger its centrality.

The closer a point is to the center point, that is, the smaller the value of the degree of centrality, the greater its influence, the closer it is to the center of the network. That is, the smaller the value of close to the centrality, the more it is in the center of the group, and the lower the degree of control by others.[1]

2.2 Player Comprehensive Score

First, we analyze the game data of all Huskies players and the characteristics they exhibited in these 38 games. From these 38 game data, sort out the game data and characteristics of the 38 players of the Huskies team.(see table 1 and table 2)

From the results of the principal component analysis, we can see Huskies's D10, F1, F4, F5, M10, M11, M12, M2, M9 belongs to the category with the lowest score. Observing their data, we can see that their average number of passes is too small, and the pass success rate is about 65% to 75%.The lowest number of fouls is 0.375, and the highest number of fouls is 1.78, concentrated around 0.7 and 1.3. The average number of free kicks is about 0.1, which is relatively small. The average number of corner kicks is close to half, which is zero. The average number of free kicks is about 20, which is less than other players.

Huskies's F3, M5, M7, and M8 can be seen in the results of the principal component analysis. Except for M7, other players belong to the lower-scoring category. It can be known from the data that the four players' appearances and average passes are smaller than others, and the corresponding appearances will also have a certain impact on the average passes. The M7 pass rate is relatively high. Compared with other players with the same average number of passes, it has the highest pass rate, of course, it is also the highest among all people.The average number of fouls is also lower, and it is less prone to fouls than others, but the average number of free kicks and corner kicks are much lower than some other players.The average point is relatively low, echoing the number of passes.

Huskies's D1, D2, M1, M13, M3, and M4 can be seen in the results of the principal component analysis. These players belong to the category of people in the score.It can be known from the data that the average number of passes and pass rate of this type of players is at a relatively high level, with strong activity and offensive capabilities, but the average number of fouls is relatively high. The average number of free kicks and average corner kicks is relatively low, while the average point score is at a high level.

Huskies's D3, D4, D5, D6, D7, D8, D9, F2, F6, G1, M6 can be seen in the results of the principal component analysis, this type of player belongs to the highest rated category.It can be seen from the statistics that their average passes and pass rate are higher, but they are still slightly inferior to the third type of players.The average number of fouls is relatively small, but the number of free kicks is much larger than other players.

Table 1: Players' match data (1)

Player	Appearances	Average passes	Pass rate	Average fouls	Average free kicks	Average corners
Huskies_D1	30	36.33333333	0.780733945	0.633333333	0.466666667	0.466666667
Huskies_D10	4	13	0.730769231	0.75	0.5	0
Huskies_D2	24	29.91666667	0.807799443	1	0.75	0.208333333
Huskies_D3	25	37.36	0.778372591	0.32	1.36	0.32
Huskies_D4	22	35.09090909	0.737046632	0.363636364	9.681818182	0.227272727
Huskies_D5	21	39.47619048	0.753920386	1.333333333	10.28571429	0.619047619
Huskies_D6	15	31.66666667	0.677894737	0.8	3.666666667	0.133333333

Huskies_D7	19	31.63157895	0.678868552	0.894736842	11.10526316	0.631578947
Huskies_D8	12	31.83333333	0.688481675	0.25	10.58333333	0.583333333
Huskies_D9	2	32.5	0.876923077	0.5	1	0
Huskies_F1	32	11.3125	0.657458564	1.3125	0.09375	0.21875
Huskies_F2	31	36.12903226	0.766964286	1.193548387	2.806451613	0.387096774
Huskies_F3	8	10	0.7	0.125	0.25	0
Huskies_F4	22	9.636363636	0.589622642	1.363636364	0.090909091	0.227272727
Huskies_F5	14	16.21428571	0.696035242	1.785714286	0.142857143	0.357142857
Huskies_F6	14	23.35714286	0.681957187	0.357142857	0.5	0.071428571
Huskies_G1	38	20.44736842	0.608751609	0.026315789	12.05263158	0.236842105
Huskies_M1	33	46.24242424	0.822411533	1.484848485	0.606060606	0.333333333
Huskies_M10	7	6.714285714	0.765957447	0.714285714	0	0
Huskies_M11	8	10.5	0.773809524	0.375	0	0
Huskies_M12	16	15.125	0.553719008	0.75	0.375	0.4375
Huskies_M13	4	25.25	0.732673267	0.75	0.5	0
Huskies_M2	7	13.85714286	0.783505155	0.714285714	0.142857143	0
Huskies_M3	30	35.8	0.825884544	1.3	0.466666667	0.2
Huskies_M4	33	20	0.753030303	1.242424242	0.606060606	0.090909091
Huskies_M5	5	11.2	0.75	0.2	0.4	0.2
Huskies_M6	27	26	0.723646724	0.666666667	5.481481481	0.407407407
Huskies_M7	2	11.5	0.956521739	0.5	1	0
Huskies_M8	15	13.73333333	0.665048544	0.733333333	0.733333333	0.2
Huskies_M9	12	16.83333333	0.643564356	0.5	0.083333333	0.25

Table 2: Players' match data (2)

Player	Average degree centrality	Average betweenness centrality	Average closeness centrality
Huskies_D1	51.1	19.8	29.33333333
Huskies_D10	17.25	19.5	31.5
Huskies_D2	43.79166667	18.125	29.5
Huskies_D3	52.76	15.6	31.8
Huskies_D4	54	16.81818182	31.54545455
Huskies_D5	58	17.57142857	32.61904762
Huskies_D6	43.53333333	18.66666667	32.86666667
Huskies_D7	45.21052632	15.89473684	32.36842105
Huskies_D8	46	15.25	32.91666667
Huskies_D9	49	19	29.5
Huskies_F1	22.28125	19.75	29.53125
Huskies_F2	57.93548387	14.61290323	36.77419355
Huskies_F3	16.125	17.25	25.625
Huskies_F4	19.09090909	19.31818182	30.04545455
Huskies_F5	27.35714286	23.78571429	27.64285714
Huskies_F6	37	17.85714286	31.64285714
Huskies_G1	25.60526316	20.18421053	29.81578947
Huskies_M1	69.27272727	16.21212121	35.06060606
Huskies_M10	10.85714286	17.85714286	28.71428571
Huskies_M11	14.375	17.375	32.875
Huskies_M12	22.0625	17.625	32.875
Huskies_M13	33.25	17.5	34
Huskies_M2	21.28571429	16.42857143	30.42857143
Huskies_M3	53.46666667	16.4	32.6
Huskies_M4	31.33333333	17.39393939	31
Huskies_M5	17	18.8	33.2
Huskies_M6	38.92592593	16.25925926	32.92592593
Huskies_M7	20	13	34.5
Huskies_M8	21.06666667	17.2	31.8
Huskies_M9	22.66666667	18.5	30.83333333

It can be found that the average closeness and average middleness of these players are relatively average, and it can be seen that there are no similar leaders in the team who have a similar relationship and coordination between players.

Then use the principal component analysis to reduce the dimension and give these 38 players a comprehensive score:(see table 3)

Table 3: Players' Comprehensive Score

rank	Ingredient 1	Ingredient 2	Ingredient 3	Ingredient 4	Ingredient 5	Total score
Huskies_D4	1.761274436	0.175548509	1.290330447	1.208200755	0.238527965	4.673882113
Huskies_D3	1.64424061	0.913485464	0.178143078	0.754132781	0.79504416	4.285046093
Huskies_D8	2.099615987	0.191288726	2.608595202	0.040920568	- 0.821965979	4.118454505
Huskies_G1	0.121270652	- 2.399426202	2.414174002	0.987397675	1.455419664	2.578835791
Huskies_M7	- 0.989509192	4.085312725	0.128895478	- 0.649481001	- 0.035319687	2.539898322
Huskies_F2	3.190505105	1.044591147	- 0.632710653	- 1.503418964	0.427382293	2.526348928
Huskies_D7	2.47524907	- 0.797063413	1.656414895	- 0.377218414	-0.85840831	2.098973828
Huskies_M6	1.406728497	- 0.101805963	0.757571347	- 0.457521148	0.447574414	2.052547146
Huskies_M1	3.524956953	0.767469394	- 2.119025431	- 0.383202788	0.246578661	2.036776789
Huskies_D9	- 0.547974948	1.441752051	- 0.859840129	2.332752156	- 0.927254142	1.439434989
Huskies_D5	3.255695261	- 0.873917255	0.172201355	0.167292293	- 1.371730263	1.349541391
Huskies_M3	1.825302537	0.675035107	- 1.804079639	0.039312395	0.608510564	1.344080962
Huskies_D1	1.2621376	- 1.089774039	- 0.856359305	1.527819725	0.254563612	1.098387593
Huskies_F6	- 0.697864841	0.323147437	0.368040696	0.225373458	0.310913781	0.5296 10532
Huskies_D2	0.429235025	- 0.105293472	-1.287 75987	1.008 60715	0.308757881	0.3535 46714
Huskies_D6	0.473714716	- 0.160403543	0.042480626	0.050641684	- 0.396654599	0.0097 78884
Huskies_F3	- 3.029762912	- 0.025738263	0.840268912	1.394742849	0.707115605	-0.11337381
Huskies_M1 1	- 2.005823591	1.407552086	0.361717573	-0.47963529	0.123048066	- 0.593141156
Huskies_M1 3	- 0.658177457	1.323966654	- 0.192224704	- 0.474896418	- 0.671550868	- 0.672882792
Huskies_M2	- 1.761053126	1.22223608	- 0.204509329	0.036172885	0.033998747	- 0.673154743
Huskies_M4	- 0.063339261	- 0.346015382	- 1.332265298	- 0.404188038	1.253614883	- 0.892193095
Huskies_M5	- 1.750213218	0.788693993	0.827111055	- 0.394400648	- 0.584782204	- 1.113591022
Huskies_M8	- 1.109490366	- 0.150266423	0.372494276	- 0.962303618	0.170152521	- 1.679413609
Huskies_M9	- 1.328228795	- 0.591465323	0.57895888	- 0.349731175	- 0.134752557	- -1.82521897
Huskies_M1 0	- 2.794618641	0.406425353	- 0.187123954	0.241220099	0.019826453	- 2.314270689

Huskies_M1 2	- 0.562695433	- 1.222766407	0.933768542	- 1.870842832	- 0.202583452	- 2.925119582
Huskies_D10	- 2.220001681	0.288952085	- 0.199742772	-0.13981301	- 0.688984519	- 2.959589897
Huskies_F1	-1.05463526	- 2.075934125	-1.10165823	- 0.678684957	0.828845509	- 4.082067064
Huskies_F4	-1.45812053	-2.06918315	- 0.623530221	- 1.283191877	0.1461 44066	- 5.287881713
Huskies_F5	- 1.438417194	- 3.046403852	- 2.130336828	0.393943702	- 1.682032267	- 7.903246439

We can get the player's comprehensive ranking, and then perform hierarchical clustering on the player to determine its type.

Spearman is used as the distance measure between samples, and weighted (weighted average fa) is used as the distance measure between classes. Divide the players into four categories to get the correlation degree map of the clustering results and the clustering tree.

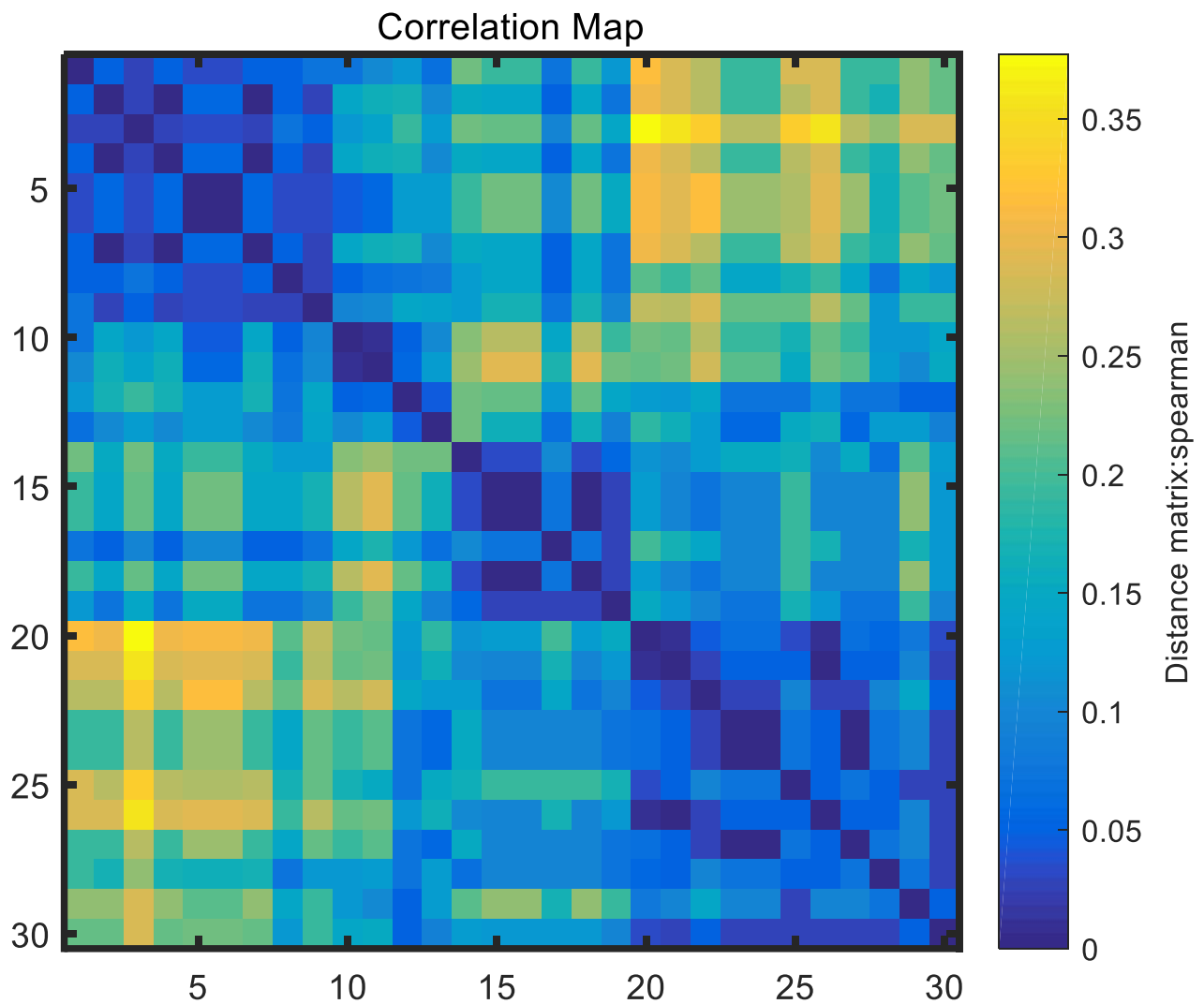


Figure 1: Correlation Map

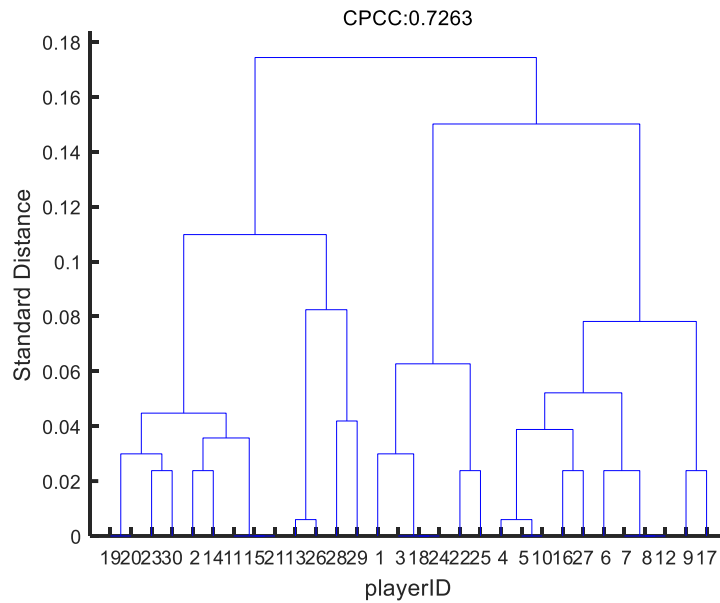


Figure 2: Cluster tree

3. Analyze opponents in 38 games

Analyze the opponent's 4,5,7,9 battle situation and configuration (lost twice) against the eight games of 4,5,7,9,22,23,26,29. Count opponents' games, passes, pass percentage, fouls, free kicks, corners, passing network density, passing network heterogeneity.[2]

Concentration of the passing network, the number of consecutive passes, the highest centrality of the individual, the number of ABAB,ABA,ABAC、ABCD、ABCB、ABCA.Then correlation analysis. Except for the opposing team during the first match with Opponent7 and the second match with Opponent9. Apart from the different settings, Opponent5's binary and ternary coordination and the number of consecutive passes in the first game are similar to those of the Huskies. However, the high-heterogeneity and low-density network structure and high individual centrality show a strong contrast with the Husky team. Explain that Opponent 5 has a leader-level hero, leading the entire team to victory. Looking back at the Huskies, though there is a better passing system in this game, but the lack of leadership of the leader (star player) leads to a disadvantage in the game.

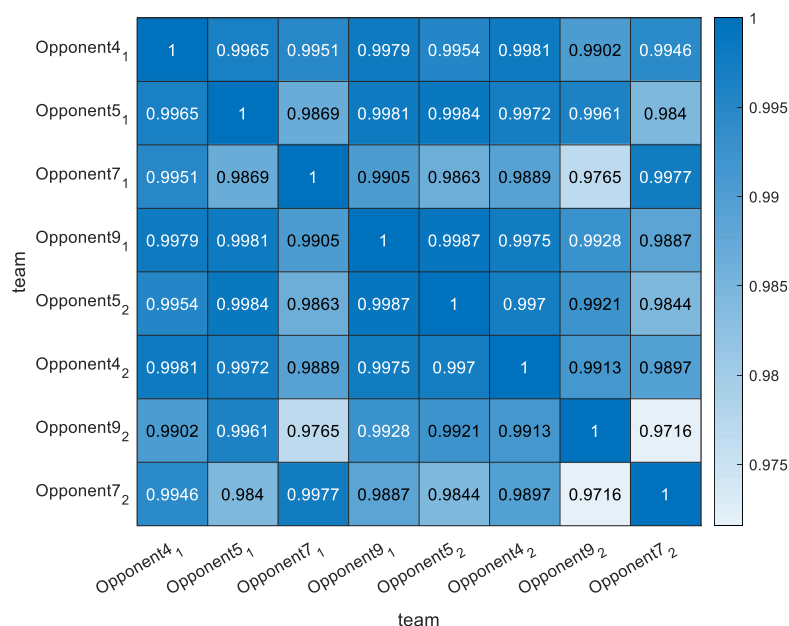


Figure 3: Correlation analysis

Opponent7 didn't show much in the seventh game, his pass ratio, heterogeneity, number of consecutive passes, and individual maximum centrality are not as high as those of the Huskies. There are no particularly good players, the formation is relatively balanced, the passing structure is simpler than the Huskies, the number of passes is far less than that of the Huskies, but the Huskies lost. The reason may be that the Husky team has a lot of passes, but it is only limited to passing the ball without making some actual offensive scores. Too many meaningless passes cause the player to lose physical strength quite quickly.

Therefore, if you want to achieve good results in the next season, for highly heterogeneous teams like Opponent5, you need to spend money to train core players.

Focus on the third and fourth types of players in the Husky team, and pay attention to the number of fouls when raising the third type of players. For the fourth type of players, it is necessary to cultivate core skills belonging to the players while ensuring a low foul rate, so that they can play stably in a downwind situation and can reverse the situation in a headwind situation.

For teams such as Opponent7, there are not many characteristics in themselves, so they only need to play normally, and they don't need to be too fancy. The reason for the failure is the abuse of tactics.

Therefore, the team must strengthen the study of tactical theory and understand the application conditions of some tactics.

4. Conclusion

In normal games, if you meet some strong opponents, if you are in a disadvantage, you can send a third offensive player, although it is easy to foul, but it has a better explosion than other players, and there may be a reversal of the situation. In the case of teams with weak examples, considering the balance of the team, you can send a second type of player, and the number of fouls is relatively small and will not lead to failure. However, the selection of players in the first category should be cautious, because they may lose the game if they are not careful. Generally, they will not consider sending them to the game unless they have to do so (when other players cannot play because of special circumstances).

References

- [1] Weihua Cao. Visualization Analysis on Passing Technique of Spanish Tiki - Taka Tactics Based on Social Network Analysis Method[J]. Journal of Chengdu Sport University, 2019(4):65-72.
- [2] Clemente F M , Couceiro M S , Martins F M L , et al. Using Network Metrics in Soccer: A Macro-Analysis[J]. Journal of Human Kinetics, 2015, 45(1):123---134.