|  |  |  |
| --- | --- | --- |
| **Problem Chosen** D | **2020 MCM/ICM Summary Sheet** | **Team Control Number** 2008495 |

Use this template to begin typing the first page (summary page) of your electronic report. This template uses a 12-point Times New Roman font. Submit your paper as an Adobe PDF electronic file (e.g. 1111111.pdf), typed in English, with a readable font of at least 12-point type.

Do not include the name of your school, advisor, or team members on this or any page.

Papers must be within the page limit specified in the problem statement.

Be sure to change the control number and problem choice above.  
You may delete these instructions as you begin to type your report here.   
  
**Follow us @COMAPMath on Twitter or COMAPCHINAOFFICIAL on Weibo for the most up to date contest information.**

**Summary**

**Have a wonderful soccer game**

This paper proposes a method, with graph theory, probability theory and calculus, to build machine learning models based on data analysis, which aims at providing strategies for soccer coach's lineup arrangement and players' training.

Firstly, the Pass Network Model can be established according to the graph theory, whose edge-weights are evaluation of coordination degree of each dyadic configurations. Pass Evaluate Index is designed for evaluate a single pass, and the summation of each pass can be defined as the edge-weights of PNM. For analysis, the adjacency matrix of N participating players within a period. Several outstanding M configurations can be found by the sort of M-element combination with the key of the sum of the sub-complete graph edge weights. What’s more, investigation of the influence of time on Pass Diversity depends on the constructed and approximate function of time and pass.

Secondly, performance indicators that reflect successful teamwork can be divided into dynamic indicators and static indicators. Static indicators include player position arrangement and line-up with which player season heatmap models and player position models can be established while the dynamic indicators include opponents’ strength, side, coach, passes, defense, attack and fail .etc. After visualized analysis of the correlation between the dynamic indicators extracted after data cleaning, and with the setting label by the goal difference, the random forest classifier is used as the machine learning model as the evaluation model of dynamic indicators. After the Grid Search used for tuning parameters, and cross-validation, the accuracy of the model achieving 80% approximately.

Thirdly, the study focuses on the role of static indicators in the performance of the team and establishes different players' value evaluation models in different positions which comprehensively consider the player’s position and technical statistical data evaluation. To optimize the value of 11-person permutation, we choose simulated annealing(SA) algorithm which searches the global optimal solution in the cousin points in the same minimized search tree as the local optimal solution has attained. The model finally gave the best starting lineup formation. In addition, we also consider the following three secondary factors: tacit understanding between players, home and away influence, and coaching arrangements. All analysis above can be concluded as comprehensive suggestion to the coach.

Finally, we use the case of the Huskies to explain group dynamics. And use the conclusions obtained by the Huskies to build a model to explain how to design a more effective team and supplement the team performance indicators.

**Contents**

**1.Introduction 1**

1.1 Background 2

1.2 Our work 3

**2.Problem restatement 4**

**3.Terminology 1**

3.1 Terms 2

3.2 Symbols 3

**4.Preparation of the Models 4**

4.1 Assumptions 5

4.2 Notations 6

**5.The Models 1**

5.1 Model 1 2

第3级 3

5.2 Model 2 2

第3级 3

5.3 Model 3 2

第3级 3

5.4 Model 4 2

第3级 3

**6.Strengths and Weaknesses 4**

6.1 Strengths 5

6.2 Weaknesses 6

**References**

# 1 Introduction

## 1.1 Background

Football has a long history. It has been loved all over the world since it was popularized. Football can be considered as the most popular sports in the world. Football, a seemingly simple sport, contains the secrets of individual ability and team cooperation. With the development of the times and the progress of science and technology, football players and coaches continue to improve in skills, showing the audience wonderful games. As we all know, a wonderful football match is inseparable from the contributions of players and teams. By studying the actions of everyone in the team, coordinating the team relationship, reasonably arranging the minutes and line-up, we can score best.

## 1.2 Problem restatement

Football is a sport suitable for all ages. Since its inclusion in international tournaments, people have created a variety of methods to evaluate the team dynamics throughout the game and over the entire season to help determine specific strategies that can improve teamwork next season. We need to use the data provided by the ICM team to build a model to solve the following four problems.

1. Consider each player as a node and create a passing network to identify dyadic, triadic and multiple configurations. We need to establish a value evaluation model of a single pass and a general evaluation model of the passing of the time structure index under the passing network.
2. To Identify performance indicators that reflect successful teamwork, we need to consider static and dynamic indicators. Establish a model of the impact of each performance indicator on successful teamwork, and use one model to encompass these four sub-models.
3. By observing and analyzing the model established in Questions 1 and 2, tell the coach that which form of structural strategy is applicable to the Huskies. Using the results of the model analysis to make suggestions for the coach to improve the team's success rate next season.
4. Use the case of the Huskies to explain the theory of group dynamics, and use the conclusion of the model established by the Huskies to explain how to design a more effective team, and supplement the team performance indicators.

# 2 Preparation of the Models

## 2.1 Data Cleaning

|  |  |  |
| --- | --- | --- |
| Data Name | Processing Type | Feature Name |
| Side | Map + Dummy | Side\_1, Side\_0 |
| Coach | Dummy | Coach\_1, Coach\_2, Coach\_3 |
| Opponent Strength | Analysis | Oppo |
| Shots | Count | Attack |
| Dribbles |
| Touch |
| Corner |
| Offside |
| Tackle | Count | Defence |
| Dispossess |
| Aerial Won |
| Interception |
| Clearance |
| Blocks |
| Saves |
| Passes | Count | Pass |
| Possession | Search + Integrate |
| Pass Success | Calculate |
| Foul | Count | Fail |
| Loss of Possession | Search + Count |

## 2.2 Processing Tools

|  |  |
| --- | --- |
| Tool | Uses |
| Visual Studio Code 1.42 | Coding, Visualization |
| IPython 3.6.8 | Run Code |
| Visio | Design Flowchart |
| Excel | Arrange Dataset |
| GitHub | Synchronization, Storing |

# 3 Establishment of PNM and Analysis of Influence Factors

为了构造结构化的传球网络，用来分析球员间传球配合的默契程度，应在不同多维度、状态变化情况下进行分析。例如从微观上两两球员之间的行为，到宏观上多个球员之间的行为；以及时间尺度从比赛中的单位时刻到整个赛季。

## 3.1 传球评价Index(PEI)

两人传球评价指数，用于评价两人配合程度。在一场球赛中，宏观来看，球员相对于球场可以视为一个个节点，球场可以视为一个网络，每一次传球可以视为节点之间的连线。以两人之间每次传球累计评价作为2人的传球评价指数。多人传球网络体系中在传球评价体系中，将三个节点连接成闭合环路，边权之和即为3人传球评价指数。

通过生活经验和数据挖掘所发现的规律可以构建PEI计算模型：

1. 传球类型权重表：



1. 计算传球或接球时分别受到的防守压力

其中，x为球员到对方球门的横坐标，与受到的防守压力成负相关

1. 两人之间的边权，即单次传球评价，为此次传球类型的权重乘以防守压力加权平均数，量化为以下公式。

根据这一传球评价指数模型，统计出一定时间范围内所有参与比赛的个球员的邻接矩阵数据。由得每两人之间所有传球价值评价总和图：



整个赛季单次评价图

## 3.2 传球网络模型构建及识别网络模式

网络上两两球员之间的联系，宏观上体现为球员间传球的评价总和。筛选两人传球评价超过一定阈值的边，运用图论的方法选择性剔除交叉边，将基于传球评价指数构建的line-up传球网络可视化，用线的深浅表示：



单场比赛的两两传球网络图

从这一模型的可视化中，我们可以直观地分析传球配合频繁和默契地球员，还可以直观的看出主力球员中多人传球配合的组合。

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| N | Players | | | | Score |
| 2 | M1 | F2 |  |  | 342.4 |
| M1 | M3 |  |  | 338.6 |
| D5 | F2 |  |  | 213.4 |
| 3 | M1 | M3 | F2 |  | 816.1 |
|  | D5 | M1 | F2 |  | 727.8 |
| 4 | D5 | M1 | M3 | F2 | 1113.5 |

## 3.3 时间尺度上传球状态波动

传球状态会在时间尺度上进行波动。定义

并以作为球队实时状态的指标。球赛刚开始时，球员身体还未warm up，导致传球概率密度较小，5—10分钟过后，传球效率逐渐提高并大致趋于稳定，即：

As the time goes，球员们体力消耗，传球密度降低，即传球数量的增速减缓（虽然一场球赛中成功传球次数仍在增加，但传球失败概率开始增加），此后传球密度呈现下降趋势，即：

纵观整个赛季中38场球赛球员们的传球概率密度，整体上看可以与单场球赛所展示出的概率密度变化趋势相同。令时间为横坐标，成功传球密度为纵坐标作图：



第1场比赛传球密度 整个赛季传球密度

总体来说，传球的密度在时间尺度上相对稳定。若使用Monte Carlo method对每次传球进行模拟，设定在上一次传球后下一次传球还需要的时间概率分布服从，其中为统计的平均传球间隔时间，则在样本规模满足时会近似与左图的分布；随着样本规模增加，当满足后则会近似与右图的分布。因此我们可以认为每一个时间点发生传球事件的概率服从。

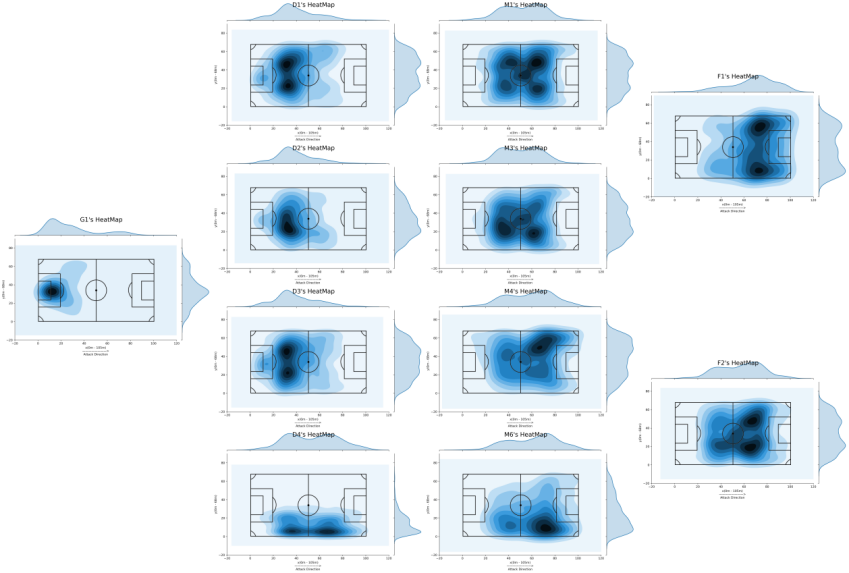
# 4 Soccer Team Indexes And Performance Prediction Based on ML

足球队中成功团队合作有许多指标，通过数据分析和实际经验，我们主要考虑以下indicators：静态指标和动态指标。首先，我们使用 评价一场比赛的球队整体发挥，作为单场比赛表现标签，定义：

## 4.1静态指标

为了考虑球员位置分布，我们采出每个球员在整个赛季中的位置坐标，做出球员运动位置的热点图，热力图每个点的值定义如下：

颜色越深则表示出现在此处的概率较大，越浅表示出现的概率越小。经过的计算，主力11人的位置热点图如下：



球员位置分布热点图

在一场球赛中，球队的阵型对团队协作起到重要作用，我们考虑在一场球赛中球员阵型，我们采取每一场比赛中每一位球员的运动坐标，采用坐标对时间积分的方法，找出每场球赛中，每一位球员平均坐标。将在数据中可以获取（球员出现在Origin/Destination）的时间点作为新的横坐标，X或Y坐标作为新的纵坐标，得出函数。我们近似认为在任意两个有记录的时间点，球员在X或Y方向上匀速移动，这样就将离散型的数据集转换为了连续性的数据集（每个）。因此平均坐标，以X坐标为例，Y坐标同理：

将这11位球员的位置标在图中绘制出每场球赛的阵型图，部分阵型图如下：



Match 1 and Match 11球员阵型图

## 4.2动态指标

动态指标包括了球队人为影响因素和在比赛里产生的技术数据：人为影响因素包括了教练、对手水平、主客场，技术数据包括了射门、传球、解围在内的各种events统计。原始的数据以单个事件作为样本的单位，而我们将其分类统计为以一场比赛为单位的动态类型数据，通过观察以新结构存储的数据，提取出其中的若干特征信息。

### 4.2.1数据清洗和特征工程

In feature engineering，为了降低特征的维度，不仅使用PCA筛选并剔除影响不显著的特征，还可以使用ChiMerge这一特征分箱的方法，将EventSubTypes分为传球，进攻，防守和Fail四个方面，与教练、主客场、对手水平一起作为一场比赛的特征。通过标准化、哑变量、结合分析等方法处理统计后的数据来量化比赛的特征：

（1）统计型数据

（2）多事件结合分析型数据

（3）One-Hot编码哑变量数据

### 4.2.2 可视化分析

分析对于影响：



主客场与得分的关系图

时的分布更多，分布更高，因此主场表现结果整体上比客场要好。

分析不同Coach的执教水平以及对于球队的指导成效：





不同教练指导下球队4种表现数据和净胜球对比图

从boxen图我们可以看出，在Coach 3指导下，球队等数据较好，其次是Coach 2和Coach 1。我们还可以得出教练们的执教风格，例如：教练1更具侵略性，防守就显得平庸；教练2强调强硬防守；教练3则较为平衡，战绩最佳。

分析、对于的贡献：



进攻、传球与净胜球之间关系图

从图中我们可以看出，在不同净胜球数下，进攻和传球大体上为线性相关，斜率为正。

与呈正相关，且分布越集中，的方差较小。我们可以得出结论：在一场球赛乃至整个赛季，越多，大概率有着更高的。

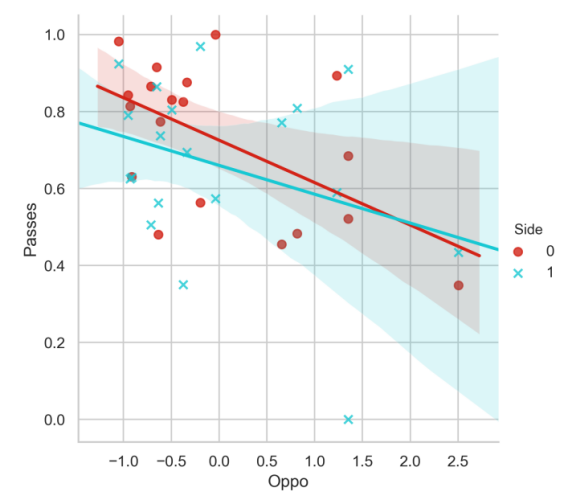
分析、对于的贡献：



防守、失败与净胜球之间关系图

与呈正相关，与呈负相关，且分布越集中，的方差较小。观察发现：图2左1的点分布在下方，因此防守不好会导致输球；右1左半边没有点，因此期望赢球则失误不能多。

以作为考察球队整体表现的positive指标，结合指标进行多角度分析：



左：进攻和防守之间关系图 右：对手水平与传球关系图

从左图中我们可以看出数据重心分布在右下角，认为整个赛季上（进攻表现）显著优于（防守表现）。从右图中我们可以看出不论是在主场还是客场，，但主场更可能有较小提升；结论是对手水平越高，我方传球率越低。

综合所有处理得到的特征，通过Pearson相关系数的计算来估计出变量间两两特征相关性。

令矩阵，得：



动态指标整合的特征相关系数矩阵

### 4.2.3 随机森林分类器模型建立和训练

我们以作为每场比赛评价标签，希望学习后的模型能够基于处理后的数据对比赛进行分类，对应到的标签。由于个特征数量较多，且与标签相关性不一，不宜采用线性模型进行分类；且样本数据数量极少，在尝试一些深度学习算法时容易有较大偏差。综上，我们选择随机森林模型建立标签分类器。

随机森林是一个包含多个决策树的分类器， 并且其输出的类别是由个别树输出的类别的众数而定。对于很多种资料，它可以产生高准确度的分类器；它可以在决定类别时，评估变数的重要性；在建造森林时，它可以在内部对于一般化后的误差产生不偏差的估计。建立随机森林分类器Random Forest Classifier的方法如下：

1. 输入特征数目，用于确定决策树上一个节点的决策结果；
2. 利用Bootstrap取样，从个训练用例中以有放回抽样的方式，取样次，形成一个训练集，并用未抽到的用例作预测，评估其误差；
3. 对于每一个节点，随机选择m个特征，决策树上每个节点的决定都是基于这些特征确定的。根据这m个特征，计算其最佳的分裂方式；
4. 每棵树都会完整成长而不会剪枝，这有可能在建完一棵正常树状分类器后会被采用。

随机森林分类器的训练后，使用网格搜索grid search进行参数调优，选定

作为参数，利用K折交叉验证验计算其accuracy score，用于评估模型准确率。

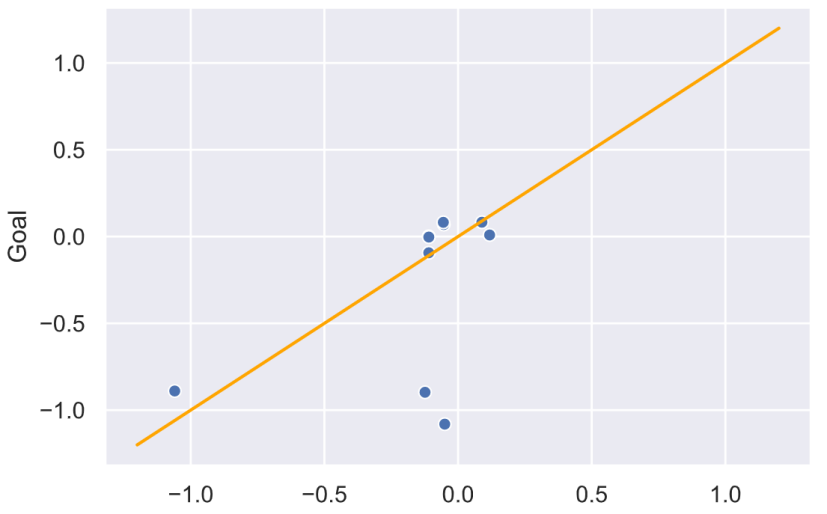


图 1模型预测交叉验证情况

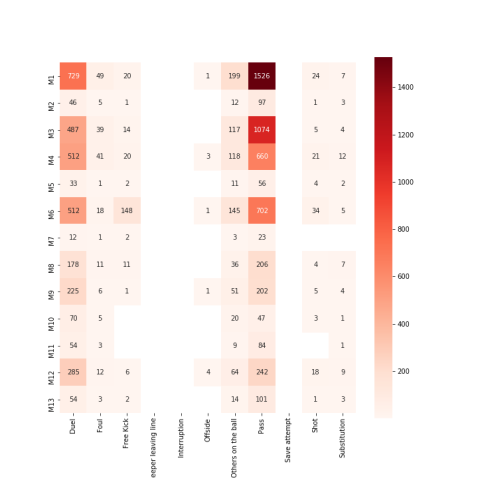
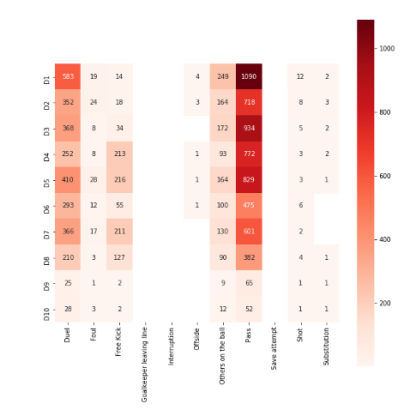
经过一定的数据调整和多次模拟结果，平均情况下得分为，最好的数据情况下可以达到的得分，在样本规模仅有的情况下，我们可以接受这一模型通过动态指标对比赛净胜球情况进行预测的准确率。

# 5 Design of Structural Strategies Driven By SA Algorithm

结构策略影响着成功的团队合作，作为一名成功的教练应该具备较好的统筹规划，协调合作，人员安排能力。我们认为，具体的结构策略应该最主要体现在以下两个个方面：球员位置安排和球队阵型。此外，还应考虑球员间默契度，主客场影响，教练安排。

## 5.1 位置evaluation engineering

在考虑球员位置安排时，需要计算守门员，前锋，中场，后卫四个位置不同球员的贡献值。我们在数据集中采集哈士奇球队30位球员的EventType，并以其为横轴，球员编号为纵轴，统计出每个球员在整个赛季中每个EventType次数，用颜色的深浅来表示次数的多少，以下分别为前锋，中场和后卫的EventTypes统计图：

前锋&中场&后卫EventTypes统计图

由以上四个图，我们可以看出F中贡献最大的是F2，其次是F1,F6,F5,F4。M中贡献最大的是M1，其次是M3,M4,M6。D中贡献最大的是D1，其次是D3,D5,D4,D2,D7,D6,D8。

我们希望能有实际的模型来对不同球员球员在不同位置表现进行量化评价。此时需要结合实际知识，分析不同位置各自的重要数据，通过不同EventTypes权重分配、结合球员各种能力performance，进行计算，作为evaluation of球队的29位球员（除守门员）分别在G,F,M位置上的表现情况。下图中，颜色越红表示越适合这个位置，反之越蓝则表示越不适合。



不同球员在不同位置评分表

## 5.2 基于SA算法优化排列组合

我们分析整个赛季38场比赛中主力阵容/首发阵容line-up，希望建立模型为教练建议最好的球队阵容该模型的目标是要找到一个最优的有序组合，使场上11人在各自位置的能力之和最大。把场上11个位置有序排列，用11位的30进制格雷码表示当前状态；例如格雷码0A1GRD739KI表示11个位置依次有第0,10,1,16,26,13,7,3,9,11,18号球员。在搜索树极为庞大、算力资源有限的情况下，我们选择模拟退火算法。模拟退火算法主要优点之一就是能以一定的概率接受目标函数值不太好的状态，且在迭代的过程中不断能够接受使目标函数向好的方向前进的解。模拟退火算法的具体步骤如下：

1. 给定冷却进度表参数及迭代初始解.以及,其中冷却进度表参数包括：控制参数的初值，衰减函数，终值以及链长度；
2. 参数时，按照如下过程作次试探搜索：
   1. 根据当前解的性质，产生一个随机偏移量，从而得到一个当前解邻域的新的试探点；
   2. 产生一个在区间上均匀分布的随机数,计算出在给定当前迭代点和温度下与接受准则相对应的转移概率：
   3. 试探搜索小于次，返回步骤1，否则进入步骤3；
3. 根据给定的温度衰减函数产生新的温度控制参数，及链长度，转入步骤2，进入下一温度点的平衡点寻优。



图 2 SA Flowchart

在实际试探搜索中，我们很可能现入局部最优，需要进行判定以退出。当前解的优化程度小于当前最优解的优化程度的时候,新解被接受的概率为,而当温度足够低的时候，较差解被接受的概率趋近于。依据最近的次搜索中都没有优化程度更高的解出现这一特征，可以根据具体问题确定阐值而后判定搜索己经进入局部最优。

## 5.3 其他结构策略因素

考虑完主要策略后，我们考虑以下四个次要影响因素：球员间默契度，主客场影响和教练安排。

首先，选择默契度高的小分队有利于提高传球和进球的效率。默契度高的小组往往配合能力较强，有助于比赛的成功。传球效率较高的球员往往适应能力较强，与其他球员的配合度也较好。

主客场因素也是必须要考虑的，有的球员适应性较强，在主场和客场都能较好的发挥出原有的水平，而有些适应性较差的球员只在主场发挥出原有的水平，环境对他的表现有较大的影响。那么在主客场时，应该安排不同的球员上场。

最后，教练安排上，整个赛季中，Coach 1,Coach 2,Coach 3,分别指导了9,5,24场比赛，通过我们在第二题中的数据分析也可得出，Coach 3的水平较高。

## 5.4 structural strategy conclusion

纵观整个模型，为了在下个赛季中提高球队成功率，我们团队给出的建议是，球队聘用Coach 3作为球队主教练，采用442的line-up，将F1,F6,F2,M3,M1,M6,D3,D1,D2,D5作为主力球员，命名该阵容为，他们的位置按下图安排：



球员阵型图

其中的一个特殊点，根据赛季数据的评价，F2球员作为一名前锋，拥有很强中场的能力，在尝试将他安排在中场时取得了显著的新最优解，这说明每个人的任一位置评价较为重要，侧面indicate我们的模型因素考虑完善。

综上，该阵型的个人能力总分为，团队配合得分为，按照进行加权平均，最后得到综合评分为。实际比赛中与此相似的阵型取得很好的战果，也验了我们的评价模型和模拟退火算法的可行性和准确性。

# 6 Model Extension Combined With Group Dynamics

In the research of huskies, we found some factors that affect the successful team cooperation, such as passing network, personal ability, coach and so on. These factors can be linked to group dynamics to analyze why these football field factors contribute to team performance. And we can explore what factors can be considered to supplement our interpretation of excellent groups and spread them into various groups in society.

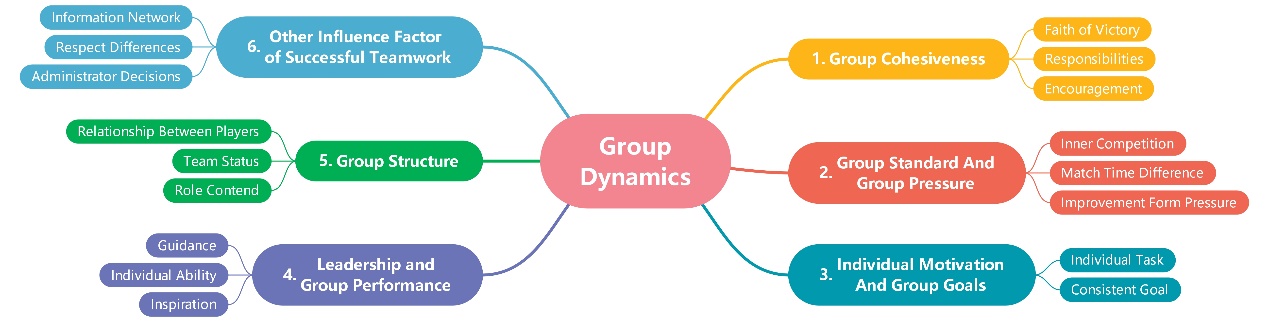


Figure 1 Group Dynamics Mind Map

## 6.1 Group and Soccer Team

Group dynamics mainly includes 5 aspects: group cohesiveness, group pressure and group standards, individual motivation and group goals, leadership and group performance and group structure.

### 6.1.1 Group Cohesiveness

Group cohesion can be regarded as the belief that the team wants to win. Group pressure comes from the outside. Improving group cohesion is a virtuous circle to inspire the team's winning belief. Responsible behavior is reflected in each player's performance. The mutual influence among the members is reflected in the mutual encouragement and progress among the players.

The Huskies is under the pressure of their opponents in the competition, but also full of victorious spirit to win. In data analysis of Huskies, each players' average acceleration, running distance and effective area of heatmap can reflect the particular player's attitude and effort on the court. In the situation of backwardness, draw and lead, the team bears varying degrees of external pressure, as the result, the data of players' attitude will fluctuate. If there is still positive data in the situation of draw and backwardness, we can consider that the group cohesion is really strong.

### 6.1.2 Group Standard and Group Pressure

Group standards can help teams and players to feel more oppressive and to compete with each other, which helps improve personal ability. From this point of view, appropriate group pressure is necessary, and only when there is competition can there be development.

For Huskies, there are differences in minutes. There is a huge gap between the core players and the edge players, which leads to the edge players in the team will be under the pressure of the core players' ability and status, but it also encourages them to strive to get playing time to prove themselves. We can analyze the evaluation trend of every game of players, especially the players who can't get stable playing time. If their evaluation can be improved in limited playing chances, we can think that the pressure brought by internal competition makes them progress in a way.

### 6.1.3 Individual Motivation and Group Goals

Group goals affect group behavior. When team goals and players' goals are consistent, players will show the strongest motivation to win and strive for goals.

Players in every position of the team have different responsibilities, and the victory is that all players have completed their tasks perfectly. Therefore, everyone is the same in the goal of victory, and everyone has certain positive expectations for their own tasks. When everyone works hard for the expectation of completing their own tasks at the same time, the group goals will reach an agreement.

### 6.1.4 Leadership and Group Performance

There are two kinds of leaders in the team, they are the coaches who provide training, pre-match guidance and post-match summary for the games, or the leaders who inspire morale, set an example and dispatch the command on the field. The ability of leaders will affect the team's performance and progress. In addition, the coach and the captain's help to the players can improve the team's vitality.

### 6.1.5 Group Structure

When a team has a stable relationship between players, it has a team structure. Stable structure is conducive to the cultivation of tacit understanding among players, so as to have a greater probability of winning the game. It can be convinced that the number of people in different positions of a team obeys the upper triangle distribution, and if the position range is too large or too small, it is not conducive to the stability of the team. In addition, if the competition in one position of the line-up is too great and the competition in other positions is very small, there will be structural imbalance and structural change.

## 6.2 Other influence factor of successful teamwork

There is also a need for close and unblocked information exchange network between teams. International teams need to ensure that the language communication between players can be smooth on and off the court; in addition, they need to carry out more friendship activities to improve the harmonious atmosphere within the team. In this way, we can communicate and cooperate effectively in the competition.

Leaders and everyone should respect the differences between others and themselves, and improve the team atmosphere by accepting the differences between people while maintaining group goals and cohesion.

Administrators should have a clear understanding of the situation and be able to make adjustments to the situations.

# Evaluation

## 7.1 Strength

* The establishment of PNM is closely related to the design of PEI. PEI comprehensively considers many aspects of each pass, quantifies the quality of the pass, and can reduce the error and variance with the actual situation. And the network model of pass, based on graph theory, intuitively describes the degree of cooperation, which is conducive to the search of multiple combinations, and the visual effect can highlight the familiar combination.
* Heatmap generation model has strong compatibility for the approximate continuity of discrete data, and can cope with the situation of too little or sparse coordinate data. Based on the visual data analysis, player position is consistent with the actual situation. In addition, for the cleaning of events data, the impact of data abnormality and missing is effectively avoided. Even though the number of samples is only Under 38, it is not easy to overfit or deviate too much. The highest accuracy of 80% after parameter optimization is enough to effectively predict the general results of the game, what means this model can make predictions for the future game, that reflects the current ability of the team, based on the recent data and give training and line up to coaches as reference.
* The static structure strategy, which should be developed by the coach, is transformed into the optimal arrangement and combination problem of 11 elements. Large-scale data supports the increase of the dimension of the evaluation index and reduces the expectation of the deviation of the particular position ability value of each player. Under the condition of limited computing power, the simulated annealing algorithm is properly used, and we manually set the starting strategy according to the actual experience and uses the imprecise individual ability evaluation index to find the arrangement and combination of 11 players. The partial optimal solution can be accepted as the global optimal solution within the threshold range when the accuracy expectation is certain.
* The models can easily correspond to the theoretical key points in group dynamics, and the additional aspects based on the existing influencing factors are also of great practical significance.

## Weakness

* There are many hyper parameters in the model, so the parameter optimization of the model has a great challenge.
* There are fewer players with data at each time, so it is impossible to test the players' ability without the ball, and it is difficult to evaluate the attack or defense from the aspect of the overall position and formation.
* There are too few samples to input the RFC model, so the training results fluctuate greatly.
* The optimal lineup obtained by the simulated annealing algorithm can only be guaranteed to be a local optimal solution, not a global optimal solution.