

The Soccer Tactics Optimization Based on Network Science

—Based on the Huskies Team of MLB

Yutong Tian¹, Yuelin Zhang¹, Renjun Tang^{1,2}

¹Westa College, Southwest University, Chongqing

²University of Tasmania, Hobart, Australia

Email: k31000277@email.swu.edu.cn

Received: Jun. 1st, 2020; accepted: Jun. 15th, 2020; published: Jun. 23rd, 2020

Abstract

In the competitive game of football, the performance of a team is often estimated simply by the addition of the strength of each player. Yet the power of teamwork is far greater than the simple addition. In order to explore the dynamic characteristics of the cooperation of the players in the team and come up with effective strategies to guide the winning game next season, we put forward 3 series of models, which are non-directive graph for team pattern, multi-index model for level assessment, and comprehensive model for the combination of the models above. For the validity of the model, we adopt the competition dataset of Huskies team in American, analyzed the situation of the teams respectively, and put forward the suggestions for the tactics and strategies next season.

Keywords

Network Science, Soccer Competition, Non-Directive Graph, Multi-Index Model

基于网络科学的足球战术优化

——以Huskies队为例

田雨桐¹, 张跃麟¹, 汤仁君^{1,2}

¹西南大学, 西塔学院, 重庆

²塔斯马尼亚大学, 霍巴特, 澳大利亚

Email: k31000277@email.swu.edu.cn

收稿日期: 2020年6月1日; 录用日期: 2020年6月15日; 发布日期: 2020年6月23日

摘要

在竞争激烈的足球比赛中，球队的战绩往往是简单地以每个球员的实力加成来估算。然而，团队合作的力量远比简单的加法要大得多。为了探究队内球员合作的动态特点，并拿出有效的策略来指导下赛季的比赛取胜，本文提出了三套模型，依次是针对球队阵型布置的无向图模型、针对球队水平评价的多指标分析的回归模型，及融合上述两套模型的综合评价模型。为了验证模型有效性，本文根据某联赛Huskies队的数据，对其本身及对手情况进行分析，为其下赛季的战术布置及策略给出建议。

关键词

网络科学，足球比赛，无向图，多指标模型

Copyright © 2020 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

在竞技团队运动中，团队合作的重要性不可否认。同时，竞赛运动的规则和场地也为我们提供了判断团队个人表现和团队整体表现的标准。因此，与在其它领域进行数据处理和分析相比，以足球为研究对象，探讨诸如团队合作和战略结构等配置方法对团队整体绩效或水平的影响要容易得多。但是，团队成员之间的互动以及每个团队成员的优劣，以及在有限时间内团队成员的默契，都是影响整体绩效的因素。

至于比赛过程中球员形态和队员配合的变化，大多数体育数据分析师通常使用录像来手动记录[1]，这既费时又不准确。因此，许多研究人员花费大量时间、精力和金钱来分析体育视频，提取和研究视频中的图像，确定运动员的运动轨迹和习惯，并使用复杂的算法来模拟和预测互动运动员之间。当然，这对于了解团队是输是赢具有重要意义，但是这种评估系统既昂贵又复杂，也使其很难被推广到足球比赛的各个方面[2]。而且，在许多著名的球队中，各自比赛的数据被保存为商业秘密，这使得我们很难从数学的角度来分类球员和球队的表现数据。即使使用数据进行数学建模，也仅限于分析团队内部固定组合[3]，并且很少研究对手策略的调整和团队整体绩效。

在量化球员表现的过程中，大多数研究人员使用传球次数的增加作为表现指标[4]。Dush 等人[5]提出了一种利用量化指标评价运动员之间互动效果的算法，获得了较好的成效；Buldú 等人[6]以巅峰时期的巴塞罗那队为例，应用网络科学的方法对团队互动的效果进行评价，得出了球员间内在的隐含关系，发掘出了巴塞罗那队的战术要领。尽管该评估方法与实际游戏结果相符，但缺乏尺寸丰富性，从而导致传球次数权重过大的问题。有鉴于此，我们的团队还在以下建模中提出了改进计划。如图 1 所示，基于上文中提到的先前研究的空白，本文将对其进行一一解决。

为了探索团队中球员合作的动态特征，并提出有效的策略来指导下赛季的获胜比赛，本文作者提出了一套用于足球战术优化的网络科学模型，并以美国职业足球联盟 Huskies 队在 2018~2019 年的比赛数据为蓝本，对模型进行效果检验。以下为本模型为解决所提出的策略。

首先，本文使用无向图模型构建矩阵来处理玩家数据并解决“最大团队”问题和阵型问题，以研究 Huskies 团队中两个、三个或三个以上频繁互动的球员的內部传球组合。在解决第一个问题的过程中，我

们使用了各种方法来可视化数据。通过编程，从时间过程，空间分布以及参与者之间的相应交互中构建问题所需的网络模型。从微观的角度，我们研究了赫斯基团队的特定游戏的区间分布，并绘制了替换前后的网络图，并观察了三合会人员的复杂交互配置。从宏观的角度，我们研究了 Huskies 团队中整体互动和形成结构的强度和频率，并在过滤后将相应的值附加到网络节点之间的连接中。

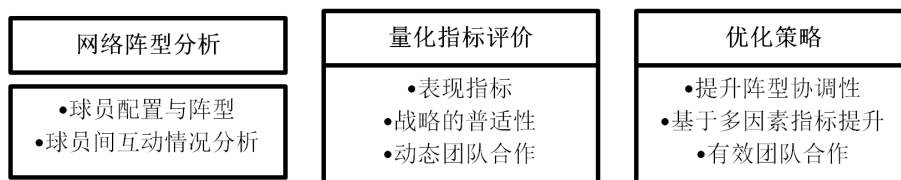


Figure 1. Problem analysis
图 1. 问题分析

其次，本文从不同的文献中选择了效果良好且相关性强的技术指标，并在量化模型中对其进行了组合和优化，以基于合理的假设对足球队的整体实力水平进行量化。从通过事件的数量到入侵指数，从加速度指数到模拟标量 H ，我们对这些指标和通过次数进行了回归分析，并获得了误差值内的有偏差结果。尔后，我们将获得的团队水平量化模型应用于 19 个对立的团队，以研究团队的策略是否通用并且得到否定的结果。

第三，本文基于上述两个数学模型总结了影响团队在团队合作和互动中的力量的主要因素。我们还对不同获胜率下的不同编队进行了统计，以便为教练提供最佳的编队安排和替换时间。然后，根据偏微分方程的模型原理对上述技术指标进行加权，得出最能代表团队互动和整体实力水平的公式。通过公式的系数，我们为教练提供了更准确的分析以及下赛季的多重胜利计划。然后，我们将其扩展到存在团队合作的其他社会领域，以讨论最佳团队球员分配策略和整体协调。

2. 模型假设及数据集分析

2.1. 符号假设

Table 1. Table of the symbol assumption description
表 1. 符号假设描述表

符号	描述
ω	总传球数
μ_p	队员平均传球数
μ_r	区域平均传球数
σ_p	队员传球数标准差
σ_r	区域传球数标准差
H	以上指标的调和平均数
λ_a	入侵指数
λ_b	加速指数
λ_d	控球率
λ_e	身体对抗成功率

本文所建立模型，对各类变量进行了相应的符号假设，其假设情况如表 1 所示。为简化陈述，本文中所有相关变量均用符号假设中的符号表示。

2.2. 条件假设

- 假设球员的个人身体及精神状态是理想的，并且所有球员处于同一状态，因此我们不考虑个人的影响，而是关注团队合作。
- 假定与匹配有关的所有事件都包含在给定的数据集中。
- 假设任何事件 P 的得分概率都在二维高斯分布的分布下，即 $P(X, Y) \sim N(\mu_1, \mu_2, \sigma_1, \sigma_2, \rho)$ ，其中 $\mu_1 = 0, \mu_2 = 50, \sigma_1 = 5, \sigma_2 = 5, \rho = 0$ 。

2.3. 数据及分析

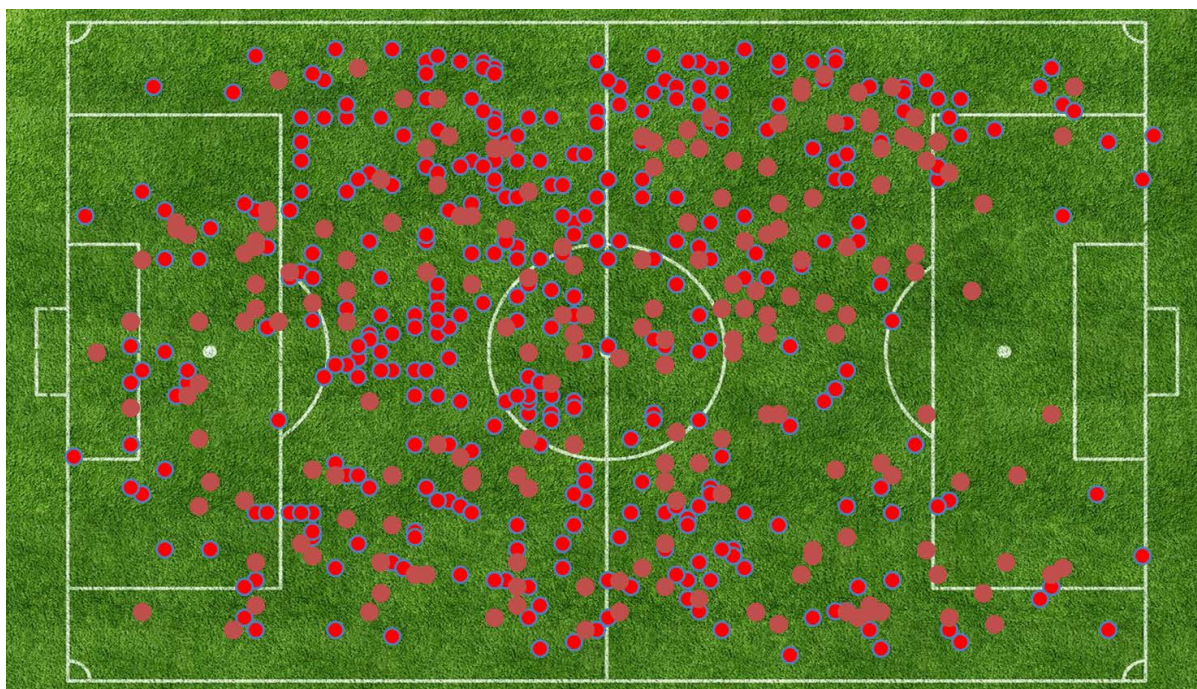


Figure 2. Field distribution of effective passing

图 2. 有效传球区域分布

在数据处理的过程中，我们收集了球队所有传球的坐标位置，筛选出球队中的有效传球点，通过传球行为在场上的分布，直观地显示出球队中传球行为的具体位置，并结合入侵指数和加速指数来分析球队的整体实力。为了说明数据集情况，我们选择了 Huskies 队数据集中的其中一场比赛，图 2 中，红点代表 Huskies 队，棕点代表对手。

此外，我们对各种事件发生的数据时空分布进行分析。如图 3 右面板所示，当比赛接近 20 分钟左右的时候，哈士奇队的活跃度较高。另外如图 3 左面板所示，射门(Shot)和传球(Pass)的时间有很大的差距，这代表着从传球到射门的转化率较低，Huskies 队的攻击力不强。

之后，我们决定以 Huskies 队的首场主场比赛为例，对数据进行筛查。据观察，如图 4 所示，Huskies 队在下半场比赛中做出了三次换人，相反，对手做出了两次换人，如下图所示。这三次换人分别发生在下半场开始后的 60 秒、90 秒和 1920 秒，都是在下半场开始后。我们以替补队员的出场时间作为比赛日

程的分界点，将比赛的半场时间划分为若干段，从而研究球员的角色调整以及球队结构和阵容的变化。此外，通过更换替补球员，还可以更新球队的网络结构，观察替补球员在球队中的作用是否相同，甚至在团队合作中的表现是否更好。

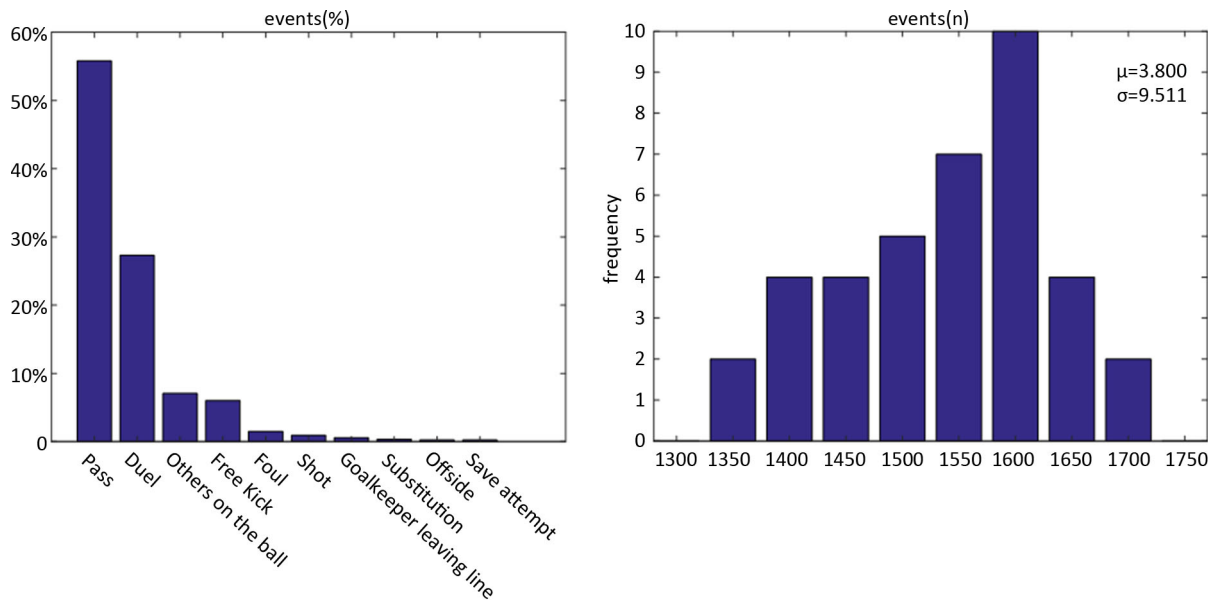


Figure 3. Field distribution of distinctive events
图 3. 不同事件的区域分布

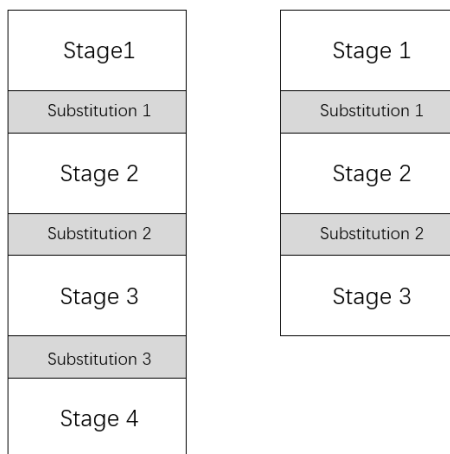


Figure 4. Schedule section
图 4. 时间阶段分布

3. 网络分析模型

3.1. 模型建立

首先，我们建立了一个无定向图模型来分析 Huskies 队的战术和模式。为此，我们假设每个球员都是一个节点，而球员两两之间的传球线路则定义为边。由于在没有数据集的情况下，球员之间的连接强度(传球频率)是未知的，所以我们将网络初始化为所有边权重均为 1 的全连接无向图，如图 5 所示。此外，基于我们对数据集的了解和对足球比赛的常识，我们发现列表中的球员分为前锋(F)、中场(M)、后卫(D)

和守门员(G)，所以我们将网络分为4层：前锋层，中场层，后卫层和门将层。此外，足球比赛中的球员限定为11人，包括守门员，因此节点设计为11个。

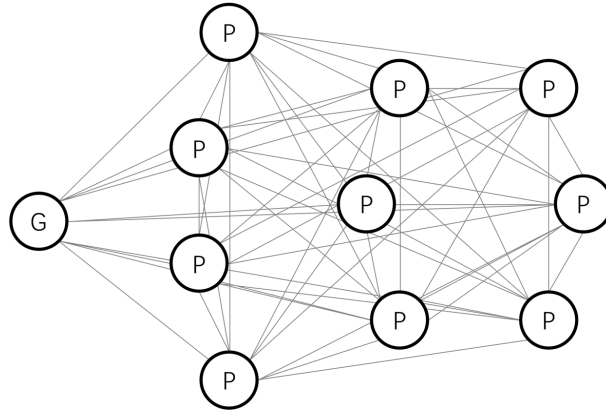


Figure 5. Idealized initial formation of the team
图 5. 球队的理想化初始阵型布局

作为无向图，邻接权重矩阵是至关重要的。第一步是建立状态向量 $A_{1 \times 30}$ ，用“1”来表示，然后用零矩阵定义不同阶段比赛的邻接权重矩阵 $W_{30 \times 30}^{stage}$ 。

首先，我们将 Huskies 队所有球员与另一个球队所有球员的传球事件进行计数，然后将它们应用到邻接权重矩阵 W^1 中。矩阵可以表示为：

$$W^{stage} = \begin{bmatrix} p^{1,1} & \cdots & p^{1,30} \\ \vdots & \ddots & \vdots \\ p^{30,1} & \cdots & p^{30,30} \end{bmatrix} \quad (1)$$

其中 $p^{x,y}$ 是球员 x 到球员 y 的传球次数，为了评估每个球员的传球次数，提出传球事件向量 $B_{1 \times 19}^{stage}$ ：

$$B^{stage} = AW^{stage} \quad (2)$$

B^{stage} 显示的是球员的传球次数，而其他不在场上的球员的传球次数为“0”。在此基础上，我们可以知道，如果某个球员在场上，则矩阵中对列或对应行向量中，至少有一个元素不为 0，而其他缺席的球员都是 0，基于这种分析，我们就可以获得球员的分布和位置。

3.2. 模型检验与结果评估

在 MATLAB 中进行编程，导入数据处理所有的传球事件并利用一个高通滤波器进行滤波后，我们根据替补球员的时间进行了赛程的划分，将 Huskies 队的阵容结构做成了四段式的队形结构，并且进行了队友间的组合使他们能更好地配合，球员组合包括三元配置和菱形配置。

如图 6(a)所示，在 Huskies 队的第一阶段中，三元配置为虚线所圈出的三个三角形，即 M1, F2, F1; D3, M2, M3; 以及 D1, M1, M3，由此图我们可以得出结论，中场球员在传球网络中起到了协调作用，因此让所有球员之间的配合更为默契。

在第二名替补出场后，Huskies 队的中场、后卫和前锋形组成了菱形配置，中场休息时由中场组成了三元配置，如图 6(c)所示。值得注意的是，在菱形队形中，任意两个顶点之间可以形成传球较多的二元队形，这增强了整体球队阵型的稳定性。

在第三名替补出场后，也是在比赛的最后阶段，Huskies 队的球员之间配合次数明显减少，有效传球也明显减少，没有形成稳固的三元或两元结构。

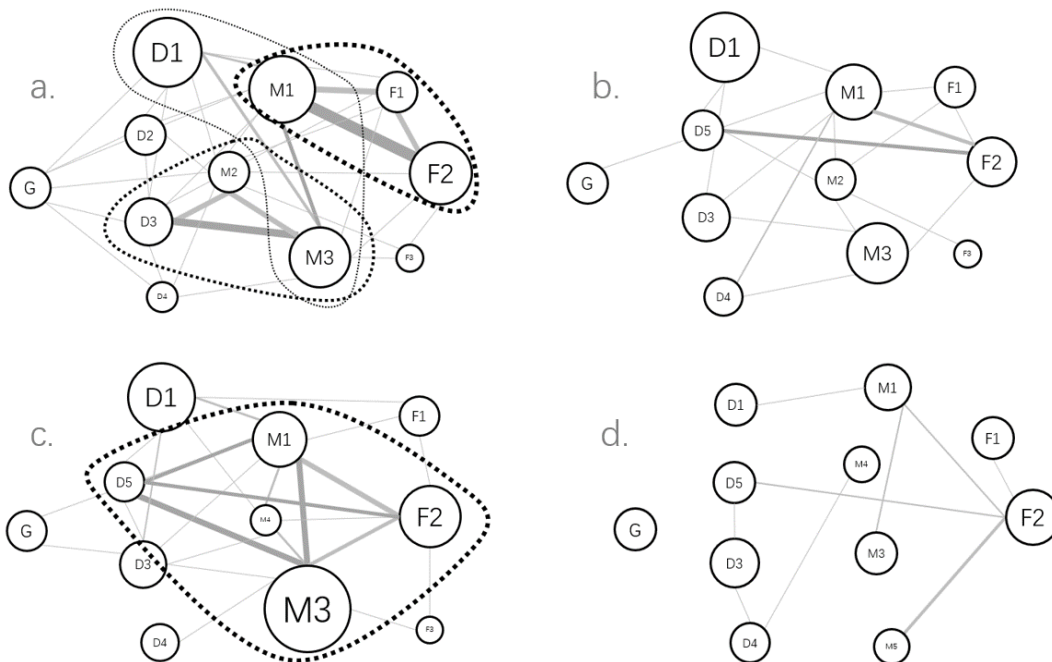


Figure 6. Huskies formation changes in a half game
图 6. Huskies 队半场比赛中的阵型变化

4. 球队绩效指标模型

4.1. 模型建立

为了评价战术和阵型的效果，我们建立了一个球队性能指数模型，共有 5 个指标：进攻指数、加速指数、多个传球数据的调和平均值、控球率和对抗率。这里将指标定义如下：

- 进攻指数： λ_A

定义 1: 进攻指数 λ_A

进攻指数 λ_A 定义为球队在对手获得控球权之前在球上发生的一系列事件。一场比赛中球队的整体进攻指数 $\bar{\lambda}_A$ 为其控球阶段的平均进攻指数：

$$\lambda_A(t) = \max P_t(x, y)$$

$$\bar{\lambda}_A = \mathbb{E}(\lambda_A(t))$$

为了计算进攻指数，该程序可以总结如下：i) 统计在控球阶段发生的每个事件，并计算从该事件发生的位置进球的概率；ii) 取这些概率中的最高者。参考 文献 5 提供了进攻指数 λ_A 的定义。为了明确每个位置进球的概率，我们建立了如图 7 所示的二维高斯分布，概率分布函数可以写为：

$$P_t(x, y) = \frac{1}{2\pi\sigma_1\sigma_2} \exp\left[-\frac{1}{2}\left(\frac{(x-\mu_1)^2}{\sigma_1^2} + \frac{(y-\mu_2)^2}{\sigma_2^2}\right)\right] \quad (3)$$

其中 σ_1 和 σ_2 是 x 和 y 的标准差， μ_1 和 μ_2 是 x 和 y 的平均值。此处需注意， x 和 y 代表事件发生的坐标。

- 加速指数 λ_B

定义 2: 加速指数 λ_B

加速指数 λ_B 定义为进攻指数与控球阶段中第一事件与最危险事件之间时间的平方之比。一支球队在一场比赛中的平均加速指数 $\overline{\lambda_B}$ 是其控球阶段的平均值。

$$\lambda_B(t) = \frac{\lambda_A(t)}{t^2}$$

$$\overline{\lambda_B} = \mathbb{E}(\lambda_B(t))$$

参考文献[5]也提供了加速度指数 λ_B 的定义。与进攻指数类似，加速指数也可以直接从文献中得到。

- 多个传球数据的调和平均值 H

定义 3: 多个传球数据的调和平均值 H

H 定义为五个指标 $(\omega, \mu_p, \mu_z, \sigma_p, \sigma_z)$ 的调和平均值。

$$H = \frac{5}{\frac{1}{\omega} + \frac{1}{\mu_p} + \frac{1}{\mu_z} + \frac{1}{\sigma_p} + \frac{1}{\sigma_z}}$$

显然，如何解决这五个动态指标的表述问题是此步的关键。基于此，我们将公式修改如下：

$$\left\{ \begin{aligned} \omega &= \sum_{k=0}^{5400} p(k) \\ \mu_p &= \frac{\omega}{n} \\ \mu_z &= \frac{\omega}{s} \\ \sigma_p &= \frac{\sum_{i=0}^n (p(i) - \mu_p)}{n(t)} \\ \sigma_z &= \frac{\sum_{i=0}^s (p(i) - \mu_z(t))}{s} \end{aligned} \right. \quad (4)$$

其中 $n(t)$ 是时间 t 时球场上的球员数量， s 是球场的面积。

- 控球率 λ_D

定义 4: 控球率 λ_D

足球比赛中的控球率 λ_D 表示的是球队中本方球员持球的总时间：

$$\lambda_D = \frac{\text{The total possession time until } t}{t}$$

- 对抗率 λ_E

定义 5: 对抗率 λ_E

对抗率 λ_E 定义为比赛中某一球队对抗动作的成功率：

$$\lambda_E = \frac{\text{The total successful times of the duel in events}}{\text{Total events in time } t}$$

4.2. 模型检验与结果分析

4.2.1. 进攻指数对比

下图 7 右边显示的是第一场比赛期间 Huskies 队的进攻指数和对手球队进攻指数的变化情况。此处将时间设置为横轴，通过分析不同时间段球队的进攻指数来讨论球队的策略是否合理。虽然 Huskies 队的进攻指数在比赛进行到 45 分钟左右时达到顶峰，但从蓝线上看，对手队也踢进了不少制胜球，其进攻指数在比赛的最后 15 分钟达到顶峰。

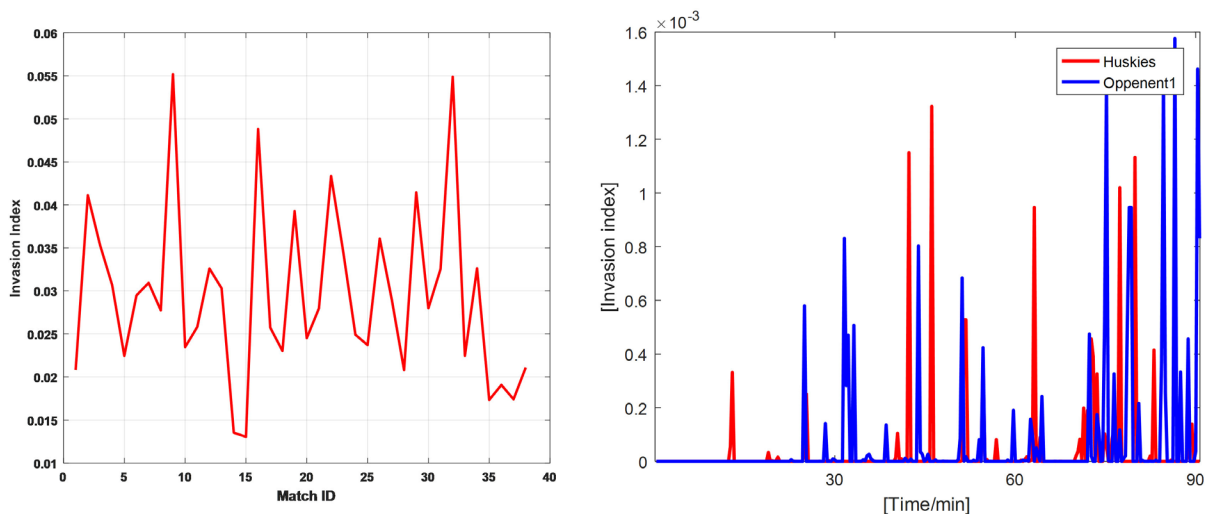


Figure 7. Left: Average Invasion index changes of this season; Right: Team invasion index contrast in Match 1

图 7. 左：本赛季平均进攻指数变化图；右：第一场比赛中球队进攻指数对比

4.2.2. 加速指数对比

在进攻指数的基础上，我们引入了加速指数来研究球队在某一时间段的进攻强度，即图 8 中的曲线斜率。在整个赛季的 38 场比赛中，我们可以看到 Huskies 队的进攻强度在前半段比较大，后期逐渐趋于缓和，这也与上面所分析的进攻指数的情况相吻合。

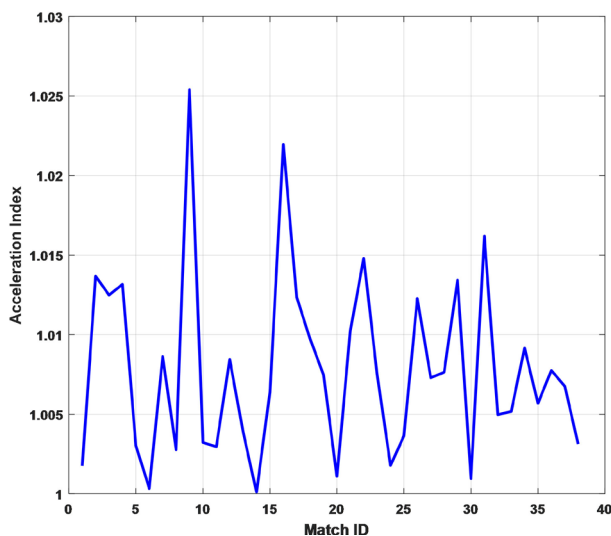


Figure 8. Average acceleration index changes of this season

图 8. 本赛季平均加速度指数变化情况

图9为H队与对手在第一场比赛中的加速指数。可以看出,Huskies队主要在比赛进行到中间时进攻,即在替补上场前和替补上场后的时间段内。除此之外,H队的进攻强度始终弱于对手。

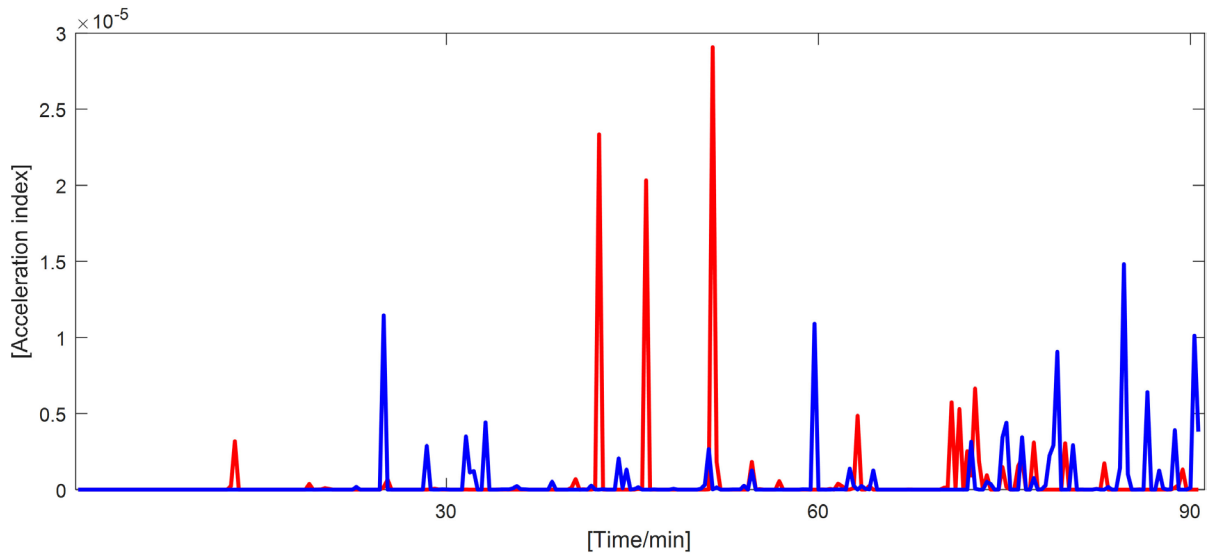


Figure 9. Team acceleration index contrast in Match 1
图9. 第一场比赛中的球队加速指数对比图

4.2.3. 球队水平比较

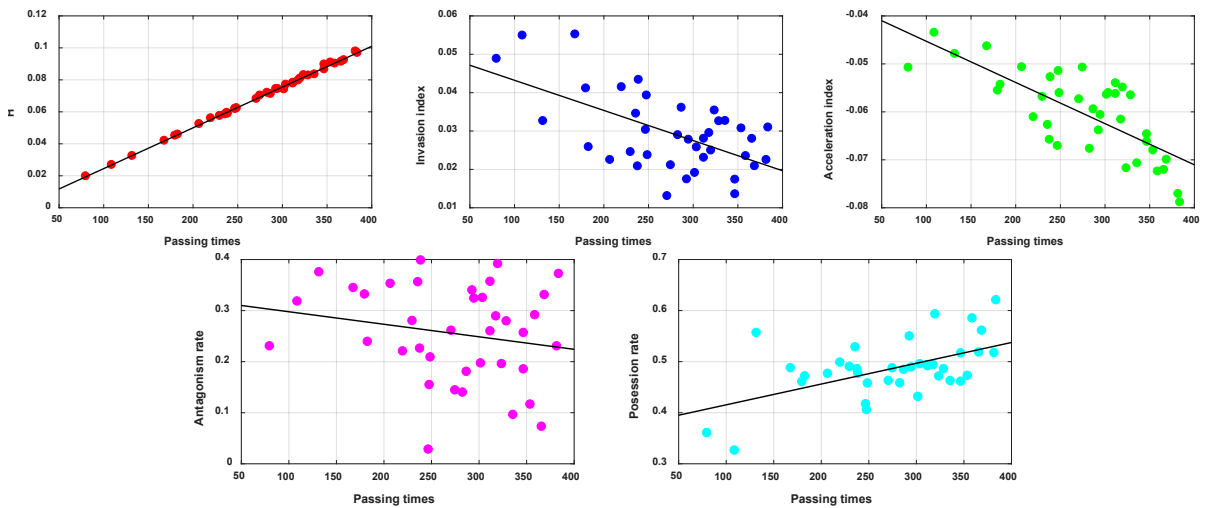


Figure 10. The correlation analysis between the performance indexes and passing times for each team of the league
图10. 联赛各队成绩指标与传球次数的相关性分析图

在评价球队的表现时,除了以上两个变化指标外,我们还根据文献得出了以下结论:一场比赛中,球队的传球次数越多,则得分越多。因此,我们取球队的总传球数、每名球员的平均传球数、球员之间传球数的方差,得到与传球数密切相关的指标——H的调和平均数[5],作为衡量球队实力的指标。

为了验证我们设置的评价指标与团队合作之间的相关性,我们通过传球次数作为横轴来表示队员之间的互动性,并观察其他四组指标之间的偏差大小,偏差越大,相关度越小。

从图10中可以看出,其他四个指标都表现出不同程度的相关性,这也为我们用这五个指标来量化和

评价球队的整体表现和球队内部的配合情况奠定了合理的基础。

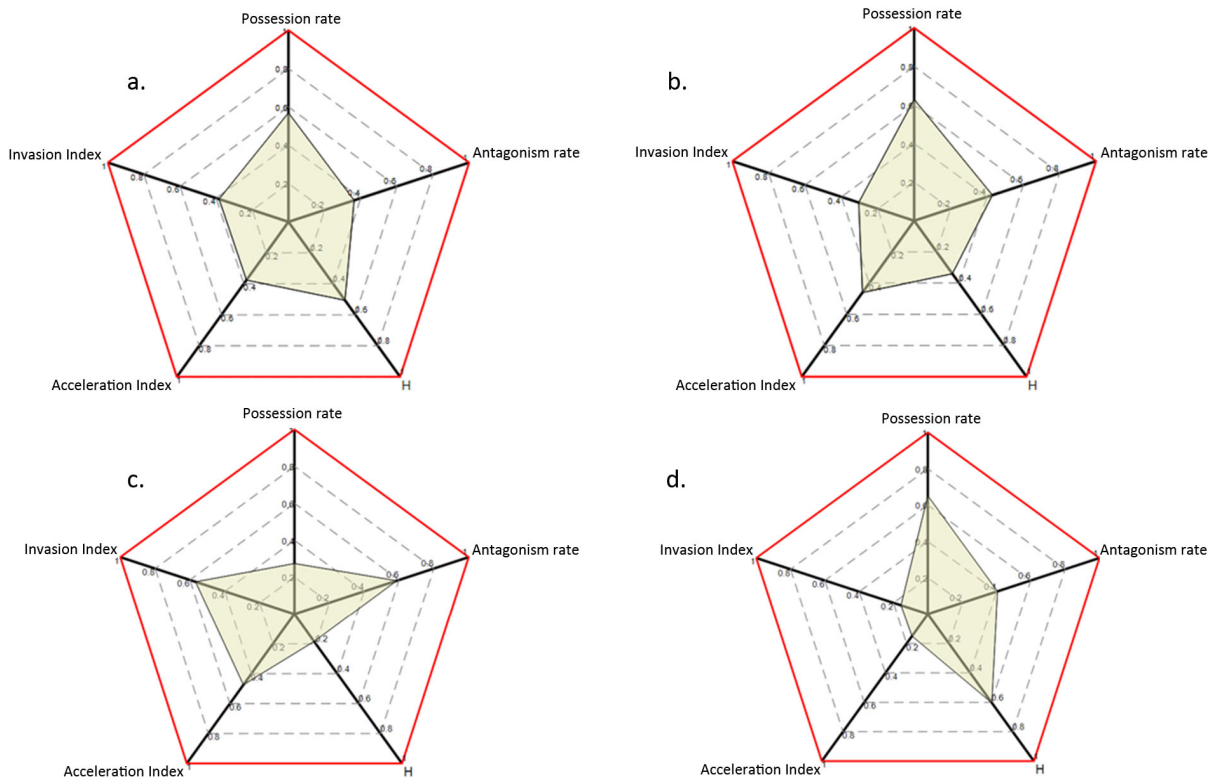


Figure 11. The radar graph of the represented teams for contrast of the indexes

图 11. 四支代表性球队用于对比实力的雷达图

图 11 中, (a)代表了 Huskies 队基于各项评价指标的综合表现, 而(b)、(c)、(d)分别代表了其对手 3、6、16 的实力情况。如果说 Huskies 队的控球策略广泛适用于各种类型的对手, 那么其实力雷达图至少应该有相同的趋势和拐点, 但实际上, 我们通过对 Huskies 队与其他 19 组对手的表现的观察, 由上图可知, 对手的实力很不均衡, 有的球队之间存在指标相差较大的问题。因此, 我们可以得出结论, 虽然 H 队各方面实力均衡, 但这种策略并不具有普遍性。显然, 对于不同水平的对手, 应及时调整球队的表现。

5. 综合加权得分模型

5.1. 模型建立

在建立模型之前, 我们首先制定了一个评估各队战术效果的方法, 如下所示。我们使用上述方法计算出本赛季使用的每一种战术的胜率, 用以评估战术和策略的有效性。同时, 我们设计了一种用加权偏导函数来直观评估效果的评分方法。

如图 10 (雷达图)所示, 面积 S 是公认的重要评估指标, 此外, 它还受到其他 5 个因素的影响: 进攻指数 μ_A 、加速指数 μ_B 、通过指数的平均值 H 、控球率 μ_D 、对抗率 μ_E 。在此基础上, 可以将 S 用这些指数来表示成一个函数。

$$S = S(\lambda_A, \lambda_B, H, \lambda_D, \lambda_E) \quad (5)$$

为了明确战术所得分, 我们将加权得分函数 C 定义如下:

$$C = C(\lambda_A, \lambda_B, H, \lambda_D, \lambda_E) \quad (6)$$

上式涉及到模型二所提出的各个评价指数。这里我们为了构造加权得分函数定义了权重系数 w_1, w_2, w_3, w_4, w_5 ，如下：

$$C = w_1 \lambda_A + w_2 \lambda_B + w_3 H + w_4 \lambda_D + w_5 \lambda_E \quad (7)$$

一个自然的问题是权重的定义。权数可以从公式(S)的部分导数中得出：

$$w_1 = \frac{\partial S}{\partial \lambda_A}, w_2 = \frac{\partial S}{\partial \lambda_B}, w_3 = \frac{\partial S}{\partial H}, w_4 = \frac{\partial S}{\partial \lambda_D}, w_5 = \frac{\partial S}{\partial \lambda_E} \quad (8)$$

但是，存在的问题是我们需要确定面积 S 的计算公式。如前所述，面积 S 是雷达图中五边形的面积，在这种情况下，可以通过面积函数求出 S 。

$$S(\lambda_A, \lambda_B, H, \lambda_D, \lambda_E) = \sin 72^\circ (\lambda_A \lambda_B + \lambda_B H + H \lambda_D + \lambda_D \lambda_E + \lambda_E \lambda_A) \quad (9)$$

事实上， S 也可以被认为是时间 t 的一个函数。当比赛持续进行时， S 也是动态的。而且，从时间的角度看，比赛中的每一个成绩指标也都是动态的。在此基础上，面积 S 和战术得分 C 与时间 t 息息相关：

$$S = S(\lambda_A(t), \lambda_B(t), H(t), \lambda_D(t), \lambda_E(t)) = S(t) \quad (10)$$

$$C = C(\lambda_A(t), \lambda_B(t), H(t), \lambda_D(t), \lambda_E(t)) = C(t) \quad (11)$$

权重，与导数一样，也涉及到时间 t ：

$$w_1 = \left(\frac{\partial S}{\partial \lambda_A} \right) (t), w_2 = \left(\frac{\partial S}{\partial \lambda_B} \right) (t), w_3 = \left(\frac{\partial S}{\partial H} \right) (t), w_4 = \left(\frac{\partial S}{\partial \lambda_D} \right) (t), w_5 = \left(\frac{\partial S}{\partial \lambda_E} \right) (t) \quad (12)$$

因此，加权得分函数可以改写为：

$$\begin{aligned} C &= \left(\frac{\partial S}{\partial \lambda_A} \right) (t) \lambda_A + \left(\frac{\partial S}{\partial \lambda_B} \right) (t) \lambda_B + \left(\frac{\partial S}{\partial H} \right) (t) H + \left(\frac{\partial S}{\partial \lambda_D} \right) (t) \lambda_D + \left(\frac{\partial S}{\partial \lambda_E} \right) (t) \lambda_E \\ &= \sin 72^\circ (\lambda_A (\lambda_B + \lambda_E) + \lambda_B (\lambda_D + H) + H (\lambda_D + \lambda_B) + \lambda_D (H + \lambda_E) + \lambda_E (\lambda_A + \lambda_D)) \end{aligned} \quad (13)$$

简而言之，利用上述性能指标可以得到 C 。总分 C' 最终可以通过计算极限获得：

$$C' = \lim_{t \rightarrow 5400} C(t) \quad (14)$$

根据以上分析，在具体的比赛中，我们可以不考虑对手的情况来评估 Huskies 队的表现情况。

5.2. 模型评估与结果分析

在本节中，我们采用加权得分模型对 Huskies 队在整个赛季的表现进行分析，如图 12 所示。 C 值越高，表示 Huskies 球队表现的越好。根据这个规律，我们可以得出，Huskies 队的得分情况一般反映了 Huskies 队的表现情况。比如说，第 32 场比赛，Huskies 队一败涂地，其得分也较低。同样的，第 36 场比赛中 Huskies 队伍是大胜，其得分也较高。不过，有一场比赛似乎与真正的进球情况并不对应。值得注意的是，比分只是反映球队战绩的一个维度，其并没有考虑对手的情况。一旦对手的实力大大超过了 Huskies 队，那么无论 Huskies 队的表现有多精彩，进球差可能也是巨大的。总之，我们可以得出结论，此分析模型在具体比赛中是可以使用的。

至于对教练的建议，在不参考对手的情况下，我们呼吁教练可以用模型中得分高的战术作为参考，但是这只是基于球队自身情况而得出的结论，因为我们已经明确了这个模型并不是每个对手都能通用的，所以教练可以把得分模型和其他模型结合在一起进行决策。

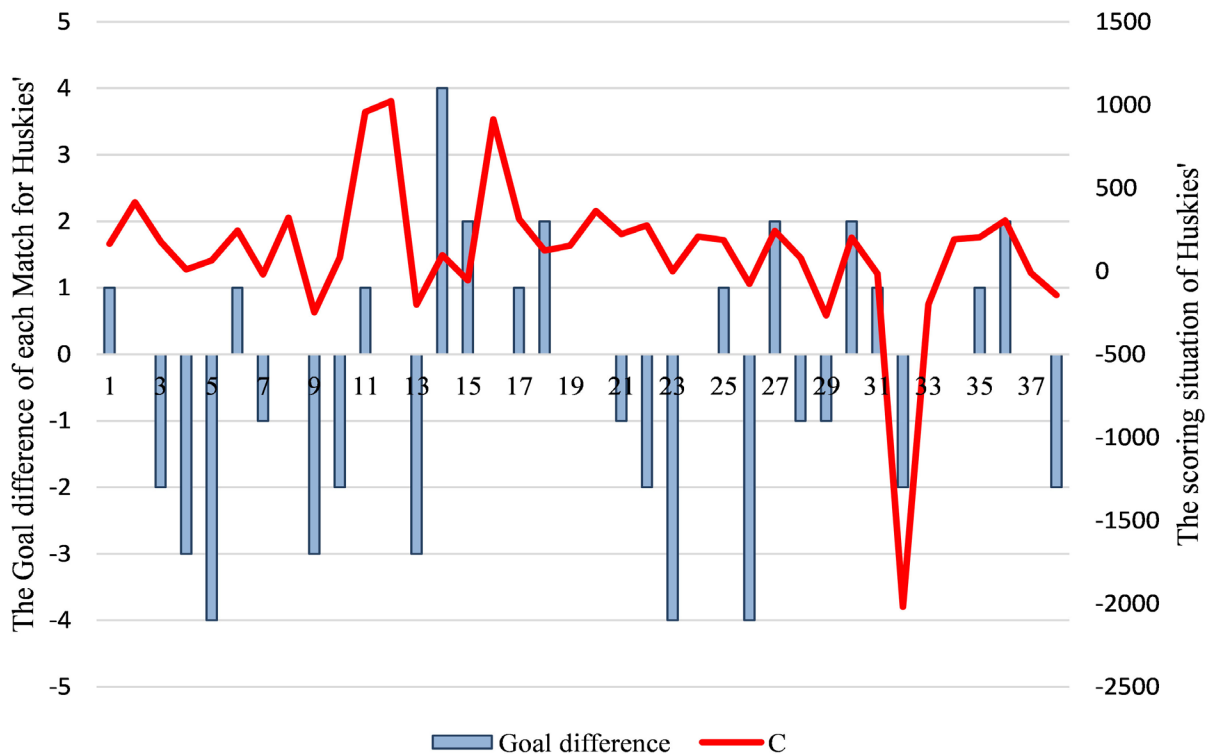


Figure 12. The contrast between matches situation of Huskies in the season and its performance scoring situation
图 12. 本赛季 Huskies 队的比赛情况与战绩得分情况的对比

5.3. 战术与球员配置建议

根据我们的网络模型分析，我们计算出各队的战术变化情况和相应的胜率如表 2 所示。在每场比赛开始和结束的大部分时间内 Huskies 队和对手的阵型都进行了调整，表 2 中的空白部分是指：通过位置坐标分析发现在比赛的第一阶段球队并没有确定具体的阵型，这可以被认为是处置失误或者是已经对球员的位置进行了安排，但球员并没有认真执行。

Table 2. The winning rate of the Huskies when adopting the corresponding tactics structure
表 2. Huskies 队采取相应战术结构时的胜率

		Huskies 队										
开始时的战术		334	424	433	442	451	532	541				
胜率		0	0	0.64	0.4	0	0	0				
结束时的战术		361	424	432	433	441	442	451	460	523	532	541
胜率		0	0.67	0	0.4	0	0.23	0	0	0	0.25	0.5

Table 3. The winning rate of the Huskies when opponents adopted the corresponding tactics structure
表 3. 当对手采用相应的战术结构时，Huskies 队的胜率

		对手队										
开始时的战术		333	352	433	441	442	451	460	532	541		
Huskies 队的胜率		0	0	0.6	0	0.38	0.38	0	0	0		
结束时的战术		342	343	423	432	433	441	442	451	522	532	541
Huskies 队的胜率		0.5	0	0	0.67	0	0	0.67	0.125	0	0	0.17

同时,我们建立了 Huskies 队应用各种阵型时,对阵各个对手的胜率统计情况,胜率情况如表 3 所示。通过观察表 3 中胜率大于 0.5 的比赛对应的球员阵容,我们发现,后卫的数量最好控制在 4 人左右。如果后卫人数过多,前锋力量不足,不便于进攻;如果后卫人数过少,则容易失去防守,使对方队伍获得优势。

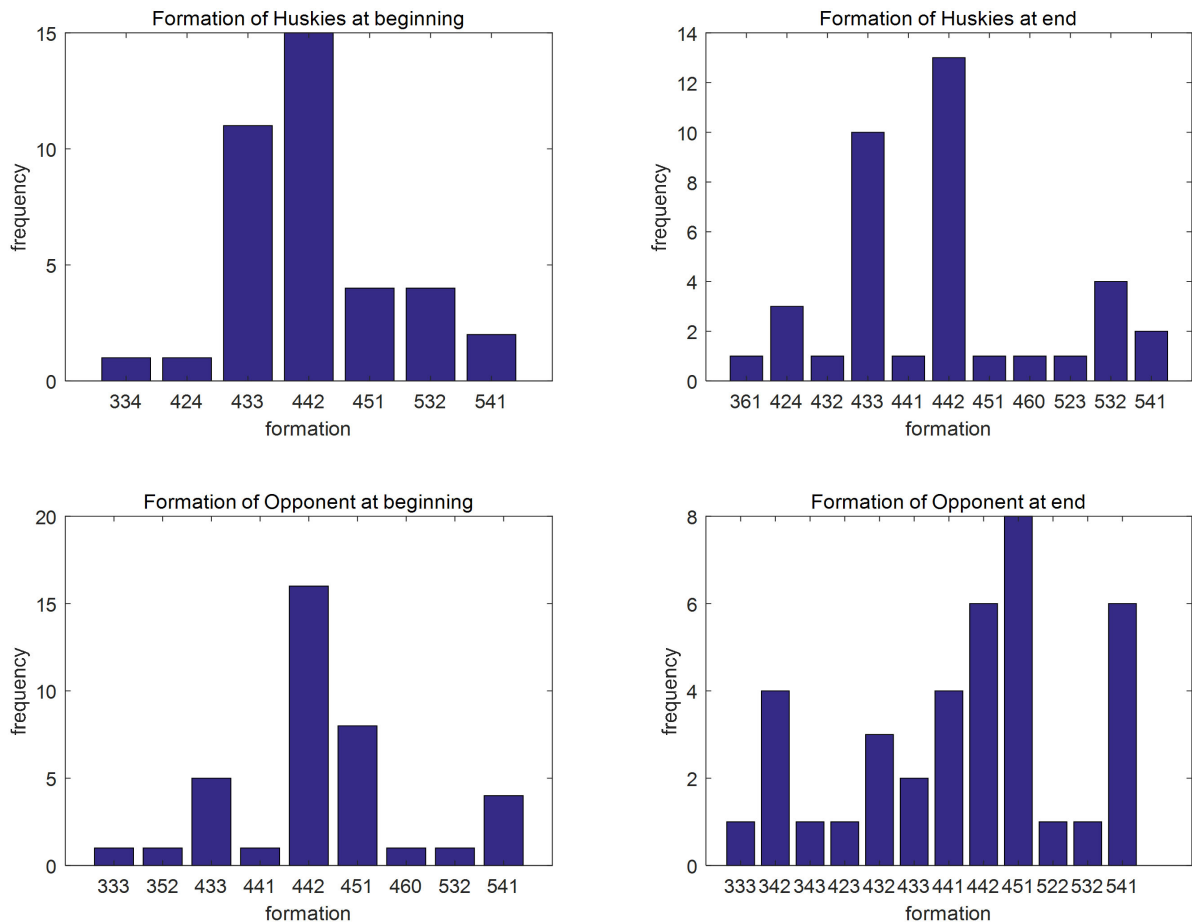


Figure 13. The formation (Tactics) frequency using in the season

图 13. 阵型(战术)在本赛季使用的频率

Huskies 队最常见的阵型显然是“442”,这与上面提出的“四后卫”阵型一致。第二种最常见的阵型是“433”。虽然对手的阵型明显多变,但从图 13 右下角的横轴上可以看出,尤其是在下半场和比赛快结束时,对手开始不断变换阵型,中场球员的数量增加,有效传球增多,全队的凝聚力增强,这就意味着 Huskies 队教练在下半场要根据球员的实力和对手的战术来考虑让 Huskies 队在阵型上做更多的改变(见图 14)。

6. 结论

我们通过无向图模型建立了传球网络,从数据中研究了 Huskies 队的三元配置和二元配置,发现菱形配置等结构在比赛中也起到了重要的作用。以上是从微观的角度进行了初步分析。从整队和长周期(整个赛季)来看,我们还对 Huskies 队与其他 19 个对手的比赛过程进行了对比。这在下面的两个模型中可以看到。

之后,我们通过设定入侵指数、加速指数等指标,进一步对球队的整体实力水平进行量化评价。通过入侵指数的变化,我们可以实时分析出球队的攻击力,通过对比加速指数,我们可以发现球队阵容的问题。我们还可以巧妙地利用雷达图中的属性和区域来判定 Huskies 队的表现水平和其他 19 个对手的实力。通过与 19 个对手的对比,我们证明了 Huskies 队的阵型结构策略并不是普遍适用于所有对手。

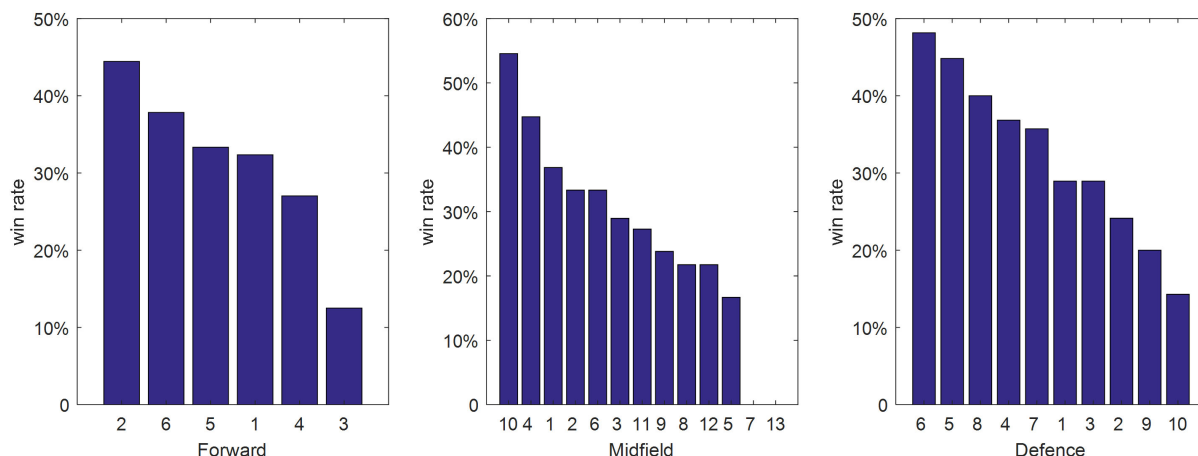


Figure 14. The winning rate of Huskies when the corresponding player is on the ground
图 14. 当相应球员在场时，Huskies 队的胜率

通过统计分析和对比 Huskies 队与对手的阵型胜率，我们对 Huskies 教练员提出了一个建议：在保持阵型灵活多变的基础上，尽量保持中后卫和中场的 4 人均衡配置，呈现出整体的“443”阵型。因为我们发现，有针对性地破解对手的阵型，是 Huskies 队胜利的关键，所以一个阵型不能使用时间太长，要及时调整和变化，才能有效地战胜对手。同时要保持严格的替补队员的选择，在上半场要教会队员们储备力量。

团队合作给了我们一定的启示，人与人之间的沟通和互动不仅在生活的方方面面十分重要，在体育运动中更是彰显其重要性。社会系统中充满了更多无形的指标，包括心理健康、团队成员荣誉感的变化、个人的收获等。量化和评价一个高效的团队，不仅仅是为了赢得比赛，更重要的是团队成员从团队合作中获得的优势互补。

致 谢

本文作者向 Huskies 队比赛数据集的提供者：美国数学及其应用联合会表示感谢！

参考文献

- [1] Pappalardo, L., et al. (2019) A Public Data Set of Spatio-Temporal Match Events in Soccer Competitions. *Scientific Data*, **6**, 1-15. <https://doi.org/10.1038/s41597-019-0247-7>
- [2] Taki, T. and Hasegawa, J. (2000) Visualization of Dominant Region in Team Games and Its Application to Teamwork Analysis. *Proceedings Computer Graphics International 2000*, Geneva, 19-24 June 2000, 227-235. <https://doi.org/10.1109/cgi.2000.852338>
- [3] Gürsakal, N., Yılmaz, F., Çobanoğlu, H. and Çağlıyör, S. (2018) Network Motifs in Football. *Turkish Journal of Sport and Exercise*, **20**, 263-272. <https://doi.org/10.15314/tsed.465664>
- [4] Cintia, P., Giannotti, F., Pappalardo, L., Pedreschi, D. and Malvaldi, M. (2015) The Harsh Rule of the Goals: Data-Driven Performance Indicators for Football Teams. 2015 *IEEE International Conference on Data Science and Advanced Analytics (DSAA)*, Paris, 19-21 October 2015, Article ID: 7344823. <https://doi.org/10.1109/dsaa.2015.7344823>
- [5] Duch, J., Waitzman, J. and Amaral, L. (2010) Quantifying the Performance of Individual Players in a Team Activity. *PLoS ONE*, **5**, e10937. <https://doi.org/10.1371/journal.pone.0010937>
- [6] Buldú, J., Busquets, J., Echegoyen, I. and Seirullo, F. (2019) Defining a Historic Football Team: Using Network Science to Analyze Guardiola's F.C. Barcelona. *Scientific Reports*, **9**, Article ID: 13602. <https://doi.org/10.1038/s41598-019-49969-2>