**问题一**

**introduction：**文章通过三个维度，分别是队伍整体传球能力，球员个人吸引力，以及球员之间的距离来定义传球网络。

**假设：**

**①在构建passingnetwork时，忽略fullevents中的除passing外其余情况。**足球中的某些战术会涉及到故意制造犯规使球出界，在此简化这部分控球技巧。

**②传球中球的运动轨迹为一条从传球点到接球点的直线。**

**③在小队配合中，只考虑最近发生传球动作的三个及三个以下的球员。**足球运动球员位置时刻在发生变化，常见的合作单元大小一般为二人或三人，所以我们仅对二元配合以及三元配合进行讨论，微观讨论某一小部分球员运动能更好地反映小队配合能力。

**④在小队配合中，从一次传球到下一次传球的时间间隔小于五秒。**

**符号及定义：**

|  |  |
| --- | --- |
| **符号** | **定义** |
|  | 队伍某场比赛传球总次数 |
|  | 某场比赛中队伍里有上过场的队员的总人数 |
|  | 传接球种类数 |
|  | 队伍某场比赛完成人配合以及 人配合的总次数 |
|  | 第个球员某场比赛接到的总传球数 |
|  | 第个球员某场比赛传出的总传球数 |
|  | 第 个球员某场比赛接到第 种传球总数 |
|  | 第 个球员某场比赛传出第 种传球总数 |
|  | 第个球员某场比赛累计控球时间（单位：分钟） |
|  | 第个球员某场比赛对其他球员的平均吸引力 |
|  | 第个球员某场比赛在场上的平均控球位置 |
|  | 第 种传出的传球种类的权重 |
|  | 第 种接到的传球种类的权重 |
|  | 第 种传出的传球种类的权重 |
|  | 第 种接到的传球种类的权重 |
|  | 某场比赛中队伍的传球网络的质心坐标 |
|  | 坐标 到坐标 的欧氏距离 |

方法：

对于队伍整体传球能力，分别通过队伍传球总数以及队伍全场配合次数来评判其好坏。附件中已经给出了整个队伍全场传球次数。赛季配合总次数则根据数据在满足假设③和假设④的前提下进一步筛选，分别得到了和。此项指标在宏观的角度评价了一个队伍传球能力的好坏。站在球员个人的角度，统计出了球员与球员之间的传接球次数作为传球网络中的边的权重，而使用万有引力模型来描述该球员的重要程度，同时也是传球网络结点大小的重要参考。

首先假设我们是 Huskies 队伍的分析员，针对某一场比赛，我们进行分析，对于该场比赛上场的各位球员的个人吸引力，其实这里的吸引力可以代指球员的个人能力，首先通过附件直接得到了Huskies队伍第 个球员在该场比赛接到的传球数为，传出的传球数为 ，累计有效控球时间，显然，若一名球员在整个赛季中接到的和传出的传球数与 多，控球时间长，则可以当然其吸引力大。由于传接球有不同的种类，传接球的总次数由不同类型的传接球构成：

其中 代表的总的传球种类数， 代表的是第 个球员， 代表的是第 种传球。

针对不同的种类传接球，其难度是不相同的，自然其对吸引力大小的影响程度不同，所以针对不同种的传接球次数 ， 应该赋予不同的权重：

由于权重的计算，如果只针对某一场比赛带有随机性，所以为了降低这种随机性，我们统计整个赛季的各类传球次数：

表1Huskies 队伍该赛季传出的传球统计

|  |  |  |
| --- | --- | --- |
| 对应符号 | 传球种类 | 传球个数 |
|  | Head pass |  |
|  | Simple pass |  |
|  | Launch |  |
|  | High pass |  |
|  | Hand pass |  |
|  | Smart pass |  |
|  | Cross |  |

表2Huskies 队伍接到的传球统计

|  |  |  |
| --- | --- | --- |
| 对应符号 | 传球种类 | 传球个数 |
|  | Head pass |  |
|  | Simple pass |  |
|  | Launch |  |
|  | High pass |  |
|  | Hand pass |  |
|  | Smart pass |  |
|  | Cross |  |

首先统计该场比赛的各种类传接球频率，而后取倒数之后再归一化得到：

最后计算得到权重为：

表3 Huskies队在传出传球计量时各类传球权重

|  |  |  |
| --- | --- | --- |
| 对应符号 | 传球种类 | 权重 |
|  | Head pass | 0.047 |
|  | Simple pass | 0.003 |
|  | Launch | 0.126, |
|  | High pass | 0.045 |
|  | Hand pass | 0.210 |
|  | Smart pass | 0.365 |
|  | Cross | 0.205 |

表4Huskies队在接到传球计量时各类传球权重

|  |  |  |
| --- | --- | --- |
| 对应符号 | 传球种类 | 权重 |
|  | Head pass | 0.047 |
|  | Simple pass | 0.003 |
|  | Launch | 0.126 |
|  | High pass | 0.046 |
|  | Hand pass | 0.209 |
|  | Smart pass | 0.364 |
|  | Cross | 0.205 |

透过权重我们可以通过经验结合之前统计的各类传球数统计推断这是合理的分配，首先观察传球种类数目分布图：

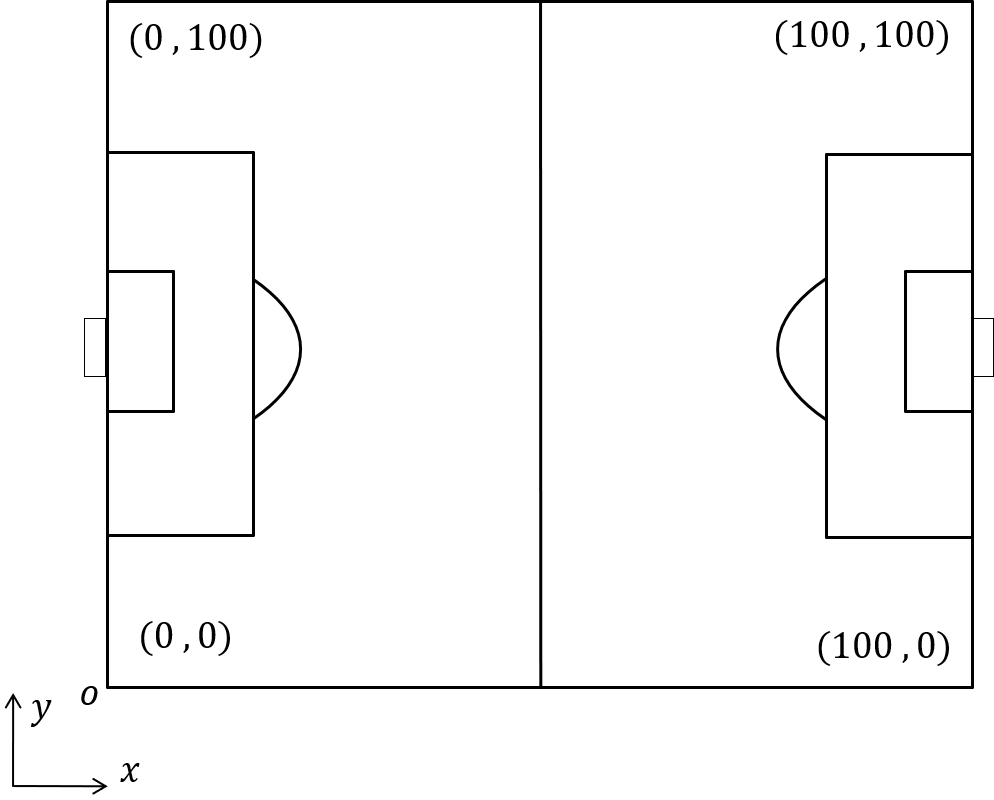
图1 Huskies 赛季传出的各类传球数目统计图

图2 Huskies整个赛季接到的各类传球数目统计图

例如，Smart pass 比较少发生，我们可以认为 Smartpass发生需要球员有很强的个人能力，无论是接球方还是传球方都有着能力的高度要求，所以这样的球比较有分量是正确的，所以赋予较高的权重，而 Simplepass是最常发生的传球，这一类传球显然并不要求球员有很强的个人能力，所以自然赋予较低的权重。

根据上述重重分析，为了综合考虑上述的球员持球时间 ，传接球次数 ，，按传球种类划分的传接球次数 ， ，则球员的吸引力定义为：

而后球员的位置通过整场比赛的接球时的位置取平均得到 ，这里要对 进行说明，首先虽然足球的上下半场会交换位置，但是考虑到数据简化，所以我们始终将 原点放置在左下角，越往右 坐标越大，即越靠近对方那一边，反之， 坐标越小，即越靠近我方这一边，而当面朝对方球门时，始终将右手边作为 轴正方向，左手边作为 轴负方向，再考虑到场地大小的波动，我们将足球场地按比例划分为 等分，即 与 的取值范围为：



将每个球员作为一个结点，则结点在图中的位置就是 ，而我们将吸引力大小作为结点的半径大小，而后通过统计每场比赛的球员之间的传球，绘制出结点与结点之间的连线，即传球网络的边，边的深浅与结点之间或者说球员之间传球次数线性相关，边的深浅也体现了边所带有的权重。由此，得到了一张传球网络图。通过使用 Huskies队伍第三场比赛的数据，绘制图如下：

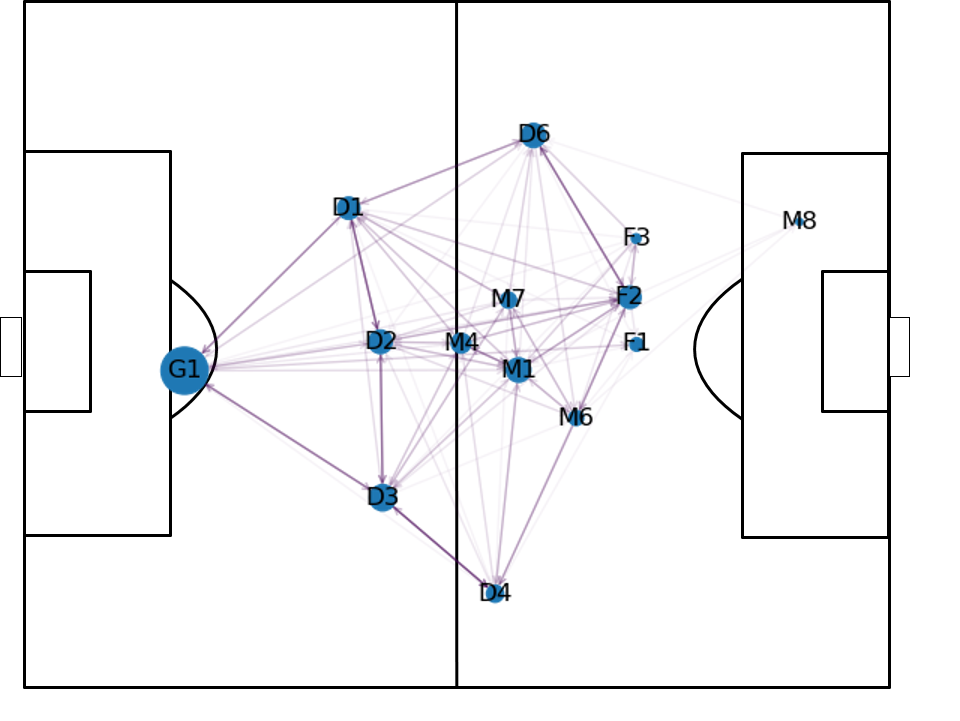


图3第四场比赛 Huskies 队伍的传球网络

通过对整个赛季的所有队伍进行统计，我们共得到了38个Huskies队伍的传球网络和38个OP队伍的传球网络。综合比赛结果和进球数，分别从三个角度对传球网络进行度量。

一、传球网络质心。

在传球网络中，球员的位置通过接球时的距离取算术平均后直接获得，由此可以猜想整场比赛中球员的接球位置能够反映出球队在该场比赛中的表现情况，因此我们定义了一个名为网络质心的参数，用来衡量整支队伍在球场上的相对位置。

类比物理学中质心的计算公式：

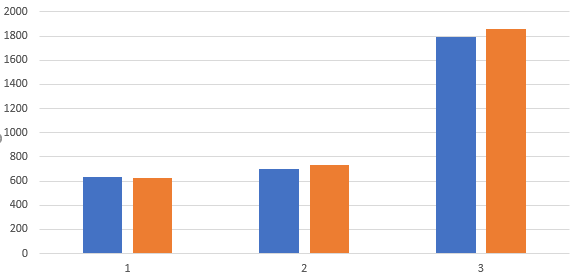
式中是第个物体的质量，是考虑到的所有结点的总质量，将质量这一概念用球员吸引力替换：

式中为该场比赛中的出场人员数量，为球员吸引力的总和。

而后我们为了综合考虑 坐标与 坐标，我们通过计算质心到原点的距离来分析：

根据坐标的定义，坐标越靠近时代表越靠近本方球门，坐标越靠近时代表靠近对方球门。则质心到原点的距离越大，则能从一定程度上反映球队倾向于出现在对方的半场，这说明球队更趋向于积极进攻，这应该在一定程度上能够说明球队的进攻组织能力，进攻组织能力当然是团队合作能力的一部分。据此，我们统计了Huskies队伍和它的对手在整个赛季38场比赛中的网络质心。

分析结果如如图4：



图四 网络质心到原点距离的对比图

图中，蓝色柱体代表Huskies队的网络质心求和，橙色则代表他的对手的网络求和。情况1统计了Huskies负于他的对手的总共15场比赛，而情况2则表示胜于他的对手的总共13场比赛。情况3统计了所有比赛Huskies和对手的网络质心求和。

由图中可以看出：Huskies队的网络质心值稍微小于他的对手，在赢球的情况下，这项指标将超过他的对手，而在输球的情况下，这项指标低于他的对手。说明网络质心和对局输赢之间存在正相关关系，通过整个赛季的网络质心分析，可以看见Huskies队的网络质心值稍微小于他的对手，这也能够说明Huskies队在本赛季赢球占比不是太大。

二、配合情况

通过统计一场比赛可以得到球员之间完成二人配合和三人配合的次数。在传球网络中，直观地可以通过边的颜色来反应球员之间的配合情况。需要注意的是，配合的情况往往与当场采取的策略有关系。

三、质心离散度

在得到网络质心后，球员位置通过球员吸引力加权后与网络质心乘上全队平均吸引力的标准差，即：

四、传球成功率

在足球比赛中经常会发生球权转换，显而易见，团队配合默契度高的队伍相应的在传球时也会表现出高的成功率，从而给队伍的进攻以及进球得分做好辅助。因此，一个队伍在比赛时会倾向于保持一个高的传球成功率。综上，我们对传球成功率作出狭义和广义两个定义。

狭义定义为：在上场队员作出传球动作（pass）后，本队队员成功接到球的概率。由此定义并结合球队的数据，分母取本队队员所有的传球动作（pass），从fullevents中直接筛选本队数据；分子取成功的传球动作，从passingevents中直接筛选本队数据。

广义定义为：当上场队员掌握球权时，在其所做出的动作（all event types）后，球权仍然被本队队员掌握的概率。由此定义并结合球队的数据，分母取本队队员掌握球权时的所有动作，从fullevents中的TeamID项目中筛选本队数据；分子取本队队员掌握球权并作出动作后球权仍在本队时的所有动作，从fullevents中筛选本队数据，若该动作后的TeamID仍与本队相同，则判断该动作为掌握球权成功。

五、加权平均传球路径

第三问：传球落点分析