**h**

**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 |

模型建立：

模型一一：传球网络模型

我们使用网络科学中的有向节点图来构建传球网络。首先将球员当做网络图中的节点，根据物理学中的万有引力公式定义了球员个人吸引力，以此作为衡量节点大小的依据。将球员之间的传球当做边，并根据传球次数的多少赋予边一定权重，最后根据球员整场比赛的算术平均坐标作为节点图中的坐标。

对于队伍整体传球能力，分别通过队伍传球总数评判其好坏。附件中已经给出了整个队伍全场传球次数。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：

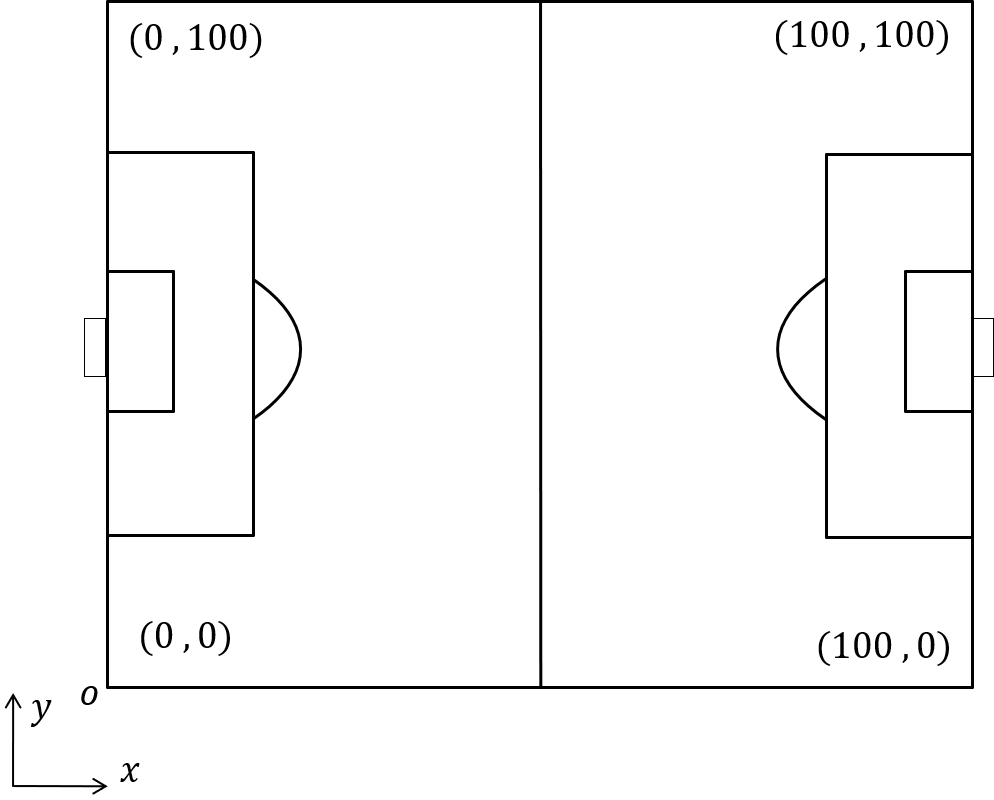
图1 Huskies 赛季传出的各类传球数目统计图

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，边的颜色深浅与结点之间或者说球员之间传球次数线性相关，边的颜色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：

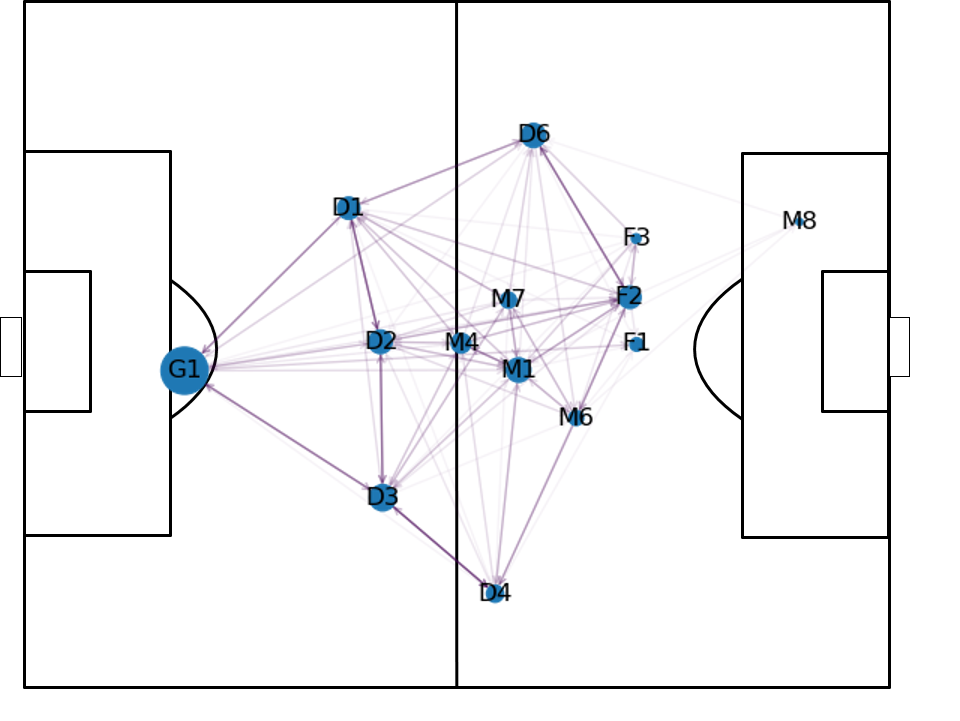


图3第三场比赛 Huskies 队伍的传球网络

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。传球网络图可以直观的看到球员之间的互动情况以及某个球员在整场比赛中的重要性等指标综合比赛结果和进球数。

模型二：团队合作评估模型

首先由于团队合作能力是一个非常复杂的概念，所以想要直接对其进行评估是一件非常复杂的事情，因为其受到了很多因素影响，所以我们决定采用模糊评价的方法来对团队合作这个概念进行评估，首先团队合作的评价集被划分为好与适中与坏，当然与评价集直接挂钩的第一级评价指标为传球数、射门数、进球数、进球率、控球率、输赢率这些直接反映团队配合能力的指标，这些指标可以作为球队的绩效评估，当然能够作为团队合作能力配合的指标，比较足球是一项多人体育竞技项目，再顶尖的球员都无法一个人赢得胜利，而据此第二级指标，还加入了传球网络的质心到圆心距离，传球网络离散程度，传球网络平均传球距离，最长传球路径等指标，这样看来第一级指标更加倾向于基于足球技术层面分析，而第二季指标则更倾向于基于数据层面来分析的指标。

首先直观上来说，第一级的评价指标是几乎和团队配合能力直接相关联的，那么我们考虑将第一级指标作为评估第二级指标是否有效的金标准。而对于总体的团队配合能力，我们使用比较分析法，通过获取赛季所有队伍的平均水平来对某只队伍进行评价，具体来说为计算得到赛季所有队伍的各项指标的平均数，以这些指标为基准评估出来的团队配合能力应该是适中的层次，然后将要评价的队伍与适中层次的这些指标比较，各项指标如果更优就是好，更差就是坏。

首先说明一点，我们将 Huskies 队伍简单命名为 Husk ，由于只有 Huskies 队伍与其他队伍对战的数据，所以如果单看某一只对手球队的话，没有意义，因为整个赛季也只会与 Huskies 队伍比赛一两场，没有太多的统计意义，所以我们将所有的对手球队结合起来求取平均来分析，我们将对手球队综合成一个队伍来看，所以将 Opponentx (x 为 1 ~ 19) 即与 Huskies 对战的其他 19 只队伍的综合命名为 Oppo 。

1. 第一级评价指标

首先统计 38 场比赛中，Husk 队与 Oppo 的所有队比赛的输赢情况，即胜局情况与平局情况和负局情况，进球数和射门数，进球率，由于有主客场对于这些指标的影响比较大，所以就将这几个指标分为主客场来讨论分析：



图4 (a) 主场时 Husk 与 Oppo 的输赢情况;(b) 客场的时候 Husk 与 Oppo 的输赢情况

(c) 主场的时候 Husk 与 Oppo 的射门和进球情况; (d) 客场的时候 Husk 与 Oppo 的射门传球情况; (e) 主场的时候 Husk 与 Oppo 的进球率; (f) 客场的时候 Husk 与 Oppo 的进球率;

从上述指标分析，从输赢情况开始，可以发现无论是主场还是客场， Husk 都比 Oppo 的情况略差，所以从输赢的情况来看， Husk 的团队合作能力应该低于平均水平。

从进球数和射门数分析，进球数的多少与射门数的多少都是是团队进攻组织能力的体现，当然进攻组织能力也是团队合作能力的其中一个部分，这样看来 Husk 的进攻组织水平低于平均水平，这也体现了 Husk 的团队合作能力应该低于平均水平。

从进球率上来看，Husk 的机会把握能力还是不错的，这个方面上来看 Husk 的团队合作能力略高于平均水平，不过考虑到进球率受球员个人因素影响较大，所以对团队合作能力贡献小，进球率高与其说是团队合作能力的体现，到不说主要是球员的个人能力的体现。

从上面的角度上来看，大致可以确定 Husk 是一支中等偏下的队伍，团队组织能力在联盟中可能处于中下水平。

而后我们从传球数，传球成功率，控球率的角度进行分析，此时主客场对指标的影响较小，所以此时可以不考虑主客场：

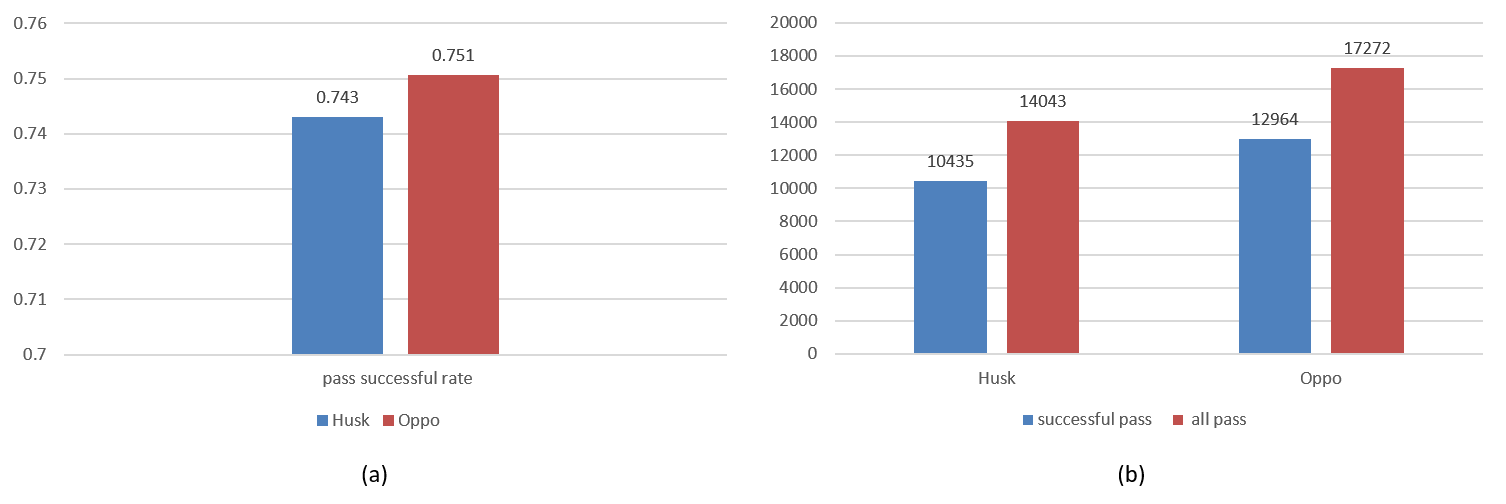


图5 (a) Husk 队伍与 Oppo 队伍的传球成功率比较; (b) Husk 队伍与 Oppo 队伍的成功传球数与总传球数

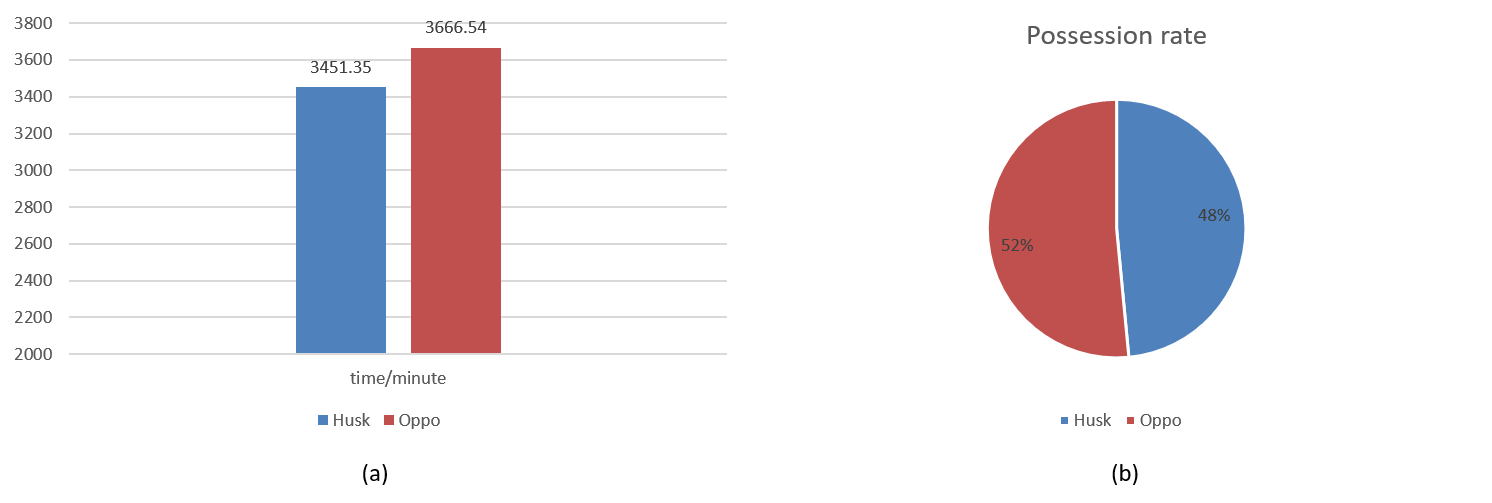


图6 (a) Husk 队伍与 Oppo 的控球时间比较; (b) Husk 队伍与 Oppo 的控球率比较

传球是组织进攻的重要手段，也是团队配合的高度体现，通过比较发现 Husk 队伍的传球数目和成功传球数，以及成功传球率都弱于 Oppo 所代表的平均水平，这也反映了 Husk 队伍是一支团队组织能力偏弱的队伍，而控球率虽然更多的反映的是团队的综合能力，但是毕竟合作能力也属于团队综合能力的一部分，所以我认为，通过控球率的比较，再次确认了 Husk 队伍团队组织能力低于平均水平这一事实。

二、第二季评价指标

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。

考虑到 坐标指示的是在场上的左右手边，而这对于我方敌方并没有很明显的影响，所以不考虑。

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。则质心的 坐标越大，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Husk队伍和 Oppo 队在整个赛季 38 场比赛中的平均网络质心 坐标：

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。

需要注意的是，传球网络质心位置并不直接与足球比赛的输赢有关，输赢往往取决于球队的具体策略以及球员当场发挥情况。但是质心横坐标位置距离原点的远近可以看出球队对场上局势是否掌握有可以帮助取得优势的主动权

二、配合情况

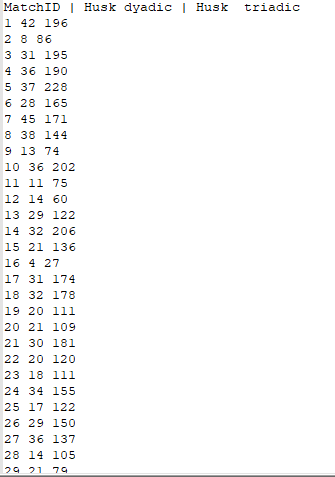
球员之间的配合情况往往是一个球队的默契程度的直观表现。足球比赛中的配合往往根据球员所属定位的不同而有所差别。例如在某些场合中，中场和前锋的配合能给对面的防线带来巨大的压力，又或者是依靠几个后卫之间的腾挪移动组成了铁桶一般的防线来阻挡对面进攻。

我们分别从宏观上的赛季配合总次数以及微观上的球员一场比赛中的配合次数来衡量一支队伍的配合情况。

赛季配合总次数根据数据在满足假设③和假设④的前提下进一步筛选，筛选方法为：

对二人配合，认为球从一个球员传出到另一个球员后，再将球传回视为完成一次二人配合》

对于三人配合，认为将球从一个球员传出后经过第二个球员传出，最终到达第三个球员视为完成一次三人配合，据此分别得到了和，如图



。在一场比赛中，本方和对方往往都有特定的传球图案（passingmotifs）。特备是足球运动，球员之间的互动被认为是一个复杂的网络，通过提取出传球网络，可以将网络进一步细分成子图，该子图在网络中的出现频率要比在随机网络中更高（5），If we want to disrupt the opponent's game format, we must first be aware of the pass motifs that the team often uses，（13）在我们所构建的传球网络中，通过构建与球员在场上位置有关的passingmotifs来识别足球比赛中常用到的配合。使用以下算法遍历传球网络：

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。直接定义传球成功率：

由于真实情况比理想情况要复杂的多，在某些特定场合，球员做出某些动作后（如为了拦截进攻而故意制造的出界球）也应该在宏观上被视为一次成功控球。对此，对传球成功率分别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。

由此，我们得到了Huskies队和敌队的传球成功率统计图。如图：

可以看到，在广义和狭义两个角度，Huskies的传球成功率均值都略低于对手。

五、加权平均传球路径

第三问：传球落点分析