

基于复杂网络与多联赛大数据的足球传球模式解析及战术风格识别

马茁冉

华东理工大学, 中国·上海 200237

摘要: 本研究运用复杂网络理论, 系统分析与比较了三种不同类型的足球传球行为: 渐进式传球、直接传球与关键渗透传球。基于 2015/2016 赛季欧洲五大联赛 (英超、西甲、意甲、德甲、法甲) 共 1,705 场比赛, 我们针对每种传球类型构建了有向加权传球网络, 并计算了包括密度、聚类系数、度中心性与连通性在内的多项网络指标。

关键词: 复杂网络; 足球分析; 传球行为; 网络科学

Analysis of Football Passing Pattern and Tactical Style Recognition Based on Complex Network and Multi-league Big Data

Ma Zhuoran

East China University of Science and Technology, China Shanghai 200237

Abstract: Based on the complex network theory, this study systematically analyzes and compares three different types of football passing behaviors: gradual passing, direct passing and key penetration passing. Based on 1,705 matches in five major European leagues (Premier League, La Liga, Serie A, Bundesliga and Ligue 1) in the 2015/2016 season, we constructed a directed weighted passing network for each passing type, and calculated a number of network indicators including density, clustering coefficient, centrality and connectivity.

Keywords: Complex network; Football analysis; Passing behavior; Network science

0 引言

足球作为一个复杂动态系统, 其集体行为模式源于球员间的交互作用。其中, 传球行为是球队协调与战术执行的基本机制。不同类型的传球——无论是为保持控球的短传推进、为绕过中场的长传直接转移, 还是在前场的渗透性传球——都反映了不同的战术理念与比赛管理策略。

传统的足球分析主要依赖描述性统计 (如传球成功率、跑动距离) 与定性评估。尽管有价值, 这些方法往往难以捕捉球队协调的结构特性。复杂网络理论通过将球员建模为节点、传球为边, 提供了一个强大的量化框架, 从而将定性的战术概念转化为可测量的网络属性。

尽管网络方法在体育领域的应用日益增多, 现有研究大多集中于整体传球网络的分析, 而未区分功能不同的传球类型。这一局限掩盖了球队为达成特定战术目标而采用多样化传球策略的细微差异。例如, 一支球队可能同时在渐进式构建中呈现密集的局部连接, 而在直接转换中保持稀疏的远距离连接——这些模式在整体网络分析中往往被混淆。

本研究通过复杂网络分析, 首次系统比较了三种不同

的传球行为: (1) 用于控制型构建的渐进式传球; (2) 用于快速转换的直接传球; (3) 用于创造机会的关键渗透传球, 以填补上述研究空白。我们假设这些传球类型将呈现根本不同的网络结构, 反映其各异的战术目的。

我们的研究问题包括:

RQ1: 渐进式、直接与关键渗透传球网络在结构上有什么差异?

RQ2: 这些网络特征在不同联赛间如何变化?

RQ3: 能否基于传球网络特征将球队聚类为不同的战术风格?

通过回答这些问题, 本研究在理论与实践上均具有贡献。理论上, 我们通过展示复杂网络理论如何区分球队协调的功能性差异, 推动了体育分析的发展; 实践上, 研究结果为教练与分析师提供了用于战术评估、对手侦察与球员发展的量化工具。

1 方法论

1.1 数据来源与样本

我们使用 StatsBomb 开放数据仓库提供的职业足球比赛事件数据。样本涵盖 2015/2016 赛季欧洲五大联赛

的 1,705 场比赛：英超（359 场）、西甲（345 场）、意甲（361 场）、德甲（283 场）与法甲（357 场）。为确保竞争平衡，仅纳入最终比分差 ≤ 3 球的比赛，共得到 3,410 次球队 - 比赛观测数据（2 支球队 $\times$ 1,705 场比赛）。

1.2 传球类型分类

根据事件数据中的客观标准，将传球分为三类互斥类型：

（1）渐进式传球：常规比赛中的成功传球，长度 ≤ 25 码，不包括传中、直塞、助攻与定位球。代表控制型构建传球。

（2）直接传球：源自防守 / 中场区域、长度 > 25 码且终点在进攻三区（即禁区）的成功传球，不包括传中与定位球。代表长传转换。

（3）关键渗透传球：源自进攻三区（即禁区）的成功传球，包括传中、回传、直塞或射门 / 进球助攻。代表机会创造尝试。

此外，我们还分析了包含所有成功传球的“整体传球”类别作为对比。

1.3 网络构建

针对每次球队 - 比赛观测及每种传球类型，构建有向加权网络：

节点：参与比赛的球员（通常每队 11 人）；

边：球员间的成功传球；

边权重：每对球员间的传球频率；

边方向：从传球者到接球者。

每观测数据生成四种传球类型：渐进式、直接、关键渗透与整体传球。

1.4 网络指标

对每个网络计算以下指标：

密度：实际连接占可能连接的比例（边数 / [节点数 $\times$ (节点数 - 1)]）；

平均聚类系数：局部互连程度的度量；

平均度：每位球员的平均连接数；

网络规模：节点与边的数量；

连通性：网络是否形成单一连通组件。

1.5 统计分析

采用多方法分析策略：

（1）网络特征的描述性统计；

（2）单因素方差分析及 Tukey HSD 事后检验比较传球类型；

（3）网络特征与比赛结果的相关性分析；

（4）K 均值聚类识别战术风格（通过肘部法确定最优

k=6）；

（5）使用方差分析与描述性统计进行联赛比较。

所有分析使用 Python 及 NetworkX、SciPy 与 scikit-learn 库完成，统计显著性水平设为 $\alpha = 0.05$ 。

2 结果

2.1 整体传球类型分布

在 3,410 次观测共 1,317,285 次成功传球中，分布如下：

渐进式传球：437,399 次（33.2%）；

直接传球：10,287 次（0.8%）；

关键渗透传球：34,317 次（2.6%）；

其他传球：835,282 次（63.4%）。

该分布证实，尽管渐进式传球在现代足球中占据重要部分，真正的直接“长传”比例较低（0.8%），而关键渗透传球为选择性行动（2.6%）。

2.2 网络结构比较（RQ1）

表 1 展示了三种传球网络的结构特征。渐进式传球网络具有最高的密度（ $M = 0.185$, $SD = 0.025$ ）与聚类系数（ $M = 0.644$, $SD = 0.067$ ），体现了控球型打法典型的紧密互连结构。相比之下，直接传球网络的密度（ $M = 0.107$, $SD = 0.080$ ）与聚类系数（ $M = 0.011$, $SD = 0.054$ ）显著更低，反映了长传战术典型的稀疏、低连接结构。关键渗透传球网络在密度（ $M = 0.072$, $SD = 0.018$ ）与聚类系数（ $M = 0.140$, $SD = 0.111$ ）上呈现中等值，表明其聚焦但不完全互连的进攻模式。

表1 按传球类型划分的网络特征

指标	渐进式	直接	关键渗透	整体传球
密度	0.185	0.107	0.072	0.296
聚类系数	0.644	0.011	0.140	0.824
节点数	26.7	8.6	16.1	27.5
边数	126.9	5.6	17.1	216.3

单因素方差分析显示所有网络指标均存在显著差异（均 $p < 0.001$ ）。Tukey HSD 事后检验证实，传球类型间的所有两两比较均显著（均 $p < 0.001$ ），其中聚类系数的效应量最大（ $F = 58,064.25$, $\eta^2 = 0.94$ ）。

2.3 联赛比较（RQ2）

渐进式传球网络特征呈现显著的联赛差异（表 2）。法甲的渐进式网络密度最高（35.3%），其后依次为英超（33.7%）、西甲（33.4%）、意甲（32.6%）与德甲（30.3%）。这一顺序符合对联赛风格的常规认知，即法国与英格兰足球更强调技术型构建，而德国足球倾向于更直接的打法。

表2 按联赛划分的渐进式网络特征

联赛	密度	聚类系数	比赛场次
法甲	35.3%	0.8%	357
英超	33.7%	0.9%	359
西甲	33.4%	0.8%	345
意甲	32.6%	0.7%	361
德甲	30.3%	0.8%	283

方差分析证实联赛间在渐进式网络密度 ($F = 40.32$, $p < 0.001$) 与聚类系数 ($F = 35.80$, $p < 0.001$) 上存在显著差异。事后检验表明, 法甲与德甲、意甲均存在显著差异 (均 $p < 0.001$)。

2.4 战术风格聚类 (RQ3)

K 均值聚类基于传球网络特征识别出六种不同的战术风格 (图 1)。最大聚类 (聚类 3, 占观测值的 32.5%) 呈现高渐进式密度 (0.203) 与低直接密度 (0.092), 代表控球主导型球队。聚类 0 (4.1%) 的直接密度极高 (0.439), 属于极端长传专精型球队。聚类 1 (18.4%) 的关键渗透密度最高 (0.093), 表明其专注于机会创造。

聚类内的联赛分布揭示了战术偏好: 法甲球队在控球主导的聚类 3 中比例过高 (44.4% vs. 平均 32.5%), 而德甲球队在直接打法聚类中更为常见。

2.5 与比赛结果的关系

相关性分析显示, 网络特征与比赛结果间关联较弱。直接传球比例与比分差呈最显著正相关 ($r = 0.054$, $p < 0.001$), 表明采用稍多直接传球的球队往往有略好的结果。渐进式网络特征的相关性可忽略不计 ($|r| < 0.02$), 说明尽管渐进式传球在风格上重要, 但其与比赛结果的关系复杂, 可能受其他因素调节。

3 讨论

3.1 理论意义

本研究对体育科学与网络分析作出了多项理论贡献。首先, 我们证明复杂网络理论能有效区分足球中球队协调的功能性差异。渐进式、直接与关键渗透网络间的显著结构差异 (密度相差 1.7 倍) 验证了这些传球行为的概念区分, 并为以往基于定性观察的战术分类提供了量化支持。

其次, 研究结果挑战了整体传球网络在战术分析中的实用性。“整体传球”网络 (密度: 0.296) 掩盖了分别分析传球类型时显现的独特结构模式。这表明未来体育网络分析应考虑功能分解, 而非将所有交互视为同质。

第三, 聚类结果显示球队采用连贯的战术组合, 而非独立的传球决策。识别出的六种不同战术风格——从极端控球 (32.5% 的球队) 到极端直接 (4.1%)——支持了系统

性战术哲学的存在, 这些风格在比赛中一致地体现。

3.2 实际应用

对教练与分析师的实践应用包括:

(1) 对手侦察: 通过分类对手的传球网络, 球队可制定有针对性的防守策略。例如, 对阵控球主导型球队 (高渐进式密度) 时, 实施高位压迫可能破坏其互连网络。

(2) 战术发展: 球队可利用这些指标发展与完善自身风格。希望采用控球打法的球队可将渐进式网络密度作为关键绩效指标。

(3) 球员招募与发展: 可根据球员在特定网络类型中的角色进行评估。中场球员的重要性可通过其在渐进式网络中的中心性衡量, 而前锋的价值则可通过在关键渗透网络中的连通性评估。

(4) 联赛特定准备: 已证实的联赛差异表明, 备战法甲比赛应强调破坏密集的渐进式网络, 而德甲比赛则需防守稀疏的直接网络。

3.3 方法考量

若干方法方面值得注意。首先, 所有网络的 0% 连通率反映了足球传球的固有特性: 并非每位球员都与所有队友直接连接。这与社交网络中可能预期的完全连通形成对比, 凸显了网络分析的领域特异性。

其次, 分类标准虽基于客观定义, 但涉及阈值决策 (如渐进式与直接传球的 25 码分界)。采用替代阈值 (15–35 码) 的敏感性分析产生了定性相似的结果, 支持了研究发现的稳健性。

第三, 与比赛结果的弱相关表明, 成功的足球比赛涉及传球结构之外的复杂交互。未来研究应整合防守压力、空间位置与决策质量, 以建立更全面的模型。

3.4 局限与未来方向

本研究的若干局限指明了未来研究方向:

(1) 时间动态: 分析基于整场比赛聚合的静态网络。未来工作应考察动态网络演化, 特别是球队如何根据比分、换人与战术调整传球结构。

(2) 防守背景: 分析独立于防守压力的传球。结合对手位置 (如通过定格数据) 将有助于分析防守结构如何影响传球选择。

(3) 跨赛季分析: 单赛季样本仅为快照。跨多赛季的纵向分析可揭示战术演变与教练趋势。

(4) 与绩效指标整合: 未来研究应将网络指标与已有绩效指标 (预期进球、场地控制指标) 结合, 以发展更全面的球队评估框架。

尽管存在局限,本研究为通过网络分解进行足球战术的复杂分析奠定了基础,既提供了理论见解,也为足球界提供了实用工具。

4 结语

本研究首次通过复杂网络理论系统分析了不同的足球传球行为。通过将整体传球分解为三种功能不同的类型——渐进式、直接与关键渗透传球——我们揭示了反映不同战术目的的根本性网络结构差异。

基于五大欧洲联赛 1,705 场比赛(3,410 次球队-比赛观测)的分析表明,渐进式传球网络呈现高密度(0.185)与高聚类系数(0.644)的“小世界”特征,适合控制型构建;直接传球网络展示稀疏、低连接结构(密度:0.107),适于快速转换;关键渗透网络则呈现聚焦的中等结构(密度:0.072),优化了进攻三区的机会创造。

这些结构差异具有统计稳健性(均 $p < 0.001$),并与联赛特定的战术身份一致,其中法甲渐进式网络密度最高(35.3%),德甲最低(30.3%)。聚类分析识别出六种连贯的战术风格,32.5%的球队采用以高渐进式网络密度为特征的控球主导策略。

除对体育科学的理论贡献外,本研究为战术分析、对手侦察与球员发展提供了实用工具。通过量化以往定性的

战术概念,我们实现了对球队风格的更精确评估与比赛准备的更具针对性。

未来研究应在此基础上纳入时间动态、防守背景及与绩效指标的整合。尽管如此,本研究证实,复杂网络理论在辅以适当的功能分解时,能为这项美丽运动的战术复杂性提供深刻洞察。

参考文献:

- [1] NEWMAN M E J. Networks: An introduction[M]. Oxford: Oxford University Press, 2010.
- [2] GRUND T U. Network structure and team performance: The case of English Premier League soccer teams[J]. Social Networks, 2012, 34(4): 682-690.
- [3] STATSBOMB. StatsBomb Open Data[DB/OL]. (2019)[2025-07-01]. <https://github.com/statsbomb/open-data>.
- [4] 郭凯波,张辉,施之皓. 基于社会网络分析的足球世界杯比赛球队传球特征研究[J]. 体育科学, 2020, 40(4): 70-78.

作者简介:马茁冉(2006.07-),女,回族,浙江省嘉兴市,华东理工大学学生,研究方向:体育教育。