结合知识图谱的篮球赛事新闻可视化

吉娜烨^{1,2)}, 高勇^{1,3)}, 赵友兵^{1,2)*}, 俞定国^{1,2)}, 褚少微^{1,2)}

¹⁾ (浙江传媒学院智能媒体技术研究院 杭州 310018)

²⁾ (浙江省影视媒体技术研究重点实验室 杭州 310018)

³⁾ (浙江大学软件学院 杭州 310058)

(zyb@cuz.edu.cn)

摘 要: 体育赛事新闻的可视分析涉及体育赛事分析和新闻可视化, 对快速生成赛事新闻报道、增强赛事新闻表现力、辅助赛事分析有重要作用。针对篮球比赛的重要赛事数据, 如篮球比赛球员数据、球队数据和比赛的时间、空间数据, 辅以知识图谱丰富背景知识, 设计了交互式可视化工具。设计的交互可视工具面向数据新闻的制作, 以体育赛事新闻的自动报道为需求, 除了包括基础数据的概览和统计视图外, 还包括赛事背景数据的知识图谱可视设计、赛事时空数据可视视图, 让篮球爱好者和普通大众读者拥有多维度的视角去了解篮球赛事的各种信息。实现的可视化系统, 通过篮球爱好者、普通大众读者和专业体育新闻记者等用户的评测, 在 5 分制的量化评分下, 各可视模块的均分基本在 4 分以上, 证明了知识图谱用于篮球赛事可视化兼具生动性和知识性, 可以增强篮球赛事新闻的表现力。

关键词: 篮球赛事; 知识图谱; 数据新闻

中图法分类号: TP391.41 **DOI:** 10.3724/SP.J.1089.2021.18590

Knowledge Graph Assisted Basketball Sport News Visualization

Ji Naye^{1,2)}, Gao Yong^{1,3)}, Zhao Youbing^{1,2)*}, Yu Dingguo^{1,2)}, and Chu Shaowei^{1,2)}

¹⁾ (Intelligent Media Institute, Communication University of Zhejiang, Hangzhou 310018)

²⁾ (Key Lab of Film and TV Media Technology of Zhejiang Province, Hangzhou 310018)

³⁾ (School of Software Technology, Zhejiang University, Hangzhou 310058)

Abstract: Visual analysis of sports event news involves sports event analysis and news visualization, which plays an important role in quick generating of event news reports, enhancing the expression of event news, and assisting event analysis. A knowledge graph based visual analytics method is proposed for basketball data news generation. The interactive visualization tools are designed for basic matches and team data of basketball games, such as basketball player data, team data and spatio-temporal data, supplemented by knowledge graph, which enriches background knowledge. Besides the overview and statistical view of basic data, it also includes the visual design of knowledge map of the background data for the game and the visual view for the spatio-temporal data of the game, so that basketball fans and ordinary readers could benefit from a multi-dimensional perspective to enjoy the various aspects of the basketball game. Therefore, the visualization system can increase the richness of sports event news, and serve basketball fans and ordinary readers better. Through the evaluation by basketball fans, general readers, and professional sports journalists, most of the average scores of each visual module are above 4 points basically under a 5-point quantitative

收稿日期: 2020-08-18; 修回日期: 2020-10-29. 基金项目: 浙江省重点研发计划(2019C03131); 浙江省基础公益研究计划(LGF20F-020001). 吉娜烨(1987—), 女, 博士, 讲师, CCF 会员, 主要研究方向为人工智能、可视化、计算机视觉、计算机图形学; 高勇(1996—), 男, 硕士研究生, 主要研究方向为数据可视化; 赵友兵(1976—), 男, 博士, 讲师, 论文通讯作者, 主要研究方向为信息可视化和可视分析; 俞定国(1976—), 男, 博士, 教授, 硕士生导师, CCF 会员, 主要研究方向为媒体人工智能应用、媒体融合; 褚少微(1982—), 男, 博士, 副教授, CCF 会员, 主要研究方向为人机交互。

scoring system. It is proved that the knowledge graph used in basketball event visualization has both vividness and knowledge, which enhances the expressiveness of basketball event news.

Key words: basketball games; knowledge graph; data journalism

随着计算力和互联网的发展,当今社会已经进入大规模生产数据、消费数据和传播数据的“大数据”时代.在传媒行业,大数据是2012年以来国内新闻传播学界所热议的话题.相较于传统的新闻报道,现在的新闻从业者不仅专注于新闻事件和现场内容报道,同样也可以利用数据挖掘分析等技术发现新闻背后的关系,并在报道新闻中采用数据可视化和可视分析技术^[1-2].

在众多数据新闻类别中,体育新闻是受众量较多的新闻类别之一.得益于数据采集技术的提升,不论是奥运会、世界杯和世锦赛等国际重大体育比赛,还是中超和CBA等国内联赛,每时每刻都有大量的赛事数据产生.结合赛事数据挖掘分析与数据可视化技术驱动生成数据新闻,已经成为世界各大媒体常用的体育赛事新闻报道手段.本文旨在研究数据新闻中的体育领域,以美国职业篮球联盟(national basketball association, NBA)赛事为例,对如何结合赛事数据获取和相关可视化技术进行研究和探索,以服务于新闻可视化的数据视觉分析.本文设计的NBA赛事数据可视化分析系统^①可作为赛事数据新闻的辅助报道手段.该系统将收集的大量赛事数据以二维或三维图表进行视觉呈现,为普通受众和篮球爱好者提供兼具知识性和生动性的赛事数据分析,在一定程度上也可以辅助篮球专业人员赛事决策.

1 相关工作

由于体育比赛中产生的赛事时刻数据具有易收集和数量大的特点,对体育数据进行可视分析以辅助新闻呈现是目前数据新闻的热点研究方向之一.对体育赛事数据的可视化分析通常有2个基本方向,分别是面向比赛结果的统计型数据和面向比赛过程的时空数据^[3].

其中,统计数据的可视化相对较为直观,既能直观反映比赛方的技术特点与风格,又具备易记录和易判断的特点.而时序数据和空间数据的可

视化较为复杂和多样化,一直是新闻可视化和可视分析的热点^[4].已有不少从各类体育赛事数据和球员运动轨迹中获取信息来设计可视方案和工具的研究工作.代表性的工作有:针对网球比赛的数据可视化系统TenniVis^[5],将摄像机捕获的数据通过扇形的可视化编码展现一场网球比赛过程;为冰球射门坐标创建了一个径向热图方法^[6],使用环坐标分析NHL(北美冰球联盟)球队最常用于射门的距离;通过弧线图描述美式橄榄球比赛的运动序列^[7];对乒乓球的对阵模式进行了可视分析^[8].上述体育赛事可视化工具引入时空数据后,可以帮助体育专家进行更深入的赛事分析.

此外,足球运动可视化更重视时空数据背后的关系分析.SoccerStories^[9]通过使用热力图和关系图展示球员传球的联系.针对足球运动模式的视觉分析系统^[10]将可视化技术与数据挖掘技术相结合,实现了足球运动模式的交互式识别;结合足球比赛的视频和运动数据,该系统还对比赛中的球员运动轨迹和进攻防守状态进行可视分析,以辅助团队的运动决策^[11-12].Bialkowski等^[13]将足球的空间数据进行算法再分析,发现球员角色和球队信息.足球赛场可视化工具ForVizor^[14]可对足球阵型的时空变化进行可视分析.

而在本文重点关注的NBA篮球赛事数据可视化方面,Beshai设计了一个交互网站buckets^②,方便用户查看近年来NBA各球员的投篮热点等投篮信息.NBA可视化工具GameFlow^[15]从赛季、比赛和节3个层次对NBA比赛进行可视分析.BKViz^[16]对源自Basketball-Reference的时序数据进行可视分析,提供了2015—2017年的单场NBA比赛的详细数据分析和可视化展示.BasketView^[17]使用了流状图和时间线等对NBA篮球比赛的赛季、单场比赛和重要事件进行了可视化.

综上所述,现有的对体育赛事特别是NBA篮球可视化更多面向专家分析工具,以新闻报道面向的普通受众为目标的篮球赛事可视化仍值得深入研究,特别是如何能帮助用户快速、深入地获得

① https://github.com/siniangy/425_frontEnd/

② <http://buckets.peterbeshai.com>

球赛信息,这正是本文的工作。

除了传统的可视化技术之外,本文还采用一种新兴且热门的技术——知识图谱,以辅助体育赛事数据分析。知识图谱由 Google 公司于 2012 年提出^①,其数据的可视表达一般有节点链接图、热图和邻接矩阵等形式。目前常见的知识图谱可视化的研究有大规模知识图谱可视方法和异质网络可视分析等^[18]。

结合知识图谱中强大的背景知识,并以合适的可视化形式呈现用户感兴趣的球员、球队背景及历史比赛信息等,可有力地支持篮球赛事新闻的深度阅读。而当前还未有针对体育赛事的专用知识图谱,且并未有知识图谱在体育赛事可视化中的应用。

2 数据描述

本文篮球赛事新闻可视化使用了常规统计数据 and 比赛时空数据。数据来源于 stat-nba.com 和 Basketball-Reference 网站。

2.1 常规统计数据

篮球赛事的常规统计数据包括球员统计数据、球队统计数据和比赛统计数据。

球员的常规统计数据具体包括球员姓名、是否首发、出场时间、投篮命中率、投篮命中个数、投篮出手次数、3 分投篮命中率、3 分投篮命中个数、3 分投篮出手次数、罚球命中率、罚球命中个数、罚球出手次数、真实命中率、篮板总数、前场篮板数、后场篮板数、助攻个数、抢断个数、盖帽个数、失误个数、犯规个数以及本场得分共 22 个字段属性。

对阵双方的关键信息数据具体包括球队标志地址、球队名称、球队赛前战绩、球队总得分、球队主客场所属、球队第 1 节得分、球队第 2 节得分、球队第 3 节得分、球队第 4 节得分共 9 个字段属性。

球队的常规统计数据具体包括球队出场人数、总投篮命中率、总投篮命中数、总投篮出手次数、总 3 分命中率、总 3 分命中数、总 3 分出手次数、总罚球命中率、总罚球命中个数、总罚球出手次数、真实命中率、总篮板总数、总前场篮板数、总后场篮板数、总助攻个数、总抢断个数、总盖帽个数、总失误个数、总犯规个数以及本场总得分共 20 个字段属性。

2.2 球队知识图谱信息

体育赛事新闻报道除了传达常规赛统计信息,还需要充分结合体育背景知识传达新闻要素。因此,除赛事数据外,本文还使用知识图谱获得球队和球员的背景知识。具体地,将球队名称作为关键词传入 CN-DBPedia 百科知识图谱接口或自建的篮球专有知识图谱。关键词传入后会返回球队的拓展词条信息,具体包括中文队名、英文队名、所属运动项目、所属赛事、所属地区、球队成立时间、球队主场馆、主场馆容纳人数、球队现任主教练、球队知名球星、球队所获荣誉、球队加入 NBA 联盟时间、球队所在城市、球队所在部区、球队所在分区和球衣颜色共 16 个属性字段。

2.3 比赛时序数据

除常规统计数据以外,篮球比赛还有更为重要的时空数据,记录了球赛的动态时空进程。本文所使用的时序性数据来源于 stat-nba.com,空间投篮坐标位置来源于 Basketball-Reference 网站。

比赛时序性数据(play-by-play data)记录了本场比赛中球员的有效操作事件(如投篮、犯规、失误等)以及事件的发生时刻。源网页中时序数据的格式包括事件发生时刻、球队 A 发生事件描述内容、比分变化、球队 B 发生事件描述内容共 4 个字段内容,如表 1 所示。

表 1 时序数据字段属性

时序数据字段	属性
时间	11:27:00
球队 A 事件详情	T·Mozoov makes 2 pt layup from 1 pt (assist by W.Chandler)
比分变化	2-0
球队 B 事件详情	

2.4 比赛空间数据

本文中比赛空间数据指投篮事件坐标数据。通过分析 Basketball-Reference 的 shot-chart 页面中相关比赛的超文本标记语言(hyper text markup language, HTML)结构信息,再爬取并处理投篮位置数据。具体地,包括投篮发生事件序号、投篮相对位置坐标、投篮发生事件描述、投篮命中或投失共 4 个字段内容。以 2015-01-02 骑士队对阵黄蜂队的比赛中,骑士队在比赛第 3 节 9:15:00 这一时刻发生的事件为例,其空间数据各字段如表 2 所示。

① <https://blog.csdn.net/eli00001/article/details/64905724>

表 2 空间数据字段属性

空间数据字段	属性
投篮事件序号	75
投篮相对位置坐标	top: 209px; left: 60px
投篮发生事件描述	Kevin Love made 3-pointer from 25 ft
投篮状态	· (表示命中)

3 可视化任务

篮球赛事不仅具有统计信息和时间进程,还具有空间属性。由于各个球员的在场位置通过语言描述不易表述清楚,要在短时间内获得对篮球赛事的直观了解,篮球球场的可视化不可缺少。因本文的可视化工作服务于篮球赛事新闻,经需求分析设计了以下任务。

(1) T1. 篮球比赛统计数据可视呈现

该任务针对第 2.1 节获得的各项统计数据,如一场比赛各阶段的得分数据、比赛的总体各项指标等统计数据,采用适合新闻可视呈现的方式进行可视化。

用户可以在首页了解到本场比赛对阵双方的概况,包括主客场信息、赛前战绩、本场比赛节比分等宏观信息,使用户首先判断本场比赛是否含有自己感兴趣的主队和球员,然后再决定是否进一步了解本场赛事的相关可视化分析和数据新闻。

表格式的统计数据可视化模块可方便用户查看本场比赛球员和球队常规统计数据和进阶数据。此外,另有设计可视化组件呈现比赛球队之间关键数据项的对比,可使用户对球员、球队的当场比赛表现有宏观了解。

(2) T2. 球员赛事数据和背景数据可视呈现

在篮球新闻中,球员的在场信息和背景资料是新闻报道的重点。基于此,篮球赛事新闻可视化中,球员信息可视化是不可缺少的部分,包括查看球员的贡献、通过知识图谱获取球员资料等。

通过球员本场比赛与生涯平均表现的数据对比,可以反映球员本场比赛的表现是否积极。本文根据获取到的球员本场比赛的统计数据以及生涯统计数据,设计了球员表现对比可视化组件,可以帮助用户从新颖角度了解球员——其对球队胜负的作用是正反馈还是负反馈。

此外,用户在阅读球赛新闻时往往对本场比赛中表现最佳的球员数据比较关心。可视化模块应该对于比赛中的关键数据(得分、篮板、助攻)表现最好的球员提供相关的可视化展示。分析球员

之间的助攻关系可以让用户从不同角度了解本场比赛中对战双方球队的比赛风格。基于时序数据特点,系统应设计对应的可视化组件展示本场比赛球员之间的助攻关系,从而以不同视角辅助用户了解球员之间的化学反应。

(3) T3. 篮球赛事比赛进程绘制

该任务对篮球赛事进程实现可视化,包括各场比赛的分数变化、得分动作等。当今,各大体育门户网站对于 NBA 赛事时序数据的处理一般都以简单的文字形式呈现,没有对时序数据进行深入挖掘。为了让用户更直观、更详细地了解比赛实况,可视化系统应设计基于时序数据挖掘分析的可视化组件,方便用户以交互动态图表的方式了解本场比赛的时序事件详情和事件发生流程。

(4) T4. 篮球赛事球场可视化

将赛事时空数据中的空间坐标数据与时序数据事件相结合可以为用户了解比赛提供空间加时间的高维融合视角。基于时序数据和空间坐标数据格式特点,系统应设计相应的可视化组件展示本场比赛球员的投篮位置,并结合时序事件的投篮坐标、投篮时间进行可视化尝试。

4 篮球赛事新闻统计数据可视化设计

基于统计数据设计的可视化视图主要包括概览视图和统计视图。概览视图布局主要展示了球队的基本信息以及单节比分差异;统计视图布局主要包括以表格形式展示的球队和球员的比赛统计数据以及基于统计数据设计的对比可视化组件。

4.1 概览视图布局设计

概览视图以可视化形式呈现新闻表达所不可或缺或文本和图像信息,分为位于中间的球队比分以及位于两边的球队信息概览 2 部分。球队信息概览包括球队名称、球队主客场信息、球队赛前战绩以及球队标志共 4 个字段。在概览视图中球队在比赛各节的比分差异除了以表格方式呈现之外,还以柱状图形式呈现。图 1 所示为柱状图形式显示的比分比较,其中,柱高表示每节的比分,不同颜色代表不同球队。

4.2 统计视图布局设计

统计数据可视化视图如图 2 所示。针对可视化任务 T1 和 T2,本系统还分别设计了球员统计数据表格可视化组件(如图 2a 所示)、球队统计数据柱状图可视化对比组件(如图 2b 所示)以及关键数据表现最佳球员表格可视化组件(如图 2c 所示)。球

图1 球队比分概览视图

萨克拉门托国王

基础数据 进阶数据

球员	首发	时间/min	投篮/%	命中	出手	3分/%	命中	出手	罚球/%	命中	出手	真实命中/%	篮板	前场	后场	助攻
鲁迪-盖伊	1	33	58	7	12	0	0	0	87	7	8	67	6	2	4	5
达伦-科里森	1	39	53	8	15	50	1	2	87	4	4	62	4	0	4	0
德马库斯-考辛斯	1	27	57	8	14	0	0	0	100	3	3	62	7	2	5	4
德里克-威廉姆斯	1	31	60	6	10	50	3	6	100	2	2	78	5	1	4	3
本-麦克莱默	1	38	66	6	9	50	1	2	50	1	2	70	4	1	3	2
卡尔-兰德里	0	27	25	1	4	0	0	0	75	6	8	53	9	3	6	1

a. 球员统计数据表格可视化组件

b. 球队统计数据柱状图可视化对比组件

各项最佳

	萨克拉门托国王	明尼苏达森林狼
得分	鲁迪-盖伊 21	安德鲁-威金斯 27
篮板	卡尔-兰德里 6	戈尔吉-吉昂 9
助攻	达伦-科里森 5	蔡察斯-巴丁格 2

c. 关键数据表现最佳球员表格可视化组件

图2 统计数据可视化

员统计数据可视化组件以表格形式展示了球员本场比赛的常规数据和进阶统计数据。常规统计数据源自 stat-nba.com 数据分析网站,包括球员姓名、得分、助攻数等共 22 个为新闻报道所关注的字段属性。进阶统计数据源自 Basketball-Reference 数据分析网站,包括球员姓名、出场时间、有效投篮命中率(effective field goal, eFG)等共 16 个新闻报道所关注的字段属性。

5 篮球赛事知识图谱数据可视化设计

5.1 球队知识图谱可视化

在上述直观可视化的基础上,本文结合知识

图谱对球队信息进行背景知识的载入与可视化。可视化组件渲染数据来源之一 CN-DBpedia 是由复旦大学知识工厂实验室研发并维护的大规模通用领域结构化百科图谱^[19]。该知识图谱是从百科纯文本 HTML 页面中提取的目标实体、实体属性以及实体关系,经数据过滤、数据清洗和数据融合推断等操作后,最终形成实体-关系-实体的三元结构化数据。数据以资源描述框架(resource description framework, RDF)的规范存储格式存入数据库。

图谱数据可视化的整体流程如图 3 所示。用户查询的实体作为关键词输入,通过调用开源应用程序接口(application programming interface, API)获得三元组数据,经过过滤以及格式处理之后,处

理完毕的数据与球员数据可视化组件(主要展示组件方式为力导关系图)做适配渲染,从而完成球队知识图谱数据可视化渲染.图4所示为丹佛掘金队的知识图谱可视化.

图3 知识图谱数据可视化流程

图4 球队知识图谱可视化

5.2 球员综合知识可视化

在上述球队基础知识可视化的基础上,本文结合知识图谱的拓展数据进行背景信息的载入与可视化.在概览视图中,用户单击球队名称,球队相关知识图谱数据以力导图形式呈现.用户如果单击球员姓名,球员姓名会作为一个关键字传入CN-DBPedia百科知识库和自建的篮球专用知识图谱库^①.

本文构建的篮球专用知识图谱属于特定领域知识图谱,着重体现NBA赛事新闻的特点.根据从虎扑篮球网站(<https://nba.hupu.com/>)以及百度百科上爬取的相关数据,从中抽取1985—2019年共43510场比赛,最终构建的知识图谱共有5893个节点数据.其中,现役球员节点数量有620个,退

役球员节点3407个,球队节点数量有30个,共包含3个实体类、4种关系和27个属性.实体类定义为篮球赛事中的3个重要概念,分别为球队、球员和比赛.对于各实体定义了阅读NBA赛事新闻时用户感兴趣的基本信息.例如,针对球队实体,定义了名称、别名和分区等属性.在属性定义的基础上,定义了突显新闻特点的实体之间的关系,包括当前效力、交手、历史交手和参与比赛关系这4种.

从对应知识图谱库返回的数据经过过滤,选出生地、出生时间、效力球队、球队比赛位置4个字段的基本信息,以雷达图呈现.图5所示为国王队的鲁迪·盖伊的相关数据.

图5 球员综合知识可视化

5.3 球员助攻关系可视化

从知识图谱中可获得的信息除了球队、球员实体及相应的属性之外,还可得到实体与实体之间的各类关系,通过对这些关系可视化可以辅助赛事新闻的阅读.例如,使用节点链接图表示球队中球员之间的助攻关系,链接由助攻球员指向被助攻球员.如果助攻次数大于1,则在链接中间位置显示助攻次数,这样可以帮助用户快速了解本场比赛的助攻关系和助攻次数.当用户鼠标移至某球员节点时只显示与该球员相关的助攻关系.

以掘金队对阵公牛队的比赛为例,移动鼠标至球员威尔森·钱德勒,最终渲染结果如图6所示.

^① <http://www.openkg.cn/dataset/basketballkg>

可以看到威尔森·钱德勒收到了泰·劳森、内特·罗宾逊、肯尼思·法里德和阿隆·阿弗拉罗 4 名队友送出的 7 次助攻,并分别给提莫菲·莫兹戈夫和肯尼思·法里德送出了一次助攻。图 6 中灰色部分是他们其他球员的助攻情况,当鼠标移至对应的球员节点时,与其相关的助攻情况会高亮显示,不相关的则呈灰色显示。

图6 球员助攻关系可视化

6 篮球赛事时空数据可视化设计

6.1 时序视图布局

针对可视化任务 T3, 本文设计 TimeEvents 视图对源自 stat-nba.com 的时序数据进行可视化, 如图 7 所示, 系统基于 EChart^① 的柱状图组件设计了时序数据柱状图组件, 渲染时序数据中的事件详情。以丹佛掘金队对阵芝加哥公牛队的比赛为例,

时序数据柱状图组件包括事件详情(EventsDetail)、分差(PointDiff)以及视图缩放(DataZoom)3 个可视化元素。

EventsDetail 视图元素以鼠标悬浮的方式显示时序数据详情。单条时序数据在前端经过过滤, 比分变化抽取, 将球队 A 事件详情和球队 B 事件详情属性字段进行字符串连接, 最终提示框显示结果包括当前时刻球队比分和当前事件详情 2 个字段的信息。如图 7 所示, 当前比分掘金队 10:8 领先于公牛队, 掘金队的提莫菲·莫兹戈夫犯规, 公牛队的保罗·加索尔第 1 罚命中、第 2 罚打铁。

比赛分差是比较球队表现的很重要维度, 在 PointDiff 视图元素中, 设计使用 Y 轴的正负轴差异表示 2 个队的分差, 柱状图 Y 轴刻度属性大于 0 说明丹佛掘金队领先于芝加哥公牛队; Y 轴刻度属性小于 0 说明丹佛掘金队落后于芝加哥公牛队。分差的绝对值由柱状图偏离水平中线的高度编码, Y 轴的刻度显示当前时刻领先或落后的幅度。

时序数据柱状图渲染组件中位于右上角的文字提示表示时序数据的事件数量。例如, 从图 7 右上角可知, 当前未经过滤的事件总数有 323 个。考虑事件数量和可视化组件设计美学的平衡, 在底部通过设置 DataZoom 属性实现区域缩放, X 轴默认显示 30% 的事件数量。设置 X 轴主轴过滤模式为 filter, 辅轴 Y 轴过滤模式为 empty, 使拖动 DataZoomX 改变 X 轴过滤柱形时, Y 轴的范围也随之改变。

6.2 球场赛事可视化

针对可视化任务 T4, 本文设计 SpatialEvents 视图, 结合时序数据和空间数据的结构特性做相关可视化组件设计。空间数据主要指投篮事件坐标数据。以掘金队对阵公牛队的比赛为例, 渲染投篮事件位置和状态采用 2 种可视化方案: 圆点编码投篮

图7 时序数据柱状图渲染组件

① <https://echarts.baidu.com/>

事件映射方案和热力图编码投篮事件映射方案。

圆点编码方案根据空间坐标数据结构特点,投篮状态用颜色圆点编码表示(红色代表投篮失败,绿色代表投篮成功)。鼠标悬浮至圆点,组件会以弹窗文字呈现投篮事件详情这一字段。具体视图如图 8 所示。

投篮点分析

小节 投篮点

图 8 圆点编码投篮事件可视化

系统使用 Echarts 的 heatmap.js 实现热力图可视化组件的绘制。在可视化模块设计中考考虑到投

篮位置与二维直角坐标系的适配。采用热力图编码的投篮事件可视化如图 9 所示。

图 9 热力图编码投篮事件可视化

结合投篮事件坐标和投篮事件视频摘要的可视化方法,在投篮事件坐标圆点中,对于存在精彩镜头的事件做了圆点尺寸和颜色 2 个维度的编码区分。考虑色彩的对比度,选择了与红色和绿色对比度较深的紫色和深蓝作为视频入口节点编码颜色。图 10 展示了基于圆点编码方案进行的视频映射,以精彩镜头短视频丰富了球场赛事的可视化形式。

图 10 圆点编码方案视频映射可视化

7 用户反馈

为了评估可视化组件设计对于辅助篮球赛事新闻可视化目标完成所起的作用,进行了可视化系统的用户评测。评测过程一共邀请了 18 名评测人员,其中有 2 名一线体育新闻记者,11 名 NBA 篮球爱好者,5 名仅对 NBA 赛事有所耳闻的普通观众。评测人员在经过简单指导后很快熟悉了可视化系统。他们在使用系统后,针对可视化任务中的

T1~T4 完成了量化评估,分值为 1.00~5.00 分(1.00=效果不好,5.00=效果极好)。其中,T1 任务分别评估了第 4.1 节的概览视图可视化方案(记为 a)和第 4.2 节的统计数据视图可视化方案(记为 b);T2 任务分别评估了第 5.1 节的球队知识图谱可视化方案(记为 c)、第 5.2 节的球员综合知识可视化方案(记为 d)和第 5.3 节的球员助攻关系可视化方案(记为 e)。分别统计了 3 类用户对各可视化方案的打分平均值和总均分,如表 3 所示。从评分结果

看,大部分参与者认为使用可视化方式可以更方便、直观地获取篮球赛事信息,明显优于传统的以文本为主的新闻报道。

表3 用户对可视化方案的评分

用户群体	T1		T2			T3	T4
	a	b	c	d	e		
专业记者	4.00	4.00	2.50	3.00	2.50	3.00	3.00
NBA爱好者	4.45	4.00	3.80	4.10	3.90	4.00	4.18
普通用户	4.80	4.60	4.80	4.20	4.40	5.00	4.40
汇总	4.50	4.22	3.94	4.00	3.89	4.17	4.28

各可视化方案中,评测用户对概览视图的设计效果最为满意,给出了4.50的平均分。对于统计数据视图,部分评测者认为表格型的数据显示方式虽然可以清晰地反映数据数值,但创新性稍显缺乏。

大部分评测用户认为结合知识图谱后的可视化效果较为新颖实用。其中,球员综合知识可视化结果最被认可,原因是力导图的呈现方式大家较为熟悉,辅以文字的简短背景介绍比较友好。但大多数评测用户对球员助攻关系的可视化满意度不高,有NBA爱好者认为球员间助攻关系可视化功能相对鸡肋,各条连线感觉稍显杂乱,并不易懂,观感效果不好。

针对可视化任务T3的TimeEvents视图,大部分评测者认为时序数据的交互图表设计能够让他们发现比赛中有趣的信息。

对于服务于任务T4的SpatialEvents视图,有一些评测用户提出了一些建议,有NBA爱好者认为界面坐标显示颜色过于花哨,还有NBA爱好者建议比赛关键球要在投篮坐标数据里特别标出。

总体而言,对NBA赛事不非常关注的普通观众对本文的篮球赛事新闻可视化系统非常满意,认为用当前的可视化系统辅助报道一场NBA比赛,可以快速获取最有价值的关注点,并能激起用户兴趣。而平时经常打篮球的NBA爱好者则认为结合知识图谱的可视化对球赛新闻报道的补充最有意义,建议添加更多的球员信息,如身高和体重,再配以照片头像会更受欢迎。

但是,目前的篮球赛事新闻可视化系统在专业体育新闻记者眼里功能还略显单一,其打分也是3类用户中最低的。专业体育新闻记者对当前的可视化系统提出了如下建议:添加更专业的数据对比;提供高阶数据的释义;提供单个球员的投篮点可视化图;提供更多时序事件,如犯规等的可视化;提供更多基于多场比赛的球员数据比较。

8 结 语

数据新闻综合使用图表、视频、文字等多维度的报道方式取代了传统新闻中冗长复杂的文字描述,使新闻报道更具有时效性、生动性和知识性,是当前传媒行业突围的重点方向之一。本文针对新闻行业中热门中的体育新闻领域,以NBA赛事为例,设计了面向体育赛事报道的数据可视分析系统,为篮球爱好者提供兼具知识性和生动性的赛事数据分析,为篮球专业人员提供辅助的赛事决策参考。结合知识图谱的可视化为赛事新闻的各类受众提供了快速、直观的阅读感受,具有丰富背景知识的可视化与交互操作提高了新闻的阅读深度与沉浸感。

本文针对NBA赛事数据的特点分别提出了基于赛事统计数据和时空数据的不同模块可视化方案,可以帮助用户全面了解赛事事件的发生流程。此外,针对本文提出的可视化设计目标,对若干用户进行调研,收集了专业体育新闻记者、具备NBA知识和不具备NBA相关知识的各类用户对NBA赛事数据可视化系统是否达到预期目标的量化评分,以及对可视化系统使用的整体反馈。从反馈中可以看出,知识图谱对于篮球赛事的新闻可视化起到了很好的作用,但限于目前知识图谱库中知识的不完备性,很多用户感兴趣的扩展性知识都还未能呈现。今后应进一步完善知识图谱的丰富度,提出更恰当的适用于知识图谱的可视化方案,并提高可视化呈现的美观度。

参考文献(References):

- [1] Chen Wei, Shen Zeqian, Tao Yubo. Data visualization[M]. 2nd ed. Beijing: Publishing House of Electronics Industry, 2019: 67-70(in Chinese)
(陈为, 沈则潜, 陶煜波. 数据可视化[M]. 2版. 北京: 电子工业出版社, 2019: 67-70)
- [2] Thomas J J, Cook K A. A visual analytics agenda[J]. IEEE Computer Graphics and Applications, 2006, 26(1): 10-13
- [3] Weber W, Rall H. Data visualization in online journalism and its implications for the production process[C] //Proceedings of the International Conference on Information Visualization. Los Alamitos: IEEE Computer Society Press, 2012: 349-356
- [4] Engbrechtsen M, Kennedy H, Weber W. Data visualization in Scandinavian newsrooms: emerging trends in journalistic visualisation practices[J]. Nordicom Review, 2018, 39(2): 3-18
- [5] Polk T, Yang J, Hu Y Q, et al. TenniVis: visualization for tennis match analysis[J]. IEEE Transactions on Visualization and Computer Graphics, 2014, 20(12): 2339-2348
- [6] Pileggi H, Stolper C D, Boyle J M, et al. SnapShot: visualiza-

- tion to propel ice hockey analytics[J]. *IEEE Transactions on Visualization and Computer Graphics*, 2012, 18(12): 2819-2828
- [7] Owens S G, Jankun-Kelly T J. Visualizations for exploration of American football season and play data[C] // *Proceedings of Workshop on Sports Data Visualization*. Los Alamitos: IEEE Computer Society Press, 2013: 1-15
- [8] Wu Y C, Lan J, Shu X H, *et al.* iTTVis: interactive visualization of table tennis data[J]. *IEEE Transactions on Visualization and Computer Graphics*, 2018, 24(1): 709-718
- [9] Perin C, Vuillemot R, Fekete J D. SoccerStories: a kick-off for visual soccer analysis[J]. *IEEE Transactions on Visualization and Computer Graphics*, 2013, 19(12): 2506-2515
- [10] Stein M, Janetzko H, Lamprecht A, *et al.* Bring it to the pitch: combining video and movement data to enhance team sport analysis[J]. *IEEE Transactions on Visualization and Computer Graphics*, 2018, 24(1): 13-22
- [11] Janetzko H, Sacha D, Stein M, *et al.* Feature-driven visual analytics of soccer data[C] // *Proceedings of the IEEE Conference on Visual Analytics Science and Technology*. Los Alamitos: IEEE Computer Society Press, 2014: 13-22
- [12] Stein M, Häußler J, Jäckle D, *et al.* Visual soccer analytics: understanding the characteristics of collective team movement based on feature-driven analysis and abstraction[J]. *ISPRS International Journal of Geo-Information*, 2015, 4(4): 2159-2184
- [13] Bialkowski A, Lucey P, Carr P, *et al.* Discovering team structures in soccer from spatiotemporal data[J]. *IEEE Transactions on Knowledge and Data Engineering*, 2016, 28(10): 2596-2605
- [14] Wu Y C, Xie X, Wang J C, *et al.* ForVizor: visualizing spatio-temporal team formations in soccer[J]. *IEEE Transactions on Visualization and Computer Graphics*, 2019, 25(1): 65-75
- [15] Chen W, Lao T Y, Xia J, *et al.* GameFlow: narrative visualization of NBA basketball games[J]. *IEEE Transactions on Multimedia*, 2016, 18(11): 2247-2256
- [16] Losada A G, Therón R, Benito A. BKViz: a basketball visual analysis tool[J]. *IEEE Computer Graphics and Applications*, 2016, 36(6): 58-68
- [17] Hung S, Xu J Y, Gao X F, *et al.* BasketView: interactive visualization of NBA games[C] // *Proceedings of International Conference on Data Science and Information Technology*. New York: ACM Press, 2018: 11-17
- [18] Wang Yongchao, Luo Shengwen, Yang Yingbao, *et al.* A survey on knowledge graph visualization[J]. *Journal of Computer-Aided Design & Computer Graphics*, 2019, 31(10): 1666-1676(in Chinese)
(王勇超, 罗胜文, 杨英宝, 等. 知识图谱可视化综述[J]. *计算机辅助设计与图形学学报*, 2019, 31(10): 1666-1676)
- [19] Liu Qiao, Li Yang, Duan Hong, *et al.* Knowledge graph construction techniques[J]. *Journal of Computer Research and Development*, 2016, 53(3): 582-600(in Chinese)
(刘峤, 李杨, 段宏, 等. 知识图谱构建技术综述[J]. *计算机研究与发展*, 2016, 53(3): 582-600)