



从大数据中寻找提升空间:数据挖掘在羽毛球比赛技战术分析应用中的个案研究

何施,赵文作,刘博,黄志剑*

摘要:本研究拟从羽毛球运动员 X 的角度出发,通过建立关联规则模型对收集的数据进行分析,尝试在众多数据中寻找 X 的提升空间。通过文献资料的方法、访谈法制作羽毛球比赛技战术多维度信息记录表,利用编码规则采集整理数据,并用数理统计的方法进行分析。使用 SPSS 21.0 对数据进行初步统计及一致性检验,后通过使用 SPSS Modeler 18.0 对各项统计指标进行关联规则分析。通过分析可以发现运动员 X 的优势技战术有发网前短球、起跳腾空点杀、后退高远球、上步挑球等,且最后一拍直接得分的情况下相对稳定的落点位置是 9→7→4;X 相对较薄弱并有待提高的技战术有发后场球以及网前球的处理等。另外,通过比分形势与连续得失分、局分与得失分的关联规则,反映了 X 在处理困难局面上的心理坚韧性品质尚有较大的提升空间。

关键词:羽毛球;技战术;数据挖掘;关联规则

中图分类号:G80-05 文献标志码:A 文章编号:1006-1207(2021)04-0008-06

DOI:10.12064/ssr.20210402

Seeking Improvement with Big Data: A Case Study on the Application of Data Mining in Technical and Tactical Analysis of Badminton Games

HE Shi, ZHAO Wenzuo, LIU Bo, HUANG Zhijian*

(School of P. E., Hubei University, Wuhan 430062, China)

Abstract: From the perspective of the badminton player X, this study intends to analyze the collected data by establishing association rules, and to seek improvement for X from the numerous data. This study uses the methods of literature review and interview to make a multi-dimensional table recording the information of badminton competition techniques and tactics, uses the coding rules to collect and systematize the data, and uses the method of mathematical statistics to analyze the data. SPSS 21.0 is used for preliminary statistics and consistency test, and then SPSS modeler 18.0 is used to analyse association rules of various statistical indicators. The analysis shows that the badminton player X has advantageous techniques and tactics such as net shot service, takeoff kill, long high shot, and lifting. The landing position 9→7→4 is relatively stable when the player scored in the last stroke. Relatively X has weakness in techniques and tactics yet to improve including the handling of backcourt shots and net shots. In addition, according to the association rules of score situation, consecutive scores and losses, and game scores and losses, X still has much room for improvement in tenacity when facing ordeals.

Key Words: badminton; techniques and tactics; data mining; association rule

羽毛球是我国竞技体育传统优势项目之一,在国际上享有很高的声誉。但在近年来相关的重大赛事中,我国选手的成绩显得不如人意,足以说明现阶段

中国羽毛球正面临着严峻挑战。面对严峻挑战要想取得好成绩,就必须要有先进的技术手段和思路。羽毛球比赛的技战术具有不确定性、多变性、复杂性等特

收稿日期:2020-09-09

第一作者简介:何施,女,在读硕士研究生。主要研究方向:运动人体科学。E-mail:2318460231@qq.com。

*通信作者简介:黄志剑,男,博士,教授,硕士生导师。主要研究方向:运动心理。E-mail:psychlab@qq.com。

作者单位:湖北大学体育学院,湖北武汉430062。



点^[1],而传统羽毛球技战术分析主要是描述性的统计分析^[2-3],各因素之间是相对独立的,无法还原运动员在某一时段的连续动作状态。教练员往往凭借经验及描述统计数据来安排日常训练、指挥临场决策。

数据挖掘兴起于20世纪末,一般是指利用计算机等工具从大量、零散、不完全并且随机的数据中通过特殊算法提取或“挖掘”出隐藏在其中的信息和知识的过程^[4]。数据之间的关联是数据挖掘中存在的一种非常重要的可被发现的知识,如果给定两个或多个数据集并从中发现存在某种频繁出现的项集模式就称为关联规则^[5]。目前,关联规则广泛应用于超市、银行、网购等领域,在体育领域主要应用在网球^[6]以及其他隔网类项目中^[7],在羽毛球项目中的应用尚属少见。羽毛球的竞赛规则和技战术与网球有相似之处,所以网球技战术的关联规则研究对羽毛球有参考性。本研究尝试使用关联规则中的Apriori算法在众多数据中寻找羽毛球运动员X的提升空间^[8]。

1 研究方法

1.1 研究对象

某省运动健将级羽毛球单打女性运动员1名,年龄为20岁,球龄10年(以下简称为“X”)。通过教练员收集X在2017—2018年国家羽毛球队集训期间参加的6次国际赛事,共21场比赛的视频资料,围绕这些比赛进行技战术多维度信息采集,通过建立关联规则模型对X在比赛中的技战术以及决策问题进行挖掘分析。

1.2 研究方法

1.2.1 文献资料法

研究通过检索多个数据库,查阅了解相关文献,为研究的分析提供可靠的理论支持和技术借鉴。

1.2.2 专家访谈法

研究前期与羽毛球国家队、省队等多名教练员及科研人员进行现场访谈、电子邮件沟通。根据教练员及专业科研人员的意见和建议制作出羽毛球比赛技战术多维度信息记录表(以下简称“信息记录表”),反复修改后的信息记录表得到一致认可。

1.2.3 数理统计法

利用编码规则采集、整理数据,并用数理统计法进行分析。使用SPSS 21.0处理数据及一致性检验,后通过使用SPSS Modeler 18.0对各项统计指标进行关联规则分析。

1.3 研究工具

1.3.1 羽毛球比赛技战术多维度信息记录表

在制作信息记录表时,考虑到影响羽毛球比赛成绩的因素有很多,将这些因素视为自变量,比赛的得失分视为因变量,将运动员在比赛中的每一个得失分定义为一个单元并以此为单位进行多维度关联规则分析。同时,将影响羽毛球比赛成绩的因素分为比分因素、时间因素、空间因素、技战术因素四大类,每一个因素由多项内容组成,这些因素共同作用最后形成运动员的得分或失分结果。

1.3.2 羽毛球场地区域编码规则

羽毛球运动属于隔网类运动项目,整个场地被高1.524 m的球网均分为2个半场。将本研究的研究对象X称为本方运动员,与X同场竞技的运动员称为对方运动员。本方运动员所在半场定义为A区,对方运动员所在半场定义为B区。

羽毛球场较大,羽毛球及运动员的运动轨迹较复杂,与教练员沟通后将羽毛球的每一个半场分为9个位置,如图1所示。球网到前发球线的前场区域从左至右(面向球网)分别编码为1、2、3,前发球线到单打底线均分为中场和后场2个部分,中场从左至右依次编码为4、5、6,后场从左至右分别编码为7、8、9,其中左路、中路和右路的分界线与单打左右边线的三等分线重合。

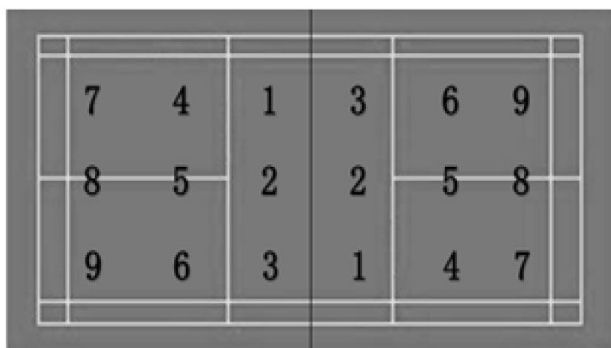


图1 场地位置编码图

Figure1 Diagram of Venue Locations

羽毛球的发球规则要求发球运动员和接发球运动员都必须站在斜对角发球区内发球和接发球。根据接发球运动员的站位推测发球运动员发球的落点,图2表示从另外半场的右区发到该半场斜对角左发球区的落点图。其中,网前内角短球编码为1,网前外角短球编码为2,后场内角高远球编码为3,后场外角高远球编码为4,网前中路短球编码为5,后场中路高远球编码为6。

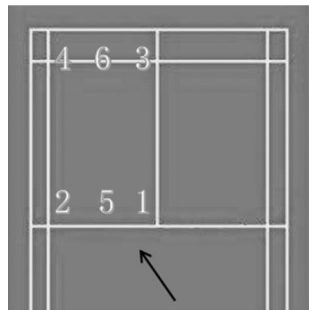


图2 发球区域落点编码图

Figure2 Diagram of Landing Points in the Service Area

1.4 对比赛视频的数据编码

视频材料共记录 X 在 6 次国际赛事中与来自不同国家的 21 名对手之间的 21 场,共 48 局比赛,共 1 644 个球,视频时长 16 时 33 分 25 秒。其中, X 胜 18 场,负 3 场,胜率为 85.7%, 2:0 胜 14 场, 2:1 胜 4 场, 0:2 负 1 场, 1:2 负 2 场。

根据编码规则对 X 的比赛视频进行数据编码^[9]。所有编码由 3 名研究生独立完成,正式编码前随机选取一场比赛录像,3 人分别对该段录像独立观察并进行编码,然后对 3 人的编码进行两两比较,使用 SPSS 21.0 计算出 Cohen's kappa 系数对 3 人的编码结果进行评分者一致性信度检验^[10]。结果显示,评分者一致性强度在一般至强的范围内,3 人可以独立对其他的视频资料进行编码。

1.5 关联规则

关联规则是形如 $A \rightarrow B$ 的蕴涵式, A 被称为关联规则的前项,它可以是一个或多个项集; B 被称为关联规则的后项,它一般是一个单独的项集^[11-12]。对于事务集 D, $A \in D, B \in D, A \cap B = \Phi$, 即满足每次商品交易都与唯一一个标识符 TID (Transaction ID) 对应^[13-14]。在已有研究中,常用支持度、置信度 (Confidence) 来描述关联规则的属性^[15], 本研究新增提升度 (Lift) 和部署能力 (Deployment) 2 个指标^[14-16], 目的是为了更进一步验证规则的有效性并尝试挖掘更多的隐藏信息。

支持度 (Support) 描述了 A 和 B 这 2 个项集在所有事务中出现的概率; 置信度是指在给定项集 A 的前提下, 项集 B 也同时出现的概率; 提升度反映了相比于所有事物, 项集 B 受到项集 A 的影响程度; 部署能力是已经满足项集 A, 但是还没满足项集 B 的事务比例。

支持度 (S)、置信度 (C)、提升度 (L) 以及部署能力 (D) 的公式如下:

$$S_{A \rightarrow B} = \frac{N(A \cap B)}{N} \quad C_{A \rightarrow B} = \frac{N(A \cap B)}{N(A)} = \frac{S_{A \rightarrow B}}{S_A}$$

$$L_{A \rightarrow B} = \frac{C_{A \rightarrow B}}{S_B} = \frac{N(A \cap B)}{N(A)} \bigg/ \frac{N(B)}{N} \quad D_{A \rightarrow B} = S_A - S_{A \rightarrow B}$$

其中, $N(A \cap B)$ 表示同时包含前项 A 和后项 B 的事务数量, N 表示总的事务数量, $N(A)$ 表示包含前项 A 的事务数量, $N(B)$ 表示包含前项 B 的事务数量, 条件支持度 $S_A = N(A)/N$, 结果支持度 $S_B = N(B)/N$ 。

当关联规则满足最低条件支持度阈值和最小规则置信度阈值时则认为该规则是有趣的, 这些阈值是根据挖掘需要人为设定^[17]。为了挖掘尽可能多的关联规则, 本研究默认最低条件支持度为 0%、最小规则置信度为 0.01%。

2 研究结果

研究从 X 的角度出发, 通过建立关联规则模型对收集的数据进行分析, 主要分析 X 在比赛中的比分形势与连续得失分、局段与得失分、多种技战术的使用与直接或间接得失分以及最后 2 拍与最终落点之间的关联规则。

2.1 比分形势与连续得失分之间的关联规则

将比赛中连续得 3 分及 3 分以上的情况定义为连续得分, 连续失 3 分及 3 分以上的情况定义为连续失分。将比分形势看作关联规则的前项, 连续得失分则是关联规则的后项, 对比分形势与连续得失分进行关联规则建模分析得出表 1 数据。

表 1 比分形势与连续得失分之间的关联规则

Table1 Association Rules between the Score Situation and Consecutive Scoring

关联规则(形如: A→B) (比分形势→连续得失分)	支持度 /%	置信度 /%	提升度	部署 能力
比分领先→连续得分	14.11	22.10	1.43	49.76
比分平分→连续得分	0.55	4.59	0.30	11.37
比分落后→连续得分	0.79	3.27	0.21	23.42
比分领先→连续失分	2.49	3.90	0.62	61.37
比分平分→连续失分	0.43	3.57	0.56	11.50
比分落后→连续失分	3.41	14.07	2.22	20.80

分析发现在比分领先的形势下, 比分与连续得分成正向影响且此情况下支持度和置信度相对较高; 当比分落后时, 连续失分的支持度和置信度也相对偏高, 前项对后项的影响程度也为正向。部署能力栏的数据表明, 在比分形势领先或落后的情况下, 如果降低 X 的失误可以有更多的连续得分或更少的连续失分发生。这些分析说明比分形势对 X 的技战

术合理表现有较强的影响,同时为后期的运动心理干预提供了明确的数据支持。

2.2 局段与得失分之间的关联规则

羽毛球比赛多为21分制,根据教练员的建议将每局的21分分为4个局段:0~6分为开局,7~11分为中局前段,12~16分为中局后段,17~21分为局末。规定在领先一方达11分时中场休息,所以在11分处将中局分为前后两段。

表2 局段与得失分之间的关联规则

Table2 Association Rules between Match and Game Scoring

关联规则 (局段→得失分)	支持度 /%	置信度 /%	提升度	部署能力
开局→得分	17.94	59.00	1.03	12.47
中局前段→得分	14.48	53.85	0.94	12.41
中局后段→得分	14.72	58.31	1.02	10.52
局末→得分	9.98	57.14	1.00	7.48

由表2可知,X在开局和中局得分的频率相对较高,在局末得分频率较低,得分情况下局段的置信度由高到低依次是开局、中局后段、末局和中局前段。除中局前段以外,其他局段均与得分成正相关,部署能力相对较高的是开局和中局前段。由此可见X在开局状态较好且有很大潜在得分能力;相较而言在局末有较多失分,在日常训练中应该多加注意该段的心理变化及技战术使用。

2.3 发球落点与发球是否直接得分之间的关联规则

发球技术的使用在羽毛球比赛中是极其重要的技术之一。高质量的发球会给对手增加难度,迫使对手只能被动地做防守性的回击,甚至会直接致使对手接发球失误;质量差的发球会使对手获得进攻机会,而自己处于被动。因此发球质量的好坏直接影响比赛的主被动^[18]。如图2所示将接发球区分为6个落点并进行编码,同时将发球4拍之内发球方得分的情况称为发球直接得分(表3)。

表3反映了发球落点与发球直接得失分之间的关联规则,表中各项支持度普遍较小,说明在高水平的羽毛球比赛中发球直接得分的情况较少发生。但是,比较表中数据可以发现,相对而言,X在发网前中路短球直接得分的置信度最高且提升度为正向影响,所以在平时的训练中应该多加练习并在比赛中大胆尝试使用。

表3 发球落点与发球是否直接得分之间的关联规则
Table3 Association Rules between the Landing Point of Service and the Direct Score of Service

关联规则(发球落点区域 →发球是否直接得失分)	支持度 /%	置信度 /%	提升度	部署能力
落点1→发球直接得分	3.09	16.47	0.97	15.66
落点2→发球直接得分	1.76	18.60	1.10	7.72
落点3→发球直接得分	2.32	18.58	1.09	10.14
落点4→发球直接得分	1.65	17.44	1.03	7.83
落点5→发球直接得分	2.98	21.43	1.26	10.92
落点6→发球直接得分	5.18	14.42	0.85	30.76

2.4 手法技术与得失分之间的关联规则

根据羽毛球相关教材和教练员的建议将手法分为高远球、吊球、杀球、抢搓、推球、勾球、扑球、抽球、挑球、挡球等^[18]。同时,教练员根据以往经验认为X的杀球技术和其他运动员的区别主要是擅长使用点杀,于是又将杀球分为点杀和长杀2种手法(表4)。

表4 手法与得失分之间的关联规则

Table4 Association Rules between Strokes and Scoring

关联规则 (手法→得失分)	支持度 /%	置信度 /%	提升度	部署能力
高远球→得分	6.75	63.07	1.10	3.95
吊球→得分	7.97	61.21	1.07	5.05
抢搓→得分	3.28	51.43	0.90	3.10
推球→得分	0.79	27.08	0.47	2.13
勾球→得分	2.49	49.40	0.86	2.55
扑球→得分	1.34	44.00	0.77	1.70
抽球→得分	3.41	60.22	1.05	2.25
挑球→得分	10.71	60.90	1.07	6.87
点杀→得分	5.72	64.38	1.13	3.16
长杀→得分	2.25	55.22	0.97	1.82
挡球→得分	1.95	62.75	1.10	1.16

对每回合最后一拍手法与得失分进行数据挖掘建模分析得出表4规则。根据表4指标分析发现,在比赛中挑球的使用频率和赢球率都较高,点杀、高远球等手法赢球率也相对较高。因此在平时的训练中应该对挑球、点杀、高远球等优势手法多加巩固,对其他得分率较低的手法多加练习。

2.5 步法技术与得失分之间的关联规则

在教练员的建议下,将羽毛球步法大致分为:上网步法、后退步法、两侧移动步法、起跳腾空步法及其他步法五大类(表5)。



表 5 步法与得失分之间的关联规则

Table5 Association Rules between Footwork and Scoring

关联规则 (步法→得失分)	支持度 /%	置信度 /%	提升度	部署 能力
上网步法→得分	18.61	52.13	0.91	17.09
后退步法→得分	19.34	56.08	0.98	15.15
两侧移动步法→得分	6.63	62.64	1.10	3.95
起跳腾空步法→得分	6.87	64.57	1.13	3.77
其他步法→得分	5.66	65.96	1.15	2.92

通过对步法与得失分进行数据挖掘建模分析得出数据,如表 5 所示。结果显示,上网步法和后退步法的支持度最高,置信度相对偏低;两侧移动步法和起跳腾空步法的置信度相对较大并且对得分成正向影响。

2.6 手法、步法技术与得失分之间的关联规则

羽毛球运动中有“七分步法,三分手法”的说法,两者相辅相成,缺一不可。对手法、步法和得失分进行数据挖掘建模分析,设定最低条件支持度为 10%、最小规则置信度为 50%,通过筛选满足条件的有 4 种关联(表 6)。

表 6 手法、步法和得失分之间的关联规则

Table6 Association Rules among Strokes, Footwork and Scoring

关联规则 (步法→手法→得失分)	支持度 /%	置信度 /%	提升度	部署 能力
后退步法→高远球→得分	10.22	86.49	2.55	1.60
后退步法→杀球→得分	5.43	51.52	1.52	5.11
上网步法→挑球→得分	14.16	75.57	2.32	4.58
起跳腾空步法→点杀→得分	5.01	50.00	4.15	5.01

表 6 中上网挑球得分的支持度和置信度都较高,后退高远球得分的置信度高达 86.49%,在接下来的比赛和训练中应该多加练习并使用。后退杀球和起跳腾空点杀得分的部署能力相对较高,所以应该加强练习。

2.7 最后两拍与最终落点之间的关联规则

在 X 最后一拍直接得分且球的最终落点在对方有效区的范围内,对最后 2 拍与最终落点进行数据挖掘建模分析。设定最低条件支持度为 5%、最小规则置信度为 40%,通过筛选满足条件的有 3 种关联规则(表 7)。其中,最终落点为球最后落在对方有效区触地的位置,最后一拍位置为 X 最后击球时的站位,倒数第 2 拍位置是对方运动员最后击球时的站位。

表 7 最后 2 拍与最终落点之间的关联规则

Table7 Association Rules between the Last Two Shots and the Final Landing Point

关联规则(倒数第两拍位置 →最后一拍位置→最终落点)	支持度 /%	置信度 /%	提升度	部署 能力
位置 9→位置 7→位置 4	3.68	71.43	5.40	1.47
位置 7→位置 9→位置 6	3.68	66.67	4.90	1.84
位置 3→位置 7→位置 9	2.57	50.00	2.83	2.57

由于羽毛球线路具有复杂性、多变性、不确定性等特点,所以最后 2 拍与最终落点关联规则的支持度普遍偏小,从置信度的角度来看,筛选出来的 3 个关联规则的置信度都高于 50%,前项对后项都是正相关影响,部署能力值都偏低。

在羽毛球场地中还原最后 2 拍的路线如图 3 所示。其中,虚线表示与对方运动员最后一拍击球的路线图,实线表示 X 最后一拍击球的路线图。图 3 中的线从粗到细分别表示关联规则由强到弱。

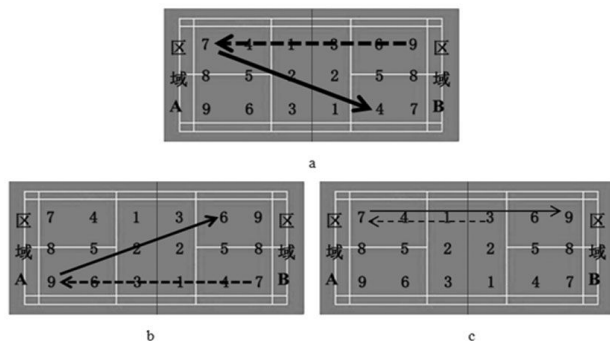


图 3 最后 2 拍路线图

Figure3 Route Map of the Last Two Shots

3 讨论

3.1 关联规则在羽毛球比赛技战术分析中的应用

灵活多变是羽毛球项目的主要特征之一,这也使得对比赛发展规律、运动员技战术特点的描述与把握变得愈发复杂和困难。与传统的描述性统计相比,本研究尝试使用数据挖掘的方法对优秀羽毛球运动员比赛视频进行编码并用关联规则中的 Apriori 算法挖掘分析影响羽毛球比赛的多维度因素并寻找这些因素之间潜在的关系和相互作用。

在实际比赛过程中,影响比分、局分和场分的因素非常复杂。本研究在与专业教练员、运动员及科研人员深入研讨基础上设计编制信息记录表共包含时间、空间、技术和比分等 4 个方面,二十多项因素。这些因素在比赛中往往是以不同的组合方式和显著程度共同影响 X 的运动表现。



对 X 的运动表现有利的情况主要有:(1) 比分或局段领先、发网前中路短球、最后一拍中使用上网步法与挑球手法相结合或后退步法与高远手法相结合、最后 2 拍路线为 9→7→4 时的得分效率较高;(2) 连续得分或是在开局和中局得分频率较高,此时 X 的状态较好且存在潜在得分能力,这便是所谓的顺风局;(3) 网前中路短球虽发得少,但却是 X 发球直接得分率最高的发球技术;(4) 最后一拍中使用上网挑球或后退高远对 X 来说,使用后的得分率也相对较高。此外,分析最后 2 拍得分率最高的路线 9→7→4,可根据 X 的技战术还原为对方的一个高远球从区域 B 的位置 9 打到区域 A 的位置 7, X 接过对手的高远球使用起跳点杀技术将球击落在区域 B 的位置 4。X 有待提升的方面有:面对比分或者局势的落后,应加强心理建设;对得分率较高的技战术要加强巩固并大胆使用。

运用大数据的思路,使用数据挖掘的方法通过积累大量比赛数据并发现这些复杂现象中存在的内在关系与规律,是从更科学客观的角度认识优秀运动员比赛技战术特点和致胜规律的一条可行途径,同时也应注意以下几点。

(1) 明确研究对象,长期系统收集有针对性的数据有利于解决实际问题。本研究的初衷是使用大数据的研究思路为教练员与运动员提供切实可行的比赛技战术运用与得失分规律分析,属于问题导向的应用性研究。随着今后 X 的比赛场次不断增加,将积累更多比赛视频资料和数据资料,这对于扩大数据分析的方法与内容,提高数据分析的准确性和可靠性是相当有利的。同样在不断的数据积累过程中,教练员和运动员可以在实际比赛中检验本研究提供的规律与建议,双方相互促进,形成长期有效的合作。对于向顶尖水平冲击的少数优秀运动员而言,个人长期数据的意义可能优于横向的众人数据。

(2) 合理运用新技术、新指标,提升数据的利用率和解释度。在已有研究中,常用支持度、置信度来描述关联规则的属性,本研究新增提升度、部署能力 2 个指标来尝试挖掘运动员在比赛中更多的隐含信息。如果说支持度和置信度指标反映了某一规则在已有数据中的发生情况,是运动员已完成的过往表现,那么提升度和部署力指标则更倾向于反映该运动员在某一规则上的改进空间。这对于处于技战术上升期的运动员而言,无疑具有更明确的指导意义。

(3) 加入比赛维度,从心理学的角度分析运动员比赛技战术规律。已有的比赛技战术分析研究大多局限于以单个比分为单位,对运动员的技战术和得

失分效率进行分析。本研究在数据编码过程中增加了比赛维度,从局分、比分、局段等具体因素描述,反映运动员每一个得失分发生时的比赛背景。这些背景条件对运动员在比赛过程中的心理具有显著影响,同时也在一定程度上反映了运动员在不同比赛局面下的心理特点。以 X 为例,从结果分析可知,该运动员在比分领先情况下的连续得分,比赛落后下的连续失分情况较多发生,反映了运动员处理困难局面上的心理坚韧性品质尚有较大的提升空间。

总的来说,利用先进技术手段和思路,简单、方便、快捷地将运动员在赛场上的优势和不足用可视化的数据呈出来,一方面为提高羽毛球比赛中的技战术运用提供解决思路,另一方面为运动员、教练员提供科学的技战术改进建议。

3.2 研究存在的问题

从数据挖掘的角度来看,本研究获取的数据仍相对较少,得出的结论具有局限性。在条件允许的情况下,可以做优秀运动员的纵向跟踪分析,累积数据并分析,争取在更深层次上揭示运动员与不同对手,尤其是某几位实力较强的主要对手比赛时的技战术特征,以便得到更加具有针对性的比赛技战术建议。

4 结论与建议

4.1 结论

通过数据分析,可以总结运动员 X 的优势技战术有:发网前短球、起跳腾空点杀、后退高远球、上步挑球等,最后一拍直接得分的情况下相对稳定的落点位置是 9→7→4。X 相对较薄弱并有待提高的技战术有:发后场球以及网前球的处理等。另外,通过比分与连续得失分、局分与得失分的关联规则可以发现运动员 X 的临场表现很容易受比分和局面因素的干扰。

以关联规则为主的数据挖掘方法在运动员比赛技战术大数据分析中具有较高的应用价值。提升度和部署力指标对于认识发现运动员的潜在提升空间有一定指导意义。

4.2 建议

在技术方面,X 应该加强网前球的处理练习;在战术方面,X 应该重视发球技术,掌握主动权;在临场决策方面,应该适当对 X 进行心理干预。

在数据挖掘方法的应用方面,对庞大复杂的比赛技战术数据的合理编码与分析需要从实际问题 and 需求出发,注意分析思路和视角的选择与设定,以更好地服务运动训练、竞赛实践。(下转第 25 页)

**参考文献:**

- [1] 冯本余,任沛尧,闫艳.第28-33届世界蹦床锦标赛竞争格局分析与中国蹦床发展策略研究[J].体育科技文献通报,2020,28(1):70-73.
- [2] 方伯生.优秀男子蹦床运动员预跳高度与运动成绩的关系[J].中国体育教练员,2016,24(1):43-44.
- [3] 李健.高度分规则实施前、后世界蹦床网上个人项目比赛成绩变化特征研究[J].中国体育科技,2015,51(3):36-40.
- [4] 朱礼金.训练学视角下蹦床比赛制胜规律的哲学思考与层次解析[J].中国体育科技,2013,49(5):35-42+59.
- [5] 汪乃媛.蹦床运动新规则下制胜因素分析:以2013年第十二届全运会男子蹦床比赛为例[J].体育科技,2016,

37(5):7-8+10.

- [6] 陆怡敏.2007体操世界杯上海站女子各单项决赛得分统计与分析[J].体育科研,2008,29(5):77-79.
- [7] 徐慧明,李艳翎.蹦床比赛动作难度、类型与稳定关系的研究[J].北京体育大学学报,2004,27(11):1578-1580.
- [8] 王乐军,王钰婷,吴昊,等.“位移新规”实施后我国优秀蹦床运动员比赛成绩的灰色关联度研究[J].体育科研,2018,39(2):100-104.
- [9] 郭奕海,张予南,郭小洋.世界优秀女子蹦床运动员比赛成绩构成特征的分析[J].北京体育大学学报,2006,29(1):128-129+132.

(责任编辑:刘畅)

(上接第13页)

参考文献:

- [1] 刘仁健.羽毛球[M].北京:科学出版社,2010.
- [2] 戴金彪.当今羽毛球男子双打技战术运用情况的分析[J].上海体育学院学报,1998(2):36-40.
- [3] 李芳.羽毛球女子双打风格的探讨[J].中国体育科技,1999,35(9):28-29+39.
- [4] Fayyad U M, Piatetsky-Shapiro G, Smyth P. Knowledge discovery and data mining: Towards a unifying framework [C]. Proceedings of the International Conference on Knowledge Discovery and Data Mining, 1996.
- [5] 赵会群,孙晶,花勇民,等.数据挖掘技术在体育比赛技战术分析中的应用研究[J].北京体育大学学报,2008,31(5):712-715.
- [6] 梁成军,虞丽娟,王继红.网球技战术关联分析数据挖掘研究[J].计算机工程与应用,2012,48(7):246-248.
- [7] 虞丽娟,张辉,凌培亮.隔网对抗项目技战术分析的系统研究[J].体育科学,2008,28(10):41-48.
- [8] Agrawal R, Srikant R. Fast algorithms for mining association rules [C]. Proceedings of the 20th International Conference on Very Large Data Bases, 1994.
- [9] 刘星玉,陈滔,余曦.“2017年顶级羽毛球赛”中外男子单打运动员得失分特征比较研究[J].广州体育学院

学报,2018,38(3):93-96.

- [10] 唐万,胡俊,张晖,等.Kappa系数:一种衡量评估者间一致性的常用方法(英文)[J].上海精神医学,2015,27(1):62-67.
- [11] 贝里,利诺夫.数据挖掘-客户关系管理的科学与艺术[M].袁卫,等,译.北京:中国财政经济出版社,2004.
- [12] 程红霞.基于关联规则的数据挖掘算法研究[J].电脑知识与技术(学术交流),2007,1(3):593+609.
- [13] Badminton Single Matches [C]. Proceedings of 1st Joint International Pre-Olympic Conference of Sports Science & Sports Engineering, 2008.
- [14] 张浩彬.小白学数据挖掘与机器学习:SPSS Modeler案例篇[M].北京:电子工业出版社,2018.
- [15] 徐阳.关联规则挖掘在羽毛球临场战术分析系统中的研究与应用[D].青岛:中国海洋大学,2014.
- [16] 李钦.羽毛球女单运动员技战术统计及关联规则分析[D].西安:陕西师范大学,2019.
- [17] 蒋盛益,李霞,郑琪.数据挖掘原理与实践[M].北京:电子工业出版社,2011.
- [18] 张勇.羽毛球[M].北京:北京体育大学出版社,2003.

(责任编辑:刘畅)