

华东师范大学校园足球研究 & 发展中心科研成果

足球运动

数据化初探

刘 宇 李 群 柳志刚 编著



上海教育出版社
SHANGHAI EDUCATIONAL
PUBLISHING HOUSE

足球运动数据化初探

足球运动

数据化初探

——球

华东师范大学校园足球研究 & 发展中心科研成果

足球运动

数据化初探

刘 宇 李 群 柳志刚 编著



上海教育出版社
SHANGHAI EDUCATIONAL
PUBLISHING HOUSE

图书在版编目(CIP)数据
足球运动数据化初探 / 刘宇,李群,柳志刚编著. —上海:
上海教育出版社,2015.6
ISBN 978-7-5444-6382-9

I. ①足... II. ①刘...②李...③柳... III. ①足球运动 —数据管理
IV. ①G843.3-39

中国版本图书馆CIP数据核字(2015)第128151号

责任编辑 季陆生 王嫣斐
封面设计 王 捷

足球运动数据化初探
刘 宇 李 群 柳志刚 编著

出 版	上海世纪出版股份有限公司 上 海 教 育 出 版 社 易文网 www.ewen.co
地 址	上海市永福路 123 号
邮 编	200031
发 行	上海世纪出版股份有限公司发行中心
印 刷	上海信老印刷厂
开 本	787×1092 1/16 印张 8.5 插页 3
版 次	2015 年 6 月第 1 版
印 次	2015 年 6 月第 1 次印刷
书 号	ISBN 978-7-5444-6382-9/G·5237
定 价	45.00 元

(如发现质量问题,读者可向工厂调换)

主 编 刘 宇 华东师范大学

李 群 华东师范大学

柳志刚 上海体育学院

副主编 虞轶群 华东师范大学

袁 晶 上海理工大学

陈 刚 上海交通大学

闫 燕 上海体育学院

部义峰 江苏师范大学

第一章 计算机软件与网络环境下的足球比赛分析探究 1

第二章 中国女足与亚洲主要对手同场对抗比赛跑动类体能对比分析..... 11

第三章 中国女子足球队体能训练过程研究..... 23

第四章 优秀女子足球运动员跑动能力与比赛效果关联研究..... 44

第五章 中超联赛裁判员犯规判罚与助理裁判员越位判罚行为研究..... 58

第六章 中超联赛助理裁判员越位判罚特征研究..... 71

第七章 足球运动员的体能研究范式与体能训练研究..... 82

第八章 青少年足球天才识别方法外文文献研究进展综述 104

第九章 足球运动员运动表现与球队控球率的互动关系研究 117



第一章 计算机软件与网络环境下的足球比赛分析探究

【摘 要】随着计算机自动化位置记录系统的问世,现代足球比赛数据记录与分析变得越来越容易。然而,如何从大量的数据中获取有价值的情报信息仍然是一个难题。本文结合足球比赛数据挖掘——简化和运用软件系统原理——模式识别,对计算机网络环境下的足球比赛分析软件系统机理进行探究,主要目的是让足球从业人员更好地了解和应用足球比赛计算机分析软件,也为我国计算机行业人员更好地自主研发足球计算机比赛分析软件系统提供参考依据,从而打破以德国等西欧国家为首的对足球比赛分析计算机产业的垄断。

我们通过分析得出,计算机软件系统计算、分析足球比赛的原理主要基于以下两种思路:数据简化和模式(图形)识别。第一步,借助 SOCCER 软件,将某队球员的位置数据标准为类似于进攻或者防守的战术类型的数据集合,然后进行简化,即将球员们在球场上形成的坐标图变为与之中心相关的几何队形,又称为球场位置定位模式。第二步,这些模式源于自组织的神经网络 DyCoN,会自动生成一些集群的集合,每个集合都包含各种相应形式的队形、阵型。基于这些信息,DyCoN 和 SOCCER 软件的足球比赛分析工作原理如下:沿时间轴,不同的战术组合的相对应位置数据被上传至网络,网络能够识别与时间相对应的球队阵形。第一次定量分析的结果是队形的频数分布,结合球场位置信息将会产生一个特定的球场场景频数分布。如果再匹配时间信息,可以获得比赛进程和战术场景互动的分析。同时,SOCCER 软件通过使用附加的语义动作评估,不仅可以提供定量分析结果,而且提供定性的结果,如比赛动态场景和战术的分析等。

【关键词】模式分析 数据 统计 阵型 轨迹

Abstract: With the advent of computer automatic location record system, the modern football match data record and analysis is becoming more and more easy. However, how to obtain valuable information from large amounts of data information is still a problem. This article is based on the football match, simplification and principle of software



system, data mining, pattern recognition, with football game analysis of the computer network environment mechanism of the software system, and the main purpose is to provide football professionals with a better understanding and application of computer analysis software; to provide a reference basis for the development of our country's computer industry personnel football computer game analysis software system, so as to break the monopoly on football game analysis of the computer industry by Germany and other western European countries.

Through our analysis, computer software department statistics and analysis of the principle of the football match is mainly based on the following two ideas: data reduction and model (graphical) recognition. First step, with the aid of SOCCER software, the location data of a team player is standardized as the offensive or defensive tactics of types of data collection, and then it is simplified, figuring into the geometric formations related to the center of the coordinates, also known as the course position location model. The second step, the model is the result of self-organizing neural network DyCoN, which will automatically generate a collection of cluster, and each set contains various forms of corresponding formation. Based on these information, the principle of work of DyCoN and SOCCER football game analysis is as follows: along the time axis, different combinations of corresponding location data are uploaded to the network, and the network can identify the team formation corresponding time. Quantitative analysis for the first time as a result of the formation of the frequency distribution, combining with the field location information would create a scenario specific field frequency distribution. If combined with the match time information, we can get the game process and the analysis of the interaction of tactical scenarios. At the same time, with the use of additional semantic action evaluation of SOCCER software, it can not only provide quantitative analysis results, but also qualitative results, such as dynamic scene and the tactics analysis, etc.

Key words: model analysis data statistics formation trajectory

1. 前言

20 世纪 70 年代开始,足球从业人员借助计算机的辅助对足球比赛进行分析,人们的许多想法逐渐被付之行动,却由于可用数据的缺乏而无法实现效益。与此同时,自动位置记录系统的诞生使数据采集逐渐变得相对容易,因而借助计算机收集比赛数据进而分析比赛由梦想变为现实。随即而来的操作瓶颈是如何从大量的可用数据中



选取有用信息为教练员指挥比赛和制订赛场优胜策略服务。我们分析认为,驾驭通过计算机系统获取的比赛统计数据的前提条件是要厘清软件系统的工作原理,只有如此,足球教练员及相关工作人员才能更好地解读和运用采集的足球比赛数据为球队取得赛场优胜服务。

本文研究的灵感、研究基础、出发点和原始资料均源于历史上最早的一项计算机足球比赛分析系统——它自问世以来先后在欧洲和全世界各地被广为报道,这是德国的一项基于计算机网络足球比赛分析的研究基金会项目^[1~5]。该系统的工作原理基于两种思路:数据简化和模式(图形)识别。

这个程序始于位置数据的准备和预处理,借助 SOCCER 软件工具完成,具体步骤如下:

第一步,借助 SOCCER 软件工具,将某队队员的位置数据精简转变为进攻或者防守的战术类数据,然后进行标准化,将队员们在球场上形成的星座图变为与之中心相关的几何坐标图形,即球场位置化定位模式。

第二步,这些模式被自组织的神经网络 DyCoN 获取并解码,形成一些集群类集合,每个集合都包含相对应模式类型的图形。基于这些信息,借助 DyCoN 和 SOCCER 软件工具分析比赛的工作流程如下:沿时间坐标轴,相互交集的战术组合的位置数据集被上传至计算机网络,计算机能够识别与时间相关的相应形式的队形、阵型;第一次数量分析的结果是形成球队队形的频数分布,再结合球场队员位置信息将会产生一个特定的球场位置化的频数分布。最后再结合时间信息,将会形成一个整个过程和战术集合互动导向的分析结果。当然,在 SOCCER 软件系统的支撑下,也可以进行如语义评价和比赛进程的动画模块的动态类定性分析。

第三步,由 DyCoN 成分组成的轨迹分析能够进行比赛的战术分析,其中包含互动和位相阶段的分析。这种方式,特别是在长期互动模式中,一些隐蔽的或有创造性的球员的战术行动一样能够被成功识别和分析。

2. 网络软件背景下的足球比赛分析研究

2.1 基于 SOCCER 软件系统的数据简化

在第一个处理步骤中,位置数据可被转化为用于统计、绘制图形和战术策略分析的信息。

主要的数据处理原理是将球员们的星座坐标图转化为一组标准化的队形和球场上对应位置的图,如图 1-1 所示,一组以建立时间为主的星座坐标图位置数据,能够被分为它的位置(即星座坐标图的中心)和它典型的队形(即与中心相关星座坐标图

的标准化)。

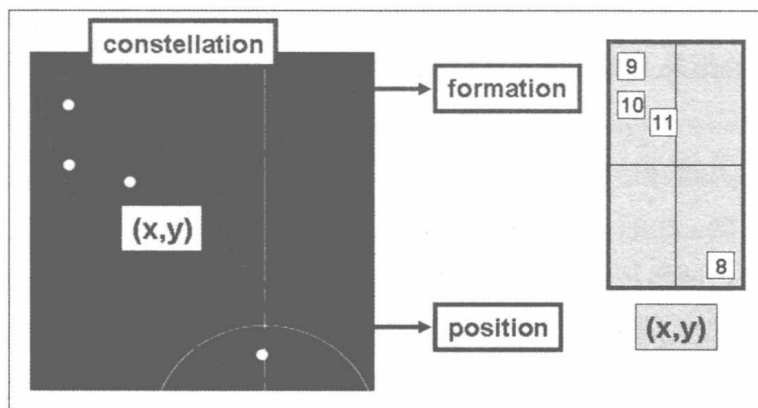


图 1-1 坐标点数据生成位置和队形

为了达到这个目的,SOCCER 为记载球员位置数据和选择有战术意义的阵型或队形提供下载界面(详见图 1-2)。常态化的数据被录入生成一个档案文件或者数据库。

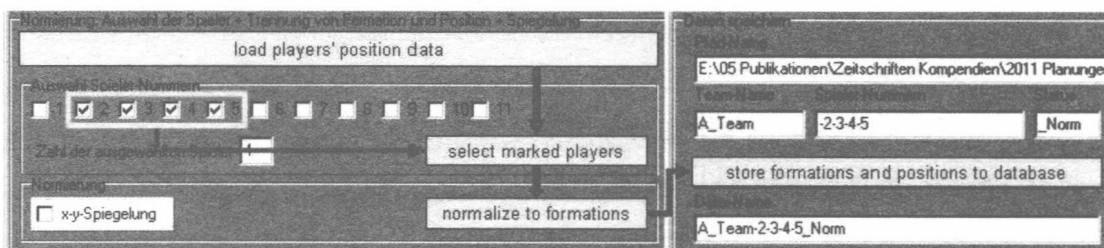


图 1-2 SOCCER 软件数据预生成界面

当然,显而易见的客观事实是,常态化队形的数量远小于可能形成的星座坐标图的数量。因此,数据简化模式允许丢失少部分非常态化队形的战术组合数据信息,因为真正的比赛进程通过使用每个位置存储的时间列序而能够被轻易重组和识别。

2.2 基于 DyCoN 的模式分析

通过采用自组织的神经网络 DyCoN,不同类型的队形被识别并且被分为集群,如图 1-3 所示,不同的集群由不同颜色表示。在每个类型的代表集群中,不同的神经元(着色的小方格)代表相应类型的队形,如图中的四名防守队员的四种不同的防守站位队形“线”。

如果队形数据被输入到预先编程的网络环境中,类型、阵型(尤其是特定的形状)将被网络识别并且能够与相应的位置和时间数据重组,将会最大程度呈现出一个比赛

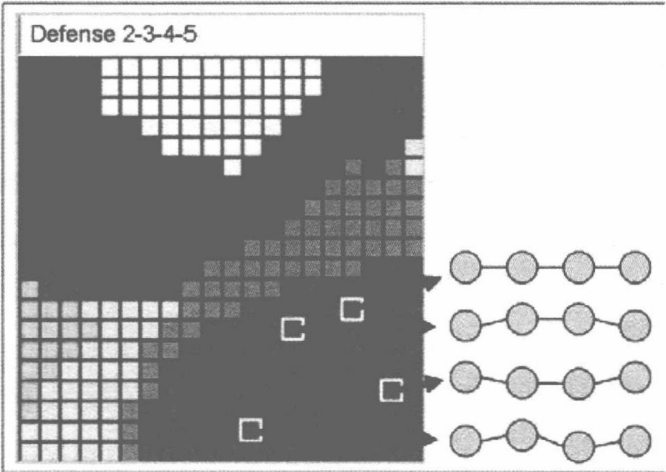


图 1-3 不同颜色的队形集合代表不同的对应队形形状

模式和比赛进行过程,不过重要的队形信息仍然会被完好无缺地保存。

那些以时间序列为主线的队形类型和位置记录,充分考虑到了时间和空间分布的分析诉求,远优于那些泛泛而谈的战术层面和阶段的交互互动过程。

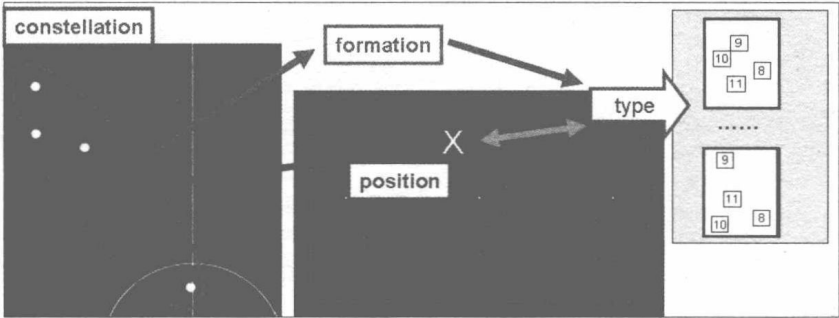


图 1-4 基于网络队形识别和位置、时间信息的重组图示

上述经过筛选、压缩和简化的足球比赛数据信息能被用于生成如图 1-5 所示的比赛数据库,时间、比赛阵型模型和位置可由一个手动的语义评价完成。

time	type	position	action / result
...			
HT1 34:15	7	(45,23)	pass / successful
...			

图 1-5 人工手动语言定义评估生成比赛数据库和比赛分析结果

这些数据信息随后将作为为实际的比赛分析而建立的数据库文本,具体操作本文

将在稍后进行详细论证。

2.3 SOCCER 软件背景下的足球比赛分析程序

接下来的比赛分析被大致分为几个阶段:处理特定空间和时间内相对应的阵型频数以及相应互动的频数分布分析。比赛过程分析主要是针对比赛过程的动态研究,特别是处理比赛过程中相对应的阵型的仿真和动态画面模拟。战术分析则主要是将比赛作为一个整体进行处理,主要目的是为了识别特别的局部小组配合行为和成功互动的过程。

2.3.1 分布分析

图 1-6 所示的分布分析图演示了一个典型的比赛特征情况:在一场对抗竞技的比赛中,A 队以队形 4 进行进攻,B 队则相应地以队形 3 进行防守。分布模型显示这个特定的互动场景在相关的半场比赛时间内发生了 523 次(即 523 秒)。

总体而言,模型分布验证并显示了双方球队的各自攻防队形和各自相应的互动分布。

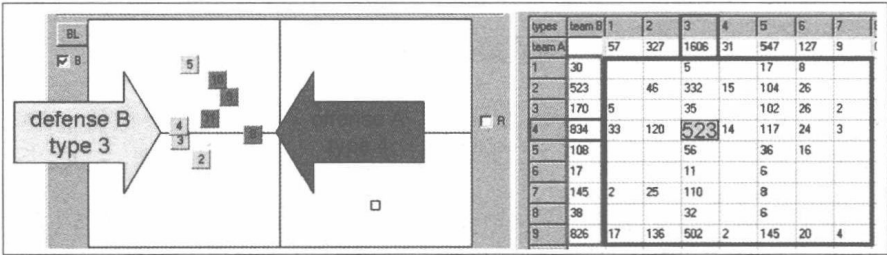


图 1-6 队形和阵型互动分布模型

SOCCER 软件中的分布模型是相互的,也就是说点击其中一个条目就会显示球场方格上(具体地点)相应事件的分布(可以被任意设置)。例如,图 1-7 显示了“防守阵型 3 VS 进攻阵型 1”的 132 次互动在球场上的分布。点击其中一个场上对应模拟方格区域就会显示该区域其他余下条目的分布。

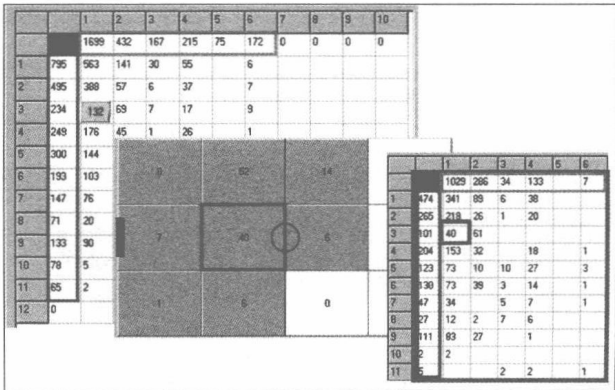


图 1-7 相互影响阵型分布的交互演示



数据统计分析有助于对常规的、罕见的或者非常显著的比赛情景的首次识别。为了更好地识别队员们在比赛进程中所扮演的角色,通常需要将数据统计分析与动态比赛视频录像分析相结合。

2.3.2 进程分析

点击分布模型中任何一个条目同时也会显示比赛过程中该地点或时间段对应互动的列表、场景。如图 1-8 中所示,点击互动条目将会激活一个比赛数据统计序列中的列表——其中第 8 个发生于第 1374 秒到 1384 秒。当然,这个序列被激活后,将产生球场上球队整体重心向球场中心移动的动态画面。

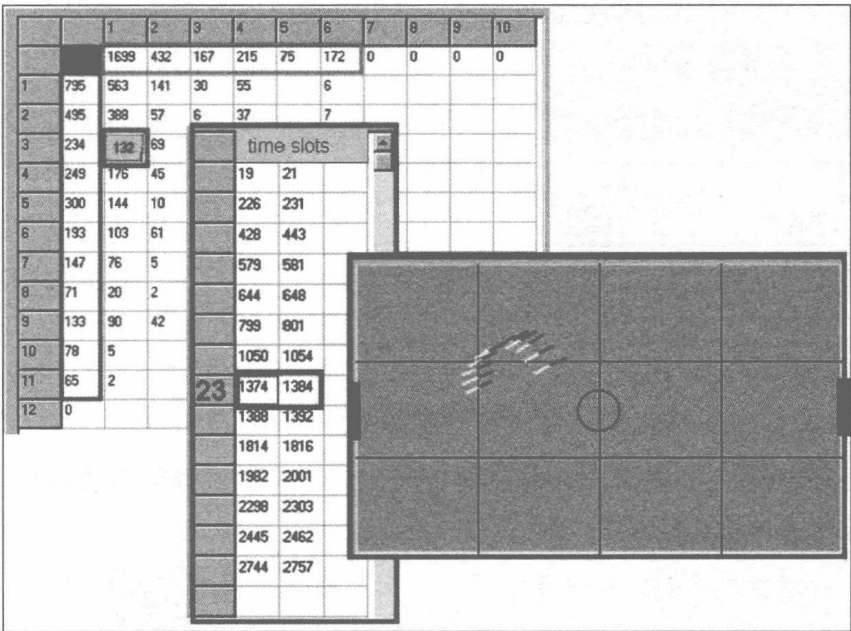


图 1-8 交互动演示图:互动的序号→与选择部分对应的进程列表→球队重心排列

2.3.3 定性与定量相结合的分析

如上文提到的(详见图 1-5),队形数据能被语义数据表示如技术、战术方面的成功。接下来我们举两个例子演示如何借助 SOCCER 比赛分析软件对成功进行评估。

在第一个例子中,对一名球员的成功评估取决于他的战术位置和他当前队形情况下的行动。在图 1-9 中标记了选择的球队(ITA—意大利)、球员(10 号)和队形(2 号,与移动方向一起显示于最左边的图解中)以及选择的动作(第 6 个)、战术位置(第 9 个)。右边的图示显示了当时有 27 个事件,其中 24 个是成功完成的。

第二个例子则是处理在给定阵型互动中如何评估一个球队的成功。图 1-10 从左至右显示了一个球队被评估的互动数字,以绝对数表示的消极互动和百分比。关注右图可以发现,看起来 A 队在应用队形 3 VS 3 的互动中存在严重的问题,绝对数数值

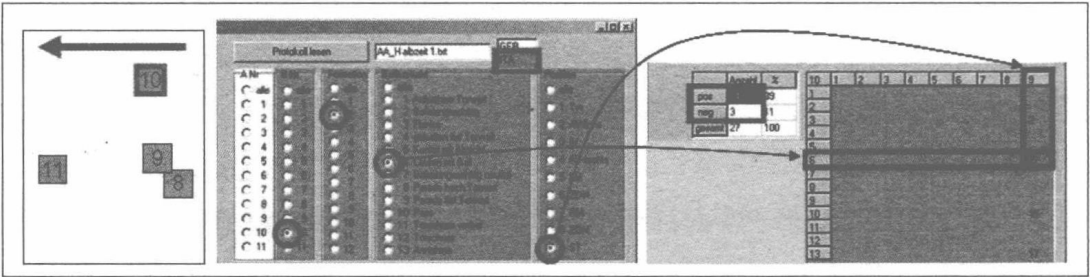


图 1-9 选择球队、阵型、球员、战术位置和球员成功动作信息结果的表格图

非常小,从而极大地减弱了其重要程度的显著性。5 VS 5 也很消极但却不那么重要,然而 5 VS 2 则是个很大的弱点,尽管消极的百分数仅仅只有 16%。

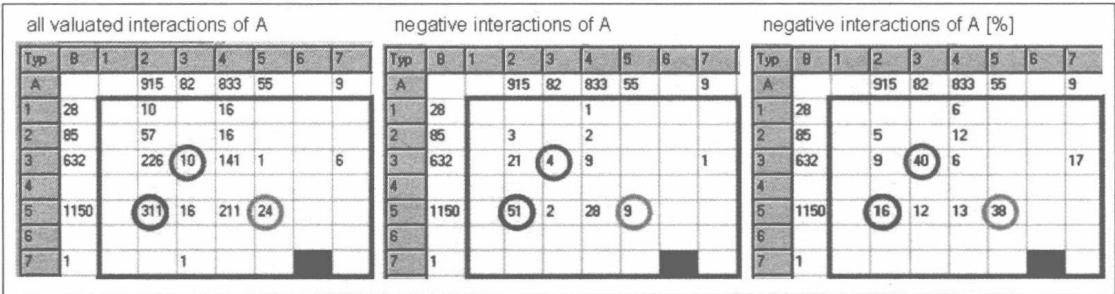


图 1-10 阵型互动环境下球队成功评估模型

但是务必切记,上文提出的分析只是个例子,如果评估数据可用,那么就可以从任意一个角度完成评判及分析应用。

2.3.4 网络背景下的战术轨迹分析

最后一种由 SOCCER 提供的分析模式是基于网络环境下的战术运用轨迹追踪分析。其工作思路及原理是在每个时间截点球队的战术组织的队形是确定的、可以识别的,因而能分别用不同颜色对其进行相应的号码编译。在图 1-11 中,网络识别了 10 种进攻队形。左图显示了 A 队和 B 队是如何在第 16 min 的 60 s 内变换他们的队形的,即 A 队和 B 队在第 16 min 的战术组织及变换过程轨迹。同理,用同样的方法也能显示整场比赛或者如图 1-11 所示的半场比赛中某一特定的时间段的球队战术变换轨迹(这里是 15~30 min),显示了 A 队和 B 队的战术行为轨迹完全不同,折射出双方球队不同的战术理念。

类似的轨迹图是否能够真正体现并且能多大程度上清晰地体现出取样球队的战术组织状况,取决于队伍的战术素养和评判人员的专业知识背景,例如由教练员之类的专业人员来进行评判所获得的结果和信息价值是远高于普通人群的。

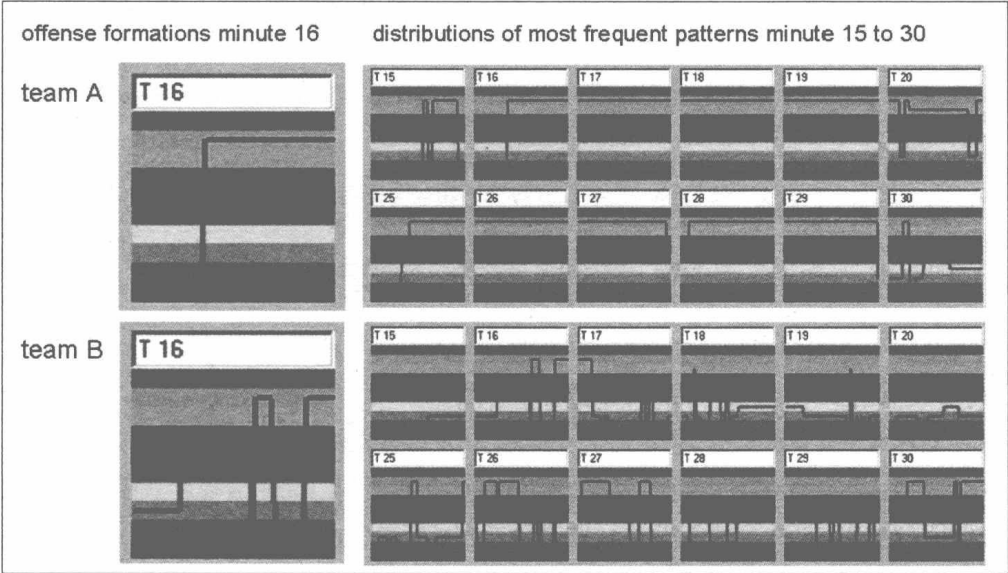


图 1 - 11 基于时间序列的球队阵型轨迹过程图

3. 总结和展望

图 1 - 12 中所示的分析进程原理,在很大程度上可以支持我们对计算机网络在足球比赛分析中运用的展望,旨在证明此类基于网络识别模式和传统统计方法结合的足球比赛分析方式是完全可行的。

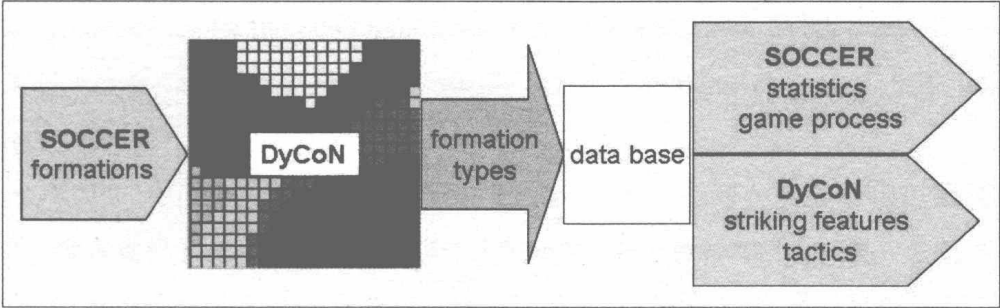


图 1 - 12 SOCCER 软件理念:网络识别模式与传统数据统计方法相结合

我们完全有理由认为,本文的研究和讨论结果是有着积极的现实意义,上文讨论的分析方法有助于合理且成功地对足球比赛进行分析。我们也坚信随着计算机科学的日新月异,更新颖的思路和更先进的比赛分析方法将会很快问世。

参考文献

[1] Grunz A., Memmert D. , Perl J. Analysis and Simulation of Actions in Games by Means of Special



- Self-Organizing Maps[J]. International Journal of Computer Science in Sport, 2009, 8(1): 22 - 37.
- [2] Grunz A., Endler S., Memmert D., Perl J. Netz-gestützte Konstellations-Analyse im Fußball [Net-based Constellation Analysis in Soccer] [J]. Darmstadt, Workshop on Computer Science in Sport, 2010.
- [3] Memmert D., Perl J. Game Creativity Analysis by Means of Neural Networks[J]. Journal of Sport Science, 2009, 12(7): 139 - 149.
- [4] Memmert D., Perl J. Analysis and Simulation of Creativity Learning by Means of Artificial Neural Networks[J]. Human Movement Science, 2009, 2(28): 263 - 282.
- [5] Perl J., et al. Computers in Sport[M]. Wit Press, 2008.
- [6] 刘宇.足球运动竞赛情报基本内容的初步研究[J].中国体育科技,2012(5).
- [7] 刘宇.足球运动竞赛情报内容的理论框架构建与实证分析[J].上海体育学院学报,2012(6).
- [8] 刘宇.足球教练员如何利用比赛视频[J].中国体育教练员,2012(1).



第二章 中国女足与亚洲主要对手同场对抗比赛跑动类体能对比分析

【摘要】 本文主要目的是通过对中国女子国家足球队与亚洲主要对手同场对抗比赛中的跑动类体能数据进行对比分析,找出现在女子足球比赛的跑动距离与速度的规律,发现我国女足国家队队员与日韩等亚洲主要竞争对手的差距,以期为今后训练指明方向,并为中国女足竞技水平的迅速提高提供参考依据。

【关键词】 中国女足 跑动距离 高强度跑 冲刺跑

Abstract: The main purpose of this article is, by analyzing and comparing the Chinese National Women's Soccer Team and Asia's main rival players in the physical data of running during matches, to find the rules of running distance and speed of women's soccer matches, so as to find the gap between Chinese National Women's Soccer Team and Asia's major rivals such as Japan and South Korea, indicating the direction for the future training, and providing a reference for the rapid improvement of China's women's soccer competitive level.

Key words: Chinese National Women's Soccer running distance high intensity running sprints

众所周知,中国国家女子足球队曾在世界杯和奥运会等大赛上取得过辉煌的成绩,曾一度占据世界霸主的地位。然而,近些年来,由于种种原因,中国女足竞技成绩严重下滑,甚至连世界大赛的参赛资格都无法获得,作为国人,尤其是足球从业人员应当痛定思痛,分析原因找出差距,卧薪尝胆,力争早日重返世界之巅。

良好的竞技成绩和对抗中的技战术水平的展现都离不开充沛的体能基础。足球比赛中的体能又以跑动类为主。本文主要目的是通过对中国国家女子足球队与亚洲主要对手同场对抗比赛中的跑动类体能数据进行对比分析,找出现在女子足球比赛的跑动距离与速度的规律,发现女足国家队队员与日本、韩国、澳大利亚(澳大利亚已加入亚足联)等亚洲主要竞争对手的差距,以期为今后训练指明方向,并为中国女足竞技水平的迅速提高提供参考依据。