

Characterizing different team sports using network analysis

Florian Korte^{1,2,*} & Martin Lames²

1 Center for Digital Technology and Management, Munich, Germany

2 Department of Sport and Health Sciences, Technical University of Munich, Germany

* Corresponding author: Center for Digital Technology and Management, Marsstraße 20-22, 80335 Munich, Germany,

Tel: +49 (0) 89-28928163, Fax: +49 (0) 89-28928459

Email: korte@cdtm.de

COMMENTARY

Article History:

Submitted 12th October 2017

Accepted 14th March 2018

Published 25th April 2018

Handling Editor:

Ernst-Joachim Hossner

University of Bern, Switzerland

Editor-in-Chief:

Martin Kopp

University of Innsbruck, Austria

Reviewers:

Reviewer 1: Filipe Manuel Clement

Instituto Politecnico de Viana do

Castelo, Portugal

Reviewer 2: Anonymous

Citation:

Korte, F., & Lames, M. (2018). Characterizing different team sports using network analysis. *Current Issues in Sport Science*, 3:005. doi: 10.15203/CISS_2018.005

Introduction

Matches, or games, in team sports can be seen as complex dynamic systems (Glazier & Davids, 2009). The frequent interaction of various players is an integral part of any team sports match (Passos, Araújo & Volossovitch, 2016). Hence, a team must be regarded as more than the sum of its parts, and the secret to successful performance is believed to lie in the collective action of team members (Grund, 2012). Understanding the patterns of play is important to deduce the nature of the sport. Moreover, the individual contribution of each player to the organizational process is highly relevant to revealing how

ABSTRACT

Team sports are complex dynamic systems based on the frequent interaction of various players. Recently, social network analysis has been introduced to the study of sports dynamics in order to quantify the involvement of individual players in the interplay and to characterize the organizational processes used by teams. Nonetheless, only a limited set of team sports has been assessed to date, and the focus of most studies has been on the application of small sets of network metrics to a single sport. Our study aims at comparing the network patterns of different team sports in order to contribute to the understanding of their underlying nature. It considers three invasion games, namely professional matches from basketball, football and handball. By applying relevant centrality measures and minimum spanning trees a first comparison between the nature of interplay in various team sports is offered as well as a deeper understanding of the role of different tactical positions in each sport. The point guard in basketball, defensive midfielder in football and center in handball are identified as the most central tactical positions. Direct interplay is most balanced in football followed by basketball and handball. A visualization of the basic structure of interplay for each sport is achieved through minimum spanning trees.

Keywords:

social network analysis – team sports – interaction matrices – minimum spanning trees

a team functions (Vilar, Araújo, Davids, & Bar-Yam, 2013). The complexity of matches and team dynamics makes breaking down such patterns difficult, creating an ongoing challenge for performance analysis in team sports.

There is an increasing interest in applying Social Network Analysis (SNA), a method that exploits familiar performance variables such as passes, in order to detect patterns in the interplay of teams (Clemente & Martins, 2017). Network approaches focus on breaking down the web of interactions in systems of multiple agents also referred to as nodes (Passos et al., 2011). Traditional application areas of this method can be found in biological (e.g. spread of diseases) and sociological (e.g. acquaintance net-

works) contexts. In sports, the frequent interaction between a limited set of players, e.g. through the passing of a ball qualifies network theory as a powerful performance analysis method. Clemente, Martins, Wong, Kalamaras and Mendes (2015b) analyze professional football matches by applying SNA. On a micro level, i.e. focusing on the prominence of individual players in a team, the authors identify the position of the central midfielder as the most prominent player in their study, as midfielders are responsible for building offensive lines of attack. Pena and Touchette (2012) detect certain cliques within football teams that interact more frequently than others. This is in line with another micro-level study by Gama et al. (2014), who find that only a subset of players in football teams is responsible for the majority of interaction and thus shaping the pattern of play. On both, micro and macro level, i.e. focusing on the collective organization of a team, Duch, Waitzman and Amaral (2010) identify a strong connection between several network measures and traditional performance indicators whereas Grund (2012) connects the distribution of individual networks measures to performance outcomes. In his macro-level analysis, the author finds that successful teams in football demonstrate a more balanced interplay.

In basketball, SNA has been applied in professional and amateur settings. Fewell, Armbruster, Ingraham, Petersen and Waters (2012) and Clemente, Martins, Kalamaras and Mendes (2015a) identify the Point Guard as the dominant player structuring plays for the team.

However, the set of sports that SNA has been applied to has been limited so far. Moreover, the focus of most studies has been on the application of small sets of network metrics to a single sport. Our study aims at comparing the network patterns of different team sports in order to contribute to the understanding of their underlying nature. It considers three invasion games, namely professional matches from basketball, football and handball. The overarching task of each team trying to collectively outperform or -score its opponent unites these popular team sports. However, as they differ in their environmental constraints (e.g. areas, rules), different interaction patterns are needed in order to succeed (Araújo & Davids, 2016).

SNA enables us to investigate the resulting complex webs of interaction between the players in the different sports. To ensure a thorough analysis, individual and team metrics are applied alongside the computation of minimum spanning trees, a network technique that facilitates an intuitive visualization of the strongest relationships in complex networks revealing the basic structure of the sports.

In combination with the macro-level analysis, i.e. applying team metrics, this assesses the overall interaction patterns. The micro-level analysis, i.e. applying individual metrics, is specifically targeted at revealing the dominant tactical positions in terms of their involvement in the interplay for each sport and who are responsible for structuring these patterns. The combined analysis enables us to break down the complex organizational processes within teams and thus contributing to the understanding of the underlying nature of basketball, football and handball.

To our knowledge, this is the first study that attempts a comparison of different team sports applying SNA. Furthermore, it is the first analysis that takes handball into consideration along with football and basketball and applies minimum spanning trees in the context of team sports.

Hence, this study breaks down the underlying complexity of team sports by characterizing and quantifying individual and team performance through SNA.

Methods

Samples

For each sport, eight knockout round matches in the men's competition at major professional tournaments are considered for analysis, minimizing the home/away bias (Courneya & Carron, 1992). For basketball and handball the knockout stages at the Rio 2016 Summer Games Olympics tournaments are recorded and analyzed. For football, the authors consider the last eight matches from the knockout stage of the FIFA World Cup 2014 tournament. A total of 16 adjacency matrices for each sport are generated, capturing the interaction between players of each team. A total of 4059 passes are analyzed in basketball, 6934 in football and 8054 in handball.

Procedure

In order to apply SNA, adjacency matrices capture the passing distribution seen in every analyzed match. The matrices are constructed from a set of nodes and edges for every team respectively. Players represent nodes such that the number of passes between them defines the edge weight. The overall match-based interaction matrix per team is a result of an aggregation of the units of attack defined as the moment from ball recovery until possession is lost (Passos et al., 2011).

The tracking process for basketball and handball games was executed through video analysis applying the software Dartfish®. The passing distribution at the FIFA World Cup 2014 tournament was provided in the official FIFA match reports on their website (www.fifa.com/worldcup/archive/brazil2014). In a thorough post-match analysis players were assigned to their respective tactical position to ensure the comparability between teams and focus on the tactical aspects of each sport. In line with O'Donoghue (2009), we acknowledge the increasing complexity of tactical roles in team sports, i.e. forwards taking on defending tasks in football. Players might temporarily occupy different areas on the pitch and fulfill different tasks which can be acknowledged as part of the role repertoire of the different tactical positions, especially in football. Eventually, this is part of why we see complex webs of interaction in team sports and why we expect that this finds its expression in the results of our analysis. The definition of tactical roles for the three sports is displayed in Table 1.

Table 1: List of tactical roles in basketball, football and handball

Basketball	Football		Handball
Center (C)	Defensive Midfielder (DM)	Offensive Midfielder (OM)	Center (C)
Point Guard (PG)	Goalkeeper (GK)	Right Central Defender (RCD)	Left Back (LB)
Power Forward (PF)	Left Central Defender (LCD)	Right Defender (RD)	Left Wing (LW)
Shooting Guard (SG)	Left Defender (LD)	Right Forward (RF)	Pivot (P)
Small Forward (SF)	Left Forward (LF)	Right Midfielder (RM)	Right Back (RB)
	Left Midfielder (LM)		Right Wing (RW)

Following the codification for each tactical position ensured that frequent substitutions of players lead to a reassignment of the given tactical positions. Predominantly, substitutions lead to a direct replacement for the corresponding tactical position, meaning the player who was codified to a specific position was replaced by his substitute. However, substitutions occasionally implied the reassignment on multiple positions, mostly in basketball and handball. To detect these changes, each unit of attack was considered separately. Tracking and codification processes were executed by researchers with more than ten years of experience in the sports described. In order to ensure the reliability of the study, Cohen's kappa and Gwet's AC1 inter-rater statistic were computed in a two-stage process (Gwet, 2001). In a first step, the agreement on the occurrence of passes was analyzed using Gwet's statistics. In a second step, the agreement on passer and pass receiver was tested applying Cohen's Kappa. 12.5% of the overall data were tested for reliability purposes. The Kappa (Gwet, 2001) values were above 0.94 (0.85) respectively for each sport, ensuring the reliability of the data.

Network Metrics

For the 16 adjacency matrices in each sport a set of individual- and team-related centrality network metrics are computed. The analysis was carried out using the software Matlab® and the visualization of networks was generated by applying Cytoscape®.

Centrality calculations allow a quantification of the influence of tactical positions on their team's interplay as well as the balance of influence between players overall. To account for the nature of the sports, metrics that consider weighted directed graphs were applied. This allows for a breakdown of the connection between any two players in both passing directions. For individual (or micro-level) analysis weighted in-/out-degree, weighted betweenness and weighted closeness were computed. For team (or macro-level) analysis, the corresponding centralization values were calculated. These metrics are explained in detail in the following.

Individual Metrics Weighted in-degree (C_{wid}), also referred to as Prestige, is the sum of the incoming weighted edge values of a node. Hence, these metrics capture the number of successfully received passes of a player and a high value is often taken as a first indicator for the prominence of a particular player (Clemente et al., 2015b). Team members appear to trust this player, when in possession, to positively contribute to the team's performance and therefore target him more frequently than others. Weighted out-degree (C_{wod}), also referred to as Centrality, is the sum of outgoing weighted edges of a node. In the context of sports, (C_{wod}) is the number of completed passes of a player and a high value is often associated with a high contribution to ball circulation (Clemente et al., 2015b).

We also calculate the ratio C_{wid}/C_{wod} to assess a potential deviation between the share in pass reception and execution. A player with a higher reception than execution share, i.e. a value above 1, could indicate a player who rather finishes attacks. He frequently receives the ball from team members to execute shots on target rather than passing on. The opposite, i.e. a value below 1, might be a player who initiates attacks.

Weighted betweenness (C_{wb}) assesses how often a node is on the shortest path between two other nodes (Wassermann & Faust, 1994). A modified version of the standard computation of C_{wb} according to Newman (2001) is applied, which is more suitable for team sports since it favors strong connections rather than penalizing them. It measures how often a player is in between the most frequent passing connections of any other two players, thus functioning as a bridging unit (Pena & Touchette, 2012). As this implies a certain level of dependency on that particular player to ensure ball circulation it can be considered as a playmaker indicator.

Weighted closeness (C_{wc}) addresses how well connected a node is to all other nodes, directly or indirectly, within a network following Freeman (1978) and Opsahl, Agneessens and Skvoretz (2010). In a nearly complete network, i.e. in which almost every node is connected to each other, the metric can be seen as a more sophisticated approach to the weighted degree computations as the distribution of weights between other nodes is taken into account. In team sports, C_{wc} describes the

how well a player directly or indirectly interacts with all other team members on the field. Hence, a player with high weighted degree values but comparatively low weighted closeness value might only interact strongly with a subset of his team members.

Team Metrics Centralization measures are concerned with the distribution of the individual metrics in a network. Following Freeman (1978) and Wasserman and Faust (1994), weighted in-degree centralization (C_{WIDC}) captures the deviations from all in-degree values to the highest value in the network adjusted by the number of passes and the number of players. This adjustment in the computation allows a comparison between different sports. Weighted out-degree centralization (C_{WODC}), weighted betweenness centralization (C_{WBC}) and weighted closeness centralization (C_{WCC}) is calculated accordingly.

By construction, all centralization values are bounded between 0 and 1. A network is regarded as highly centralized, i.e. a value close to 1, when the score of a particular node clearly outweighs the scores of all others and rather decentralized, i.e. a value close to 0, when the scores are similar among all nodes (Grund, 2012). In a sports context, C_{WIDC} and C_{WODC} scores can be seen as indicators for the balance of direct interplay in a team. C_{WBC} and C_{WCC} scores signal how balanced the influence on the overall interplay is within the team, considering direct and indirect connections. In general, high values could imply that interplay depends on only a subset of players.

For reasons of comparability between different matches, we normalized all centrality values by the total scores of the respective metrics following Leydesdorff (2007). The values themselves have no direct relevance. Relative comparisons between the different values of a respective metric for the tactical positions were highly crucial.

Visualization

A more intuitive visualization of the underlying structure of the networks was allowed for by computing minimum spanning trees (MSTs) for each sport. MSTs are meant to provide a revelation of the strongest relationships in complex networks (Mantegna, 1999). As a visualization method, they reduce the complexity of connected graphs of n nodes with up to $n(n-1)$ connections to the strongest $n-1$ edges under the side condition that each node is still contained. According to Araújo and Davids (2016), sport teams demonstrate a task-specific organization to reach a common goal under certain constraints. In past studies, MSTs have been applied to visualize how sets of team members organize themselves to form an effective collective organization for a specific task (Lappas, Liu, & Terzi, 2009; Li & Shan, 2010). Hence, we apply MSTs to trace how teams consisting of a limited set of players organize their interplay in order to achieve group success. The method reduces the complex network of passes to the most basic structure presenting the most intensive connections under the consideration of all players.

As MSTs are only applicable to undirected graphs, the total passing intensity between pairs of players is considered in their construction. Reducing the amount of edges and thus complexity of the otherwise nearly complete networks, offers an alternative perspective on the pattern of interplay of the different team sports and hierarchical structure of weighted graphs (Gower & Ross, 1969).

Statistical Procedures

The authors of this paper utilized multiple one-way ANOVA to test for statistical differences between the centrality levels of the tactical positions within each sport, and between the analyzed sports. The assumption of normality for dependent variables was tested using Kolmogorov-Smirnov tests (p -value < .05). The assumption of homogeneity for groups' variances was examined by using Levene's test. There were no violations of either normality or homogeneity. Pairwise comparisons were established by running Bonferroni post-hoc tests. The statistical analyses were all conducted at a significance level of $p < .05$ using Matlab®. Following Ferguson (2009) and Clemente and Martins (2017), η^2 is reported to interpret the effect size according to the following criteria: no effect ($\eta^2 < .04$); small effect ($.04 \leq \eta^2 < .25$); moderate effect ($.25 \leq \eta^2 < .64$); strong effect ($\eta^2 \geq .64$).

Results

The tests found statistical differences in the dependent variables for all centrality measures applied for the three team sports considered in this study. The η^2 values reported in Table 2 almost all demonstrate moderate to strong effects sizes for the multiple one-way ANOVA in this study.

Individual Parameters

Table 3 shows the descriptive statistics and post-hoc results for tactical positions in basketball. The PG position is assigned the highest values for all centrality metrics and is significantly more central than every other tactical position. For weighted betweenness, the normalized value of the PG is 0.87 and thus more than ten times higher than the next ranked tactical position. There is no value assigned here for the forward positions implying that no strongest connection between any two players on the team runs via those tactical positions. In general, the other four tactical positions demonstrate similar values and no statistical differences are found between them for the other metrics applied in this study.

The C_{WID} / C_{WOD} ratios are shown in Figure 1. Notable in the ratio revealed is the relatively low value for the center position. Here, the share in pass completion rate outweighs the share in pass reception.

Table 2: Effect size values η^2 for multiple one-way ANOVA

	Basketball	Football	Handball		All
C_{WID}	.59 (moderate)	.23 (small)	.92 (strong)	C_{WIDC}	.89 (strong)
C_{WOD}	.46 (moderate)	.27 (moderate)	.92 (strong)	C_{WODC}	.81 (strong)
C_{WB}	.87 (strong)	.32 (moderate)	.91 (strong)	C_{WBC}	.89 (strong)
C_{WC}	.72 (strong)	.44 (moderate)	.93 (strong)	C_{WCC}	.83 (strong)

No effect ($\eta^2 < .04$); small effect ($0.04 \leq \eta^2 < .25$); moderate effect ($.25 \leq \eta^2 < .64$); strong effect ($\eta^2 \geq .64$)

Table 3: Descriptive statistics and post-hoc results for basketball

	PG	SG	SF	PF	C
C_{WID}	0.30 (0.04) _{all}	0.20 (0.03) _{PG}	0.16 (0.03) _{PG}	0.16 (0.02) _{PG}	0.17 (0.02) _{PG}
C_{WOD}	0.28 (0.04) _{all}	0.18 (0.04) _{PG}	0.17 (0.03) _{PG}	0.16 (0.03) _{PG}	0.21 (0.03) _{PG}
C_{WB}	0.87 (0.20) _{all}	0.07 (0.19) _{PG}	-	-	0.05 (0.10) _{PG}
C_{WC}	0.27 (0.02) _{all}	0.19 (0.03) _{PG}	0.17 (0.02) _{PG}	0.17 (0.02) _{PG}	0.19 (0.02) _{PG}

Subscripts indicate to which tactical positions given value is statistically different for $p < .05$, e.g. *PG*: given value is statistically different to the value of the point guard; *All*: value is statistically different to all other tactical positions.

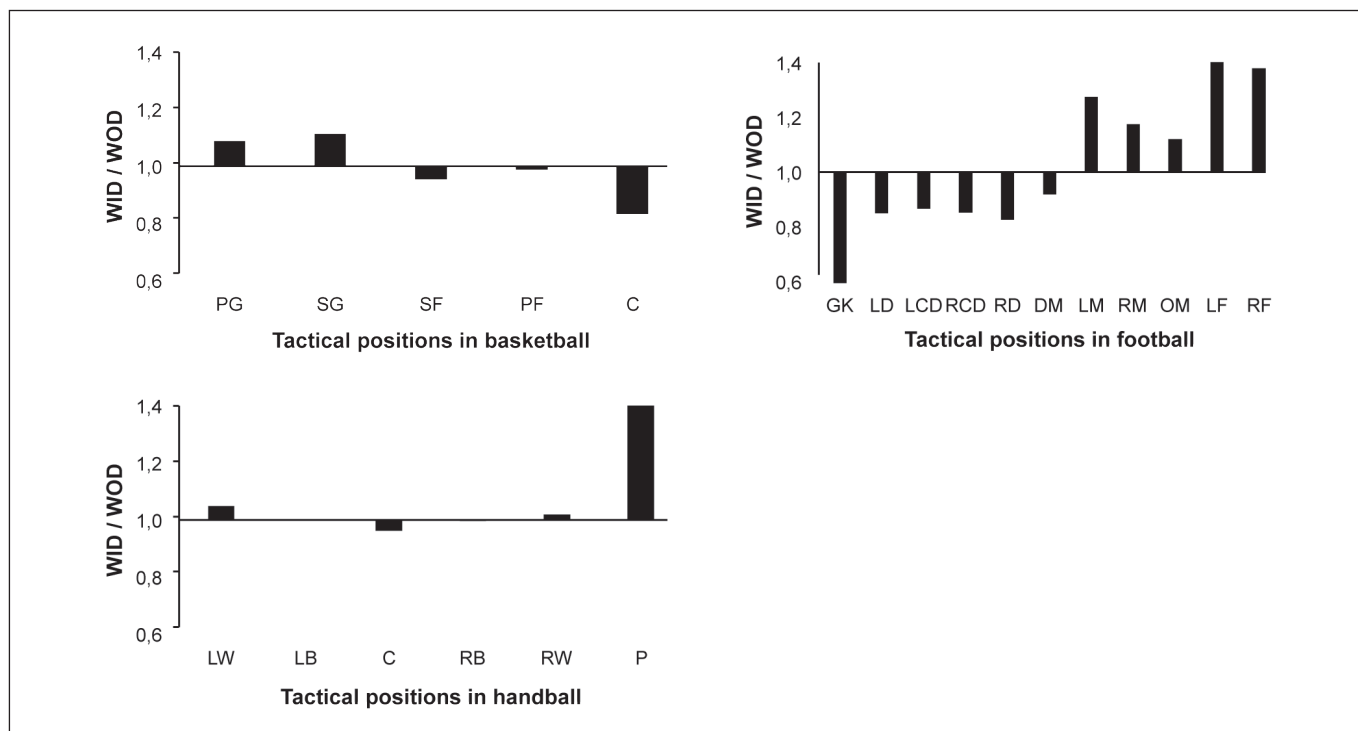


Figure 1: WID/WOD ratios for basketball, football and handball

Table 4: Descriptive statistics and post-hoc results for football

	GK	LD	LCD	RCD	RD	DM	LM	RM	OM	LF	RF
C_{WID}	0.03 (0.01) _{all}	0.08 (0.02) _{GK}	0.09 (0.02) _{GK}	0.09 (0.02) _{GK}	0.10 (0.02) _{GK}	0.12 (0.03) _{GK}	0.10 (0.02) _{GK}	0.11 (0.02) _{GK}	0.11 (0.02) _{GK}	0.09 (0.04) _{GK}	0.09 (0.02) _{GK}
C_{WOD}	0.06 (0.02) _{mult}	0.10 (0.02) _{mult}	0.10 (0.02) _{mult}	0.11 (0.02) _{mult}	0.12 (0.02) _{mult}	0.13 (0.02) _{GK,LM,Fs}	0.08 (0.02) _{mult}	0.09 (0.02) _{mult}	0.10 (0.02) _{mult}	0.06 (0.03) _{mult}	0.06 (0.02) _{mult}
C_{WB}	0.00 (0.01) _{mult}	0.08 (0.09) _{mult}	0.12 (0.06) _{mult}	0.13 (0.07) _{mult}	0.18 (0.14) _{mult}	0.18 (0.10) _{all-CDs,OM,RD}	0.05 (0.05) _{mult}	0.07 (0.07) _{mult}	0.11 (0.09) _{mult}	0.05 (0.08) _{mult}	0.03 (0.06) _{mult}
C_{WC}	0.06 (0.01) _{all}	0.09 (0.01) _{mult}	0.09 (0.01) _{mult}	0.10 (0.01) _{mult}	0.10 (0.01) _{mult}	0.11 (0.01) _{GK,Fs,LD,LF,LM}	0.09 (0.01) _{mult}	0.10 (0.01) _{mult}	0.10 (0.01) _{mult}	0.08 (0.02) _{mult}	0.08 (0.01) _{mult}

Subscripts indicate to which tactical positions given value is statistically different for $p < .05$, e.g. *GK*: given value is statistically different to the value of the goalkeeper; *All*: value is statistically different to all other tactical positions; *All-“tactical position(s)”*: value is statistically different to all other tactical positions *except* the listed ones; *Mult*: value is statistically different to various tactical positions that are not part of further analysis in this study; *Fs* includes *LF* and *RF*; *CDs* includes *LCD* and *RCD*.

Table 5: Descriptive statistics and post-hoc results for handball

	GK	LW	LB	C	RB	RW	P
C_{WID}	-	0.04 (0.02) _{C,Bs}	0.23 (0.02) _{all-RB}	0.36 (0.03) _{all}	0.26 (0.02) _{all-LB}	0.05 (0.02) _{C,Bs}	0.05 (0.02) _{C,Bs}
C_{WOD}	0.01 (0.00) _{all-LW,P}	0.04 (0.02) _{C,Bs}	0.23 (0.02) _{all-RB}	0.38 (0.03) _{all}	0.26 (0.02) _{all-LB}	0.05 (0.02) _{C,Bs,GK}	0.03 (0.01) _{C,Bs}
C_{WB}	-	0.03 (0.04) _{C,Bs}	0.23 (0.05) _{all-RB}	0.45 (0.05) _{all}	0.25 (0.10) _{all-LB}	0.02 (0.02) _{C,Bs}	0.02 (0.02) _{C,Bs}
C_{WC}	0.04 (0.01) _{all}	0.14 (0.02) _{all-PRW}	0.18 (0.01) _{all-RB}	0.18 (0.01) _{all-Bs}	0.18 (0.01) _{all-C,LB}	0.15 (0.01) _{all-LW}	0.13 (0.02) _{all-LW}

Subscripts indicate to which tactical positions given value is statistically different for $p < .05$, e.g. *C*: given value is statistically different to the value of the center; *All*: value is statistically different to all other tactical positions; *All-“tactical position(s)”*: value is statistically different to all other tactical positions *except* the listed ones, e.g. *All-C*: given value is statistically different to all other values but the one of the center; *Bs* includes *LB* and *RB*.

The corresponding results for football matches under investigation can be seen in Table 4. The DM position scores the highest C_{WID} and C_{WOD} values, meaning that this position had on-average the highest number of successfully received and executed passes. Statistically significant differences can only be shown in comparison with the GK position for C_{WID} and certain attacking positions for C_{WOD} additionally. DM is also leading the C_{WB} scores followed by the RD and central defender positions. Their respective values are significantly different to the values of the other tactical positions; whereas the C_{WC} values are similar between all tactical roles apart from the GK. The C_{WID}/C_{WOD} ratios in Figure 1 show values below 1 for defensive positions and above 1 for offensive positions, especially strikers.

For handball, C is significantly more central than all other tactical positions based on C_{WID} , C_{WOD} and especially C_{WB} . The C_{WB} values indicate that C frequently functions as the bridging unit between other tactical positions. Table 5 shows that the remaining back positions (LB and RB) have similar values for each metric and are significantly different to all other tactical positions for C_{WID} , C_{WOD} and C_{WB} . The same applies for the wing positions (LW and RW). However, their values fall into the same category with the pivot position. The GK values are neglecting low and ranked last for the considered metrics. The C_{WID}/C_{WOD} ratios in Figure 1 reveal a high value above 1 for the point. Its share in pass reception outweighs share in pass completion.

Team Parameters

The descriptive statistics and post-hoc results for the team metrics in Table 6 show that the considered sports have significantly different values for almost all centralization measures

Table 6: Descriptive statistics and post-hoc results for team metrics

	Basketball	Football	Handball
C_{WIDC}	0.13 (0.05) _{FB,HB}	0.05 (0.02) _{BB,HB}	0.24 (0.04) _{BB,FB}
C_{WODC}	0.10 (0.04) _{FB,HB}	0.05 (0.01) _{BB,HB}	0.25 (0.03) _{BB,FB}
C_{WBC}	0.89 (0.15) _{FB,HB}	0.22 (0.09) _{BB,HB}	0.35 (0.06) _{BB,FB}
C_{WCC}	0.13 (0.03) _{FB,HB}	0.05 (0.01) _{BB}	0.05 (0.01) _{BB}

Subscripts indicate to which team sport given value is statistically different for $p < .05$, e.g. *FB*: given value is statistically different to the value in football.

calculation, we were able to follow Freeman’s definition in our between each sport. As the highest values were unique in every computations. The average C_{WIDC} and C_{WODC} values are highest for handball, followed by basketball in second place. This order for first and second rank switches between these two sports for C_{WBC} and C_{WCC} . Football has the lowest average values for all team metrics employed in this study.

Visualization

Figure 2 displays the aggregated passing distribution of all matches in each sport and the corresponding MSTs next to that on the right-hand side. As edge weights were unique in each network, the resulting MSTs are unique as well (Li, Hou & Sha, 2005). The tree representing the passing network in basketball shows a typical star network topology with the PG as the central node to which all other tactical positions are connected. The topology of the handball MST has a strong resemblance with the tactical formation of the sport. The C position emerges as the centrally located node connected to the pivot and back positions who themselves are adjoined to the wings. No

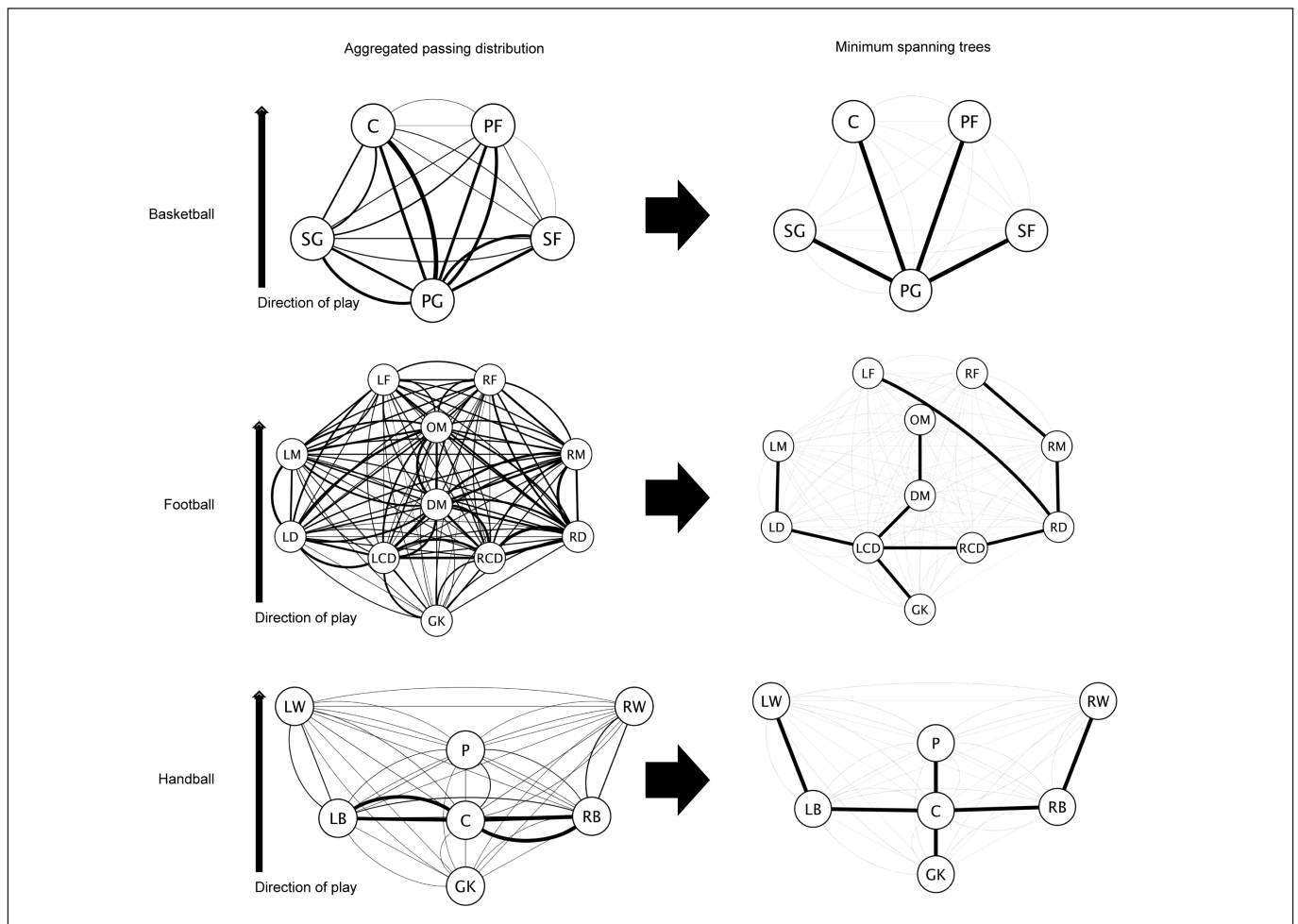


Figure 2: Visualization of aggregated passing distribution and MSTs for basketball, football and handball

distinct shape can be taken from the football MST. However, defensive positions are centrally located, and the tree displays three clusters in the longitudinal direction. Apart from the direct connection between the RD and LF, tactical positions are subdivided into left, central and right areas of the pitch and were shown as directly connected.

Discussion

The aim of this study was to characterize and compare the complex interactions visible in team sports. Network properties aid in breaking down this complexity and assessing the overall cooperation or collective organization of players and their individual contribution to a team's interaction. This is known to be vital in the analysis of team sports (Vilar et al., 2013).

This research study was conducted using passing data from several matches of major professional tournaments in basketball, football and handball. Of course, team interactions might also take other forms than passing events to express the relationship between players, e.g. the communication between the players on the field. Although there is no doubt on the importance of these forms of interaction, we assess direct passes between players as the most relevant form of interaction to characterize collective organization in team sports (Grund, 2012). The resulting analysis of our study reveals statistical differences in the pattern of play between different sports and the tactical positions therein with moderate to strong effect sizes.

The results of the individual metrics identified the DM as the most prominent player in football. He and the central defenders who act as the bridging players, as revealed by their leading C_{WB} scores, secure the ball circulation. The MST topology supports this line of argument, as these positions are centrally located within the tree, implying a strong contribution to the interaction in the sport. A centrally located player in the MST indicates a close connection or interaction with team members supporting the argument that he is a vital part in forming the collective organization of his team. There are several reasons why the RD position is also ascribed a central role to in this study according to the network metrics. First, 50% of all attacks on average were built via the right wing in comparison to 31% via the left wing. Second, the RD was among the top 3 pass executors in 10 out of 16 networks confirming the involvement of that position in building attacks via the right wing. Third, renowned players such as Philipp Lahm took on the RD position during the tournament. He alone produced 10-20 deliveries or solo runs into the attacking third per game in comparison to 2-5 for his counterpart on the LD position. This supports the dominant role of the RD and strong connection to forward positions visualized through the connection in the MST. However, the similar C_{WC} scores suggest that all players in general are equally strongly connected with each other, directly or indirectly, implying that a quick ball circulation from any player to another is given in football, in line with previous studies (Pena & Touchette, 2010).

In basketball the central role of the PG becomes obvious looking at the C_{WB} scores. A majority of the strongest connections between positions run via the PG, identifying him as the bridging player between tactical positions in basketball. The star network topology of the MST with the PG situated in the center visualizes these findings. The dominant role of this tactical position is also in line with several previous studies (Clemente et al., 2015a; Fewell et al., 2012).

In handball, the C_{WB} results suggest a central role of the C position in facilitating the ball and structuring the interplay in that sport. The C_{WC} metric evaluates how closely a player is connected with *all* other players. The fact that the corresponding C_{WC} share is less than half as high (0.18 to 0.45) suggests that C predominantly interacts with a subset of players i.e. the back positions. The C_{WID} and C_{WOD} scores support the argument that the back positions are the dominating players here.

A deeper role division can be taken from the reported C_{WID}/C_{WOD} ratios. In football, the ratios indicate a subdivision between attacking and defensive roles. The defensive roles show higher C_{WOD} than C_{WID} values, thus ratios below 1, as they initiate plays while attacking roles rather finish them. This observation is not made in the other two sports. Solely in the case of handball, the P has a relatively high C_{WID}/C_{WOD} ratio as that player is mostly targeted to finish attacks rather than initiating them. Apart from these indications, a clear division into distinct roles is not visible in either basketball or handball. Although we analyzed matches from tournaments at the highest professional level, differences in C_{WID} and C_{WOD} values might also be ascribed to limited technical abilities to a certain extent. Whereas in basketball (13.5 turnovers against 253.7 passes for a 94.9% passing success rate on average per match for each team) and handball (10.8 turnovers against 503.4 passes for a 97.9% passing success rate) this aspect might be considered rather negligible, the passing success rate in football for the considered matches is only at 76.5%. Therefore, technical limitations might add to the high ratios of C_{WID} to C_{WOD} in football for some players.

The results of the team metrics show that general interplay is most balanced between players in football based on the distribution of all individual metrics among tactical positions. As the DM and RD have relatively high C_{WB} scores in comparison to the other tactical positions, the corresponding C_{WBC} value is slightly higher than for the other team metrics in football. This could mean, that although interplay is quite balanced, there is a tendency towards a few players having a stronger influence on the structuring of the interplay.

The interplay in basketball was demonstrated to be more unbalanced than in football. Although pass reception and execution were equally distributed between most tactical positions, the PG leads both categories significantly also resulting in higher C_{WIC} and C_{WOC} values than in the case of football. The bridging player characteristic of the PG also explains the high C_{WBC} score of 0.89. In fact, in 9 of the 16 networks in basketball the C_{WBC} score takes on the maximum value of 1. This implies that every strongest connection between any two players in these mat-

ches involved the PG confirming the dominant role of this player in facilitating the interplay.

The most unbalanced interplay between tactical positions in this study can be seen in handball according to the distribution of the direct interplay captured in the C_{WIC} and C_{WOC} scores. However, the low C_{WCC} score suggest that, similar to football, all players in handball, are quite equally strongly connected, directly or indirectly, with each other. The low direct involvement of the GK in the interplay is partly offset by the consideration of indirect connections in this metric.

The topology of the MSTs, which reduces the complexity to the most intense connections between players, offers a richer insight into certain patterns of play. For handball, the patterns in question perfectly resemble the basic order of the tactical line-up. This suggests that interplay is quite structured and pre-defined and therefore that the central role of the three back positions is primarily a result of their tactical position in a quite static basic order. They are crucial for the ball circulation and structure the collective organization of the team in order to score. In football, we have similar findings, however, less strong. Here a longitudinal clustering, meaning a subdivision into attacking wings, is visible. The basic order of the tactical positions appears to foster a stronger interplay of certain dyads e.g. between wing defenders and wing midfielders.

In basketball, the central role of the PG in structuring the offensive plays outweighs any other potential cluster formation of tactical positions, resulting in the star network topology of the MST. According to Bonanno, Caldarelli, Lillo and Mantegna (2003) this kind of topology is an argument for a clear hierarchical structure, i.e. that the PG has a strong impact on structuring the interplay of his team. Teammates continuously bring the PG into possession to initiate and structure plays (Bourbousson, Poizat, Saury & Seve, 2010).

The main limitation seen in this research study was related primarily to the sample size of the data utilized. Moreover, matches from only one major tournament are considered in each sport. In order to generalize the results for each sport, a larger sample across different occasions would be needed. Besides, definitions of tactical positions in football are approximations in some instances by combining data on tactical lineups and positional data provided by FIFA (www.fifa.com/worldcup/archive/brazil2014). There is an overall consensus on the definition of tactical roles in previous studies focusing on basketball and especially handball induced by its quite static formation (Cardinale, Whiteley, Hosny, & Popovic, 2017; Fewell et al., 2012; Karcher & Buchheit, 2014). However, in football, we acknowledge that tactical roles are a more complex factor. Here, we believe that temporarily occupying different areas on the pitch and fulfilling different tasks, i.e. a striker who takes on defending tasks, can be acknowledged as part of the role repertoire of players in football. Eventually, this is why we are faced with such complex webs of interaction in which different tactical positions interact with each other and that network analysis is able to capture for the purpose of our study.

Moreover, it is important to make two remarks regarding the application of weighted closeness in this study. First, one could argue that the nearly completeness of the present networks in this study, in which almost all players are directly connected with each other, mostly account for the similar C_{WC} scores in football. However, in basketball, for example, we find statistical differences especially with regard to the PG while having complete networks in every analyzed match exclusively. We claim that in weighted networks, in comparison to unweighted networks, strong indirect connections might dominate weak direct connections and thus weaken the influence of the level of completeness in a network to a certain degree.

Second, only 13 of the 16 analyzed networks could be considered in the one-way ANOVA of the C_{WC} scores in handball, as the GK was not involved in any interplay in some matches. However, as the metric analyzes the connection with *all* players in the network and cannot consider disconnected components by definition, we had to drop three networks (Opsahl et al., 2010). This stresses the low involvement of the GK in building attacks in handball.

Nevertheless, this study contributes to the understanding of the nature of team sports and the respective involvement of the different tactical positions within each sport. This identifies SNA as a powerful tool not only to break down the performance of a single sport but also to allow a profound comparison between the styles of interaction in team sports.

Conclusion

The aim of this study was to characterize the nature of team sports and the role of their respective tactical positions.

By applying methods from social network analysis it was possible to break down the complexity of a handful of popular sports, by quantifying and intuitively visualizing roles of players and overall team interaction. Thus, this is the first study that compares the network patterns of different team sports. Moreover, MSTs are applied for the first time in a team sports context which in particular turn out to be effective in breaking down the complexity of almost complete networks.

Ultimately, the analysis revealed significant findings, on the prominent tactical positions for building attacks in the three sports discussed: in basketball, this dominant tactical position tended to be the PG, in football the DM and C in handball. The general pattern of play appears to be significantly more unbalanced in handball than in basketball and football. As a final takeaway, the study indicated strong findings that the level of fixedness in the basic order of the tactical positions in the sports influences the prominence levels of players.

We chose three popular invasion games in this study to offer a first comparison between the network properties of team sports. However, as we assess the outlook of this method as fruitful, more team sports should be incorporated in future studies to further examine and characterize the different dynamic

systems present in team sports. Moreover, individual modifications of traditional network metrics may lead to an even more accurate quantification of performance in each sport.

Funding

The authors have no funding or support to report.

Competing Interests

The authors have declared that no competing interests exist.

Data Availability Statement

All relevant data are within the paper.

References

- Araújo, D., & Davids, K. (2016). Team synergies in sport: Theory and measures. *Frontiers in Psychology, 7*:1449. doi: 10.3389/fpsyg.2016.01449
- Bonanno, G., Caldarelli, G., Lillo, F., & Mantegna, R. N. (2003). Topology of correlation-based minimal spanning trees in real and model markets. *Physical Review E, 68* (4), 046130.
- Bourbousson, J., Poizat, G., Saury, J., & Seve, C. (2010). Team coordination in basketball: Description of the cognitive connections among teammates. *Journal of Applied Sport Psychology, 22* (2), 150–166.
- Cardinale, M., Whiteley, R., Hosny, A. A., & Popovic, N. (2017). Activity profiles and positional differences of handball players during the world championships in Qatar 2015. *International Journal of Sports Physiology and Performance, 12*, 908–915.
- Clemente, F. M., Martins, F. M. L., Wong, P. D., Kalamaras, D., & Mendes, R. S. (2015a). Midfielder as the prominent participant in the building attack: A network analysis of national teams in FIFA World Cup 2014. *International Journal of Performance Analysis in Sport, 15*, 704–722.
- Clemente, F. M., Martins, F. M. L., Kalamaras, D., & Mendes, R. S. (2015b). Network analysis in basketball: Inspecting the prominent players using centrality metrics. *Journal of Physical Education and Sport, 15*(2), 212.
- Clemente, F. M., & Martins, F. M. L. (2017). Network structure of UEFA Champions League teams: Association with classical notational variables and variance between different levels of success. *International Journal of Computer Science in Sport, 16*(1), 39–50.
- Courneya, K. S., & Carron, A. V. (1992). The home advantage in sport competitions: A literature review. *Journal of Sport and Exercise Psychology, 14*(1), 13–27.
- Duch, J., Waitzman, J. S., & Amaral, L. A. N. (2010). Quantifying the performance of individual players in a team activity. *PLoS one, 5*(6), e10937.
- Ferguson, C. J. (2009). An effect size primer: A guide for clinicians and researchers. *Professional Psychology: Research and Practice, 40*(5), 532.
- Fewell, J. H., Armbruster, D., Ingraham, J., Petersen, A., & Waters, J. S. (2012). Basketball teams as strategic networks. *PLoS ONE, 7*(11), e47445.
- Freeman, L. C. (1978). Centrality in social networks conceptual clarification. *Social Networks, 1*(3), 215–239.
- Gama, J., Passos, P., Davids, K., Relvas, H., Ribeiro, J., Vaz, V., & Dias, G. (2014). Network analysis and intra-team activity in attacking phases of professional football. *International Journal of Performance Analysis in Sport, 14*(3), 692–708.
- Glazier, P. S., & Davids, K. (2009). Constraints on the complete optimization of human motion. *Sports Medicine, 39*(1), 15–28.
- Gower, J. C., & Ross, G. J. (1969). Minimum spanning trees and single linkage cluster analysis. *Applied Statistics, 18*, 54–64.
- Grund, T. U. (2012). Network structure and team performance: The case of English Premier League soccer teams. *Social Networks, 34*, 682–690.
- Gwet, K. (2001). *Handbook of inter-rater reliability: How to estimate the level of agreement between two or multiple raters*. Gaithersburg, MD: STATAXIS.
- Karcher, C., & Buchheit, M. (2014). On-court demands of elite handball, with special reference to playing positions. *Sports Medicine, 44*, 797–814.
- Lappas, T., Liu, K., & Terzi, E. (2009, June). Finding a team of experts in social networks. In *Proceedings of the 15th ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 467–476). ACM.
- Leydesdorff, L. (2007). Betweenness centrality as an indicator of the interdisciplinarity of scientific journals. *Journal of the Association for Information Science and Technology, 58*, 1303–1319.
- Li, N., Hou, J. C., & Sha, L. (2005). Design and analysis of an MST-based topology control algorithm. *IEEE Transactions on Wireless Communications, 4*, 1195–1206.
- Li, C. T., & Shan, M. K. (2010, August). Team Formation for Generalized Tasks in Expertise Social Networks. In *Proceedings of the 2010 IEEE Second International Conference on Social Computing* (pp. 9–16). IEEE Computer Society.
- Mantegna, R. N. (1999). Hierarchical structure in financial markets. *The European Physical Journal B-Condensed Matter and Complex Systems, 11*(1), 193–197.
- Newman, M. E. (2001). Scientific collaboration networks. II. Shortest paths, weighted networks, and centrality. *Physical review E, 64*(1), 016132.
- O'Donoghue, P. (2009). *Research methods for sports performance analysis*. London: Routledge.
- Opsahl, T., Agneessens, F., & Skvoretz, J. (2010). Node centrality in weighted networks: Generalizing degree and shortest paths. *Social Networks, 32*(3), 245–251.

- Passos, P., Davids, K., Araújo, D., Paz, N., Minguéns, J., & Mendes, J. (2011). Networks as a novel tool for studying team ball sports as complex social systems. *Journal of Science and Medicine in Sport, 14*(2), 170–176.
- Passos, P., Araújo, D., & Volossovitch, A. (2016). *Performance analysis in team sports*. London: Routledge.
- Pena, J. L., & Touchette, H. (2012). A network theory analysis of football strategies. *arXiv preprint arXiv:1206.6904*.
- Vilar, L., Araújo, D., Davids, K., & Bar-Yam, Y. (2013). Science of winning soccer: Emergent pattern-forming dynamics in association football. *Journal of Systems Science and Complexity, 26*(1), 73–84.
- Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications* (Vol. 8). Cambridge: University Press.