Journal of Physical Education and Sport ® (JPES), 16(3), Art 130, pp. 823 - 830, 2016
online ISSN: 2247 - 806X; p-ISSN: 2247 – 8051; ISSN - L = 2247 - 8051 © JPES

Original Article

Social network measures to match analysis in soccer: A survey

FILIPE MANUEL CLEMENTE[1,2], FERNANDO MANUEL LOURENÇO MARTINS[2,3], RUI SOUSA MENDES[3], FRUTUOSO SILVA[2,4]
[1]Instituto Politécnico de Viana do Castelo, Escola Superior de Desporto e Lazer, PORTUGAL
[2]Instituto de Telecomunicações, Delegação da Covilhã, Portugal
[3]Polytechnic Institute of Coimbra, ESEC, Department of Education, IIA, RoboCorp, ASSERT, PORTUGAL
[4]University of Beira Interior, PORTUGAL

Published online: September 26, 2016
(Accepted for publication August 20 2016)
DOI:10.7752/jpes.2016.03130

Abstract:
In this study we will present the most common and adequate network measures to analyse graph properties and to inspect the prominence of each player on a soccer team. Both approaches can provide a range of useful information. This kind of analysis will help to identify the prominence of players and also characterize the collective organization and patterns of the teams. General measures and centrality levels will be described and the applications will be discussed. By using social network analysis will be possible to quantify the structure of play and predict some behaviour. Such methodology will add new options to the field of match analysis.
Keywords: graph theory; network analysis; team sports; match analysis.

Introduction

A *team sport* involves players working together to achieve a common goal, thus, within a team, intragroup relationships are important and include decisive indicators such as cohesiveness and hierarchies among players (Lusher, Robins, & Kremer, 2010). One of the main challenges in team sport is that the teammates must coordinate their efforts and actions to avoid and overcome the opponent's strengths (Deleplace, 1995), a team should work together to recapture and conserve territory in order to earn points (Gréhaigne, Bouthier, & David, 1997). However, it is not enough to have a lot of very good players if they cannot play together. Thus we have the impression that a team is more than the sum of its parts (Grund, 2012). In sum, a team of experts it is not necessarily an expert team (Bourbousson, Poizat, Saury, & Seve, 2010).

In team sport there may be problems related to the organization, for example players must accept a move from an individual to a collective project (Gréhaigne et al., 1997). In that sense, macroscopic patterns of behaviour emerge from teammates interactions at a microscopic level of organization, thus the individual actions are strongly associated with collective project, and the inverse is a dynamic and complex point-a-view (Davids, Araújo, & Shuttleworth, 2005). This becomes an assessment problems of players in team sport such as (Gréhaigne, Godbout, & Bouthier, 1997): (i) intervening elements that are not only numerous but also interacting; (ii) the rapport of strength plays an important role and may vary in different opposition contexts; (iii) the members of a team are interdependent; and (iv) the player must be assessed within a system (the team) that has its own coherence. In fact, the assessment of individual players must be understood in a global concept of the team. *Match analysis* has been used in the last decades to provide coaches the best possible information about the individual and collective organization of players and teams (Clemente, Couceiro, Martins, Mendes, & Figueiredo, 2013a).

Starting from this concept of match analysis and its utility for coaches, this article aims to discuss an integrated approach that analyses individual performances within the collective organisation and allows identification of the collective characteristics that determine the style of play and the specific patterns and signatures of teams. This article also aims to be a survey tool for sport researchers and coaches that integrates several techniques using a single approach, *social network analysis*, to inspect individual and collective organization in team sports.

The analysis to the overall team: Looking for the graph theory characteristics

The analysis at different network properties can provide valuable insight into the organization of teammates' network, identifying some global characteristics of the overall players' interactions. In the following, will be provided a short description about the main properties that are commonly analysed in network with application for team sports analysis.

For such application and analysis the SocNetV (Kalamaras, 2014) free software will be used. This software is a graphical application for the analysis and visualization of social networks. It allows the researcher to load

Corresponding Author . FILIPE MANUEL CLEMENTE, E-mail: filipe.clemente5@gmail.com

formatted network data such as sociomatrices, analyze the social and mathematical properties of the corresponding social networks in the form of mathematical graphs, and compute basic graph properties, such as density, diameter and clustering coefficient, as well as more advanced structural measures such as centrality and prestige indices, which were used in this survey.

Total Links

The first metric used in this study called *total Links.*

In our case considering each element a_{ij} of the adjacency matrix was the number of interactions (passes) from player i to player j and, in terms of the corresponding weighted digraph (sociogram) produced by SocNetV, it was represented by a directed line (arc) between node n_i and node n_j.
The sum of the elements of each row of the adjacency matrix $\sum_{\substack{i=1 \\ j\neq i}}^{n} a_{ij}, j = 1, \ldots, n$ was the total number of passes from player i to all its other teammates. The sum of all elements $a_{ij}(i \neq j)$ of the adjacency matrix, $L_D^W = \sum_{i=1}^{n}\sum_{\substack{j=1 \\ j\neq i}}^{n} a_{ij}$, is the *total Links* (passes) between each team's players. In the corresponding weighted directed graph, this number is the total arcs between all nodes.
If from of the interaction between nodes result a undirected weighted graph, the *total Links* is obtained $L_G^W = \sum_{i=1}^{n}\sum_{\substack{j=1 \\ j> i}}^{n} a_{ij}$. In the corresponding weighted graph, this number is the total lines between all nodes.

Considering only one interaction between two nodes, if exist, each element a_{ij} of the adjacency matrix is 1 or 0. If from of the interaction between nodes result a undirected graph or digraph, we obtain the *total distinct links*.
One can presume that a higher than the average total Links index of a team is an indicator for stronger cooperation between the players of that team. It might also be correlated with a higher probability of its players to interact successfully with each other, which might result in longer ball possession, better performance and generally strong collective organization against the opponent team.

Graph Density

In graph theory, the density of a (directed) graph is the proportion of the maximum possible lines (or arcs) that are present between nodes. The graph density shows how sparse or dense a graph is according to the number of connections per node (Pavlopoulos et al., 2011). Because a graph is consisted of a finite number of nodes (denoted by n), in the case of an undirected graph there can be at a maximum possible distinct pairs between nodes and $n^{(n-1)}/_2$ possible distinct links (divided by because the link (i,j) is the same to (j,i) and we do not want to count that twice). The density Δ of the graph is defined as the ratio of the distinct total lines present L_G^W to the maximum possible number of links that could exist:

$$\Delta = \frac{L_G^W}{n(n-1)/2} \text{ or } \Delta = \frac{2\times L_G^W}{n(n-1)}. \quad (1)$$

In the case of ordered relations, as in the teammate interactions, the possible distinct directed links in a digraph of *n* nodes is $n(n-1)$ so the density is computed by:

$$\Delta = \frac{L_D^W}{n(n-1)}. \quad (2)$$

In both cases, the density is a ratio having a minimum of zero (no lines/arcs present) and a maximum of (all lines/arcs are present).

As network interpretation the density measure the overall affection between teammates. In the case of binary graphs or digraphs a density value close to indicates that all teammates strongly pass each other, while a density value of 0.5 suggests the presence of more ambiguous relationship (Horvath, 2011). In the case of weighted digraphs the bigger values represents a higher affection among team-members if the network density is big as well. In the case of small network density and bigger weighted density it is ambiguous the conclusion, because in a digraph with 11 nodes only three members can be highly connected and the others non-connected. Considering weighted graphs or digraphs, the density is obtained from the expressions (1) and (2), respectively. In this case, the density is a ratio having a minimum of zero (no lines/arcs present) and a maximum of ∞.

Graph Distance

In graph theory, two nodes are *connected* if there is a sequence of nodes and their adjacent links (*walk*) from one to the other. In the example above, nodes 3 and 9 are connected through the walk 3-1-4-3-1-5-9. If a walk is consisted only of distinct nodes and lines, then it is called a *path*. In the same example above there is the path 3-1-5-9 from node 3 to node 9.

The *geodesic distance,* $d(n_i, n_j)$ between two nodes, n_i and n_j is the length of shortest path between them and in cases that no path was generated it is possible to set $d(n_i, n_j) = \infty$ assuming that the nodes are so far between each other so they are not connected (Pavlopoulos et al., 2011).

In a unweighted graph or digraph, a geodesic distance between n_i and n_j is

$$d(n_i, n_j) \quad \min_{ijh}(a_{ih} + \cdots + a_{hj}), \tag{3}$$

where *h* are intermediary nodes on paths between node n_i and n_j, and a_{uv} are elements of the corresponding adjacency matrix (Opsahl, Agneessens, & Skvoretz, 2010; Rubinov & Sporns, 2010).

In particular, considering *only* unweighted graph the the distance between n_i and n_j is equal to the distance between n_j and n_i; $d(n_i, n_j) = d(n_i, n_j)$ (Wasserman & Faust, 1994). Let us provide an illustrative example in the following Figure 1.

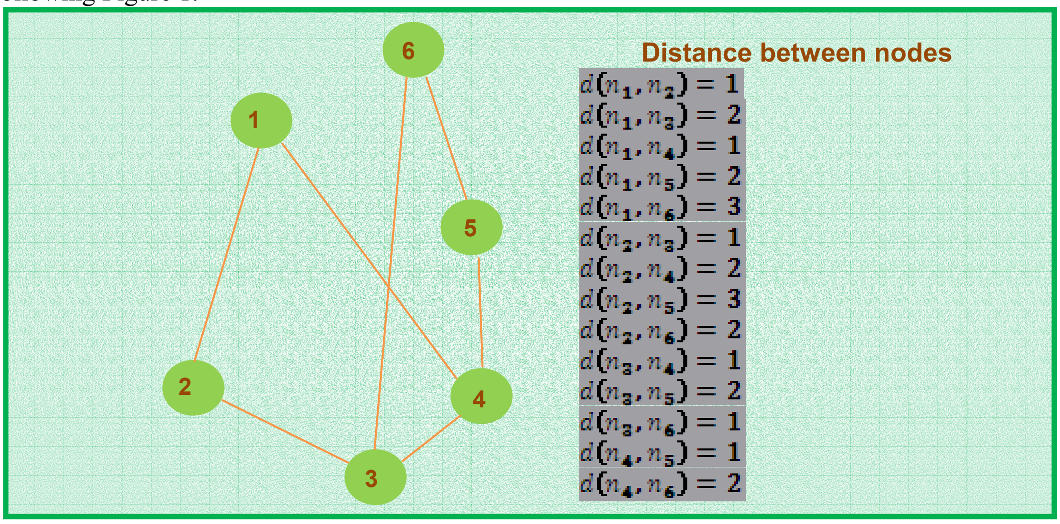

Fig. 1. Geodesic distance between nodes of a undirected and unweighted graph.

The average path length of a unweighted graph is defined to be the average and maximum value of $d(n_i, n_j)$ taken over all pairs of distinct nodes, $n_i, n_j \in V(G)$ which are connected by at least one path, thus the average path length of a network is the average number of edges between teammates, which must be crossed in the shortest path between any two nodes (Pavlopoulos et al., 2011):

$$\bar{d} = \frac{2}{N(N-1)} \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j>i}}^{n} d(n_i, n_j). \tag{4}$$

where $d(n_i, n_j)$ is the minimum distance between nodes n_i and n_j. If $d(n_i, n_j) = \infty$, then we consider that $d(n_i, n_j) = \max(d(n_i, n_j) + 1)$ (Pavopoulos, 2011). In similar for unweighted digraph we obtain

$$\bar{d} = \frac{1}{N(N-1)} \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j\neq i}}^{n} d(n_i, n_j). \tag{5}$$

In a weighted graph or digraph, a geodesic distance between n_i and n_j is

$$d^w(n_i, n_j) = \min_{ijh}\left(\frac{1}{a_{ih}} + \cdots + \frac{1}{a_{hj}}\right). \tag{6}$$

where *h* are intermediary nodes on paths between node n_i and n_j, and a_{uv} are elements of the corresponding (weighted) adjacency matrix (Opsahl et al., 2010).

The average path length of a weighted graph or digraph is determined by (Rubinov & Sporns, 2010).

$$\bar{d} = \frac{1}{N(N-1)} \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j \neq i}}^{n} d^w(n_i, n_j). \qquad (7)$$

As network interpretation one can define that lowest average values represents that teammates have a great connection between them and greater average values suggests that teammates are more disconnected between them.

The distance can be also used to characterize the spatio-temporal relationship between players. In the case of distance between teammates, would be very interesting to identify how close or how far way are the players from their teammates. This would help to characterize the defensive process (keeping the block of pressure) and the attacking process (exploring the width and length) (Costa, Garganta, Greco, Mesquita, & Seabra, 2010).

Clustering Coefficient

The Clustering Coefficient, introduced by (Watts & Strogatz, 1998), quantifies how close a node and its neighbors in a undirected and unweighted graph are to become a clique (a complete subgraph). Watts and Strogatz used the local version of Clustering Coefficient to determine whether a graph is a small-world network (a network of small average distance but relative large number of cliques).

In the case of undirected and unweighted graphs the (local) Clustering Coefficient of a node is obtain the following expression (Fagiolo, 2007; Rubinov & Sporns, 2010):

$$C_i = \frac{\sum_{j \neq i} \sum_{h \neq i,j} a_{ij} a_{ih} a_{jh}}{k_i(k_i - 1)} = \frac{(A^3)_{ii}}{k_i(k_i - 1)} \qquad (8)$$

where a_{uv} are elements of the corresponding adjacent matrix *A* and $k_i = \sum_{j \neq i} a_{ij}$.

In the case of weighted undirected graphs the (local) Clustering Coefficient of a node is obtain the following expression (Fagiolo, 2007; Rubinov & Sporns, 2010):

$$C_i = \frac{\sum_{j \neq i} \sum_{h \neq i,j} a_{ij}^{\frac{1}{3}} a_{ih}^{\frac{1}{3}} a_{jh}^{\frac{1}{3}}}{k_i(k_i - 1)} = \frac{\left(A^{\left[\frac{1}{3}\right]}\right)^3_{ii}}{k_i(k_i - 1)} \qquad (9)$$

where a_{uv} are elements of the corresponding (weight) adjacent matrix *A*, $k_i = \sum_{j \neq i} a_{ij}$ and we define $A^{\left[\frac{1}{3}\right]} = \left[a_{ij}^{\frac{1}{3}}\right]$, thus the matrix obtained from A by taking the 3rd root of each entry.

In the case of directed and unweighted graphs the (local) Clustering Coefficient of a node is obtain the following expression (Fagiolo, 2007; Rubinov & Sporns, 2010):

$$C_i = \frac{\sum_j \sum_h (a_{ij} + a_{ji})(a_{ih} + a_{hi})(a_{jh} + a_{hj})}{2\left[\left(k_i^{out} + k_i^{in}\right)\left(k_i^{out} + k_i^{in} - 1\right) - 2\sum_{j \neq i} a_{ij} a_{ji}\right]}, = \frac{(A + A^T)^3_{ii}}{2\left[\left(k_i^{out} + k_i^{in}\right)\left(k_i^{out} + k_i^{in} - 1\right) - 2\sum_{j \neq i} a_{ij} a_{ji}\right]} \qquad (10)$$

where a_{uv} are elements of the corresponding adjacent matrix *A*, $k_i^{out} = \sum_{j \neq i} a_{ij}$ and $k_i^{in} = \sum_{j \neq i} a_{ji}$ shut that k_i^{out}, k_i^{in} are outdegree and indegree of node n_i.

In the case of directed and weighted graphs the (local) Clustering Coefficient of a node is obtain the following expression (Fagiolo, 2007; Rubinov & Sporns, 2010):

$$C_i = \frac{\sum_j \sum_h \left(a_{ij}^{\frac{1}{3}} + a_{ji}^{\frac{1}{3}}\right)\left(a_{ih}^{\frac{1}{3}} + a_{hi}^{\frac{1}{3}}\right)\left(a_{jh}^{\frac{1}{3}} + a_{hj}^{\frac{1}{3}}\right)}{2\left[\left(k_i^{out} + k_i^{in}\right)\left(k_i^{out} + k_i^{in} - 1\right) - 2\sum_{j \neq i} a_{ij} a_{ji}\right]}, = \frac{\left[A^{\left[\frac{1}{3}\right]} + (A^T)^{\left[\frac{1}{3}\right]}\right]^3_{ii}}{2\left[\left(k_i^{out} + k_i^{in}\right)\left(k_i^{out} + k_i^{in} - 1\right) - 2\sum_{j \neq i} a_{ij} a_{ji}\right]} \qquad (11)$$

where a_{uv} are elements of the corresponding (weighted) adjacent matrix *A*, $k_i^{out} = \sum_{j \neq i} a_{ij}$ and $k_i^{in} = \sum_{j \neq i} a_{ji}$ shut that k_i^{out}, k_i^{in} are outdegree and indegree of node n_i and $A^{\left[\frac{1}{3}\right]} = \left[a_{ij}^{\frac{1}{3}}\right]$.

Thus the local Clustering Coefficient measures the degree of interconnectivity in the neighborhood of a node. The higher it is, the closer this node and its neighbors are to become a clique.

In the present survey, it was used a variant of the global version of Clustering Coefficient which measures the level of clustering in the whole network. This variant is the network average of the local clustering coefficients (Rubinov & Sporns, 2010):

$$\overline{C} = \frac{1}{n}\sum_{i=1}^{n} C_i. \quad (12)$$

where $n = |V|$ is the number of vertices and we consider of C_i based on the type of graph that will be performed (undirected, directed and weighted).

For the case of weighted undirected graphs the higher the clustering coefficient of a player, the higher is the affection between team-members (Horvath, 2011). In the case of $0 \leq a_{ij} \leq 1$ implies that $0 \leq C_i \leq 1$.

Testing the centralities of players: How prominent is a player?

The present section aims to show how players can be ranked or sorted according to their properties within the graph. In team sports it is important to detect central players or intermediate players that affect the patterns of play and the collective organization of team. Thus, a set of centrality metrics to inspect the prominence of players can be following found.

Degree Centrality (out-degree)

A centrality node means that player must be the most active in the sense that they have ties to other players in the network graph (Wasserman & Faust, 1994).

A centrality measure for a player should be the degree of the node, k_i, thus it is possible to define $C_D(n_i)$ as a player-level degree centrality index in the case of undirected and unweight graphs as (Wasserman & Faust, 1994):

$$C_D(n_i) = k_i = \sum_{j=1}^{n} a_{ij} = \sum_{j=1}^{n} a_{ji}. \quad (13)$$

Besides this algorithm, another can be used as a standard measure:

$$C'_{(D)}(n_i) = \frac{k_i}{n-1}, \quad (14)$$

For the case of direct and unweight graphs, we can obtain:

$$C_{D-out}(n_i) = k_i^{out} = \sum_{j=1}^{n} a_{ij}. \quad (15)$$

Besides this algorithm, another can be used as a standard measure:

$$C'_{(D-out)}(n_i) = \frac{k_i^{out}}{(n-1)^2}, \quad (16)$$

that is the proportion of nodes that are adjacent to n_i. $C'_{(D)}(n_i)$ is independent of n, and thus can be compared across networks of different sizes (Wasserman & Faust, 1994).

The case of undirected and weight graphs as (Opsahl et al 2010):

$$C_D^w(n_i) = k_i^w = \sum_{j=1}^{n} a_{ij} = \sum_{j=1}^{n} a_{ji}. \quad (17)$$

Besides this algorithm, another can be used as a standard measure:

$$C'^w_{(D-out)}(n_i) = \frac{k_i^{w^{out}}}{\sum_{i=1}^{n}\sum_{j=1, j\neq i}^{n} a_{ij}}, \quad (18)$$

that is the proportion of weights of nodes that are adjacent to n_i.

The case of directed and weight graphs as (Opsahl et al., 2010):

$$C_{D-out}^w(n_i) = k_i^{w^{out}} = \sum_{j=1}^{n} a_{ij}. \quad (19)$$

Besides this algorithm, another can be used as a standard measure:

$$C'^w_{(D-out)}(n_i) = \frac{k_i^{w^{out}}}{\sum_{i=1}^{n}\sum_{j=1, j\neq i}^{n} a_{ij}}, \quad (20)$$

that is the proportion of weights of nodes that are adjacent to n_i.

In this analysis the players with very high degree centrality are called hubs since they are connected to many other neighbours (Pavlopoulos et al., 2011).

As network interpretation a player with a high degree centrality level is “where the action is” in the network (Wasserman & Faust, 1994). Thus, such player with high level is in direct contact with the majority of their teammates and also is adjacent with the remaining players.

Closeness Centrality

This closeness centrality indicates the nodes that can interact quickly with other nodes of the network by their proximity on distance (Pavlopoulos et al., 2011).

In case of a unweighted graph or digraph (Wasserman & Faust, 1994), the closeness can be obtain:

$$C_{(C)}(n_i) = \left[\sum_{\substack{j=1 \\ i \neq j}}^{n} d(n_i, n_j)\right]^{-1}. \quad (21)$$

Another algorithm can be used to produce comparisons between unweighted graphs or digraphs with different sizes:

$$C'_{(C)}(n_i) = \frac{n-1}{\left[\sum_{\substack{j=1 \\ i \neq j}}^{n} d(n_i, n_j)\right]} = (n-1)C_c(n_i). \quad (22)$$

In particular case of this index for undirected graphs, the range between and and can be viewed as the inverse average distance between node i and all of the other nodes (Wasserman & Faust, 1994). Generally, a decrease in closeness centrality is caused by of an increase in distance between pathways (Pavlopoulos et al., 2011).

In case of a weighted graph or digraph (Opsahl et al., 2010), the closeness can be obtain:

$$C^{w}_{(C)}(n_i) = \left[\sum_{\substack{j=1 \\ i \neq j}}^{n} d^{w}(n_i, n_j)\right]^{-1}. \quad (23)$$

Another algorithm can be used to produce comparisons between weighted graphs or digraphs with different sizes:

$$C'^{w}_{(C)}(n_i) = \frac{n-1}{\left[\sum_{\substack{j=1 \\ i \neq j}}^{n} d^{w}(n_i, n_j)\right]} = (n-1)C^{w}_{(C)}(n_i). \quad (24)$$

This measure can be very useful to study the spatio-temporal relationship between teammates, particularly using the distance as the variable. The closeness centrality may be a very good topic to identify the coverage provided by the teammates to the colleagues with ball (attacking coverage) or to the colleagues in direct marking to the opponent with ball (defensive coverage).

Betweenness Centrality

Betweenness centrality measures the intermediate nodes between neighbours. Without such nodes the distance between the neighbours turns bigger. In that sense, betweenness centrality shows important nodes that lie on a high proportion of paths between other nodes in network (Pavlopoulos et al., 2011).

Consider the case of unweighted and undirected graphs. For distinct nodes $n_i, n_j, n_k \in V(G)$, let g_{ij} be the total number of shortest paths between n_i and n_j and $g_{ij}(n_k)$ be the number of shortest paths from n_i to n_j that pass through n_k. Furthermore, for $n_k \in V(G)$, let $V(n_i)$ denote the set of all ordered pairs, (n_i, n_j) in $V(G) \times V(G)$ such that i, j, k are all distinct (Pavlopoulos et al., 2011). In that sense, the betweenness centrality is calculated as:

$$C_b(n_k) = \sum_{(n_i, n_j) \in V(n_k)} \frac{g_{ij}(n_k)}{g_{ij}}. \quad (25)$$

It has a minimum of zero, attained when n_k falls on geodesics (Wasserman & Faust, 1994). Its maximum is achieved when the node falls on all geodesics. Thus, this metric can be useful to identify the players that generates shortest paths within the team, being the linkage nodes between teammates.

In the case of unweighted and undirected graphs, we standardize it just like the other actor centrality indices (Wasserman & Faust, 1994):

$$C'_b(n_k) = \frac{C_b(n_k)}{(n-1)(n-2)}. \quad (26)$$

For the other cases, weighted or/and direct graphs, the betweenness centrality is calculated the similar form shut that the path lengths are computed on respective weighted or direct paths.

Degree Prestige (in degree)

The degree prestige measure the in-degree of each node, which can be denoted by k_i^{in} or $k_i^{w^{in}}$. In the case of soccer the prestigious players can be how receive more passes from their teammates, so can be defined on directed and unweight graphs as (Wasserman & Faust, 1994):

$$P_D(n_i) = k_i^{in} = \sum_{j=1}^{n} a_{ji} \, , \quad (27)$$

In order to standardize the group size n, it is possible to compute as follows:

$$P'_D(n_i) = \frac{k_i^{in}}{(n-1)^2}, \quad (28)$$

that is the proportion of nodes that are adjacent to n_i.

The case of directed and weight graphs as (Opsahl et al 2010):

$$P_D^w(n_i) = k_i^{w^{in}} = \sum_{j=1}^{n} a_{ji}. \quad (29)$$

In order to standardize the group size n, it is possible to compute as follows:

$$P'^{w}_D(n_i) = \frac{k_i^{w^{in}}}{\sum_{i=1}^{n}\sum_{\substack{j=1 \\ j\neq i}}^{n} a_{ij}}, \quad (30)$$

that is the proportion of weights of nodes that are adjacent to n_i.

As network interpretation the largest index represents the most prestigious nodes, thus in the case of passing sequence analysis represents the player that received more ball from their teammates. The maximum prestige occurs when all passes are performed for the same player.

Looking to the future: An easy, quick, and low-cost approach for practical research applications in sports

Network analysis based on graph theory metrics can be a very useful and user-friendly solution when used as match analysis techniques in team sports. The analysis is not limited to examining the pass interactions between teammates. Many more possibilities can be found using network techniques. Here are some possible directions for further research using graph theory metrics.

Other important tactical behaviour that can be assessed using graph theory techniques can be the switching positions between teammates (Costa et al., 2010) (Duarte, Araújo, Correia, & Davids, 2012). This behaviour emerges in a match according to the tactical roles of each player. For example, the lateral defender goes to the side to confront an attacking player, and in the moment of the opponent's counter-attack, a teammate must switch positions to fill the space of their lateral defender teammate who is too far away from their original position. This kind of tactical behaviour between teammates can be also understood as a connection, thus a network analysis can be carried out. Besides a good source of information for researchers who try to understand the frequency of this switching behaviour, and what kind of players are the most likely to perform the behaviour, network analysis predicts useful information for coaches. Players can be characterized by their switching behaviour, which can be optimized during training sessions. Moreover, coaches may find these techniques helpful in examining the opposition – the information can be used to identify players that change position frequently and exploit free spaces in a given moment of counter-attack.

Finally, network analysis can be useful for players marking an opponent. This behaviour is a defensive strategy for players without possession of the ball make to avoid the attacker who receives the ball (Bangsbo & Peitersen, 2002). The connection between these players is a connection that emerges throughout a match, thus they can be viewed as a specific kind of network. This behaviour can be studied to discover the regularity of this kind of connection and to determine how players vary their behaviour based on the opponent's strategy (Clemente, Couceiro, Martins, & Mendes, 2015). Moreover, such analysis can be useful to understand how certain kinds of movements should be performed by attackers to avoid the man-defender system.

Acknowledgements

This work was supported by the FCT project PEst-OE/EEI/LA0008/2013. This article was made in the scope of uPATO project granted by Instituto de Telecomunicações.

References

Clemente, F. M., Couceiro, M. S., Martins, F. M. L., & Mendes, R. S. (2015). Using Network Metrics in Soccer: A Macro-Analysis. *Journal of Human Kinetics*, *45*, 123–134. doi:10.1515/hukin-2015-0013

Costa, I. T., Garganta, J., Greco, P. J., Mesquita, I., & Seabra, A. (2010). Influence of Relative Age Effects and Quality of Tactical Behaviour in the Performance of Youth Football Players. *International Journal of Performance Analysis in Sport*, *10*(2), 82–97.

Fagiolo, G. (2007). Clustering in complex directed networks. *Physical Review E*, *76*(2), 026107. doi:10.1103/PhysRevE.76.026107

González-Víllora, S., García-López, L. M., & Contreras-Jordán, O. R. (2015). Evolución de la toma de decisiones y la habilidad técnica en fútbol / Decision Making and Technical Skills Evolution in Football pp. 467-487. *RIMCAFD*, *59*(2015), 467–487. doi:10.15366/rimcafd2015.59.005

González-Víllora, S., Serra-Olivares, J., Pastor-Vicedo, J. C., & da Costa, I. T. (2015). Review of the tactical evaluation tools for youth players, assessing the tactics in team sports: football. *SpringerPlus*, *4*(1), 663. doi:10.1186/s40064-015-1462-0

Horvath, S. (2011). *Weighted Network Analysis: Applications in Genomics and Systems Biology*. New York: Springer.

Kalamaras, D. (2014). Social Networks Visualizer (SocNetV): Social network analysis and visualization software. *Social Networks Visualizer*. Homepage: http://socnetv.sourceforge.net .

Opsahl, T., Agneessens, F., & Skvoretz, J. (2010). Node centrality in weighted networks: Generalizing degree and shortest paths. *Social Networks*, *32*(3), 245–251. doi:10.1016/j.socnet.2010.03.006

Pavlopoulos, G. A., Secrier, M., Moschopoulos, C. N., Soldatos, T. G., Kossida, S., Aerts, J., … Bagos, P. G. (2011). Using graph theory to analyze biological networks. *BioData Mining*, *4*(1), 10.

Peña, J. L., & Touchette, H. (2012). A network theory analysis of football strategies. In *arXiv preprint arXiv* (p. 1206.6904).

Rubinov, M., & Sporns, O. (2010). Complex network measures of brain connectivity: uses and interpretations. *NeuroImage*, *52*(3), 1059–69. doi:10.1016/j.neuroimage.2009.10.003

Serra-Olivares, J., González-Víllora, S., García-López, L. M., & Araújo, D. (2015). Game-Based Approaches' Pedagogical Principles: Exploring Task Constraints in Youth Soccer. *Journal of Human Kinetics*, *46*(1). doi:10.1515/hukin-2015-0053

Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications*. New York, USA: Cambridge University Press.

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of "small-world" networks. *Nature*, *393*, 440–442.