APPLYING NETWORKS AND GRAPH THEORY TO MATCH ANALYSIS: IDENTIFYING THE GENERAL PROPERTIES OF A GRAPH

Filipe Manuel Clemente, Fernando Manuel Martins and Rui Mendes

Polytechnic Institute of Coimbra, Coimbra College of Education, Department of Education, Portugal

Correspondence: Filipe.clemente5@gmail.com

INTRODUCTION

The structure of football teams can be defined as a collective organization that emerges during the match (Gréhaigne, Bouthier & David, 1997). Thus, more than a sum of individual performances, a team is a whole that can be seen as a complex network (Grund, 2012). To understand the network structure there are some possibilities that comes from graph theory (Pavlopoulos et al., 2011). These possibilities can be defined by network metrics that identify some general properties of the graph and also allows identifying the prominence of each player for the team's network (Wasserman & Faust, 1994). In that sense, this study will apply some general techniques to identify the graph properties of Spain team during FIFA World Cup 2014. The objective of such analysis is to understand the variance of team's properties during the three matches on competition.

METHOD

Sample

Three official matches from Spain in FIFA World Cup 2014 tournament were analysed in this study. A total of 290 adjacency matrices were generated based on the teammates' interactions and then converted in 3 network graphs. A total of 1154 passes were analysed.

Observation and Coding

To perform the network analysis it is necessary generate an adjacency matrix. Such matrix represents the connections between nodes (player) with an adjacency node (teammate). The criteria to define the connection were the pass between players. An adjacency matrix per unit of attack was generated, that corresponds to the first pass after recover the ball until the last before to lose her. After each match, an overall adjacency matrix was build corresponding to the sum of all adjacency matrices per unit of attack. Per each pass between

nodes it was give a code of 1 (one) and for no passes between teammates it was give a code of 0 (zero).

Network Analysis

Once the 3 overall adjacency matrices were generated based on passes between teammates, they were imported into Social Networks Visualizer (SocNetV) to be analysed. Four network metrics were used to analyse the general properties of graphs: i) total links; ii) density; iii) diameter; and iv) clustering coefficient. The total Links measure is the absolute number of the total interactions conducted between teammates during the match. Whereas the total Links metric is an absolute number of interactions from one player to another, the density of the team's network is a relative index that also measures the overall affection between teammates. In graph theory, the density of a (directed) graph is the proportion of the maximum possible lines that are present between nodes. While the first two metrics (total links L and density Δ) focus on the number of links inside a given social network, the *diameter d* of the corresponding graph is related to the distance between nodes. The *diameter* of a graph is the maximum distance (the length of the largest geodesic) between any two connected nodes and is computed by the formula: *diameter = maximaxj d(i,j)*. Finally, The Clustering Coefficient, quantifies how close a node and its neighbours in a graph are to become a clique (a complete sub graph).

RESULTS

Besides the graph properties analysed in this study, it was also collected the regular notational information about the goals scored, ceded, shots and shots on goal. The results can be seen in the following table 1.

TABLE 1
Notational information and values of general graph properties per match

Team	Score	Goals Scored	Goals Ceded	Shots	Shots on Goal	Total Links	Density	Cluster Coefficient	Diameter
Match 1	Lose	1	5	9	6	93	0.845	0.843	2
Match 2	Lose	0	2	15	9	87	0.791	0.793	2
Match 3	Win	3	0	11	8	95	0.864	0.852	2

It is possible to observe the highest values of total links (95), density (0.864) and cluster coefficient (0.852) were found in the match 3 that corresponds to the unique match won by Spain during the tournament.

DISCUSSION

This case study showed that the greatest values of graph properties emerged in the last match. In the case of total links and density such greatest values can represent a great possibility to distribute the ball by the overall connection between all players. In fact, the ability to decentralize the patterns of passes may be a specific characteristic to increase the variability of action and decrease the exposure to the opponent (Gréhaigne et al., 1997). In fact some studies showed that highest density values lead to best team's performances and, in other hand, small values of density and bigger values of centralization lead to worst team's performance (Grund, 2012). No differences were found in diameter. In fact, the diameter represents the smallest path between pair of players. Generally, there are always a shortest path because the players are all connected with their teammates, thus this indicator maybe not the best indicator to be used in the future network studies. Finally, the greatest clustering coefficient was found in the last match, thus revealing the tendency to team acts as one. Lowest values of clustering coefficient suggest that team deal with many clusters within team, thus there are some players that not connect too much with other. Briefly, the general graph properties reveals that the last match corresponds to the best collective performance of the team and, in other hand, the second match may correspond to the worst collective performance of the team by the lowest values found in such game.

This study focused in the general network metrics to find some team's properties during the FIFA World Cup 2014. The results suggested that it is possible to use the graph theory approach as a good solution to perform the match analysis and increase the information available for coaches. Despite of this important contribution, the general properties of the graph can be not enough to understand the cooperation dynamic of the team. Thus, future studies should consider individual centrality metrics to identify the prominent players inside the team to identify how player contributes and influences the general cooperation of the team.

ACKNOWLEDGEMENT

This work was supported by FCT project PEst-OE/EEI/LA0008/2013.

REFERENCES

Gréhaigne, J. F., Bouthier, D., & David, B. (1997) Dynamic-system analysis of opponent relationship in collective actions in football. *Journal of Sports Sciences*, 15, 137-149.

Grund, T. U. (2012) Network structure and team performance: The case of English Premier League soccer teams. *Social Networks*, *34*, 682-690.

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

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