Journal of Physical Education and Sport ® (JPES), 16 Supplement issue 2, Art 162, pp. 1023 - 1028, 2016
online ISSN: 2247 - 806X; p-ISSN: 2247 – 8051; ISSN - L = 2247 - 8051 © JPES

Original Article

Network measures and digraph theory applied to soccer analysis: Midfielder is the key player in youth teams

PATRÍCIA OLIVEIRA[1], FILIPE MANUEL CLEMENTE[1,2], FERNANDO MANUEL LOURENÇO MARTINS[1,3]
[1]Instituto de Telecomunicações, Delegação da Covilhã, PORTUGAL
[2]Instituto Politécnico de Viana do Castelo, Escola Superior de Desporto e Lazer, PORTUGAL
[3]Instituto Politécnico de Coimbra, Escola Superior de Educação, Departamento de Educação, IIA, RoboCorp, UNICID, PORTUGAL

Published online: Octomber 30, 2016
(Accepted for publication September 20, 2016)
DOI:10.7752/jpes.2016.s2162

Abstract
Graph and digraph theories have been used to test the relationships between teammates and the network properties of team sports. Nevertheless, no studies in young soccer teams have been found, as far we know. Therefore, the objective of the study was to apply network measures to identify centrality levels of young soccer players during official matches and analyse the variance between tactical positions and tactical line-ups. Seventy young soccer players from under-10 competitive level were observed during 10 matches. Significant statistical differences were found between players' positions in IDC (p = 0.001; *ES* = 0.090; *minimum effect*); ODC (p = 0.001; *ES* = 0.156; *minimum effect*); and BC (p = 0.001; *ES* = 0.110; *minimum effect*) variables. No significant statistical differences were found between 1-3-2-1 and 1-2-3-1 line-ups for %IDC (p = 0.113; *ES* = 0.056; *minimum effect*), %ODC (p = 0.126; *ES* = 0.048; *minimum effect*) and %BC (p = 0.204; *ES* = 0.035; *minimum effect*). This study found that midfielder is the key position on the field, being a linkage player to attacking building.
Keywords: applied mathematics; graph theory; social network analysis; performance analysis; football.

Introduction

Match analysis requires serious information about the team's dynamics (Vilar, Araújo, Davids, & Button, 2012). The dynamic can be observed in different scales and levels from sub-phases until the overall spectrum of cooperation/opposition in the full match (Travassos, Davids, Araújo, & Esteves, 2013). In the specific case of soccer there are two main relationships that can be studied: the teammates network and the rapport of strength against the opponents (Gréhaigne, Bouthier, & David, 1997). Network have been the most studied in the literature, as far we know (Cotta, Mora, Merelo, & Merelo-Molina, 2013; Grund, 2012; Peña & Touchette, 2012). The ability to play as a unit and to cooperate to achieve the collective goals of the team is one of the main reasons that may justify the research performed in network in the last few years (Clemente, Martins, & Mendes, 2016).

Generally, the network structure of soccer teams have been researched by using three different methods: i) temporal pattern analysis (Bloomfield, Jonsson, Houlahan, & Donoghue, 2005); ii) neural networks (Dutt-Mazumder, Button, Robins, & Bartlett, 2011); and iii) social network analysis based on graph and digraph theories (Grund, 2012; Pedro Passos et al., 2011). The temporal analysis focus the analysis in specific events that conduct to a perception about patterns of interactions, specifically attacking buildings (Bloomfield et al., 2005). The neural network analysis seek to estimate the interactions by using adaptive techniques (Memmert & Perl, 2009). Both approaches brings interesting information, nevertheless are too complex to represent specific and easy-to-use information for coaches, so far. In the other hand, social network analysis based on graph and digraph theories represent a user-friendly method to analysts (Clemente, Martins, Kalamaras, Wong, & Mendes, 2015; Malta & Travassos, 2014).

The research conducted in elite soccer using centrality measures (indegree prestige, degree centrality, betweenness centrality) have been found that some specific positions may have higher prominence levels during attacking building (indirect attack) (Clemente, Martins, Wong, Kalamaras, & Mendes, 2015; Duch, Waitzman, & Amaral, 2010; Peña & Touchette, 2012). In the case of indirect style, midfielders and forwards are the prominent players in specific teams (Malta & Travassos, 2014). Such information has been emphasizing the importance of these key positions during the attacking building of elite teams.

Despite of the available information in elite players, no available information in young players has been found. This is must be a concern of researchers, mainly to identify if specific positions may influence the time of ball contact and the number of actions, thus promoting a not desirable early specialization. Based on that, this study aimed to identify the prominence levels of specific positions in young soccer teams by using network measures. The influence of specific tactical line-ups in prominence levels of players was also analysed.

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

Methods

Participants

Seventy young soccer players (11.32 ± 0.87 years old; 3.41 ± 0.67 years of practice) from six under-12 teams were analysed in this study. All the participants were analysed in ten games with a minimum of 25 minutes of play per player per game. This is was the criteria to guarantee similar conditions of observation and similarities in the analysis. A total of 60 games and 7,861 passes between teammates were observed. Parents of each young player signed a free consent for the voluntary participation of children.

Procedures of observation

The ten analysed games per team were observed in the home matches to guarantee similar conditions of competition. The observation protocol was focused on the attacking building, thus all the sequences of passes (units of attack) were collected and then processed. The pass of the ball was classified as the linkage factor to build the adjacency matrices that were the basis to process the network treatment.

Players were classified following the techno-tactical assignment to positional roles: goalkeeper (GK); external defenders (ED); central defenders (CD); central midfielders (CM); and forwards (FW). The most used tactical line-up of the teams were also observed and codified for this study. In this competitive level (U12) the official games occurs in 7 vs. 7 format. For that reason, the following common tactical line-ups were identified and codified: i) 1-3-2-1; ii) 1-2-3-1. The observation and classification of positions and tactical line-ups were made by the same researcher with more than five years of experience in match analysis. A test and re-test procedure with 20 days interval was performed to guarantee the reliability of the classifications. A Kappa value of 0.88 was obtained, thus ensuring a standard margin for this type of procedures (Robinson & O'Donoghue, 2007).

Data collecting

Pass between teammates during attacking building was defined as the linkage criteria for the data collection. Individual adjacency matrix was built based on the sequences of passes between teammates with more than three passes and without any intervention of opponents. This follow the criteria of previous studies (Clemente et al., 2015; Malta & Travassos, 2014; Passos et al., 2011). In the end of each match a sum weighted adjacency matrix was generated representing the data from the full match. A total of 60 weighted adjacency matrices were collected from the 60 analysed matches.

Network Analysis

Centrality measures based on digraph theory were computed in the SocNetV software (version 1.9.) (Kalamaras, 2014). Three centrality measures were calculated per each player/position: i) indegree centrality; ii) outdegree centrality; iii) betweenness centrality. Following it is demonstrated the algorithms of each one.

OutDegree Centrality (ODC)

Let n_i be a vertex of weighted digraph *G* with *n* vertices. The standardized degree centrality index, , is the fraction of the weight of vertices that are adjacent to n_i, and is calculated as (Opsahl, Agneessens, & Skvoretz, 2010):

C'(D-out)wni=kiw-outi=1nj=1j≠inaij, (1)

where $\mathbf{k}_i^{w-out}$ is the degree centrality index of the vertex n_i and a_{ij} are elements of the weighted adjacency matrix of a *G*.

ODC index varies from 0 to 100% in which 100% represents the highest value of prominence in to pass the ball and 0% suggest the non-existence of prominence in to built the attack. This measure help coaches to identify the overall activity of each player in to start the attacking building (Clemente et al., 2016).

Betweenness Centrality (BC)

The standardized betweeness centrality (BC) index is calculated by (Rubinov & Sporns, 2010):

$$C'_b(n_k) = \frac{1}{(n-1)(n-2)} \sum_{\substack{n_i n_j \in V \\ i \neq n j \neq k}} \frac{g_{ij}(n_k)}{g_{ij}}, \qquad (2)$$

where $g_{ij}(n_k)$ is the number of shortest paths between n_i and n_j that pass trought n_k and g_{ij} is the number of shortest paths between n_i and n_j (Clemente et al., 2016).

Briefly, the BC score of each player can be explained as a measure of the relative control that player has on other players (Clemente et al., 2016). Players with great levels of BC can be considered the linkage athletes of a team, thus representing the key players that may unify the attacking process and to involve the teammates.

InDegree Centrality (IDC)

Let n_i be a vertex of weighted digraph *G* with *n* vertices. The standardized degree prestige index, $P'^{w}_{(D-in)}(n_i)$, is the proportion of the weight of vertices that are adjacent to n_i, and is calculated as (Opsahl et al., 2010):

P'(D-in)wni=kiw-ini=1nj=1j≠inaij, (3)

where k_i^{w-in} is the degree prestige index of the vertex n_i and a_{ij} are elements of the weighted adjacency matrix of a *G* (Clemente et al., 2016).

IDC index represent the level of prestige of each player. Greatest values of IDC represent the great capacity to receive more passes for the teammates, thus being considered the key player to pass the ball.

Statistical Procedures

Positions of the field (GK, ED, CD, CM and FW) and tactical line-ups (1-3-2-1 and 1-2-3-1) were classified as factors. The centrality measures of ODC, IDC and BC were the dependent variables. A two-way MANOVA was tested for both factors after the confirmation of normality and homogeneity assumptions. A two-way ANOVA per dependent variable was computed in the case of interactions. The one-way ANOVA was tested per factor. Tukey HSD was used to verify differences between factors. Effect size (ES) was tested and interpreted using the follow criteria (Ferguson, 2009): no effect (η^2 < 0.04), minimum effect (0.04 < η^2 < 0.25), moderate effect (0.25 < η^2 < 0.64) and strong effect (η^2 > 0.64). SPSS software (version 23.0, Chicago, Illinois, USA) was used to compute the statistical procedures. A statistical significance of 5% was defined.

Results

The two-way MANOVA revealed that players' positions had significant main effects on network centrality measures (p = 0.001; *ES* = 0.187; *minimum effect*). No statistical differences were found between tactical line-ups (p = 0.093; *ES* = 0.076; *minimum effect*). There was no significant interaction between players' positions and tactical line-ups (p = 0.101; *ES* = 0.056; *minimum effect*).

The one-way ANOVA tested the analysis of variance between players' positions. Significant differences were found between positions in IDC (p = 0.001; *ES* = 0.090; *minimum effect*); ODC (p = 0.001; *ES* = 0.156; *minimum effect*); and BC (p = 0.001; *ES* = 0.110; *minimum effect*) variables. The following figure 1 shows the descriptive statistics for the %ODC variable.

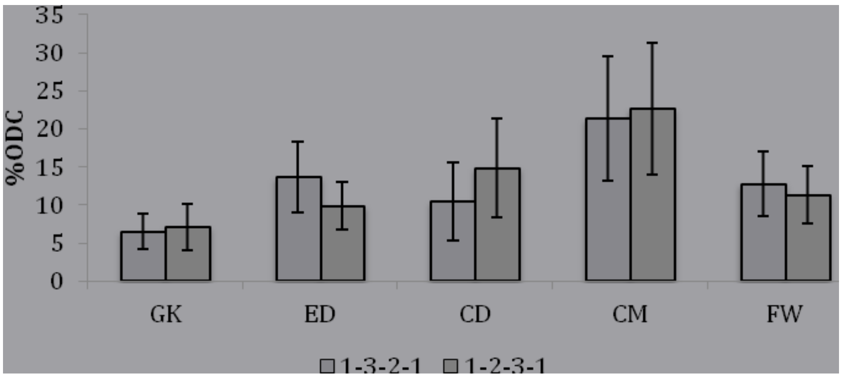

Figure 1. Mean and standard deviation of %ODC.

Post-hoc analysis for the %ODC variable revealed that CM position were statistical different from GK (p = 0.001), ED (p = 0.005), CD (p = 0.001) and FW (p = 0.001). CM was the position with highest values of %ODC. GK was also significant different from ED (p = 0.010), CD (p = 0.040) and FW (p = 0.035). No significant differences were found between ED, CD and FW in %ODC. Figure 2 shows the results for the %IDC variable.

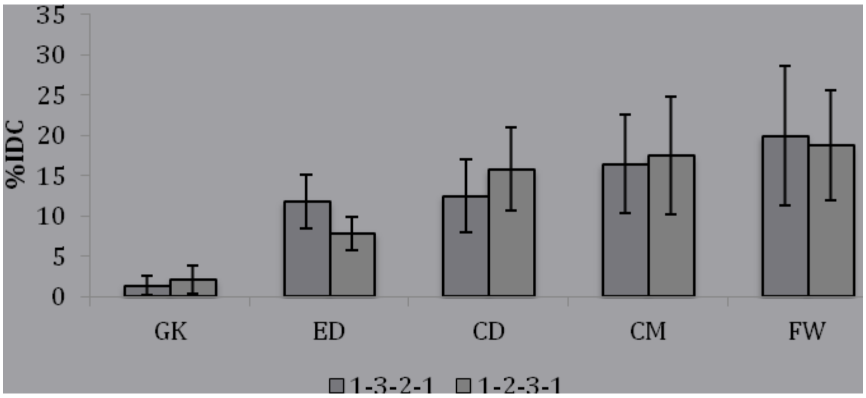

Figure 2. Mean and standard deviation of %IDC.

As possible to observe in Figure 2, there were statistical differences between GK and ED (*p* = 0.001), CD (*p* = 0.001), CM (*p* = 0.001) and FW (*p* = 0.001). FW and CM had also statistical greater values than ED (*p* = 0.040 and *p* = 0.050, respectively). No significant statistical differences were found between CD, CM and FW. Figure 3 reveals the statistical data for %BC variable.

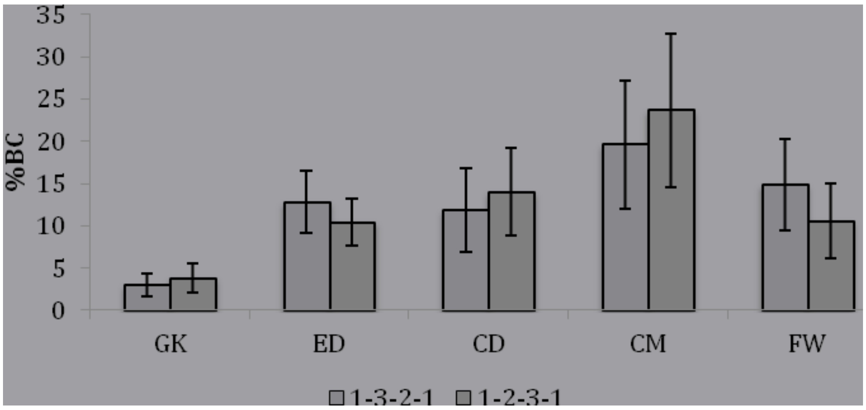

Figure 3. Mean and standard deviation of %BC.

CM position had significant greater values of %BC than GK (p = 0.001), ED (p = 0.001), CD (p = 0.005) and FW (p = 0.001). GK had also significant lower values of %BC than ED (p = 0.001), CD (p = 0.001) and FW (p = 0.001). No significant differences were found between ED, CD and FW.

No significant statistical differences were found between 1-3-2-1 and 1-2-3-1 line-ups for %IDC (*p* = 0.113; *ES* = 0.056; *minimum effect*), %ODC (*p* = 0.126; *ES* = 0.048; *minimum effect*) and %BC (*p* = 0.204; *ES* = 0.035; *minimum effect*).

Discussion

This study analysed the variance of centrality levels between tactical positions and tactical line-ups in young soccer teams. The main results revealed that the position of the field had significant statistical effects on the prominence levels of players. Previous studies in elite soccer has revealed that independently from tactical line-up, external defenders and central midfielders tend to be the most prominent players during attacking building (indirect play style) (Clemente et al., 2015; Peña & Touchette, 2012). Although without significant interaction, it was possible to verify that external defenders were most relevant than central defenders in the 1-3-2-1 line-up than in 1-2-3-1 in which occurred the inverse. This can be justified by the fact that in 1-3-2-1 the build up may start for the wings, following the principle of play of width and length (Costa, Garganta, Greco, Mesquita, & Seabra, 2010). In the other hand, in tactical line-up of 1-2-3-1 the beginning of passing sequences must be more focused in the two central defenders, thus increasing the prominence in to pass the ball and built the attacking process.

The analysis of variance made in IDC variable revealed that forwards were the most prominent players, independently from the tactical line-up. Some studies that only analysed the transition attacks (direct style of play) revealed that forwards tend to have greater values of IDC (Malta & Travassos, 2014). Nevertheless, small values of forwards in IDC have been found in the studies that analysed indirect attacking (Clemente et al., 2015). The results found in our study may suggest that in smaller formats (7 vs. 7) forwards may have a greater importance, mainly considering that is the unique player in the attacking zone. This may justify the great values

----
found. Despite of the greater average values, no significant differences were found with the outfield players. This may suggest that all teammates have greater participation in to receive the ball during attacking process.

Statistical greater values of betweenness centrality were found in midfielders. Results suggest that this strategic position lead to a linkage mission, thus unifying the sectors of play and building the dynamic of cooperation among teammates (Peña & Touchette, 2012). Similar evidences were found in previous studies conducted in elite soccer players (Clemente et al., 2015). The specific location of midfielders brings the responsibility to receive the ball from the defence and link to the attack. For that reason, it is justified the greater values obtained from these players.

This study had some limitations. All the units of attack were observed. This can be a limitation based on the fact that no accuracy criterion was adopted (such as units of attack that resulted in goal, shots or approach to score zone). Moreover, all the passing sequences were considered and no differentiation was made between direct and indirect style of attack. This may influence the results obtained in this study. Future studies must consider classifying the units of attack and comparing the successful and unsuccessful attacking buildings.

Despite the limitations, this study revealed that some strategic positions might influence the prominence and ball contacts of players. This is extremely important considering that players may have opportunities to develop their skills in game in these competitive levels. The early specialization in some tactical positions may lead to a gap of opportunities to develop specific skills in game and also during the training sessions. Future studies should consider analyse the long-term effect of early tactical specialization in the skill level of players.

Conclusion

This study aimed to analyse the prominence of young soccer players by using network measures. The main results revealed that midfielder is the prominent player in to link the teammates during the attacking process. Forwards were the key players to pass. No statistical differences were found between different tactical line-ups. The results of this study may suggest that specific tactical positions lead to greater prominence and may influence the volume of play and technical specialization of players.

Acknowledgments

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

References

Bloomfield, J., Jonsson, G. K., Houlahan, K., & Donoghue, P. O. (2005). Temporal Pattern Analysis and its Applicability in Soccer. In L. Anolli, S. Duncan, M. S. Magnusson, & G. Riva (Eds.), *The hidden stucture of interaction: From neurons to culture patterns* (pp. 237–251). Amsterdam, Netherlands: IOS Press.

Clemente, F. M., Martins, F. M. L., Kalamaras, D., Wong, D. P., & Mendes, R. S. (2015). General network analysis of national soccer teams in FIFA World Cup 2014. *International Journal of Performance Analysis in Sport*, *15*(1), 80–96.

Clemente, F. M., Martins, F. M. L., & Mendes, R. S. (2016). *Social Network Analysis Applied to Team Sports Analysis*. Netherlands: Springer International Publishing.

Clemente, F. M., Martins, F. M. L., Wong, D. P., Kalamaras, D., & Mendes, R. S. (2015). 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*(2), 704–722.

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.

Cotta, C., Mora, A. M., Merelo, J. J., & Merelo-Molina, C. (2013). A network analysis of the 2010 FIFA world cup champion team play. *Journal of Systems Science and Complexity*, *26*(1), 21–42.

Di Salvo, V., Baron, R., Tschan, H., Calderon Montero, F. J., Bachl, N., & Pigozzi, F. (2007). Performance characteristics according to playing position in elite soccer. *Int J Sports Med*, *28*, 222–227.

Duch, J., Waitzman, J. S., & Amaral, L. A. (2010). Quantifying the performance of individual players in a team activity. *PloS One*, *5*(6), e10937.

Dutt-Mazumder, A., Button, C., Robins, A., & Bartlett, R. (2011). Neural network modelling and dynamical system theory: are they relevant to study the governing dynamics of association football players? *Sports Medicine*, *41*(12), 1003–17.

Ferguson, C. J. (2009). An effect size primer: A guide for clinicians and researchers. *Professional Psychology: Research and Practice*, *40*(5), 532–538.

Gréhaigne, J. F., Bouthier, D., & David, B. (1997). Dynamic-system analysis of opponent relationship in collective actions in football. *Journal of Sports Sciences*, *15*(2), 137–149.

Grund, T. U. (2012). Network structure and team performance: The case of English Premier League soccer teams. *Social Networks*, *34*(4), 682–690.

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

Malta, P., & Travassos, B. (2014). Characterization of the defense-attack transition of a soccer team.

Motricidade, *10*(1), 27–37.
Memmert, D., & Perl, J. (2009). Game creativity analysis using neural networks. *Journal of Sports Sciences*, *27*(2), 139–149.
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., Milho, J., Fonseca, S., Borges, J., Araújo, D., & Davids, K. (2011). Tendencies of Agents in Team Sports . *Journal of Motor Behavior*, *43*(2), 155–163.
Peña, J. L., & Touchette, H. (2012). A network theory analysis of football strategies. In *arXiv preprint arXiv* (p. 1206.6904).
Robinson, G., & O'Donoghue, P. (2007). A weighted kappa statistic for reliability testing in performance analysis of sport. *International Journal of Performance Analysis in Sport*, *7*(1), 12–19.
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
Travassos, B., Davids, K., Araújo, D., & Esteves, P. T. (2013). Performance analysis in team sports□: Advances from an Ecological Dynamics approach. *International Journal of Performance Analysis in Sport*, *13*(1), 83–95.
Vilar, L., Araújo, D., Davids, K., & Button, C. (2012). The Role of Ecological Dynamics in Analysing Performance in Team Sports. *Sports Medicine*, *42*(1), 1–10.