Who are The Prominent Players in the UEFA Champions League? An Approach Based on Network Analysis

Filipe Manuel CLEMENTE1,2,* and Fernando Manuel Lourenço MARTINS2,3

1Instituto Politécnico de Viana do Castelo, Escola Superior de Desporto e Lazer, Portugal
2Instituto de Telecomunicações, Delegação da Covilhã, Portugal
3Instituto Politécnico de Coimbra, Escola Superior de Educação, Departamento de Educação, IIA, RoboCorp, UNICID, Portugal

(*Corresponding author’s e-mail: filipe.clemente5@gmail.com)

Received: 5 January 2017, Revised: 8 May 2017, Accepted: 10 May 2017

Abstract

This study aimed to analyze the centrality levels of elite football players. Tactical positions and tactical line-ups were considered factors to be used in analyzing the variance in the prominence of players, measured by social network measures. The best 16 teams from the UEFA Champions league were analyzed during the entire competition. A total of 109 matches were analyzed for this study. Significant statistical differences between positions were found in % indegree (p = 0.001; ES = 0.268, moderate effect), % outdegree (p = 0.001; ES = 0.301, moderate effect) and % betweenness (p = 0.001; ES = 0.114, minimum effect). No statistical differences between tactical line-ups in % outdegree (p = 1.000; ES = 0.001, no effect) or % indegree (p = 1.000; ES = 0.001, no effect) were found. Central midfielders had the greatest values of centrality, thus confirming their importance in the linkage process of the team. Position had great influence on the centrality levels of players.

Keywords: Applied mathematics, graph theory, soccer, football, match analysis

Introduction

Cooperation-opposition relationships may characterize the dynamics of football [1]. Contextual variables of the game constrain the occurrence of specific relationships; thus, the game can be characterized as a dynamic system with some degree of chaos [2]. For that reason, unpredictability is evidenced in the game; nevertheless, training in the sport aims to stabilize some patterns of interaction (strategy and tactics) to guarantee the best performance in different scenarios [3,4]. These patterns of play depend on the capacity of the individuals to act as a team, following specific principles of play and organization [5,6]. Based on that, a team can be classified as a network structure, constituting the players [7].

Observing, analyzing, and understanding the properties of the network are crucial tasks of match analysis [8-10]. In the specific case of the game of football, the identification of certain patterns of play will help coaches to plan training sessions and make decisions about which strategies to use. Therefore, match analysts must use different techniques to guarantee that the appropriate information will be collected, processed, and used [11].

The last decade has seen the increasing application of different quantitative and qualitative methods in match analysis [4,12,13]. Classical notational analysis has been updated with some mathematical approaches in order to change the ways in which analysts can identify the performance of a team [14]. Temporal pattern analysis [3], neural networks [15], and social network analysis [9,16] are current forms of notational analysis that guarantee that quantification can explain some occurrences of the game. Particularly, social network analysis has particularly been growing over the last 5 years, with the
application of graph and digraph theories used in the analysis of the attacking process [17-19] and to characterize the patterns of play of a team [20].

Social network analysis holds different measures that make it possible to identify general, co-dependent, and centrality properties of a network [8,21]. The centrality properties classify the prominence levels of players in the group and, for that reason, has been used in recent studies that applied network analysis to football [17-19]. One of the first network studies, which was conducted during the European Cup 2008 tournament, revealed that the network measures used to classify the most valuable and prominent players had strong associations with the point-of-view of expert coaches and analysts [18].

The most recent studies in network analysis have identified the most relevant players (positions) during attacking build-up play [17,22]. A study that analyzed attacking transition (direct style) verified that the most prominent players were the midfielders and the forwards [22]. Another study that analyzed all the matches of the FIFA World Cup 2014 verified that external defenders and, mainly, midfielders were the most prominent players during attacking build-up play (in indirect and direct style) [17].

National teams have been the main focus of network analysis in football [18,23]. No study of elite soccer teams in major international competitions has been conducted so far. Moreover, the majority of the studies have used small samples, thus not allowing the results to be generalized. For that reason, our motivation was to analyze the elite football teams from Europe that reached the latter stages of the world-famous UEFA Champions League.

Based on the above-mentioned reasons and motivations, the aim of this study was to identify the centrality levels of players during attacking build-up play during the UEFA Champions League 2015 - 2016. Particularly, our aim was to analyze the variance between players’ positions and the tactical line-ups of teams through different centrality measures that characterize the prominence of players during passing sequences. We hypothesize that midfielders will have the highest centralities in the analysis of variance between playing positions, independently from the team’s formation.

Materials and methods

Sampling

A total of 109 matches from the UEFA Champions League 2015 - 2016 were analyzed and codified in this study. The 16 best teams (which reached at least the 16th round) were analyzed in all the matches during the competition. Eight teams were analyzed over 8 matches, 4 teams over 10 matches, 2 teams over twelve matches, and 2 teams over thirteen matches (the finalists). An individual adjacency matrix per team and per match was generated; thus, a total of 109 adjacency matrices were used to compute the network measures. This study followed the ethical recommendations of the Declaration of Helsinki for human study.

Data collecting and processing

Players were codified by their positions on the basis of the tactical line-up and variations over the course of the game, as done in previous studies [24]. Four expert analysts made codifications of the positions. The reliability of the codification was tested with a 20-day interval test with 25% of the full data. Cohen’s Kappa test was executed to compare the test and the re-test. A Kappa value of 0.86 was obtained, thus ensuring enough reliability of the data for this kind of procedure [25]. The following tactical line-ups were codified: i) 1-4-3-3; ii) 1-4-4-2 classical; iii) 1-3-5-2; iv) 1-5-3-2; v) 1-4-4-2 diamond; vi) 1-4-3-2-1; vii) 1-4-2-3-1; viii) 1-4-5-1, and ix) 1-3-4-3.

Players’ positions were attributed based on the tactical line-ups. A techno-tactical assignment was adopted for the positional roles [26]. Figure 1 illustrates the space occupied by the players.
Based on Figure 1, the following positions were codified: i) goalkeeper (GK); ii) external defenders (ED); iii) central defenders (CD); iv) central midfielders (MF); v) external midfielders (EM), and vi) forwards (FW).

The linkage indicators of the network were the passes performed between teammates. The direction of passes (player A to player B is different from player B to player A) and the volume were considered in this study and, for that reason, the network measures were calculated for weighted digraphs.

The data used in this study were extracted from the official website of the UEFA Champions League (http://www.uefa.com/uefachampionsleague). The adjacency matrices of the overall passes were obtained per match and per team. The adjacency matrix identified the nodes of a graph (the players) and the arrows (in this case, the volume of passes made from one player to another). Only the positions were considered. Players were classified with a code based on their position.

Computing the centrality measures
Each adjacency matrix (per each match and per each team) was imported in Social Networks Visualizer (SocNetV, version 1.9), a software that allows visualization and analysis of social networks [27]. Three widely known actor centrality measures were used in this study: i) outdegree centrality; ii) indegree centrality; and iii) betweenness centrality. The following subsections briefly describe the interpretation and the meaning of these measures.

Outdegree centrality
The outdegree centrality (ODC) indicates the level of activity of a player in building the passing sequences. Please consider that \( n_i \) is the vertex of weighted digraph \( G \) with \( n \) vertices. The standardized degree centrality index, \( C^w_{\text{out}}(n_i) \), is the proportion of the weight of vertices that are adjacent to \( n_i \), and can be calculated as follows [28];

\[
ODC^{w}_{\text{out}}(n_i) = \frac{k^w_{\text{out}}(n_i)}{\sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}},
\]

(1)
where \( k^w_{\text{out}} \) is the degree centrality index of the vertex \( n_i \) and \( a_{ij} \) are elements of the weighted adjacency matrix of \( G \) [16]. Greater values of outdegree centrality indicate the major prominence of a position in building the passing sequences and connecting with the teammates [16]. Values of this measure are given in relative frequency from 0 to 100%.

**Indegree centrality**

The indegree centrality (IDC) measures the prestige of a player receiving the ball from their teammates. Consider \( n_i \) as the vertex of weighted digraph \( G \) with \( n \) vertices. The standardized degree prestige index, \( P'_w(\text{D-in})_w(n_i) \), can be considered as the proportion of the weight of vertices that are adjacent to \( n_i \), and is calculated as [28];

\[
IDC^w_{(\text{D-in})}(n_i) = \frac{k^w_{\text{in}}}{\sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}}
\]

where \( k^w_{\text{in}} \) is the degree prestige index of the vertex \( n_i \) and \( a_{ij} \) are elements of the weighted adjacency matrix of \( G \) [16]. Similar to ODC, the IDC\(^w\) is a relative frequency that varies from 0 to 100%. Greater values suggest a greater level of prestige of the player receiving the ball from their teammates. It can be also suggested that this player may be the star or the playmaker of the group [16].

**Betweenness centrality**

The betweenness centrality (BC) measures how often a player is situated between other teammates. For a graph \( G = (V,E) \), with \( n_i, n_j, n_k \in V, i, j, k = 1, \ldots, n \), the standardized betweenness centrality index can be calculated as follows [29];

\[
\mathcal{C}'_b(n_k) = \frac{1}{(n-1)(n-2)} \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{i \neq j \neq k} g_{ij}(n_k) / g_{ij}
\]

where \( g_{ij}(n_k) \) is the number of shortest paths between \( n_i \) and \( n_j \) that pass through \( n_k \) and \( g_{ij} \) is the number of shortest paths between \( n_i \) and \( n_j \) [16]. The index that lies between 0 and 100 % represents the capacity of a player to act as the mediator or the link of the team. Greater values of BC suggest that this player may act as an important key player in linking the sectors of a team and the teammates [16].

**Statistical procedures**

Tactical line-ups and players’ positions were classified as factors (independent variables) of this study. The centrality measures of %ODC, %IDC, and %BC were the dependent variables. The 2-way MANOVA was executed after confirmation of the normality and homogeneity assumptions. The 2-way ANOVA was tested for each dependent variable in the cases of statistical interactions between independent variables. The one-way ANOVA was also tested for each independent variable, followed by the Tukey HSD for the post-hoc procedure. The effect size (ES) was tested and interpreted using the following criteria [30]: no effect (ES < 0.04), minimum effect (0.04 < ES < 0.25), moderate effect (0.25 < ES < 0.64), and strong effect (ES > 0.64). The statistical procedures were made in SPSS software (version 23.0, Chicago, Illinois, USA) for a statistical significance of 5%.

**Results and discussion**

The 2-way MANOVA tested the interactions between tactical line-ups and positions on the field. Significant statistical differences were found for the composite of network centralities in tactical line-ups \((p = 0.001; ES = 0.010, \text{no effect})\) and positions \((p = 0.001; ES = 0.056, \text{minimum effect})\) factors.
Statistical interactions were also found between factors for the composite of network variables (Pillai’s Trace = 0.090; $F = 1.597; p = 0.001; ES = 0.030$, no effect).

The 2-way ANOVA that tested the interaction between factors per variable revealed statistical differences in % outdegree centrality (ODC) ($p = 0.001; ES = 0.038$, no effect) and % indegree centrality (IDC) ($p = 0.001; ES = 0.043$, minimum effect). No significant differences were found in the % betweenness centrality (BC) ($p = 0.603; ES = 0.017$, no effect).

The one-way ANOVA test was carried out for the positions factor. The descriptive values and the effect size can be observed in the following Table 1.

Table 1 Descriptive statistics (Mean±SD, and CI95 %) of network variables between positions.

<table>
<thead>
<tr>
<th></th>
<th>GK</th>
<th>ED</th>
<th>CD</th>
<th>MF</th>
<th>EM</th>
<th>FW</th>
<th>$p$</th>
<th>ES</th>
</tr>
</thead>
<tbody>
<tr>
<td>%ODC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b-e</td>
<td>5.47±2.44</td>
<td>9.23±2.65</td>
<td>11.05±3.59</td>
<td>11.34±4.07</td>
<td>7.66±3.54</td>
<td>5.57±2.87</td>
<td>0.301</td>
<td></td>
</tr>
<tr>
<td>b-e</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b-f</td>
<td>0.001</td>
<td>0.301</td>
<td>moderate effect</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>%BC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b-f</td>
<td>0.85±0.84</td>
<td>2.83±2.02</td>
<td>3.27±2.43</td>
<td>3.39±2.28</td>
<td>2.43±2.17</td>
<td>1.96±1.87</td>
<td>0.114</td>
<td></td>
</tr>
<tr>
<td>b-f</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b-f</td>
<td>0.001</td>
<td>0.268</td>
<td>moderate effect</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>%IDC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b-f</td>
<td>3.42±2.08</td>
<td>9.47±2.56</td>
<td>9.44±3.56</td>
<td>10.91±3.59</td>
<td>9.41±3.68</td>
<td>7.86±3.53</td>
<td>0.268</td>
<td></td>
</tr>
<tr>
<td>b-f</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b-f</td>
<td>0.001</td>
<td>0.027</td>
<td>no effect</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Central defenders and central midfielders were the positions with greater values of %ODC (11.05 and 11.34, respectively). On the other hand, GK and FW had the smallest values of %ODC (5.47 and 5.57, respectively). Excluding the GK, central midfielders had almost 104 % more prominence than FW. The analysis of the %BC also revealed that MF and CD were the positions with greater values (3.39 and 3.27, respectively) and GK and FW had the smallest values (0.85 and 1.96, respectively). Three clusters may be observed in this variable (the greater values - MF and CD, the medium values - ED and EM, and the smallest values - GK and FW). Finally, the analysis of the %IDC revealed that MC was the position with greatest value (10.91) and the smallest was found in GK (3.42).

The one-way ANOVA tested the analysis of the variance of the centralit y measures between tactical line-ups. There were no statistical differences in %ODC ($p = 1.000; ES = 0.001$, no effect) or %IDC ($p = 1.000; ES = 0.001$, no effect). Statistical differences were found in %BC ($p = 0.001; ES = 0.027$, no effect) between the 1-4-3-3 and 1-4-4-2 ($p = 0.001; 2.39$ and 3.35, respectively) and 1-4-2-3-1 ($p = 0.37; 2.39$ and 2.80, respectively) and the 1-4-4-2 and 1-4-2-3-1 ($p = 0.19; 3.35$ and 2.80, respectively).

The average of the network measures per position in different tactical line-ups can be observed in Figures 2 - 4.
Figure 2 Average of %ODC per position in different tactical line-ups.

Despite there being no differences in the network measures between tactical line-ups, it is possible to verify in Figure 2 that midfielder (MF) was the most prominent position in 1-4-3-3, 1-4-4-2, 1-5-3-2, and 1-3-4-3. In the remaining tactical line-ups, the most prominent position with %ODC was the CD.

Figure 3 Average of %BC per positions in different tactical line-ups.
Statistical differences in %BC between tactical line-ups were found, as reported earlier. In Figure 3 it is possible to verify that the greater values of %BC were obtained in the 1-4-4-2 formation. It is also possible to observe that the MF was the most prominent player in the 1-4-3-3, 1-4-4-2, and 1-5-3-2 formations. In the remaining tactical line-ups, the most prominent position was the DC.

![Figure 3](image)

**Figure 3** Average of %BC between tactical line-ups.

No statistical differences of %IDC were found between tactical line-ups. Despite that, it is possible to verify that the MC was the most prominent player in the 1-4-3-3, 1-3-5-2, 1-5-3-2, 1-4-2-3-1, and 1-3-4-3 formations. The external midfielder was the most prominent player in the 1-4-4-2 formation, and the DC was the most prominent in the remaining formations.

**Discussion**

The best 16 teams of the UEFA Champions League 2015 - 2016 were analyzed in this study. The centrality levels of players were measured by network measures, and the variation between tactical line-ups and positions were tested. The main results confirmed that midfielders were the most prominent players, as reported by previous studies on national teams. Nevertheless, centrality levels only varied between tactical line-ups in the specific case of betweenness centrality. The outcomes suggest that position may be a determinant of the centrality levels of each player.

This study focused on the passing network during attacking build-up plays and attacks. Previous studies used social network analysis to determine some patterns of interactions based on the pass [19,31]. The majority of the studies that tested the network levels of players were conducted in the context of national teams (FIFA World Cup and European Cup) [17,18,23]. Our study has focused on the world-famous UEFA Champions League and tested the different centrality levels of players. Our results revealed that midfielders and central defenders were the most prominent players in passing the ball (ODC measure). Previous studies revealed that central midfielders were the most prominent players in the build-up play of an attack during the FIFA World Cup 2014 [17] and European Cup 2008 [18]. The prominence of midfielders can be seen in the indirect style, and also in transitions or counter-attacks [22]. Interestingly, central defenders were the second most prominent players at the ODC level. This is not in line with previous studies, which showed a higher level of participation of external defenders [17,19]. This may be justified by 2 possible reasons. The specific properties of these teams may justify the use of central defenders to begin the attacking build-up play, based on the tactical line-up and model of play.
External defenders that acted in 1-3-5-2 had smaller ODC values than teams that used 4 defenders [17]. The proximity of players to the first third of the field may justify the predictability of having more possession of the ball. Another reason to justify the greater levels of central defenders in comparison with external defenders may be the phase of the game. This study included all of the passing sequences. Possession of the ball (indirect style) and transitions/counter-attacks (direct style) were included and not differentiated. The style of play may have constrained the participation level of each position in the passing sequences of the teams [22].

The prominence of each position with respect to receiving the ball from players was also studied. Our results revealed that, once again, central midfielders had the highest values. Nevertheless, external defenders, central defenders, and external midfielders had similar values, constituting a cluster. Evidence of the high impact of central midfielders in the team has been reported by previous studies [18,19]. They have the responsibility to link the sectors of the team and to create the attacking plays [32]. This may justify the great volume of passes to these players. Their great prominence in the process of linking the play can be verified by the fact that they have the highest level of betweenness centrality. Central midfielders were the most prominent players in the measure of betweenness, thus justifying the capacity to link teammates. The cluster constituted by external defenders, central defenders, and external midfielders revealed that these players have a high and similar importance in the overall connectivity of the team. Once again, these results may not clarify some specificities of the game. The prominence of external midfielders may be smaller than the forwards, or the inverse, in specific models of play. Future studies must consider conducting an in-depth analysis of the tactical models of the teams.

Forwards and goalkeepers had the smallest values in all centrality measures. The results with respect to goalkeepers are understandable. In the case of forwards, the results may reflect the specificity of the position. A forward (a striker in the majority of the cases) is the last man of the team, and does not participate in the attacking build-up play (mainly in the first phase). For that reason, in the majority of cases, they had the smallest values in the field players’ category [17]. Nevertheless, in specific analysis of the transitions, it was possible to see greater prominence [22]. Moreover, the prominence of the forwards was greater than the majority of teammates in the analysis that only analyzed attacks that resulted in goals [33]. For that reason, the prominence levels must be carefully interpreted.

This study had some limitations. The specificity of the attacking process (indirect style, transition, or counter-attack) was not considered. This limitation may mean that certain specific aspects of the interactions between teammates were not disclosed. Moreover, the specificity of the tactical behavior and model of play was also not considered. Future studies must consider crossing information concerning tactical behavior with social network measures.

Despite these limitations, this study confirmed the high participation of central midfielders in the attacking build-up play of elite teams in the UEFA Champions League. Moreover, this was the first study that analyzed a great club tournament in Europe and, for this reason, it can provide a starting point for future studies. As for practical implications, it is possible to observe that social network analysis may help coaches to identify the key players in the attacking process and to make decisions during the match or in the training sessions to optimize the team’s behavior or to avoid the strengths of the opponents. Future studies must consider crossing the network analysis with spatio-temporal measures to identify the causes for the prominence levels. Moreover, some new approaches, based on relative age effect or others, can be considered to identify the relevance of players in youth teams [34].

**Conclusions**

This study revealed that central midfielders are the most prominent players during the build-up of attacking plays among the most successful teams of the UEFA Champions League. It was also found that central defenders were more prominent in the build-up play of an attack than external defenders or external midfielders. Goalkeepers and forwards were the positions with the smallest centrality values in the team. The tactical line-up had a very small influence on the centrality levels of players. For this reason, the positions of players can be considered to be one of the greatest determinants of the network of passes.
Acknowledgements

This study was carried out within the scope of R&D Unit 50008, financed by UID/EEA/50008/2013.

References


Prominent Players in the UEFA Champions League

Filipe Manuel CLEMENTE and Fernando M.L. MARTINS

http://wjst.wu.ac.th

4. 663.


