Athens Journal of Sports - Volume 11, Issue 1, March 2024 – Pages 37-58

Analyzing the Efficiency of Passing Networks in Soccer

By Jeffrey Leela, Karim Rahaman± & Donna M. G. Comissong°*

This article explores the passing networks for the most used team formations of Manchester United and Chelsea during the 2016/2017 Premier League Season. A passing matrix is created for each team which distributes the average passes between players in a game. This facilitates the calculation of three centrality measures complementary to those previously explored by (López-Peña and Touchette 2012). These measures unlock hidden details about the strengths and weaknesses within the networks. Such include the extent to which a player stays or leaves his position, the ability of a player to affect the game through penetration, and a player's pass distribution evenness within the team. The optimal assignment for each network is also determined by applying the Travelling Salesman Problem, thereby establishing the least number of passes that keeps all players within each team connected. Useful knowledge can be obtained from this analysis - to inform coaching staff and enhance the overall level of play.

Keywords: *territorial, penetration, balance, optimal-assignment, team-connectivity*

Introduction

Network Science is one of the most dynamic fields in applied science and mathematics (Newman 2010). We are primarily concerned about the investigation of one of the most popular team sports – soccer (Sumpter 2016). This is a tactically sophisticated sport played by an intrinsically networked team of players, and it presents us with an ideal opportunity to investigate various areas of team organization. It is a well-known fact among coaching staff that success is achieved by creating the ideal team balance, rather than focusing only on the skills possessed by individual players. Indeed, achieving the correct combination of player abilities is key to proper team function - as a coordinated unit.

The ready availability of data (housed in online repositories) for all activities during a game of soccer in the major European leagues facilitates a more detailed analysis of team play. This in turn provides considerable insight into the overall behaviour of a team as a unit, based on the roles of individual players (Gudmundsson and Horton 2017). Beneath this structure, the arrangement of a team can be viewed with respect to the connection between its players, leading to a network analysis of passes presented as directed graphs, with the weights as the

*Senior Lecturer, Department of Mathematics, The College of Science, Technology and Applied Arts of Trinidad and Tobago, Trinidad.

±Senior Lecturer, Department of Mathematics and Statistics, The University of The West Indies, St Augustine Campus, Trinidad.

°Senior Lecturer, Department of Mathematics and Statistics, The University of The West Indies, St Augustine Campus, Trinidad.

https://doi.org/10.30958/ajspo.11-1-3 doi=10.30958/ajspo.11-1-3

numbers of passes between spatially embedded links. The system is dynamic, as the network continuously changes its structure with the movement of the players as the game progresses.

Key ‘topological scales’ may be pinpointed within the passing network of a soccer team: (i) the micro scale, which examines the players and their jobs within the system, (ii) the meso scale, that starts from little inner cliques showing the connection between three or four players to the structure of bigger cliques of players that operate as sub-networks within the framework (iii) the macro scale, which takes a look at the system in general.

At the topological micro scale, the significance of each player’s role has been identified by; its degree, which is the quantity of passes made by a player (Cotta et al. 2013); eigenvector centrality, a degree of significance deduced from the eigenvectors of the adjacency matrix (Cotta et al. 2013); closeness, estimating the smallest number of stages that the ball needs to experience from one player to arrive at another in the team (López-Peña and Touchette 2012); or betweenness centrality, which presents evidence for the degree to which a given player is needed for connecting the routes between any other combination of players of the team (Duch et al. 2010, López-Peña and Touchette 2012). Other measures are possible, for instance, the clustering coefficient, which calculates the resulting triangles around a player (i.e., a clique of three (3) players in a sub-network) identifying the areas of high passing intensity on the spatial map. This has likewise been utilized to assess the commitment of any given player to the enthusiasm of the passing system (López-Peña and Touchette 2012).

At the meso scale level, the investigation of system patterns has demonstrated how the excessive passes of a particular kind between gatherings of three or four players can be identified with both the achievement of a team (Gyarmati et al. 2014) and the identification of pioneers in the passing system (López-Peña and Sánchez Navarro 2015). Clemente et al. (2015) related the high assorted variety of passes between team players to the presence of sub-networks, which would influence the activities of the overall group. In a similar sense, Gyarmati and Anguera (2015) considered repetitive pass arrangements, thereby showing the grouping arrangement patterns central to a teams' playing design.

At the topological macro scale, system measurements are key to uncovering the exhibition and playing style of soccer teams. For instance, the system centroid appears to move in reverse when teams play as the visiting team (Bialkowski et al. 2014). Positional factors such as the ‘stretch index’ have been used to gauge the mean spread of the players around the centroid, while the team length and width have been utilized to estimate the effectiveness of overall team execution (Duarte et al. 2012). Duch et al. (2010) introduced a presentation metric dependent on the ‘betweenness’ of the players, which can be used to estimate the likelihood of winning a match. Other large-scale estimates, such as the ‘team average degree’ - the mean number of passes or the changeability of the players’ degrees - have been proposed for measuring team execution (Cintia et al. 2015). The average team cluster coefficient has been demonstrated to be a lot higher during a match than in comparable irregular systems, uncovering the formation of ‘player triplets’ (Cotta et al. 2013). With respect to overall player position, for maximum efficiency, it is also highly recommended to maintain ‘fair betweenness’ and ‘high closeness’ among the nodes of the passing system (Gonçalves et al. 2017).

A network is essentially a system consisting of branches and nodes. The flow through a network enters and exits each node via the connecting branches, and the resulting flow pattern can provide useful insight about the specific components of the network. In network analysis, the nodes are the junctions of the flow, and it is crucial to gauge the individual value of each node relative to the overall network (Lopez-Peña and Touchette 2012).

Players in a soccer team arranged in a specific formation can be studied using network analysis. Each player is represented by a node. The flow into each node equates to the passes each player receives with the flow out represented by the passes made by each player. These passes to and from each player can be analyzed using established measures - to judge the value of the contribution to the network by each player (López-Peña and Touchette 2012). This can inform the coaching staff as to what is working and what needs to be changed or improved. It is also important to measure the extent of the involvement of each player within the team. More involvement would be indicative of the team operating as a unit, with less reliance on individual players. The minimum number of passes used by a team to ensure that each player gets a touch could measure this (Anderson 1989, pp. 33–41). This is an application of the travelling salesman problem. Greater involvement would essentially heighten the awareness of players and keep them alert – thus providing the team with more opportunities to pounce on opponents during attacking plays, and with the heightened level of preparedness that is crucial for the effective maintenance of defensive team positions.

Literature Review

The Assignment Problem – Key References

The ‘assignment problem’ was first addressed by (Kuhn 1955), who suggested that numerical scores were accessible for the exhibition of every one of (n) people on every one of (n) employments. This quest for an assignment of persons to jobs was performed in such a manner that the (n) scores were as large as possible. The assignment of heterogeneous workers to heterogeneous jobs where workers may adopt varying strategies for seeking employment was later studied by Shimer (2005). Complex constraint programming to achieve near-optimal assignments while taking into account all available resources and positions was considered by Naveh et al. (2007). A linear assignment problem in the context of network systems to address this challenge by means of an auction algorithm was later postulated by Zavlanos et al. (2008). Two traditional issues from location theory which may fill in as hypothetical models for key strategic issues, where one allocates components in such a way that an appropriate linear or quadratic function accomplishes its minimum was implemented by Povh (2008). With respect to an important application to sport science, a framework to evaluate the capacities of baseball players in each pragmatic aspect of the game, and to create a group wherein every one of the players is allotted to positions to such an extent that the aggregate group ability is augmented was built by Britz and Maltitz (2010). A thesis written by Mansi (2011) dealt with approaches for solving the related transportation, transhipment and assignment problems as well as supply chain management.

Metrics for Network Analysis

An outline of the salient concepts of a dynamical framework hypothesis - of primordial importance in the investigation of coordination processes within and between soccer players - was presented by Davids et al. (2005). Of noteworthy mention is the situational investigation of the AC Milan Football club carried out by Papahristodoulou (2010), who formulated a quadratic assignment problem for the determination of the ideal field placement of three midfielders and three forward players. The overall significance of every player in a game of soccer was later analysed by López-Peña and Touchette (2012), utilizing player pass data collected from the 2010 FIFA World Cup.

For each team, a coordinated system can be devised, with nodes representing the players and arrows signifying passes between players. It is then possible to use a weighted graph clustering approach to study network location within a large-scale system. This approach was successfully employed to analyse client relationships within Internet communities (Liu et al. 2014). For a well-documented illustration of the efficient use of network metrics to extrapolate team characteristics to provide coaching support, see Clemente et al. (2015).

Methodology

It is crucial to this study that certain assumptions be made. This is primarily due to the lack of video data of match coverage, which limits our ability to capture the movement of players and the exact numbers of passes made and received by each player. It is however possible to source post-match data from freely available internet sources to document the pass distribution amongst players. For our purposes, the number of passes to and from each player was collected from internet sources – leading to the calculation of an average number of passes per player for a given season. Passes can be classified as forward, sideways and backwards. It is important to note that precise data on the direction of each pass was not available. Therefore, we used our intuition to distribute these passes based on reported averages, while taking into account each player's position within the network. Of course, players in general would naturally make most passes to teammates closer to them. Our analysis only considers players in the static positions that they occupy according to the team's formation, as done previously by López-Peña and Touchette (2012). Any movement must therefore be viewed as a collective shift of the entire network.

The first measure that we introduce is called 'territorial'. It measures the extent to which a player leaves or stays within his territory. The passes a particular player receives and makes can give us some understanding of the player's territorial rating. Receiving passes would coincide with a player remaining in his territory and making passes leaving his territory. The construction involves the summation of the difference between the number of passes received and made by each player $\sum_{i \neq j}(A_{ji} - A_{ij})$. A positive value indicates that the player receives more passes than the player makes, and vice versa for a negative value. We now divide the above summation by the total number of passes the player receives and makes given by $\sum_{i \neq j}(A_{ij} + A_{ji})$. We neutralize the result by adding one.

Territorial values under one will indicate that players tend to leave their position more often during play. By comparison, the smaller the territorial value under one, the more the player ventures out from his assigned team position. The result is analogous to values over one indicating that the players stay more in their position. Holding players tend to stay to a high extent within their territories. We can compare the scores for each player within a team to understand the extent to which players perform their functions within the formation in general.

$$T_i = \frac{\sum_{i \neq j} A_{ji} - A_{ij}}{\sum_{i \neq j} A_{ij} + A_{ji}} + 1 \qquad (1)$$

where A_{ij} – the number of passes from player i to player j, $i \neq j$

A_{ji} – the number of passes from player j to player i, $i \neq j$

$T_i = 1$ means the player is neutral and the result is inconclusive

Another measure we have established combines the ratios of the various categories of passes. As mentioned earlier, passes are categorized as forward, sideways and backwards. Each ratio has the higher rated pass over the lower rated one. This defines the penetration of a player in terms of passes. The formulation is given by the following equation.

$$P_i = \frac{\sum_{i \neq j} A_{ij} \text{ forward}}{\sum_{i \neq j} A_{ij} \text{ backward}} + \frac{\sum_{i \neq j} A_{ij} \text{ forward}}{\sum_{i \neq j} A_{ij} \text{ sideways}} + \frac{\sum_{i \neq j} A_{ij} \text{ sideways}}{\sum_{i \neq j} A_{ij} \text{ backwards}} \qquad (2)$$

Higher penetration scores can be interpreted as players being more inclined to move the ball towards their opponent's goal. Of course, for some players at least one of the ratios may be undefined. Players who retreat with passes a lot in comparison to going forward will have low penetration scores.

The length of a pass between two players may be classified as short, medium or long. The distance measured in metres would determine the length of a pass. A short pass can be defined as having a range less than or equal to ten metres (≤ 10 m). A medium range pass would be in the range between ten and twenty-five metres $(10\text{ m} < x < 25\text{ m})$ where x represents the pass length. A long pass would be greater than or equal to twenty-five metres ($\geq 25\,\text{m}$). To measure the balance of a player, take the ratio of short to medium to long range passes made and received by the player. This is illustrated in the following equation.

$$B_i = \frac{\sum_{i \neq j} A_{ij}}{\sum_{i \neq j} A_{ji}} \text{short} : \frac{\sum_{i \neq j} A_{ij}}{\sum_{i \neq j} A_{ji}} \text{medium} : \frac{\sum_{i \neq j} A_{ij}}{\sum_{i \neq j} A_{ji}} \text{long} \tag{3}$$

A well-balanced player in the network is one who passes and receives almost the same number of short, medium and long passes. As with penetration, at least one ratio may be undefined in some instances. Perfect balance is defined as the ratio 1:1:1. However, results such as 1:1 and 1 will also be balanced scores for players involved in two categories and one category of passes respectively. It is important to mention that our methods did not include any error measures to check the accuracy of the computed data. This can be considered in future investigations on this subject.

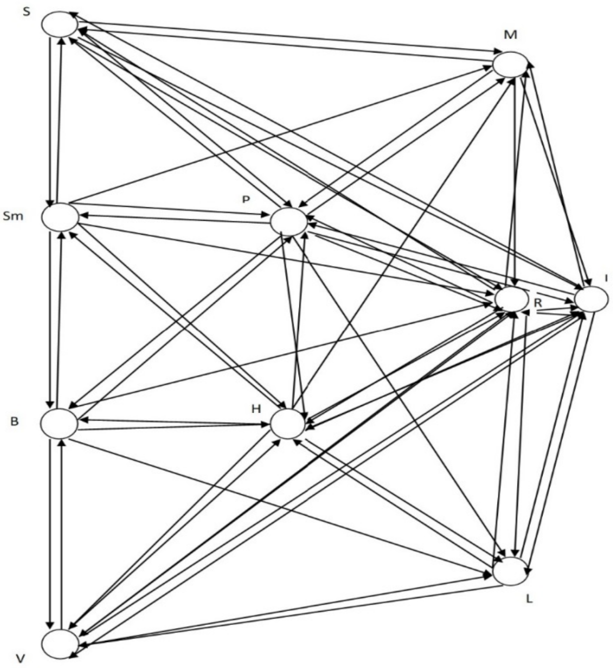
To effectively illustrate our methods, we consider the average network of two teams in the Premier League over the 2016/2017 season, for which pass data is readily available. This was a very interesting season to consider the performance of Manchester United – so soon after the appointment of the well-respected manager José Mourinho, and the recruitment of highly rated players Paul Pogba and Zlatan Ibrahimovic. In that season, Chelsea won the title, while Manchester United placed sixth with twenty-four points less than Chelsea. The two teams that we select for our analysis are Chelsea and Manchester United. Our main goal is to determine what Chelsea did effectively, and to suggest reasons for the poor league performance of Manchester United.

For the ten outfield players in each team, the average passes made by each player are first categorized as forward, sideways and backward passes – in accordance with the static team formations. Given the pass completion percentage, we are then able to create a network of directed edges for each team.

We begin by considering the pass network for Manchester United. The players selected for our analysis were the ones who were most used in the (4-2-3-1) formation used by Manchester United in the 2016/2017 Premier League Season. We then proceed to construct the network with successful passes between nodes denoted by weights and the direction indicated by arrows, noting that this is a theoretical model based on the data available and the application of our intuition. By this we mean that we take the average number of passes players make and receive, we consider the static team formation, and we then proceed to categorize the passes based on each player's position. For example, a left sided full back is likely to make more forward and sideways passes whereas a defensive midfielder more sideways and backwards passes. An attacking midfielder would tend to make more forward passes than the latter.

The websites www.whoscored.com and www.squawka.com provide the collective pass data on each player, which we categorize, based on our intuitive understanding of players, positions and team formations. Figure 1 presents the passing network for Manchester United.

Figure 1. *Manchester United's Passing Network*



Manchester United's average passes per game is given as 514.21 (Squawka 2022). Table 1 presents the average successful passes per game for Manchester United for the ten most used outfield players in the 2016/2017 Premier League Season.

Table 1. *Average Passes Completion Percentage for Manchester United*

Player	Total Passes	Pass Completion (%)
1. Shaw (S)	29.09 / 34.27	85%
2. Smalling (Sm)	22.28 / 25.28	88%
3. Bailly (B)	29.32 / 34.12	86%
4. Valencia (V)	41.29 / 48.21	86%
5. Pogba (P)	60.40 / 71.20	85%
6. Herrera (H)	57.77 / 65.97	88%
7. Martial (M)	19.68 / 24.12	82%
8. Rooney (R)	25.68 / 30.68	84%
9. Lingard (L)	25.44 / 28.92	88%
10. Ibrahimovic (Ib)	27.29 / 37.07	74%

The players in the network are Shaw (S), Smalling (Sm), Bailly (B), Valencia (V), Pogba (P), Herrera (H), Martial (M), Rooney (R), Lingard (L), Ibrahimovic (Ib). The number of passes between players was distributed accordingly and is illustrated in the passing matrix shown in Table 2.

Table 2. *Passing Matrix for Manchester United*
(A_{ij} *= passes from player "i" to player "j")*

Player	S	Sm	B	V	P	H	M	R	L	Ib
S	0	11	0	0	8	0	6	2	0	2
Sm	3	0	3	0	6	6	2	2	0	0
B	0	3	0	4	9	9	0	2	3	0
V	0	0	20	0	0	8	0	2	8	3
P	7	7	7	0	0	8	9	9	7	6
H	0	7	8	7	7	0	6	8	8	6
M	5	0	0	0	5	0	0	5	0	5
R	2	0	0	2	5	4	5	0	5	3
L	0	0	0	6	0	7	0	6	0	6
Ib	3	0	0	3	4	3	4	6	4	0

The travelling salesman problem is essentially an optimization process with far-reaching applications in the field of software engineering and operations research. It normally poses the inquiry: "Given a rundown of urban areas and the separations between each pair of urban areas, what is the briefest conceivable course that visits every city and comes back to the first city?" This is most effectively communicated as a chart portraying the areas of a set of nodes.

Dorigo and Gambardella (1997) considered a fake subterranean insect state to define and solve a 'travelling salesman problem' (TSP). Ants of the fake province can create progressively shorter feasible visits by utilizing data amassed as a pheromone trail saved on the edges of the TSP chart. Four unrelated problems that arise in the context of computer wiring, vehicle routing, clustering a data array and job-shop scheduling with no intermediate storage were explored by (Lenstra and Rinnooy Kan 1975).

We now present another important application of the travelling salesman problem. The graph that we analyse is representative of the network of passes between soccer players, as previously described. The passes between each pair of players represent the distances between the nodes. Let us consider the minimum number of passes that connects each player in the network. This translates to the passage of the football through the entire team via an optimal sequence of passes. In other words, "What sequence of passes produces the minimum number when the ball leaves a player and returns to the same player having been played to every other player in the network?" We will illustrate the process in detail using the passing matrix for Manchester United (see Table 2).

We first subtract the smallest element in each row from every other element in the row. This is followed by subtracting the smallest element in each column from every other element in that column. Now, add the minimum element in its corresponding row and column to each zero, where this minimum element excludes the zero itself. The dashes in the matrix indicate no passes between the players in the corresponding rows and columns. This is illustrated in Tables 3, 4 and 5.

Table 3. *Pass Matrix with Minimum Element of Each Row in Rightmost Column*

Player	S	Sm	B	V	P	H	M	R	L	Ib	Minimum
S	-	11	-	-	8	-	6	2	-	2	2
Sm	3	-	3	-	6	6	2	2	-	-	2
B	-	3	-	4	9	9	-	2	3	-	2
V	-	-	20	-	-	8	-	2	8	3	2
P	7	7	7	-	-	8	9	9	7	6	6
H	-	7	8	7	7	-	6	8	8	6	6
M	5	-	-	-	5	-	-	5	-	5	5
R	2	-	-	2	5	4	5	-	5	3	2
L	-	-	-	6	-	7	-	6	-	6	6
Ib	3	-	-	3	4	3	4	6	4	-	3

After subtracting the smallest element in each row, the result is presented in Table 4, where the last row contains the smallest element from each column.

Table 4. *Minimum Element of Each Column Recorded in Bottommost Row after Subtracting Minimum Element in Each Row*

Player	S	Sm	B	V	P	H	M	R	L	Ib
S	-	9	-	-	6	-	4	0	-	0
Sm	1	-	1	-	4	4	0	0	-	-
B	-	1	-	2	7	7	-	0	1	-
V	-	-	18	-	-	6	-	0	6	1
P	1	1	1	-	-	2	3	3	1	0
H	-	1	2	1	1	-	0	2	2	0
M	0	-	-	-	0	-	-	0	-	0
R	0	-	-	0	3	2	3	-	3	1
L	-	-	-	0	-	1	-	0	-	0
Ib	0	-	-	0	1	0	1	3	1	-
Minimum	0	1	1	0	0	0	0	0	1	0

Table 5. *Minimum Element of Corresponding Row/Column in Brackets Next to Each Zero*

Player	S	Sm	B	V	P	H	M	R	L	Ib
S	-	8	-	-	6	-	4	0 (0)	-	0 (0)
Sm	1	-	0 (0)	-	4	4	0 (0)	0 (0)	-	-
B	-	0 (0)	-	2	7	7	-	0 (0)	0 (0)	-
V	-	-	17	-	-	6	-	0 (1)	5	1
P	1	0 (0)	0 (0)	-	-	2	3	3	0 (0)	0 (0)
H	-	0 (0)	1	1	1	-	0 (0)	2	1	0 (0)
M	0 (0)	-	-	-	0 (1)	-	-	0 (0)	-	0 (0)
R	0 (0)	-	-	0 (0)	3	2	3	-	2	1
L	-	-	-	0 (0)	-	1	-	0 (0)	-	0 (0)
Ib	0 (0)	-	-	0 (0)	1	0 (1)	1	3	0 (0)	-

There are three zeros with the largest penalty (1) shown in Table 5 (highlighted in red); Valencia $\rightarrow$ Rooney, Martial $\rightarrow$ Pogba and Ibrahimovic $\rightarrow$

Herrera. We select Valencia $\rightarrow$ Rooney, which indicates that we cannot look at the arc Rooney $\rightarrow$ Valencia. We rewrite the matrix after deleting all entries in the row and column corresponding to Valencia $\rightarrow$ Rooney, and we repeat all these steps summarising the results in Tables 6 and 7.

Table 6. *Modified Pass Matrix after Deletions*

Player	S	Sm	B	V	P	H	M	R	L	Ib
S	-	8	-	-	6	-	4	-	-	0
Sm	1	-	0	-	4	4	0	-	-	-
B	-	0	-	2	7	7	-	-	0	-
V	-	-	-	-	-	-	-	-	-	-
P	1	0	0	-	-	2	3	-	0	0
H	-	0	1	1	1	-	0	-	1	0
M	0	-	-	-	0	-	-	-	-	0
R	0	-	-	-	3	2	3	-	2	1
L	-	-	-	0	-	1	-	-	-	0
Ib	0 (0)	-	-	0	1	0	1	3	0 (0)	-

Table 7. *Minimum Element of Corresponding Row/Column in Brackets next to Each Zero*

Player	S	Sm	B	V	P	H	M	R	L	Ib
S	-	8	-	-	6	-	4	-	-	0 (4)
Sm	1	-	0 (0)	-	4	4	0 (0)	-	-	-
B	-	0 (0)	-	2	7	7	-	-	0 (0)	-
V	-	-	-	-	-	-	-	-	-	-
P	1	0 (0)	0 (0)	-	-	2	3	-	0 (0)	0 (0)
H	-	0 (0)	1	1	1	-	0 (0)	-	1	0 (0)
M	0 (0)	-	-	-	0 (1)	-	-	-	-	0 (0)
R	0 (1)	-	-	-	3	2	3	-	2	1
L	-	-	-	0 (0)	-	1	-	-	-	0 (0)
Ib	0 (0)	-	-	0 (0)	1	0 (1)	1	-	0(0)	-

In Table 7, the zero with the largest penalty (4) is highlighted in red: Shaw $\rightarrow$ Ibrahimovic. We must therefore choose Shaw $\rightarrow$ Ibrahimovic, which would eliminate Ibrahimovic $\rightarrow$ Shaw. We then rewrite the matrix after deleting the row and column containing Shaw $\rightarrow$ Ibrahimovic. The result is illustrated in Table 8.

Table 8. *Minimum Element of Corresponding Row/Column next to Each Zero*

Player	S	Sm	B	V	P	H	M	R	L	Ib
S	-	-	-	-	-	-	-	-	-	-
Sm	1	-	0 (0)	-	4	4	0 (0)	-	-	-
B	-	0 (0)	-	2	7	7	-	-	0 (0)	-
V	-	-	-	-	-	-	-	-	-	-
P	1	0 (0)	0 (0)	-	-	2	3	-	0 (0)	-
H	-	0 (0)	1	1	1	-	0 (0)	-	1	-
M	0 (0)	-	-	-	0 (1)	-	-	-	-	-

R	0 (2)	-	-	-	3	2	3	-	2	-
L	-	-	-	0 (0)	-	1	-	-	-	-
Ib	-	-	-	0 (0)	1	0 (1)	1	-	0(0)	-

The zero with the largest penalty (2) in Table 8 is Rooney → Shaw, so we must also eliminate the entry Shaw → Rooney. Next, we rewrite the matrix deleting the entries from the row and column containing Rooney → Shaw. For simplicity, we will reduce the size of the pass matrix by removing all the players connected exclusively by dashes (i.e., delete the appropriate row/column). This procedure is summarized in Tables 9 and 10.

Table 9. *Connections between Rooney and Shaw Removed (Indicated with Dashes)*

Player	S	Sm	B	V	P	H	M	R	L	Ib
S	-	-	-	-	-	-	-	-	-	-
Sm	-	-	0	-	4	4	0	-	-	-
B	-	0	-	2	7	7	-	-	0	-
V	-	-	-	-	-	-	-	-	-	-
P	-	0	0	-	-	2	3	-	0	-
H	-	0	1	1	1	-	0	-	1	-
M	-	-	-	-	0	-	-	-	-	-
R	-	-	-	-	-	-	-	-	-	-
L	-	-	-	0	-	1	-	-	-	-
Ib	-	-	-	0	1	0	1	-	0	-

Table 10. *Reduced Matrix with Minimum Element of Corresponding Row/Column next to Each Zero*

Player	Sm	B	V	P	H	M	L
Sm	-	0 (0)	-	4	4	0 (0)	-
B	0 (0)	-	2	7	7	-	0 (0)
P	0 (0)	0 (0)	-	-	2	3	0 (0)
H	0 (0)	1	1	1	-	0 (0)	1
M	-	-	-	0 (1)	-	-	-
L	-	-	0 (1)	-	1	-	-
Ib	-	-	0 (0)	1	0 (1)	1	0(0)

As shown in Table 10, there are three zeros with the largest penalty (1) indicated in red. These correspond to Ibrahimovic → Herrera, Lingard → Valencia and Martial → Pogba. We can choose any of these, and we select Ibrahimovic → Herrera. Repeating the process, we remove the interactions between Ibrahimovic and Herrera and reduce the matrix by deleting the corresponding empty row and column. The results are recorded in Tables 11 and 12.

Table 11. *Connection between Ibrahimovic and Herrera Removed (Indicated by a Dash)*

Player	Sm	B	V	P	H	M	L
Sm	-	0	-	4	-	0	-
B	0	-	2	7	-	-	0
P	0	0	-	-	-	3	0
H	0	1	1	1	-	0	1
M	-	-	-	0	-	-	-
L	-	-	0	-	-	-	-
Ib	-	-	-	-	-	-	-

Table 12. *Reduced Matrix with Minimum Element of Corresponding Row/ Column in Brackets next to Each Zero*

Player	Sm	B	V	P	M	L
Sm	-	0 (0)	-	4	0 (0)	-
B	0 (0)	-	2	7	-	0 (0)
P	0 (0)	0 (0)	-	-	3	0 (0)
H	0 (0)	1	1	1	0 (0)	1
M	-	-	-	0 (1)	-	-
L	-	-	0 (1)	-	-	-

There are two zeros with the largest penalty (1); Martial$\rightarrow$Pogba and Lingard$\rightarrow$Valencia. We select Martial$\rightarrow$Pogba by removing the interaction Pogba$\rightarrow$Martial, deleting the row and column corresponding to Martial$\rightarrow$Pogba, and reducing the matrix appropriately. The resulting matrix is illustrated in Tables 13 and 14.

Table 13. *Connections between Martial and Pogba Removed (Indicated by a Dash)*

Player	Sm	B	V	P	M	L
Sm	-	0	-	-	0	-
B	0	-	2	-	-	0
P	0	0	-	-	-	0
H	0	1	1	-	0	1
M	-	-	-	-	-	-
L	-	-	0	-	-	-

Table 14. *Reduced Matrix with Minimum Element of Corresponding Row/ Column in Brackets next to Each Zero*

Player	Sm	B	V	M	L
Sm	-	0 (0)	-	0 (0)	-
B	0 (0)	-	2	-	0 (0)
P	0 (0)	0 (0)	-	-	0 (0)
H	0 (0)	1	1	0 (0)	1
L	-	-	0 (1)	-	-

The zero with the largest penalty from Table 14 is one (1), i.e., Lingard$\rightarrow$Valencia. We repeat the process, summarizing the results in tables 15 and 16.

Table 15. *Connections between Lingard and Valencia Removed (Indicated by a Dash)*

Player	Sm	B	V	M	L
Sm	-	0	-	0	-
B	0	-	-	-	0
P	0	0	-	-	0
H	0	1	-	0	1
L	-	-	-	-	-

Table 16. *Reduced Matrix with Minimum Element of Corresponding Row/Column in Brackets next to Each Zero*

Player	Sm	B	M	L
Sm	-	0 (0)	0 (0)	-
B	0 (0)	-	-	0 (0)
P	0 (0)	0 (0)	-	0 (0)
H	0 (0)	1	0 (0)	1

We can now eliminate any row and column containing a zero since all remaining zeros have the same penalty. We select Smalling $\rightarrow$ Martial, leading to the reduced pass matrix shown in Table 17.

Table 17. *Reduced Matrix with Minimum Element of Corresponding Row/ Column in Brackets next to Each Zero*

Player	Sm	B	L
B	0 (0)	-	0 (0)
P	0 (0)	0 (1)	0 (0)
H	0 (1)	1	1

There are two zeros with the largest penalty (1), i.e., Herrera $\rightarrow$ Smalling and Pogba $\rightarrow$ Bailly. We select Herrera $\rightarrow$ Smalling, eventually leading to the result shown in Table 18.

Table 17. *Reduced Matrix with Minimum Element of Corresponding Row/ Column next to Each Zero*

Player	B	L
B	-	0 (0)
P	0 (0)	0 (0)

Once again, all remaining zeros have the same penalty (0). We select Pogba $\rightarrow$ Bailly, and finally we are left with Bailly $\rightarrow$ Lingard.

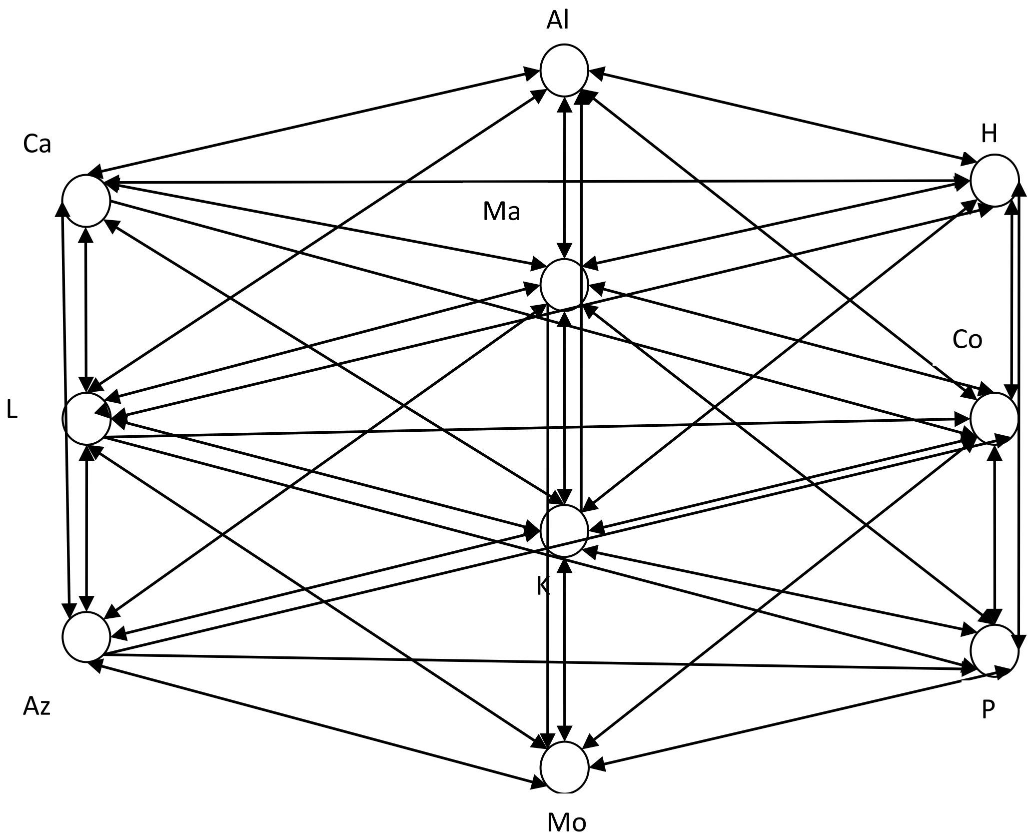
Recapping the pass sequences that we obtained, we have: Valencia $\rightarrow$ Rooney; Shaw $\rightarrow$ Ibrahimovic; Rooney $\rightarrow$ Shaw; Ibrahimovic $\rightarrow$ Herrera; Martial $\rightarrow$ Pogba; Lingard $\rightarrow$ Valencia; Smalling $\rightarrow$ Martial; Herrera $\rightarrow$ Smalling; Pogba $\rightarrow$ Bailly; Bailly $\rightarrow$ Lingard.

The minimum path of passes to keep the team connected is therefore:

Lingard → Valencia → Rooney → Shaw → Ibrahimovic → Herrera → Smalling → Martial → Pogba → Bailly → Lingard.

We now perform the same analysis to determine the minimum number of passes that keeps each player in the Chelsea Network connected. Figure 2 illustrates Chelsea's passing network. The players in the network are Cahill (Ca), Luiz (L), Azpilicueta (Az), Alonso (Al), Matic (Ma), Kanté (K), Moses (Mo), Hazard (H), Costa (Co) and Pedro (P).

Figure 2. *Chelsea's Passing Network*



Chelsea's average Passes per Game – 429.97 (Squawka 2022). As done previously for Manchester United, we begin by presenting the data for average successful passes with percentage completion and the associated pass matrix with the summarized pass distribution. This is illustrated in Tables 18 and 19 respectfully.

Table 18. *Average Passes Completion Percentage for Chelsea*

Player	Total Passes	Pass Completion (%)
1. Cahill (Ca)	45.05 / 48.89	92%
2. Luiz (L)	37.70 / 44.91	84%
3. Azpilicueta (Az)	55.18 / 63.50	87%
4. Alonso (Al)	30.90 / 39.87	78%
5. Matic (Ma)	45.77 / 52.17	88%
6. Kanté (K)	53.29 / 60.11	89%
7. Moses (Mo)	23.26 / 29.50	79%

8. Hazard (H)	40.42 / 47.97	84%
9. Costa (Co)	22.09 / 29.43	75%
10. Pedro (P)	22.17 / 26.86	83%

Table 19. *Passing Matrix for Chelsea*
(A_{ij} = passes from player "i" to player "j")

Player	Ca	L	Az	Al	Ma	K	Mo	H	Co	P
Ca	0	9	3	8	8	8	0	4	3	0
L	4	0	4	6	6	6	6	2	2	1
Az	7	10	0	0	9	9	9	0	3	7
Al	10	4	0	0	7	0	0	7	3	0
Ma	7	7	3	6	0	6	2	7	5	3
K	6	6	7	2	7	0	7	4	7	7
Mo	0	5	6	0	0	5	0	0	2	5
H	3	3	0	4	4	3	0	0	13	10
Co	0	0	0	2	4	4	2	5	0	4
P	0	0	0	0	4	4	4	4	6	0

Again, this reduces to the rotation of the football throughout the entire team through a sequence of passes that is optimal. The exact calculations as before will now be performed. Subtract the smallest element from each element of its corresponding row then do likewise for each column. After subtracting the smallest element in each row, the smallest element to subtract from each column is zero. Table 20 shows the results.

Table 20. *Pass Matrix with Minimum Element of Each Row in Rightmost Column*

Player	Ca	L	Az	Al	Ma	K	Mo	H	Co	P	Minimum
Ca	-	9	3	8	8	8	-	4	3	-	3
L	4	-	4	6	6	6	6	2	2	1	1
Az	7	10	-	-	9	9	9	-	3	7	3
Al	10	4	-	-	7	-	-	7	3	-	3
Ma	7	7	3	6	-	6	2	7	5	3	2
K	6	6	7	2	7	-	7	4	7	7	2
Mo	-	5	6	-	-	5	-	-	2	5	2
H	3	3	-	4	4	3	-	-	13	10	3
Co	-	-	-	2	4	4	2	5	-	4	2
P	-	-	-	-	4	4	4	4	6	-	4

Table 21. *Minimum Element of Corresponding Row/Column next to Each Zero*

Player	Ca	L	Az	Al	Ma	K	Mo	H	Co	P
Ca	-	6	0 (1)	5	5	5	-	1	0 (0)	-
L	3	-	3	5	5	5	5	1	1	0(2)
Az	4	7	-	-	6	6	6	-	0 (4)	4
Al	7	1	-	-	4	-	-	4	0 (1)	-
Ma	5	5	1	4	-	4	0 (1)	5	3	1
K	4	4	5	0(2)	5	-	5	2	5	5
Mo	-	3	4	-	-	3	-	-	0(3)	3
H	0(3)	0(1)	-	1	1	0(0)	-	-	10	7
Co	-	-	-	0(0)	2	2	0(0)	3	-	2
P	-	-	-	-	0(1)	0(0)	0(0)	0(1)	2	-

The penalty associated with each zero is shown above in Table 21, i.e., the sum of the minimum element in that row and column containing the zero and excluding the zero itself. The zero with the largest penalty (4) exists at Azpilicueta $\rightarrow$Costa, which is highlighted in red in Table 21. We cannot look at Costa$\rightarrow$ Azpilicueta. We rewrite the matrix deleting the row and column containing Azpilicueta $\rightarrow$Costa and repeat the steps. For brevity, we will summarize our eventual findings. After a sequence of eliminations as with the application of the algorithm on Manchester United's passing matrix, the following links are produced.

$$\begin{array}{ccc}
\text{Azpilicueta} & \rightarrow & \text{Costa} \\
\text{Hazard} & \rightarrow & \text{Cahill} \\
\text{Cahill} & \rightarrow & \text{Azpilicueta} \\
\text{Alonso} & \rightarrow & \text{Luiz} \\
\text{Pedro} & \rightarrow & \text{Matic} \\
\text{Matic} & \rightarrow & \text{Moses} \\
\text{Moses} & \rightarrow & \text{Kante}' \\
\text{Luiz} & \rightarrow & \text{Pedro} \\
\text{Kante}' & \rightarrow & \text{Hazard} \\
\text{Costa} & \rightarrow & \text{Alonso}
\end{array}$$

For Chelsea, the minimum path of passes to keep the team connected is therefore:

$$\text{Costa} \rightarrow \text{Alonso} \rightarrow \text{Luiz} \rightarrow \text{Pedro} \rightarrow \text{Matic} \rightarrow \text{Moses} \rightarrow \text{Kante}' \rightarrow \text{Hazard} \rightarrow \text{Cahill} \rightarrow \text{Azpilicueta} \rightarrow \text{Costa}$$

Results

Territorial scores T_i for the ten outfield Manchester United players are as follows:

(*where T_1= Shaw, T_2= Smalling, T_3= Bailly, T_4= Valencia, T_5= Pogba, T_6= Herrera, T_7= Martial, T_8= Rooney, T_9= Lingard, T_{10}= Ibrahimovic).

$$T_1 = \frac{\sum_{i\neq j} A_{j1} - A_{1j}}{\sum_{i\neq j} A_{1j} + A_{j1}} + 1 = \frac{(5-6)+(3-2)+(2-2)+(7-3)+(3-11)}{44} + 1$$

$$= \frac{-4}{44} + 1 = 0.909$$

In a similar way, we get: $T_2 = 1.120$, T_3 = 1.118, T_4 = 0.733, T_5 = 0.788, T_6 = 0.882, T_7 = 1.231, T_8 = 1.235, T_9 = 1.167, T_{10} = 1.069.

Valencia's territorial score (T_4 = 0.733) is the lowest by comparison, which shows that he left his position more than any other outfield player under consideration. Rooney in contrast had the highest score (T_8 = 1.235), indicating that he remained within his territory the most.

Penetrative scores P_i for the ten outfield Manchester United players (*where P_1= Shaw, P_2= Smalling, P_3= Bailly, P_4= Valencia, P_5= Pogba, P_6= Herrera, P_7= Martial, P_8= Rooney, P_9= Lingard, P_{10}= Ibrahimovic) are as follows:

$$P_1 = \frac{\sum_{i\neq j} A_{1j} \text{ forward}}{\sum_{i\neq j} A_{1j} \text{ backward}} + \frac{\sum_{i\neq j} A_{1j} \text{ forward}}{\sum_{i\neq j} A_{1j} \text{ sideways}} + \frac{\sum_{i\neq j} A_{1j} \text{ sideways}}{\sum_{i\neq j} A_{1j} \text{ backwards}}$$

$$= \frac{6+2+2+3 \text{ forward}}{11 \text{ sideways}} = \frac{13}{11} = 1.182$$

Similarly, we can show that $P_2 = 2.667$, $P_3 = 3.286$, $P_4 = 1.050$, $P_5 = 5.732$, $P_6 = 5.591$, $P_7 = 1.861$, $P_8 = 1.300$, $P_9 = 1.923$, $P_{10} = 0$.

The most penetrative player was Pogba (5.732) with Herrera second (5.591). Ibrahimovic was last, scoring zero on penetration.

Finally, we calculate the balance score B_i for each of the 10 outfield Manchester United players under consideration, where B_1= Shaw, B_2= Smalling, B_3= Bailly, B_4= Valencia, B_5= Pogba, B_6= Herrera, B_7= Martial, B_8= Rooney, B_9= Lingard, B_{10}= Ibrahimovic.

$$B_1 = \frac{\sum_{i\neq j} A_{1j}}{\sum_{i\neq j} A_{j1}} \text{short} : \frac{\sum_{i\neq j} A_{1j}}{\sum_{i\neq j} A_{j1}} \text{medium} : \frac{\sum_{i\neq j} A_{1j}}{\sum_{i\neq j} A_{j1}} \text{long}$$

$$= \frac{11+3}{3+7} \text{short} : \frac{6}{5} \text{medium} : \frac{2+2}{2+3} \text{long} = 1.400 : 1.200 : 0.800$$

Similarly, we can show that: $B_2 = 0.643$ short, $B_3 = 0.658$ short,
B_4 = 2.545 short: 1.250 medium: 1.000 long; $B_5 = 1.343$ short : 1.500 long,
B_6 = 1.071 short: 2.000 long, $B_7 = 0.714$ short : 0.476 medium, $B_8 = 0.647$ short : 0.500 long,
$B_9 = 1.118$ short : 0.333 medium, $B_{10} = 1$ short : 0.765 long

The most balanced player was Shaw with almost equal numbers of passes made and received of the three types. Ibrahimovic was second with even distribution of short passes and somewhat even long passes made and received. Unbalanced players include Smalling, Bailly, Martial and Rooney all of which had each part of their defined ratios far from one. Valencia's short and medium passing made and received were the most unbalanced in the network. However, his long-range pass distribution was balanced. Due to the fact his other types of passes were more unbalanced than any other player; he would be the most unbalanced player in the team.

The title winning team Chelsea also provides insightfulness on what they did better that Manchester United. We now provide the same analysis to possibly uncover their strengths. Recall Chelsea's passing data from Table 19.

Based on the above ten (10) outfield players the average passes made by each player are divided into forward, sideways and backward passes with pass completion percentage. We are then able to create a network of directed edges. We chose these players since they were the individuals most used in the (3-4-3) formation of the 2016/2017 Premier League season (Whoscored 2022).

We construct the network with successful passes between nodes denoted by weights and the direction indicated by arrows. Recall that this is a theoretical model based on the static data available, and that Chelsea's passing network was presented in figure 2.

The territorial scores for Chelsea (where T_1= Cahill, T_2= Luiz, T_3= Azpilicueta, T_4= Alonso, T_5= Matic, T_6= Kanté, T_7= Moses, T_8= Hazard, T_9= Costa, T_{10}= Pedro) are given by:

T_1 = 0.925, T_2 = 1.086, T_3 = 0.597, T_4 = 0.949, T_5 = 1.032, T_6 = 0.883, T_7 = 1.132, T_8 = 0.904, T_9 = 1.354, T_{10} = 1.254.

The penetration scores for Chelsea (where P_1= Cahill, P_2= Luiz, P_3= Azpilicueta, P_4= Alonso, P_5= Matic, P_6= Kanté, P_7= Moses, P_8= Hazard, P_9= Costa, P_{10}= Pedro) are:

P_1 = 2.583, P_2 = 3.625, P_3 = 2.176, P_4 = 2.643, P_5 = 2.777, P_6 = 2.914, P_7 = 2.491, P_8 =1.353, P_9 = 0.750, P_{10} = 0.833.

Chelsea's balance scores for each player, where B_1= Cahill, B_2= Luiz, B_3= Azpilicueta, B_4= Alonso, B_5= Matic, B_6= Kanté, B_7= Moses, B_8= Hazard, B_9= Costa, B_{10}= Pedro, are as follows:

B_1= 1.190: 0.846: 2.333, B_2 = 0.625: 1.333: 1.667, B_3 = 1.647 short: 2.667 medium,

B_4 = 1.333 short: 0.875 medium, B_5 = 1.056 short: 0.462 medium,

B_6 = 1.206 short: 1.091 medium, B_7 = 0.800 short: 0.875 medium, B_8 = 1.105: 1.625: 1.000,

B_9 = 0.548 short: 0.800 medium, B_{10} = 0.875 short: 1.333 medium.

The player with the smallest territorial score was Azpilicueta and the largest was Costa. Luiz was the most penetrative with his passes while Costa was the least penetrative. Kanté was balanced with Cahill, Azpilicueta, Matic, Moses and Costa unbalanced. The most balanced player in Chelsea's network was Hazard, while the least balanced player was Azpilicueta.

Revisiting our pass matrix data, the sequence = 6 + 2 + 2 + 2 + 3 + 7 + 2 + 5 + 7 + 3= 39. Based on the average passing data matrix, 39 is the optional number of passes needed to keep all players in the Manchester United network connected. Consequently, for Chelsea this sequence = 2 + 4 + 1 + 4 + 2 + 5 + 4 + 3 + 3 + 3 = 31.

Discussion

The territorial scores by comparison reveal that Chelsea's scores had a significantly higher range than Manchester United's. This suggests that Chelsea utilizes the playing space better than Manchester United. Chelsea also had four players as opposed to Manchester's three leaving their territories. However, only one of the four did this to a greater extent than the three Manchester United players. When too many players venture out of position, this could cause opposing teams to counterattack successfully. Chelsea had only one player doing this excessively, making them less susceptible to counter attacks than Manchester United.

Penetration is a very important aspect of soccer. It measures a player's ability to pierce an opponent's defence creating a goal scoring opportunity. Although Manchester United had a couple of players with higher individual penetration scores, Chelsea's all-round scores were better indicating that most of the Chelsea players contributed to penetration. Manchester relied heavily on Pogba and Herrera to provide the cutting passes. This is dangerous for a team, since the absence of such players (as a result of injury) will have a significant impact on overall team performance. Consequently, Chelsea operated more as a unit.

The ratios of short, medium and long passes to and from each player reveal better balance ratios for Chelsea than Manchester United. Although Manchester United had the overall best-balanced player in Shaw, Chelsea collectively had better scores. For instance, every player for Chelsea had at least two types of passes in their ratios which were not the case for Manchester United. In fact, two players for Manchester had only one type of passes in their distribution. This suggests that Chelsea's play had more variety. Subsequently, Chelsea's better balance resulted in a more even pass distribution. By comparison Chelsea's players' scores are generally closer in proximity to each other than Manchester United's demonstrating less disparity amongst players. Therefore, the Chelsea team was more connected than Manchester United, which justifies their respective League positions at the end of the season. It is clear from our analysis that Chelsea's all-round team performance was better than Manchester United's.

Based on the average passing data matrix, thirty-one is the optimal number of passes needed to keep all Chelsea players in the network connected. This is eight passes less than Manchester United. Chelsea's players by comparison are about twenty percent more efficiently involved in their team's play than Manchester United's players are. This heightens the alertness and sharpness of the Chelsea players, which improves all round team play. No wonder they finished significantly higher than Manchester United in the 2016/2017 Premier League Season.

Conclusion

Network analysis is an area in optimization that studies the efficiency of routes. This is done through the construction of graphs. Our passing networks are created by the formations of the teams. The outfield players constitute the nodes and the edges are the passes between them. The graphs are directed as indicated by the arrows on the passes. We are able to use the passes between players to solve the 'Travelling Salesman' problem, thereby determining the minimum number of passes required to keep the entire team involved. This is critical to the team's function as a unit. Players as a result will be more alert, knowing that they must be prepared at all times to receive passes. It enables each player to contribute to the team's all-round play and discourages the establishment of cliques. The whole is better than the sum of its parts. From the two networks studied, we saw that Chelsea had eight passes fewer than Manchester United to involve all the players. This is one of the reasons why Chelsea performed much better and finished much higher in the league table than Manchester United.

We successfully developed metrics - territorial, penetration and balance - to measure the impact of each player in the team network. These measures involved the construction of formulas using the pass data for each player. The territorial score measures the extent to which a player stays or leaves his/her position. Now, if a player's territorial score suggests the player is always out of position, this may be viewed in two ways depending on the player's position. The coach may be inclined to make an adjustment if the player is a defensive player – to prevent the team from being susceptible defensively. On the other hand, if the player has a free role to roam and create opportunities, this may be exactly what the coach wants. Similar arguments can be made for a player whose territorial score suggests the player stays in position when the coaching staff would have preferred more fluidity in team formation.

Penetration is the ratio of forward to backwards and sideways passes. This measures the extent to which a player is able to push the play forward in the opponent's defensive area. This is a crucial aspect of attacking. Players who are less penetrative tend to pass backward and sideways a lot. This decreases the time the ball spends in the opponent's defensive third of the field thus, reducing the probability of scoring a goal. Penetrative players are very creative and are considered critical in piercing defences. In the selection of a team, a coach must be very careful in maintaining the proper balance of players. A well-balanced player in the network is one who passes and receives almost the same number of short, medium and long passes. The pass classification is based on the length of the pass. A player passing and receiving the same number of short, medium and long passes is said to be in equilibrium. However, this will in most instances be impossible. Therefore, the balance score sheds some light on a player on how effective he/she performs his/her role. The position a player occupies on the field will be an indicator of the type of passes required to make. Since some players are very efficient at making passes of all lengths they can be better utilized in a variety of roles. This adds to the strength of the team. Knowing these attributes places a coach in an advantageous position in solving selection dilemmas.

References

Anderson I (1989) An Optimal Assignment Problem. *In a First Course in Combinatorial Mathematics.* 2nd ed.: 33-41. Clarendon Press; New York: Oxford University Press.

Bialkowski A, Lucey P, Peter G, Carr K, Yue Y, Matthews I (2014) Win at Home and Draw Away: Automatic Formation Analysis Highlighting the Differences in Home and Away Team Behaviours. *Presented in the Proceedings of the 8th Annual MIT Sloan Sports Analytics Conference* (*February* 28 -March 1, 2014) Boston, MA: MIT.

Britz SS, Maltitz MJ (2010) *Application of the Hungarian algorithm in baseball team selection and assignment.* Available at: https://www.ufs.ac.za/docs/librariesprovider22/mathematical-statistics-and-actuarial-science-documents/technical-reports-documents/teg419-2070-eng.pdf?sfvrsn=383cf921_0. Accessed June 1, 2022.

Cintia P, Rinzivillo S, and Pappalardo L (2015) A network-based approach to evaluate the performance of football teams. In *The Conference in Machine Learning and Data Mining for Sports Analytics Workshop* (September 11, 2015). Porto, Portugal.

Clemente FM, Santos Couceiro M, Lourenço Martins FM, Sousa Mendes R (2015) Using network metrics in soccer: a macro-analysis. *Journal of Human Kinetics* 45(1): 123–134.

Cotta C, Mora AM, Merelo JJ, Merelo-Molina C (2013) A network Analysis of the 2010 FIFA World Cup Champion Team Play. *Journal of System Science and Complexity* 26(1): 21–42.

Davids K, Araújo D, Shuttleworth R (2005) Applications of dynamical systems theory to football. In *Science and Football V: The Proceedings of the Fifth World Congress on Science and Football Conference*. Lisbon, Portugal, 2003, edited by Thomas Reilly, Jan Cabri, Duarte Araújo, London, England: Routledge

Dorigo M, Gambardella LM (1997) Ant colony system: a cooperative learning approach to the traveling salesman problem. IEEE. *Translation on Evolutionary Computation* 1(1): 53–66.

Duarte R, Araújo D, Folgado H, Esteves P, Marques P, Davids K (2012) Capturing complex, non-linear team behaviours during competitive football performance. *Journal of System. Science and Complexity* 26(1): 62–72.

Duch J, Waitzman JS, Amaral LA (2010) Quantifying the performance of individual players in a team activity. *PLOS ONE* 5(6): e10937.

Gonçalves B, Coutinho D, Santos S, Lago-Penas C, Jimnéz, S, Sampaio J (2017) Exploring team passing networks and player movement dynamics in youth association football. *PLOS ONE* 12(1): e0171156.

Gudmundsson J, Horton M (2017) Spatio-temporal analysis of team sports. *ACM Computing Surveys (CSUR)* 50(2): 1–34.

Gyarmati L, Anguera X (2015) *Automatic extraction of the passing strategies of soccer teams*. arXiv Preprint, abs/1508.02171.

Gyarmati L, Kwak H, Rodríguez P (2014) Searching for a unique style in Soccer. Paper Presented in *Proceedings of the ACM International Conference on Knowledge Discovery and Data Mining (KDD) Workshop on Large-Scale Sports Analytics*. New York, NY, September 2014.

Kuhn HW (1955) The Hungarian Method for the Assignment Problem. *Naval Research Logistic Quarterly* 2(1–2): 83–97.

Lenstra JK, Rinnooy Kan AHG (1975) Some simple applications of the travelling salesman problem. *Journal of the Operational Research Society* 26(4): 717–733.

Liu RV, Feng S, Shi R, Guo W (2014) Weighted graph clustering for community detection of large social networks. *Procedia Computer Science* 31(1): 85–94.

López-Peña J, Sánchez Navarro R (2015) *Who can replace Xavi? A passing motif analysis of football players*. ArXiv: 1506.07768 Preprint.

López-Peña J, Touchette H (2012) A network theory analysis of football strategies. In *Sports Physics Proceeding, 2012 Euromech Physics of Sports Conference*.

Mansi SG (2011) *A study on transportation problem, transshipment problem, assignment problem and supply chain management*. PhD Dissertation. Saurashtra University.

Naveh Y, Richter Y, Altshuler Y, Gresh DL, Connors DP (2007) Workforce Optimization: Identification and Assignment of Professional Workers Using Constraint Programming. *IBM Journal of Research and Development* 51(3.4): 263–279.

Newman MEJ (2010) *Networks: an introduction*. New York, NY: Oxford University Press.

Papahristodoulou C (2010) The optimal layout of football players: a case study for AC Milan. Paper Presented at *Munich Personal Research Papers in Economics (MPRA).* Munich, Germany, January 10, 2010.

Povh J (2008) Assignment problems in logistics. *Logistics & Sustainable Transport* 1(3): 15–24.

Shimer R (2005) The assignment of workers to jobs in an economy with coordination frictions. *Journal of Political Economy* 113(5): 996–1025.

Squawka (2022) *Comparison matrix*. Available at: https://www.squawka.com/en/comparison-matrix/#avg.

Sumpter D (2016) *Soccermatics. Mathematical adventures in the beautiful game, Bloomsbury Sigma*. London: Bloomsbury Publishing.

Whoscored (2022) *Football statistics*. Available at: https://www.whoscored.com/statistics. Premier League 2016/2017, Team Statistics.

Zavlanos MM, Spesivtsev L, Pappas GJ (2008) A distributed auction algorithm for the assignment problem. Paper Presented at *In Decision and Control, CDC 47th IEEE Conference*, 1212–1217. Pennsylvania, USA, December 9-11, 2008.