



# Analysis of Passing Networks in Soccer

John Laschober<sup>1</sup>, Amanda Harsy<sup>1,\*</sup>

<sup>1</sup>Department of Computer and Mathematical Sciences, Lewis University, Romeoville, IL

\*Corresponding Author E-mail:  
harsyram@lewisu.edu

## Abstract

The study of the passing networks formed during soccer matches is an area of mathematics that has exploded within the past decade as both modern technology and new-found motivation in notable organizations have provided the public with better datasets. Through the analysis of players' positioning and their passes throughout a match, better strategies can be devised for future games and training regiments can be altered to better prepare even the best of teams. The aim of the paper is to introduce an analysis application that applies basic graph theory concepts to a game of soccer. Thus, we provide further support to the claim that these graphical concepts are not only applicable but also useful to the field of sports analysis and provide worthwhile insight for organizations, coaches, and players themselves.

**Keywords:** soccer, football, passing, network, analysis, performance

---

## 1 Introduction

This paper introduces an analysis application that applies basic graph theory concepts to a game of soccer. This application can be used to create graphical visualization of the passing networks of soccer players and calculate graphical concepts like betweenness, closeness, and centrality which can then be used to help give useful insights into the game and players. This supports the claims of Yue, Duarte, Mendes, et al. that graphical concepts are applicable and useful in analyzing soccer (Yue et al. 2008; Ribeiro et al. 2017; Lusher, Robins, and Kremer 2010; Mendes, Clemente, and Mauricio 2018).

Soccer is a team-based game in which players often pass the ball between each other in order drive the current play towards the opposing team's goal and score. Soccer has a more continuous game-flow than other "passing-heavy" team sports like basketball and football which have timeouts and quarters. A traditional soccer game is made up of two 45-minute halves and at the professional level has minimal subbing. In the past, it has been challenging to analyze different factors of success in soccer. This is, in part, due to a lack of data collected but also because of the vast number of complex features which are included in the game (technical skill, individual physiological performance, team dynamics and tactics, etc.) (Rein and Memmert 2016). Because of the complexity of a soccer match, Duarte, et al. have attempted to use biological models to analyze social and interpersonal interactions between players and teams (Duarte et al. 2012). They believe that thinking of a team as a "super-organism" and modeling its behavior using a sociobiological approach can help improve the understanding of the complex behaviors of soccer teams (Duarte et al. 2012). With how vital passing is in this game, many statisticians and analysts have begun to apply network science in order to evaluate a team's performance (Ribeiro et al. 2017; Lopez Pena and Touchette 2012; Lusher, Robins, and Kremer 2010). By viewing a passing network as a weighted, directed graph, we can then look at a team and apply basic graph theory concepts onto the players and their passes, revealing statistics that might not be immediately obvious to someone watching from the sidelines.

In our time with this topic, we researched a few questions that we believed were relevant to the question at hand. We asked what a usable dataset for this kind of analysis might look like, how easy and worthwhile it was to create an analysis application, which graph theory concepts could be applied (and how they could be implemented), and whether the resulting values could be of any use to analysts. We observed how the traditional method of using a weighted adjacency matrix referring to the number of passes between players creates a weighted directed graph, visually depicting the team's strongest and weakest areas on the pitch. We also measured a value denoted as betweenness, which scores players according to how important they are in bridging together passes intended for other players, as well as

different centrality measures that denote how important a player is during passing plays and a value known as closeness that shows how far away a player is from others on average. Studying these approaches to analyzing passing networks in matches provides worthwhile insight for organizations, coaches, and players themselves.

## 2 Background

Analyzing passing networks in soccer has been done with minor variations. The research of Buldú and co. as well as the work of Einarsson both introduce passing network analysis concepts in much the same way, describing these ideas alongside formal equations (Buldú et al. 2018; Einarsson 2015). Gonçalves and co. extended this research by applying these ideas to youth association soccer matches (Gonçalves et al. 2017). They compared their results between age brackets and found that older players who passed more often led their teams to better performances. In much the same way, Peixoto and others used their research to examine elite soccer players and how differences in passing networks affected plays. They found that fewer passes led to more shots on goal which seemed to imply that breakaway plays work more often than passing plays around the net (Peixoto et al. 2017). Peña and Touchette also examined elite players, analyzing specific elite teams (such as Spain, Germany, etc.) and drew conclusions per team (Lopez Pena and Touchette 2012). Lago and Martí used 170 matches from the 2003-2004 Spanish Soccer League to determine that venue and how a team is doing during the game (winning/losing/tied) are tied to ball possession in a soccer match (Lago and Martín 2007). Yue, et al. used time series along with concepts from geometry and graph theory to help analyze ball possession along with individual and collective behaviors of players (Yue et al. 2008). Several researchers have also attempted to use graph theoretical methods to reconceptualize a soccer team as a social network (Ribeiro et al. 2017; Lusher, Robins, and Kremer 2010). Duarte, et al. suggested that viewing a team as a social network could help coaches improve their practice environments (Ribeiro et al. 2017). Lastly, work has been done by Mendes, Clemente, and Maurício in discovering how levels of variance can be found using a smaller set of data, which can then be applied to larger datasets in order to predict future outcomes or evaluate a team's performance post-match (Mendes, Clemente, and Maurício 2018).

## 3 Methods

### 3.1 The Dataset

The sole dataset used in our research is titled Magglingen2013 and was created by Martin Rumo (Rumo 2014). It is publicly available for free and was used under the Creative Commons Attribution Share-Alike license. The dataset tracks the positional data of two teams of eleven players for one half of a full length soccer match using GPS trackers. The two teams consisted of anonymous players (all members of a top U-19 Swiss soccer club), with the two teams squaring off in a game played on artificial green turf. The positional data is stored as two floating point numbers ( $x$  and  $y$ ) and is logged at 100 millisecond time-steps. The soccer ball itself is tracked through GPS in the same manner. In order to preserve anonymity, each player was assigned an ID number in order to differentiate themselves from each other. Another important variable that came into play was that each player held a "possession" variable, depicting if they were in possession of the ball at any specific time-step.

### 3.2 Application Creation

Three major programs were created in order to extract values from the Magglingen2013 dataset. One was written as a JavaScript web-based application simply for ease of use and for using HTML5's Canvas functionality to create visual graphs and to receive an array of all passes between all players. The other two files were Python scripts that evaluated two other values (closeness and betweenness)(Laschober 2019).

In order to extract an array of all of the passes between all players on each team, the code ran through the entire game and looked at the player possession variable to determine which player currently had the ball. When the ball travelled between two players, this possession variable became "null" or undefined, denoting that no player currently held the ball. When this possession variable was eventually changed to another player's ID, the code checked if this new player was on the same team as the old player. If so, this interaction was called a pass. After going through the entire game, the code produced a complete account of how many times each player ID passed to every other player ID on the same team.

The array of all player passes was manually entered into the two Python scripts so that the required data was present prior to calculating either betweenness or closeness. Ideally, a comprehensive analysis application for these kinds of datasets would not be spread out over three separate applications and would not require this manual data entry.

## 4 Concepts Applied

As noted earlier, the programs created were designed to output an array denoting who passed to whom (and how many times) throughout the second half of this soccer match. Using this data, we can view this information as a weighted directed graph (also referred to as a weighted digraph). Each vertex on the graph represents each player ID on some team. An edge between two vertices shows that, at some point in the match, a pass occurred between these two players. Since multiple passes could and probably will occur throughout a game, each of these edges has a weight associated with it, denoting how many times a pass between these two players occurred. We also analyzed the concepts of betweenness, closeness, and both in and outdegree centrality.

Figure 1: An example of a weighted digraph.

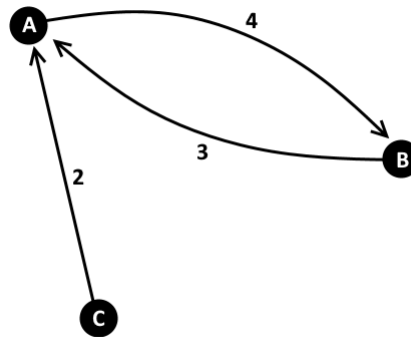


Figure 1 illustrates the basic concepts of a weighted digraph, consisting of three vertices (labeled *A*, *B*, and *C*) along with edges that have their weight denoted alongside them. *A*, *B*, and *C* represent players *A*, *B*, and *C* on the same team. Any edge with a weight (*N*) would mean that player *X* passed *N* times to player *Y*. The absence of an edge would mean that no passes occurred between the two players. Notice that the edge from *C* to *A* exists with a weight of 2, which is different from the relationship that *A* to *C* has (which is non-existent since no edge is visible between these two vertices).

### 4.1 Rendering Graphs for Visual Analysis

With our array of passing data, we attempted to emulate graphs that were present in papers by Buldú and by Gonçalves. To do so, we modified our JavaScript program to produce colored renderings of the players as numbered vertices, positioned at their average *x* and *y* positions, while we denoted passes as edges between the players (Buldú et al. 2018; Gonçalves et al. 2017). Following the methods of Buldú and Gonçalves, we denoted the weight of an edge by its thickness. Hence, a player who passes to a teammate multiple times will have created a thick edge between them, whereas a pair of players who only pass to each other once will have a very thin edge connecting them.

### 4.2 Indegree Centrality

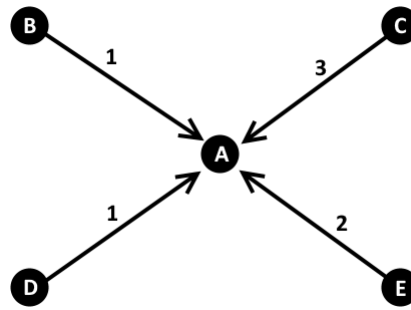
One of the concepts we analyzed in our data was Indegree Centrality for various players. Indegree Centrality for vertex *u* ( $IDC_u$ ) is a value that denotes the number of connections received for the specific node *u* within the digraph.  $IDC_u$  for an undirected graph is calculated by adding up the number of inward facing edges for any particular vertex that stem from all other vertices in the digraph. The  $IDC_u$  for the weighted digraph is calculated by adding up the weights of all inward facing edges for *u* (Peixoto et al. 2017). We can formally define  $IDC_{u_i}$  for  $u_i \in V(G)$ , the vertex set of digraph *G*, using its adjacency matrix,  $(a_{ij})$ , the matrix whose  $(i, j)^{th}$  entry is the weight of the directed edge from vertex *i* to vertex *j*. Thus for digraph *G* with vertex set  $V(G) = \{u_1, u_2, u_3, \dots\}$ , we define the indegree centrality for vertex  $u_i$  as:

$$IDC_{u_i} = \sum_{k=1}^n a_{ki}. \tag{1}$$

An example of indegree centrality is shown in Figure 2. Vertex *A* would have an  $IDC_A$  of 7, since this is the sum of the weights for all incoming edges.

In the context of this research, IDC measures the number of passes a player receives from the rest of his or her teammates. A player with a higher IDC than other players means that they were the most sought after player to receive passes throughout the match. Using

Figure 2: An example showcasing that the indegree centrality of vertex A is 7.



the array of passing data we acquired earlier, we calculated the IDC of players by looping through the array and summing the number of passes directed towards any specific player ID.

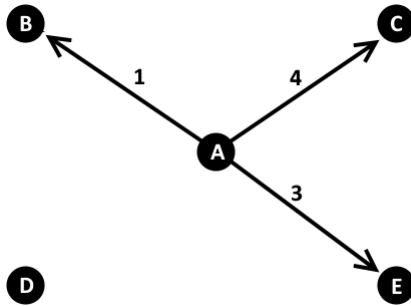
### 4.3 Outdegree Centrality

We also measured the Outdegree Centrality (ODC) of players. Outdegree Centrality of vertex  $v$  is a value given to  $v$  which is equal to the number of connections a vertex initiates within the digraph.  $ODC_u$  is calculated by adding up the number of outward facing edges for any particular vertex that stem from all other vertices in the digraph (Peixoto et al. 2017). We can formally define  $ODC_{u_i}$  for  $u_i \in V(G)$ , in a similar manner as  $IDC_{u_i}$ . Given the adjacency matrix for digraph  $G$  with vertex set  $V(G) = \{u_1, u_2, u_3, \dots\}$ , we define the outdegree centrality for vertex  $u_i$  as:

$$ODC_{u_i} = \sum_{k=1}^n a_{ik}. \tag{2}$$

Similar to the previous example for IDC, Figure 3 showcases how the outdegree of vertex A is calculated, which is done by adding up the weights of all outgoing edges originating from vertex A.

Figure 3: An example showcasing that the outdegree centrality of vertex A is 8.



Once again in the context of this research, ODC is a measure of the total number of passes that a player makes to the rest of his or her teammates throughout the match. A player with a higher ODC than other players means that they passed more often throughout the observed portion of the match. Much like IDC, we found the ODC from the array of passing data by adding the number of passes any specific ID made to other players.

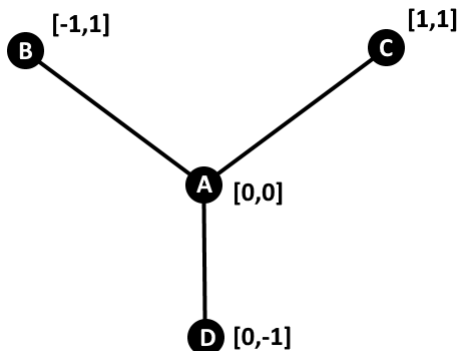
### 4.4 Closeness

We also measured the closeness of the players which shows how easily connected a vertex is to all other vertices by some form of distance. We follow (Gonçalves et al. 2017) and formally define the closeness of a vertex  $u$  in a digraph  $G$  as

$$\text{Closeness}(u) = \frac{1}{\sum_{v \in V(G)} \text{dist}(u, v)} \tag{3}$$

where  $\text{dist}(u, v)$  is the distance between vertices  $u$  and  $v$ .

Figure 4: An example showcasing that the closeness for vertex  $A$  is around 0.261.



In a soccer match, our vertices are again our players while the form of distance measured is the physical distance of the players on the pitch. For the closeness value evaluated in our research, we used the overall average  $x$  and  $y$  positions found at the end of the tracked half of the match. After having found our values of closeness, a player with a large closeness value means that they are easy to reach in terms of distance; a low closeness value means that they are farther away.

Figure 4 sets up a simple example where each vertex is associated with an  $x$  and  $y$  position. Suppose we are looking for the closeness of vertex  $A$ . Then, we would need to find the distance from  $A$  to all other vertices ( $B$ ,  $C$ , and  $D$  in this case). In this example, we use the Euclidean distance formula to find the sum of all distances between  $A$  and  $x$  where  $x \in \{B, C, D\}$  to be  $2\sqrt{2} + 1$ . Thus, the  $\text{Closeness}(A) = \frac{1}{2\sqrt{2}+1} \approx 0.261$ . Since  $A$  is situated between all of these other vertices, we should find that all other vertices have a smaller closeness value than vertex  $A$  because of the fact that their distances between all other vertices is greater compared to vertex  $A$ .

For the values found using the application we created, we acquired an average  $x$  and  $y$  position by adding up the total number of entries for positional data and then dividing by the total number of time-steps that the GPS devices had recorded. Doing this for each player individually produced these averages. Since these averages were based off of the physical soccer pitch, finding the distance between all other players was done through the use of the Euclidean distance formula. These values were then summed and their reciprocal was taken for each individual player.

### 4.5 Betweenness Centrality

The last feature we analyzed was the betweenness centrality value of each player. Betweenness centrality is a value that reveals how important each individual vertex is in completing paths between other pairs of vertices in the network. The betweenness centrality for a vertex  $u$  of digraph  $G$  is calculated through the equation

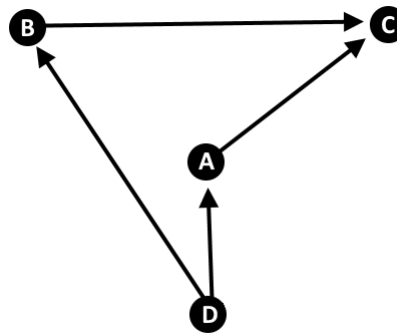
$$BC(u) = \sum_{v_i \in V(G)} \frac{p_{v_i v_j}(u)}{p_{v_i v_j}} \tag{4}$$

where  $u$ ,  $v_i$ , and  $v_j$  are distinct vertices in a digraph  $G$ ,  $p_{v_i v_j}$  is the total number of shortest paths between  $v_i$  and  $v_j$ , and  $p_{v_i v_j}(u)$  is the total number of shortest paths between  $v_i$  and  $v_j$  that pass through vertex  $u$  (Einarsson 2015).

When looking for betweenness centrality in a game of soccer, the value essentially indicates how important any given player is in completing passing plays. Players that have a high betweenness centrality are therefore identified as players who's removal from the network would result in greater difficulty for other pairs of players to pass the ball between each other (Lopez Pena and Touchette 2012).

Figure 5 sets up an example of calculating betweenness. In order to find the betweenness for a vertex, we must look at the number of shortest paths between all other vertices not including the desired vertex. Consider the vertex  $A$ . To calculate the betweenness of  $A$ , we must find the number of shortest paths between  $B - C$ ,  $C - B$ ,  $B - D$ ,  $D - B$ ,  $C - D$ , and  $D - C$ . We then account for the number of shortest paths between each pair, noting if any of these shortest paths include  $A$ . Note that the direction of arrows does matter in this case (as some relationships may not even exist, such as  $C - B$ , where no arrow ever leaves  $C$  and therefore the number of shortest paths ends up being 0). What we find for this example is that there are 4 total shortest paths and that only 1 of these includes vertex  $A$ .  $B - C$

Figure 5: An example showcasing that the betweenness for vertex A is  $\frac{1}{4}$ .



has 1 shortest path ( $B$  to  $C$ ),  $D - B$  has 1 shortest path ( $D$  to  $B$ ), and  $D - C$  has 2 shortest paths ( $D$  to  $B$  to  $C$  as well as  $D$  to  $A$  to  $C$ ). We make note of that last path since it contains  $A$ . Now, we simply follow the formula, resulting in 1 shortest path containing  $A$  divided by 4 shortest paths total. Thus, we find that the betweenness of vertex  $A$  is  $\frac{1}{4}$ .

When researching passing networks, most outside studies discussed the idea of betweenness centrality and the equation behind it, but did not go into specifics of how they personally applied these to the passing network of a soccer game. This was one of the main motivations behind the creation of our small analysis application. To calculate this, we used the array of passing data between all players and iterated over all permutations of two player combinations that did not include our designated player (as noted in the formula for betweenness centrality). Taking a pair, we used Dijkstra’s shortest path algorithm to find and store the shortest path between these two vertices (Dijkstra 1959). Having found a shortest path, we then altered the passing network itself, removing an edge if the shortest path we had found traveled over it. We then repeatedly used Dijkstra’s algorithm, reducing the number of edges in the path until we found that the graph was disconnected and that our two designated players could not actually complete a pass in this new network. Throughout all of this, we accounted for which of these shortest paths contained the player ID that we were finding the betweenness for as well as the total number of paths found. This was done for all pairs, adding to the number of paths that contained the test vertex as well as adding to the total number of paths. These two values then allowed us to produce the betweenness centrality for the player based off of Equation 4.

## 5 Results

Using the Magglingen2013 dataset, we were able to apply all of the concepts above into our analysis application in order to produce results that gave further insight into both individual players on the team as well as the team’s passing network as a whole. We started by replicating the graphs present in the papers by Buldú and by Gonçalves (see Figures 6 and 7) (Buldú et al. 2018; Gonçalves et al. 2017).

As previously mentioned, each player is represented as a vertex and placed at their average position for the whole match, while the edges connecting each pair of vertices represent a passing relationship. Thickness represents the volume of passes that occur between the connected players.

Table 1: A subset of results from the Magglingen2013 dataset

TEAM	ID	IDC	ODC	CLOSENESS	AVG X	AVG Y	BETWEENNESS
LEFT	11	13	9	0.00586	11.85	0.72	0
LEFT	12	20	11	0.00549	9.28	-3.68	0
LEFT	15	10	15	0.00429	-20.33	-1.05	0
LEFT	16	18	21	0.00675	-7.93	5.12	0.09
LEFT	19	19	23	0.00689	-8.83	-6.46	0.36
LEFT	20	8	14	0.00664	-2.99	9.56	0
LEFT	22	21	16	0.00729	7.03	3.82	0
LEFT	23	34	29	0.00873	3.70	0.28	0.55
LEFT	25	15	19	0.0052	-2.42	-10.93	0.09
LEFT	29	17	15	0.00657	3.44	-1.93	0.18
LEFT	30	10	13	0.00713	-3.42	-0.19	0.18

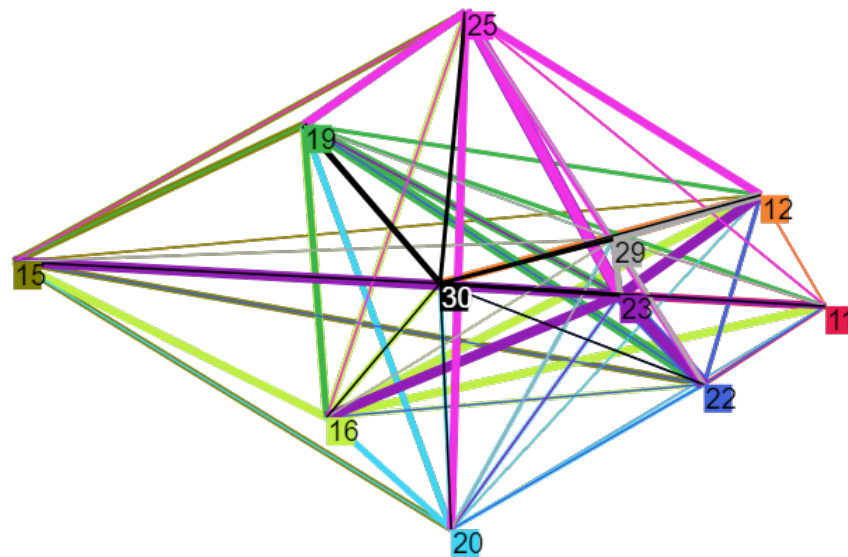


Figure 6: The passing network of the left team from the Magglingen2013 dataset.

Lastly, color indicates which vertex initiated the pass, and is the equivalent of each vertex of the same color having outwards facing arrows pointing towards the player who received the pass. This means that these figures should be viewed in the same way as a traditional digraphs.

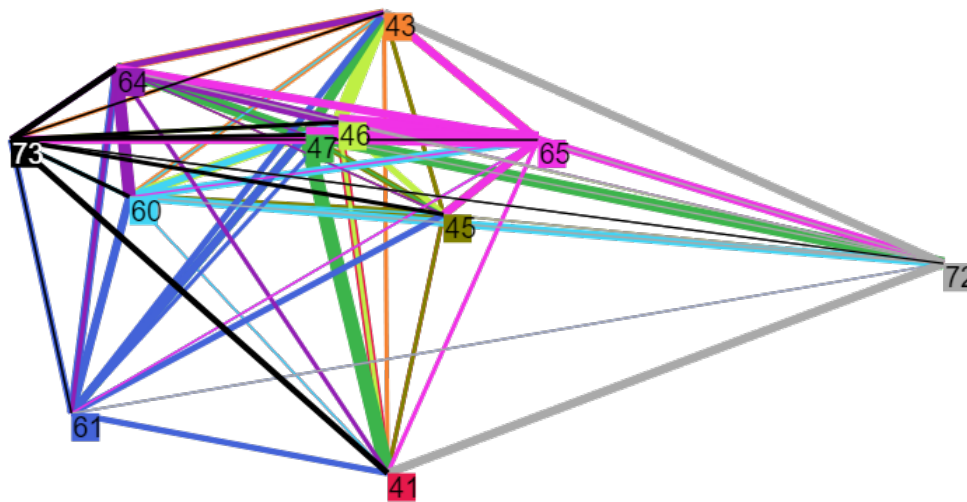


Figure 7: The passing network of the right team from the Magglingen2013 dataset.

As a tool used by either coaches, players, or any outsider hoping to better understand the data at hand, these digraphs (Figure 6 and Figure 7) immediately illustrate connections and relationships between players. Simply from a glance, one can tell which player is doing the brunt of the passes and which players are not. From here, an individual could look into the actual statistics to determine if the thickness of a passing relationship is correct or if there is something amiss between the two players at hand.

Meanwhile, the analysis application also outputs a table’s worth of data that can be evaluated separately from these rendered figures. The left team that participated in the Magglingen2013 dataset can be seen above in Table 1 while the full table can be seen in Table 2 in the Appendix. From this information, it is easy to analyze each player individually and to pinpoint problem areas that might have occurred during a game. For instance, in our subset of players, it is easy to see which players had the most passes received (player 23) or which player made the least number of passes (player 11). Average coordinates can be compared with the average coordinates of that player’s role or position on the field (midfielder, defense, etc.), which can be used to find out whether the player was straying too far from their assigned part of the field. The closeness value for a player can also reveal if that player is straying too far from the play. In the

subset of values seen in Table 1, player 15 has the smallest closeness value indicating that he or she is, on average, the farthest situated teammate throughout the observed portion of the match (however, this particular example makes sense since player 15 happens to be the left team's goalie). Lastly, betweenness can highlight players who are absolutely essential to the passing plays of the team, and special attention can be given to these individuals so that they are not blocked by players on the opposing team, disrupting the initial team's plays and flow. In this subset of data, we see that player 23 has the highest betweenness value at 0.55. Looking back at Figure 6, we also see that player 23 is in the middle of a large passing network, implying that this player is vital in what appears to be offensive plays with the surrounding teammates.

## 6 Open Problems and Future Research

One of the biggest challenges we faced during this research was the lack of publicly available datasets that contained the positional data that we needed. Regarding how this relates to the research, it would have been interesting to accumulate a season's worth of data to compare the variance of passing networks on a game-per-game basis. This idea is further researched by Mendes, Clemente, and Mauricio (Mendes, Clemente, and Mauricio 2018). With more data, it might also be possible to create some confidence interval in order to create a prediction of a passing network for a future game.

Additionally, having more data would allow us to find average values for specific positions or roles. Utilizing more datasets that include more games, one could differentiate positions such as offense, midfielders, or any other position, find average values over a large number of games, and then compare these values to one's own team, seeing if any of his or her players are not playing as well as the rest of the world. Furthermore, since we used position averaging in our closeness analysis, we could also explore different smoothing techniques in calculating the closeness of players.

Another important topic to note is that our research extracted values from one entire half of a match. It would also be beneficial to adjust the range of data that one would put into the analysis application in order to analyze important passing plays, breakaways that led to goals, or moments where the opponents had momentum. By breaking up the analysis into smaller chunks, one could possibly find other noteworthy characteristics. Finally, the research above mostly keeps both teams separate. It would be worthwhile to look into how the two teams interact with one another to see if an overlap in passing networks affects either team.

In conclusion, as data collection and analysis improves, finding innovative and creative methods to analyze and model soccer in order to understand team tactics and characteristics will continue to be useful in the field of sports analytics. Furthermore without applying graph theory concepts, the kinds of statistics that we have found would not be possible to acquire without serious effort. These types of values are seemingly hidden behind the fast-paced nature of the game and are not something one can immediately point out at any given point in a match. By using GPS tracking technology and an accompanying analysis application, these values can be pulled from the dataset at any time and can provide extra insights into how both players and the team performed. We hope that the use of our programs could be modified to help with future analysis of soccer passing networks.

## Acknowledgements

We would like to thank the Sports Special Interest Group of the Mathematical Association of America for supporting and promoting research in mathematics and sports.

## References

- Buldú, Javier M, Javier Busquets, Johann H Martinez, José Luis Herrera-Diestra, Ignacio Echegoyen, Javier Galeano, and Jordi Luque. 2018. "Using network science to analyse football passing networks: dynamics, space, time and the multilayer nature of the game." *Frontiers in psychology* 9:1900.
- Dijkstra, Edsger W. 1959. "A note on two problems in connexion with graphs." *Numerische mathematik* 1 (1): 269–271.
- Duarte, Ricardo, Duarte Araújo, Vanda Correia, and Keith Davids. 2012. "Sports teams as superorganisms." *Sports medicine* 42 (8): 633–642.
- Einarsson, Atli J. 2015. "Analyzing Soccer Strategies Using Network Theory." *Carthage College Mathematics Senior Theses*.
- Gonçalves, Bruno, Diogo Coutinho, Sara Santos, Carlos Lago-Penas, Sergio Jiménez, and Jaime Sampaio. 2017. "Exploring team passing networks and player movement dynamics in youth association football." *PloS one* 12 (1): e0171156.



Lago, Carlos, and Rafael Martin. 2007. “Determinants of possession of the ball in soccer.” *Journal of sports sciences* 25 (9): 969–974.

Laschober, John. 2019. “Soccer-Passing-Analysis (GitHub Repository). <https://github.com/johnlaschober/Soccer-Passing-Analysis>.” Accessed April 14, 2019. <https://github.com/johnlaschober/Soccer-Passing-Analysis>.

Lopez Pena, Javier, and Hugo Touchette. 2012. “A Network Theory Analysis of Football Strategies.” *Sports Physics: Proc. Euromech Physics of Sports Conference*: 517–528.

Lusher, Dean, Garry Robins, and Peter Kremer. 2010. “The application of social network analysis to team sports.” *Measurement in physical education and exercise science* 14 (4): 211–224.

Mendes, Bruno, Filipe Manuel Clemente, and Nuno Mauricio. 2018. “Variance in prominence levels and in patterns of passing sequences in elite and youth soccer players: a network approach.” *Journal of human kinetics* 61 (1): 141–153.

Peixoto, Diogo, Gibson Moreira Praça, Sarah Bredt, and Filipe Manuel Clemente. 2017. “Comparison of network processes between successful and unsuccessful offensive sequences in elite soccer.” *Human Movement* 18 (5): 48–54.

Rein, Robert, and Daniel Memmert. 2016. “Big data and tactical analysis in elite soccer: future challenges and opportunities for sports science.” *SpringerPlus* 5 (1): 1–13.

Ribeiro, Joao, Pedro Silva, Ricardo Duarte, Keith Davids, and Júlio Garganta. 2017. “Team sports performance analysed through the lens of social network theory: implications for research and practice.” *Sports medicine* 47 (9): 1689–1696.

Rumo, Martin. 2014. “Maggingen2013. <https://old.datahub.io/dataset/maggingen2013>.” Accessed April 14, 2019. <https://old.datahub.io/dataset/maggingen2013>.

Yue, Zengyuan, Holger Broich, Florian Seifriz, and Joachim Mester. 2008. “Mathematical analysis of a soccer game. Part I: Individual and collective behaviors.” *Studies in applied mathematics* 121 (3): 223–243.

## Appendix

Table 2: All outputted values from the Maggingen2013 dataset

TEAM	ID	IDC	ODC	CLOSENESS	AVG X	AVG Y	BETWEENNESS
LEFT	11	13	9	0.00586	11.85	0.72	0
LEFT	12	20	11	0.00549	9.28	-3.68	0
LEFT	15	10	15	0.00429	-20.33	-1.05	0
LEFT	16	18	21	0.00675	-7.93	5.12	0.09
LEFT	19	19	23	0.00689	-8.83	-6.46	0.36
LEFT	20	8	14	0.00664	-2.99	9.56	0
LEFT	22	21	16	0.00729	7.03	3.82	0
LEFT	23	34	29	0.00873	3.70	0.28	0.55
LEFT	25	15	19	0.0052	-2.42	-10.93	0.09
LEFT	29	17	15	0.00657	3.44	-1.93	0.18
LEFT	30	10	13	0.00713	-3.42	-0.19	0.18
RIGHT	41	20	16	0.01118	0.09	4.19	0
RIGHT	43	21	14	0.01389	0.01	-5.46	0
RIGHT	45	15	15	0.01484	1.27	-1.23	0
RIGHT	46	32	39	0.01803	-0.94	-3.16	0.25
RIGHT	47	25	27	0.01815	-1.63	-2.90	0.45
RIGHT	60	23	15	0.01374	-5.33	-1.60	0.07
RIGHT	61	11	13	0.00928	-6.55	2.94	0.07
RIGHT	64	27	24	0.01389	-5.60	-4.30	0.14
RIGHT	65	16	23	0.01303	3.25	-2.79	0.29
RIGHT	72	4	5	0.0064	11.76	-0.20	0
RIGHT	73	14	17	0.012	-7.82	-2.80	0.29