A NETWORK ANALYSIS OF THE 2010 FIFA WORLD CUP CHAMPION TEAM PLAY*

COTTA Carlos · MORA Antonio M · MERELO Juan Julián
· MERELO-MOLINA Cecilia

Abstract This paper analyzes the network of passes among the players of the Spanish team during the last FIFA World Cup 2010, where they emerged as the champion, with the objective of explaining the results obtained from the behavior at the complex network level. The team is considered a network with players as nodes and passes as (directed) edges. A temporal analysis of the resulting passes network is also done, looking at the number of passes, length of the chain of passes, and to network measures such as player centrality and clustering coefficient. Results of the last three matches (the decisive ones) indicate that the clustering coefficient of the pass network remains high, indicating the elaborate style of the Spanish team. The effectiveness of the opposing team in negating the Spanish game is reflected in the change of several network measures over time, most importantly in drops of the clustering coefficient and passing length/speed, as well as in their being able in removing the most talented players from the central positions of the network. Spain's ability to restore their combinative game and move the focus of the game to offensive positions and talented players is shown to tilt the balance in favor of the Spanish team.

Key words Complex network, complex systems, FIFA World Cup, football, La Roja, network analysis, soccer, Spanish team, sports.

1 Introduction

The hypothesis that a complex network analysis can help to understand football matches has been present for some time now[1]. Several teams, formally or informally, have performed analyses of football matches from the point of view of the network of passes formed throughout the match. One starting point was in 2004, when a competition to predict the four best-classified teams in the EuroCup that was celebrated that year was held in the Redes (Spanish word for "network") social-network mailing list (in Spanish). The results, which completely failed to predict the outcome, were published in [2].

COTTA Carlos
Department of Computer Science, University of Málaga, Spain. Email: ccottap@lcc.uma.es.
MORA Antonio M · MERELO Juan Julián · MERELO-MOLINA Cecilia
Department of Architecture and Computer Technology, University of Granada, Spain.
Email: amorag@geneura.ugr.es; jmerelo@geneura.ugr.es.

The main problem with these predictions –besides the outcome which was completely wrong (not even the two finalists, Portugal and Greece, were included in any of them) — was that they were looking at the static picture of the team as it emerged from the previous match. It is quite clear that football is a game of two teams, whose networks clash. While there must be some quantity or structural property reflecting the team's organization and playing style that is maintained from one game to the next, the other team will do its best to prevent that network to move information (namely, the football) from one part to other, resulting in a quite unpredictable result.

On the other hand, the network does not have any kind of spatiotemporal information. A network might show what one would consider a perfect structure, well formed, with short distances from goalie to forward players, but if it plays out of place or simply in its own field it will not be able to obtain a good result. Furthermore, if the network develops quite slowly with a low number of passes and low precision (low transitivity), the result will not be good either. In both cases, the static structure, while meaningful and a good qualitative description of the overall game, is unable to reflect its dynamics.

In this paper, spurred on by the victory of the Spanish selection in the 2010 World Cup, we have performed a spatiotemporal analysis of the essential games that led to victory. In this analysis, we have looked at the temporal evolution of the number of passes, at the length of the chains of passes and its transitivity (thus taking into account the effect on the opponent of simple ball losses) and analyzed the team network (considering pitch zones) at a microscopic level by identifying the most central nodes. In this way, we take into account both the complex network structure (reflected in the power-law structure of the length of the number of passes and overall network structure[3]) and the spatiotemporal nature of the game. In that sense, this paper is the first to do this kind of analysis (at the complex network level — see also [4, 5] for related work), which can later be complemented with other kinds of static micro-, macro-, and meso-measurements of the same type.

Since we are talking about sports here, and football which is claimed to be the king of all sports, in some cases we allow ourselves to get carried away by emotion and make some statements that are not purely scientific, but more in the spirit of our passion for the sport (and, even more so, for the team we are talking about). Even so, as the scientific claims we make are sufficiently supported by our data, we ask for some leniency from the reader for our language and context statements, which we feel do no harm to the scientific claims we make in this paper.

The rest of the paper is structured as follows: next, we examine the state of the art in analyses of the outcome of football matches. Then, the methodology used to extract data from the match is presented. An overall examination of the matches played by the Spanish team is subsequently performed. We close the paper with some conclusions and guidelines for future work.

2 State of the Art

Despite the huge cultural and popular interest that football has, being arguably the most popular sport (or maybe spectacle) in the world, there have not been many scientific approaches to sport performance and prediction. This was true when Onody and Castro[6] wrote their often-referenced paper in 2004, and it is still true today.

However, since then, several papers have tried to apply complex network analysis to the footballing world. The above-mentioned paper itself was seminal in its thoroughness: it analyzed the network of all Brazilian football players, and linked them if they had shared a team,

and found that several metrics (number of teams per player, number of goals per player, number of games played) follow truncated power-laws[†] or exponential distributions. This power-law behavior present in football was, later on, confirmed by Yamamoto and Yokoyama, who performed an analysis quite similar to the one we complete in this paper, analyzing several teams playing in the 2006 FIFA World Cup[7]. Their analysis detected power laws in the passes network, with exponents around 3, and made a very interesting analysis of the temporal evolution of the competitive play by looking at the number of triangles, or transitive passes, among other things. Duch, et al.[8] proved that flow centrality of the pass network can adequately qualify team and player performance; in this they coincide with the results obtained by Lee, et al.[9], who showed a moderate relationship between the flow/betweenness centralization of the whole team and the outcome, as well as a negative relationship between the out-degree centralization.

However, there has been no attempt to relate any of those quantities to performance. Could the number of teams a player has played in be related to the number of goals? In general, the prediction of performance has concentrated on the hypothesis that a team has some kind of intrinsic fitness — See [10] — the difference of which between teams affects the probability (not the certainty) of one beating the other. These studies have mainly concentrated on time series analysis[11] but not intra-game dynamics.

Since the pass data for several world-class events (Euro Cup 2004, 2008, and World Cup 2006, 2010) were made available, one of the authors has been carrying out informal analysis on the team's networks and deducting from them some kind of qualitative prediction on the result of a match. However, this analysis has not been published, and even as differences between the network qualities of different teams were appreciated, it was difficult to relate them either to the team fitness or to the match outcome. The ARSfútbol team, based in Argentina, has carried out extensive analyses of world-class events as well as local low-level football teams[12], concluding that the performance of a team is mainly related to the existence of a well-coordinated core of players (such as the players which are taken from FC Barcelona, in the Spanish national team or the set of players from Porto FC in the Portuguese selection). Even those post-hoc observations cannot be easily used for predicting performance.

In this paper, we will look at the micro-dynamics of a football team, the Spanish selection during the last World Cup in which they emerged as champions, bringing happiness to millions of Spaniards all over the world. We will try to find out which quantities made the Spanish team excellent by focusing on quantitative analysis of its game-play. This will be used as the first step for a second leg of analysis which will focus on prediction.

3 Data Extraction

In order to collect the data for the analysis, an important issue was the fact that we intended to analyze the game play from a spatiotemporal perspective. The spatial dimension tries to capture the fact that a player can act in different parts of the pitch and his role (and therefore the way he performs and interacts with his teammates) may be different in each of these zones. For example, a certain midfielder can have a defensive role when the opposing team is attacking, and have the teammates in the last defensive line as preferred targets for a pass whenever he regains control of the ball. On the other hand, this very same player can adhere to a much more combinative play when in offensive positions, interacting more with wingers and other creative midfielders. To this end, the pitch has been divided into nine zones as depicted in Figure 1: four zones correspond to defensive positions (own box, wingback lanes, and own midfield), a further four zones correspond to offensive positions (opponent's box and midfield and wingers' lanes),

---

[†]Two quantities x, y are said to be related by a power-law if $y = f(x) \propto x^k$ for some constant k[13].

and the the very central zone of the pitch which plays an important role for both initiating static attacks and pressing the opponent in defense. For the purpose of this analysis, each player is figuratively divided into 9 different players depending on the pitch zone in which they touch the ball (of course, most players only visit a limited number of zones throughout any given game).
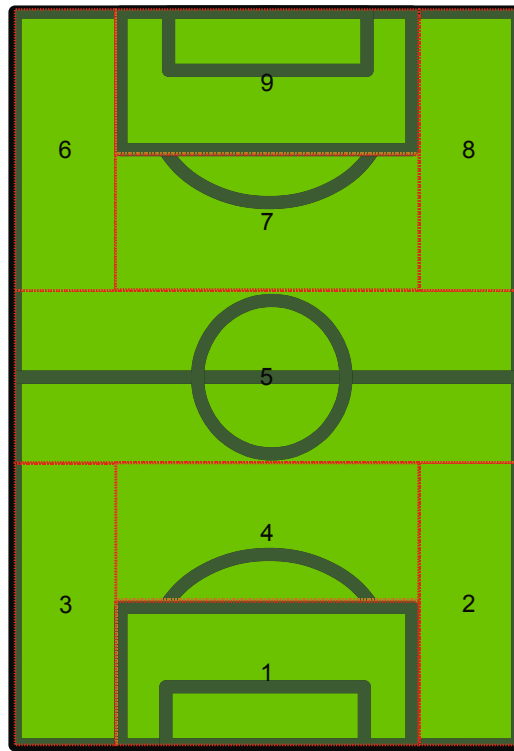

Figure 1 Model of the football field for data extraction, playing zones

As for the temporal perspective, our goal is to capture the fact that a football game can go through different phases in which a team can change its way of playing (e.g., the dominant team can become dominated even if just sporadically or the trainer may introduce tactical changes) or at least can change its effectiveness (e.g., a player that was nullified by an effective defense can resurface later on, when tiredness prevents the defenders keeping tight marks). To account for this, we also keep track of the minute in which each pass was made.

Once the kind of data we needed was defined, data extraction was done by ourselves, reviewing the knock-out games played by Spain (contrarily to regular group games in which a team can speculate with a draw or even with a somehow minimal loss in order to qualify for the next phase, knock-out games imply a win-lose situation and therefore the playing style and tactics are effectively directed at winning the game without depending on external factors[‡]). These

---

[‡]Winning the game must be interpreted here as progressing on to the next knock-out stage rather than scoring more goals than the opponent. For example, a team may decide to play very defensively in order to reach the penalty shootout if they consider they will be outplayed otherwise; be that as it may, this reflects a definite and valid style, and is anyway not applicable to the games considered in which Spain went all out to

correspond to the games against Portugal (round of 16) (not analyzed in this paper), Paraguay (quarterfinals), Germany (semifinals), and the Netherlands (final). The raw data obtained from this visual inspection consists of a list of all passes in the match, with the format

<div align="center"><half> <minute> <player> <zone> <player'> <zone'></div>

indicating which player passed to which player, in which minute, and in which pitch zones each of them were in. As mentioned before, each pair (player, zone) can be interpreted as a virtual player for the purposes of network construction. More precisely we consider a moving window of 15 minutes (which we believe is long enough to capture the state of the game at any given instant) and build a series of directed graphs $G_i(V, E_i)$ where $V = \{(p, z) \mid p \in T,\ z \in \{1, 2, \cdots, 9\}\}$, T is the set of actual players in the national team, and $(u \rightarrow v) \in E_i$ if, and only if, virtual player u passed to virtual player v within the i-th time window. Figure 2 shows an example of the networks obtained.
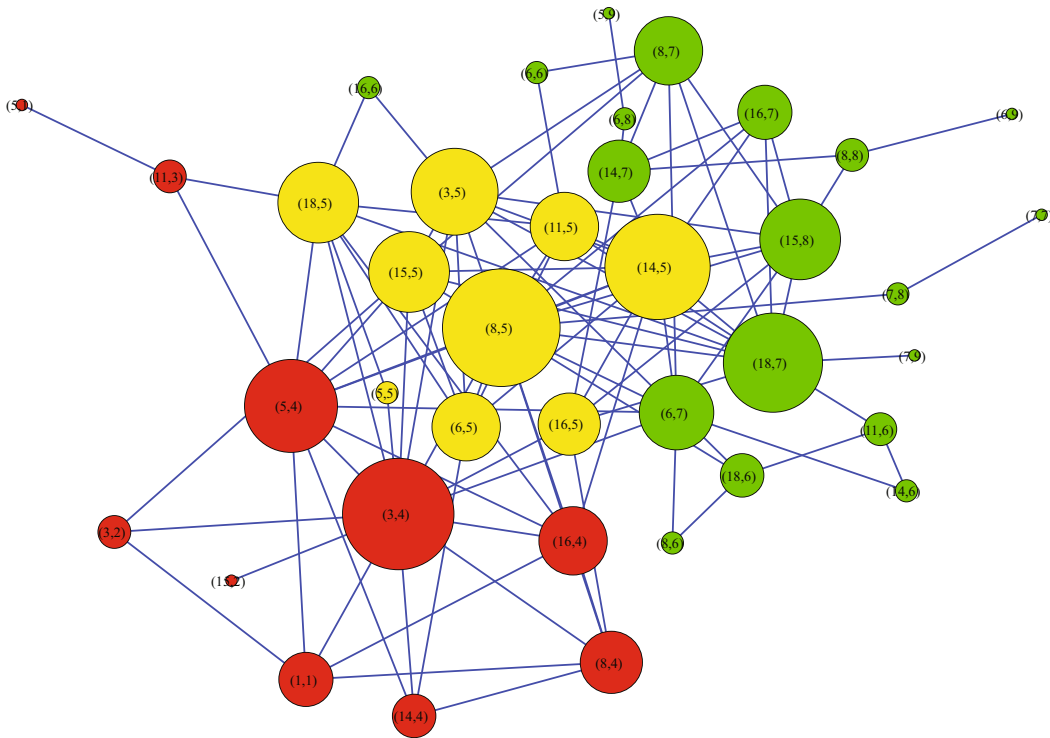

Figure 2 An example of the networks obtained, taken from the Germany-Spain match (15', 1st half). Each node is labeled with a pair (p, z) where p is the squad number of the player — see Appendix — and z is the zone of the pitch. Only those nodes corresponding to players/zones intervening in the game during the particular time window considered are displayed. Node sizes depend on degree centrality and the different shades correspond to defense (zones 1–4), midfield (zone 5) and attack (zones 6–9)

This data was extracted by visualizing the matches by the authors, which has been found to be an appropriate, and probably the only, method for gathering network data from a match[14]; data collection took approximately 4–5 times the duration of matches[§].

effectively win the game before penalty shootouts.

[§]This data is available upon request to the scientific community, with the provision that this paper is referenced.

4 Analyzing the Spanish Team

The Spanish national team entered the World Cup 2010 as one of the favorites to win the trophy. While to some extent this was not a completely new situation from a local point of view (and even a scientific point of view: Spain hosted the World Cup in 1982, and New Scientist published an article proving why they should win due to home advantage[15]), in this case this consideration was quite reasonable a priori, since Spain was the current incumbent in Europe after winning the Euro Cup in 2008 for the 2nd time in their history (the 1st time dating back to 1964 in the Euro Cup hosted precisely in Spain). This win at the European level (and even more precisely the win over Italy in penalty shootouts in the quarterfinals game) was seen by many as a turning point from the psychological point of view, finally breaking the "curse of the QFs" that had been inescapable for several other — very talented — Spanish teams. On reflection, the current generation of players was probably the most perfectly suited to this purpose, since they had been surfing on a winning wave in the national team since junior stages (Íker Casillas, Xavi Hernández, and Carlos Marchena were members of the team that won the Under-20 World cup in 1999) and were also part of the last dominant cycles of Real Madrid CF and FC Barcelona at club level (holding numerous National, European, and International trophies).

The lack of this "choking culture" that had haunted preceding generations, and finally the discovery of a playing style that players felt their own (incidentally a style that departed by a long way from the traditional "Spanish fury" that had been the national team's trademark until this generation, and arguably a finding that was in part possible due to the unfortunate series of injuries of several players) led to the current success cycle of the team. While this style –informally termed "tiki-taka"[16] (apparently, an otherwise meaningless phrase that can be translated as "touchy-touch", but which actually is very close to Basque "tipi tapa" meaning step by step) by former coach of the Spanish national team Javier Clemente in a pejorative sense and later popularized by the late Andrés Montes, a famous TV commentator, in a more neutral sense– is now considered a trademark by many in Spain, and is commonly associated with that used by FC Barcelona (a claim based on the number of players of this club in the national team), a deeper inspection indicates that there are stylistic differences not only between the national team and other clubs in Spain, but also between the Euro 2008 team and the World Cup 2010 team. The former had players such as Fernando Torres (then at Liverpool FC) or David Silva and David Villa (then at Valencia CF) at their highest performance level and played a more dynamic, faster combinative football. The latter team had several of these players at a lower performance level, some key players in midfield or defense were obviously two years older and other players more suited to passing play were incorporated. Indeed, while FC Barcelona players were a minority in the Euro 2008 team (just 3, outnumbered for example by the 4 players contributed by Valencia CF or Liverpool FC), they were more numerous in the World Cup 2010 team. The result was possession-based style in which the ball often moves rapidly but in which the net advance is slow (yet steady). In the following we will analyze some of the characteristics of this style using the network information collected as indicated in Section 3. In particular, we will pay attention to the following features:

1) Number of passes per minute: this provides an indication of the speed of the game.

2) Number of consecutive passes without losing the ball: this reflects the elaborateness of the offensive game of the Spanish team.

3) Clustering coefficient[17]: this measure is an indication of the extent to which players tend to cluster together when passing the ball. To compute this coefficient we symmetrize the network and compute the local clustering coefficient γ_v for each node v with more than two

vertices as

$$\gamma_v = \frac{N_v}{\delta_v(\delta_v - 1)/2},$$

where δ_v is the degree of node v and N_v is the number of links between nodes u, w connected to v. Subsequently, the clustering coefficient γ_G of the network G is

$$\gamma_G = \frac{\sum\limits_{v \in V \setminus V^{(1)}} \gamma_v}{\left| V \setminus V^{(1)} \right|},$$

i.e., the average clustering coefficient for all nodes with more than one neighbor — see [18].

4) Centrality: we analyze the network at a microscopic level trying to identify the most important nodes in two ways, by considering the number of passes (weighted degree centrality) and by considering the flow of the ball. Regarding this latter aspect, our analysis tries to identify which player (and in which pitch zone) is more likely to have the ball after a long sequence of passes. This is analogous to eigenvector centrality[19] computed using the power iteration method, i.e., converting the adjacency matrix in a Markov matrix and iterating from an initial vector representing a uniform distribution among all players intervening in the game; this process provides the probability of reaching a certain player by following an arbitrarily long random walk in the graph.

For the convenience of the reader, we include in Table 1 the final result and scorers of the matches considered.

Table 1 Matches analyzed in this work. For the sake of completeness we also include the final score and the player/minute in which goals were scored

Match	Result	Scorer
Paraguay-Spain	0-1	Villa (83′)
Germany-Spain	0-1	Puyol (73′)
Netherlands-Spain	0-1	Iniesta (116′)

4.1 Quarter Final Match: Paraguay-Spain

Spain reached quarterfinals after winning 1-0 (a result that would repeat itself in all remaining games) against Portugal. While the result of that last game was tight (and whether Villa, the scorer of the only goal was in off side is open to some question), from a global point of view Spain was widely regarded as having dominated the game and deservedly progressing to the quarterfinals, the historically insurmountable barrier for the national team. The game was approached by Paraguay (a strong opponent, as demonstrated by their being runner-ups in the Copa América 2011) from a tactical premise: to disrupt the flow of the ball among Spanish players. They succeeded in this for most of the game, with Spain hardly achieving more than 3–4 consecutive passes (see Figure 3–left side graphs) and only attaining a moderate pace of about 6 passes per minute (see Figure 3–right side graphs). Although Spain eventually settled into its usual combinative game (see clustering coefficients in Figure 4), the first half was very irregular for the team. In general, the clustering coefficient is not much larger than that of an Erdős-Rényi (ER) random network[20] of the same size and number of links, indicating that the interaction among players is not highly structured. No single player emerges as the most central (in terms of in-/out-degree) and the game takes place mostly in Spain's defensive zone (see Figure 5, as well as Figure 6 depicting a snapshot of the passing network during the 1st half). Indeed, an analysis of eigenvector centrality indicates that Sergio Busquets was the most central player during most of the 1st half (see Figure 7). In other words, the structure of passes
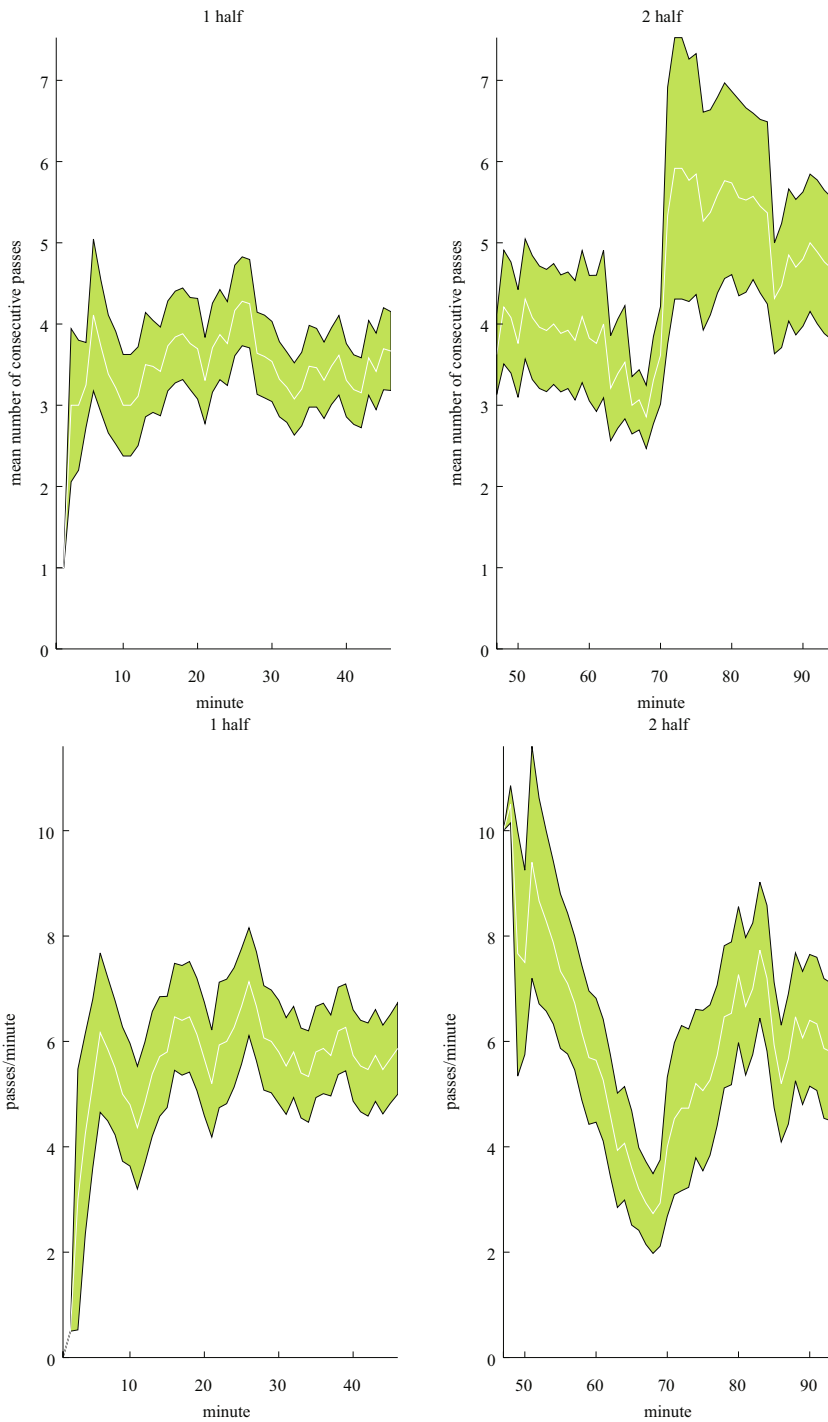
Figure 3   Paraguay-Spain. Upper graphs — Average number of consecutive passes. Lower
graphs — Passes per minute. In this figure and all subsequent ones, the white line
is the mean and the shadowed area covers one standard error of the mean above
and below the former. Data always corresponds to the Spanish national team

of the Spanish team was leading to the less creative midfielder dominatinge the circulation of the ball. This does not mean Spain was necessarily under attack, but that Paraguay's pressing avoided Spain's more creative midfielders being actively engaged in the game and settling the ball in offensive zones. The game's turning point took place when two penalty shots were failed successively first by Paraguay and then by Spain (low valley shown in Figure 3–right graph), after which Paraguay lowered their defensive strength due to tiredness, and Spain could finally enter into longer sequences of passes, with Xavi Hernández and Xabi Alonso emerging as the dominant players in the midfield.
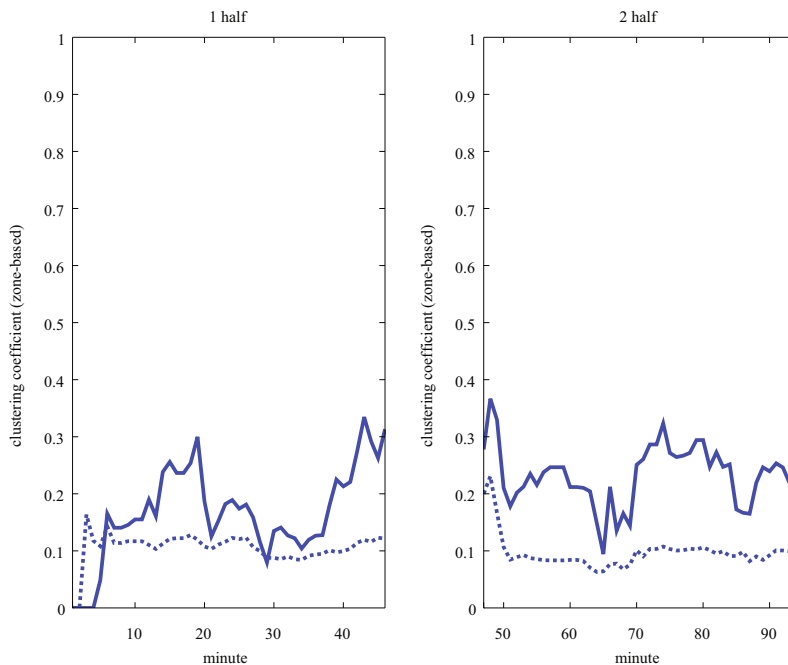
Figure 4 Paraguay-Spain. Clustering coefficient (CR). The dotted line indicates the CR of a random network with the same number of players involved and the same number of links

4.2 Semi Final Match: Germany-Spain

The victory over Paraguay in the quarterfinals meant Spain advanced for the first time in history to the semifinals[¶], and a general consensus that the team had fulfilled the expectations of the supporters. This meant that the semifinal was approached as a game in which there was nothing to lose (although it must be noted that the same could be said for Germany, who had presented a very young team and a new, less physical style, and were in a transitional year). Maybe due to this lack of pressure, Spain gave their best performance of the tournament.

The game was thoroughly dominated by Spain, who managed to engage in a fast game (well above 6 passes per minute, the ceiling reached for the greatest part of the game against Paraguay, cf. Figures 3–bottom row and 8–bottom row) and more elaborated sequences of

---

[¶]The previous best performance of Spain had been a 4th place in Brazil's World Cup 1950, an edition in which there was a small league among the top-four teams.
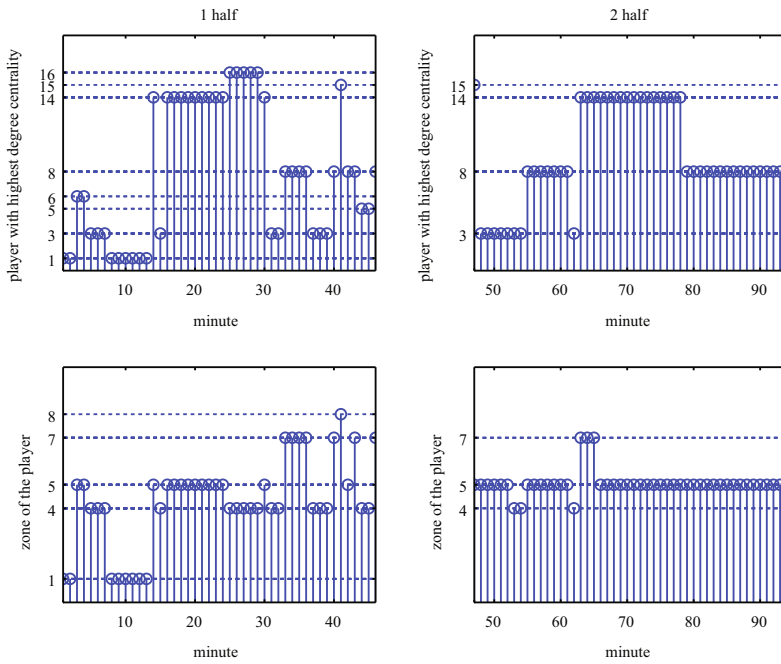
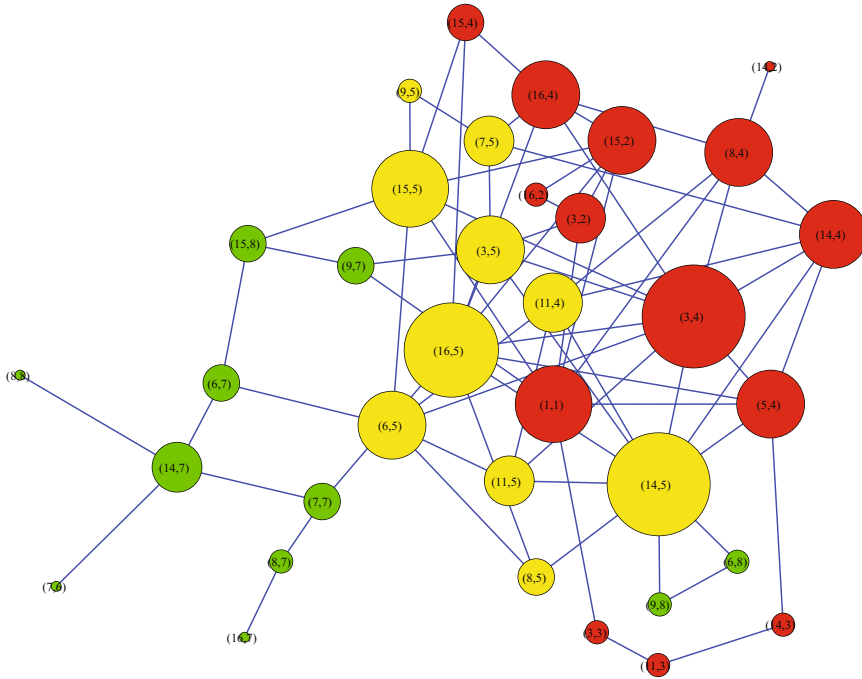**Figure 5** Paraguay-Spain. Zone-based player centrality based on number of passes

Figure 6 A snapshot of the passing network (15', 1st half) taken from the Paraguay-Spain match. Node sizes depend on degree centrality. Note the diminished role of creative midfielders such as Xavi (8) or Iniesta (6), as well as the important role of Busquets (16) in the midfield and Piqué (3) in defensive positions
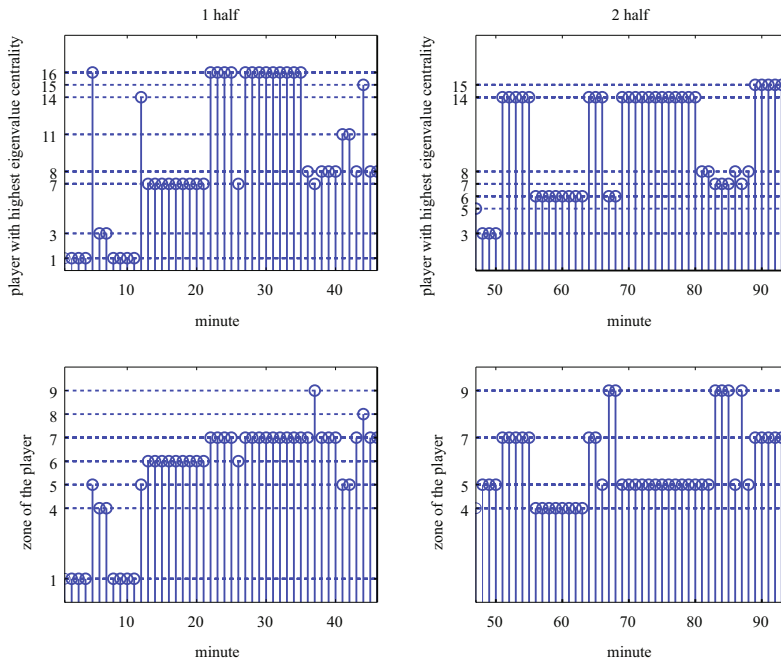
Figure 7 Paraguay-Spain. Zone-based player centrality based on random walks

passes (roughly twice as many as against Paraguay, cf. Figures 3–top row and 8–top row). It is interesting to note that the clustering coefficient was rather stable (the drop near the end of the 2nd half corresponds to the final push exerted by Germany once they were 1 goal down) and similar to that reached against Paraguay but clearly higher than that of the ER network. This indicates the underlying passing pattern was maintained in both games but was more effective — i.e., resulted in longer sequences of passes – against Germany. Note in this sense how such longer sequences of passes are correlated with shots at goal, in particular in successful teams[21]. As shown in Figure 2, the passing network looks more balanced and layered than against Paraguay at the beginning of the match, with a more important role of the most talented players in the midfield and a notable presence in offensive positions. One of the tactical surprises of the match was the inclusion in the starting eleven of Pedro Rodríguez as a false winger, playing in-between the German defense and midfield. As seen in the eigenvector centrality analysis (Figure 11), he was instrumental in the circulation of the ball during long phases of the match. Xabi Alonso and specially Xavi Hernández during the 2nd half were the clear hubs of the game (Paraguay had partly succeeded in disabling these two players, hence the more disconnected game). Actually the 2nd half saw the ball parked deeply inside the German field (Figure 10–bottom, right), indicating a clear domination of the game: sequences of up to 24 passes took place before Spain's goal (which paradoxically was a header after a corner kick).

4.3 Final Match: Netherlands-Spain

The great final took place against the Netherlands, a very talented team that featured one of the best players of the season in the midfield (Wesley Sneijder, surprisingly left out of the Ballon d'Or contest in that year despite being instrumental in his team – FC Internazionale Milano–
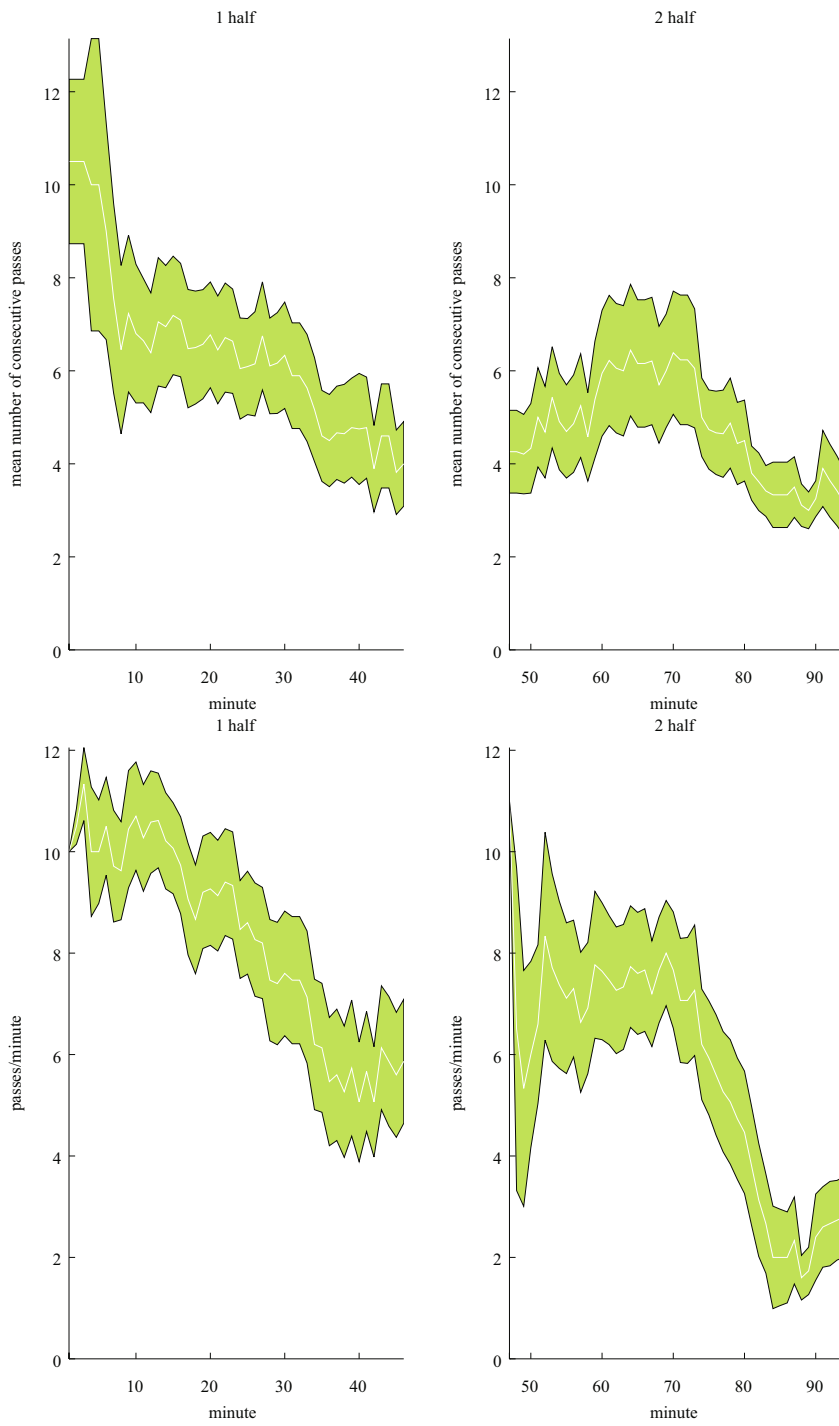
Figure 8 Germany-Spain. Upper graphs — Average number of consecutive passes. Lower graphs — Passes per minute
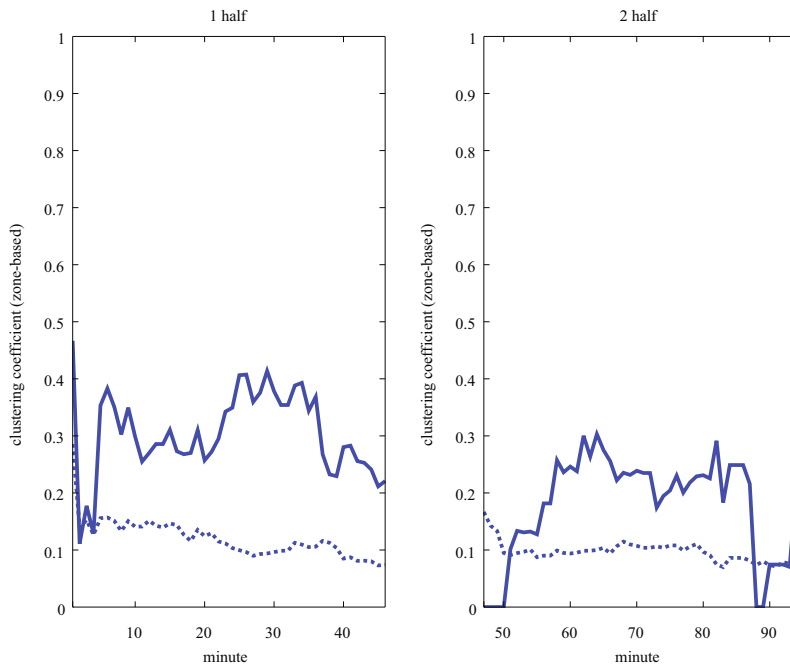
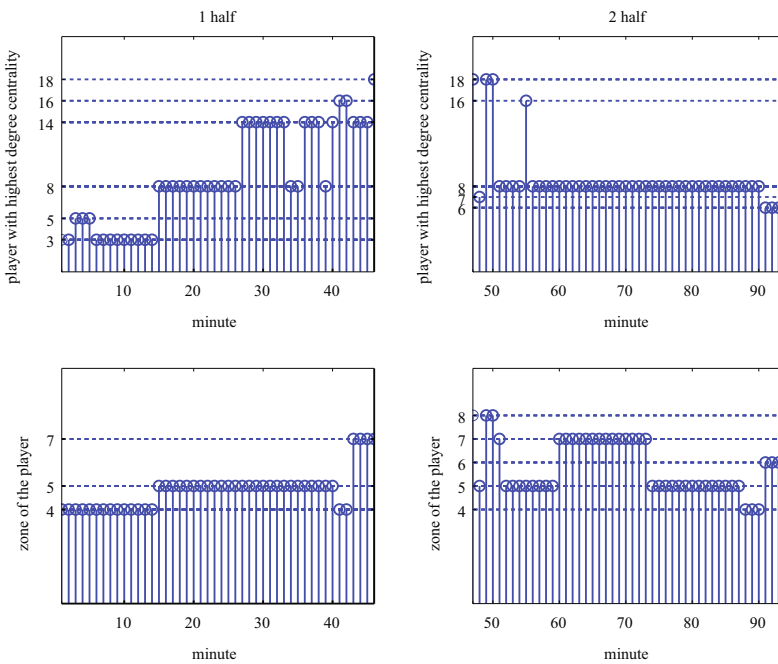**Figure 9** Germany-Spain. Clustering coefficient

Figure 10 Germany-Spain. Zone-based player centrality based on number of passes
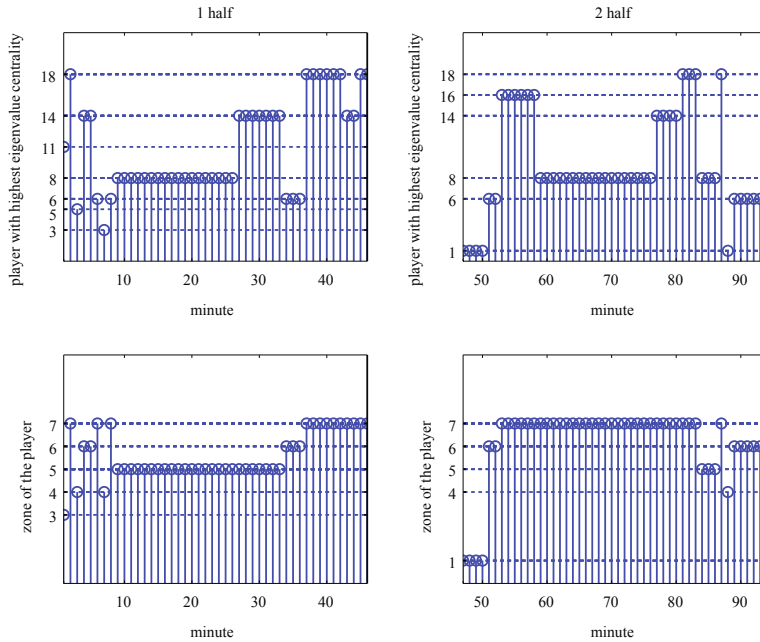
Figure 11 Germany-Spain. Zone-based player centrality based on random walks

achieving the treble) and possibly the fastest and finest winger in the whole tournament (Arjen Robben). Indeed, the connection between these two players could have very well tipped the balance in favor of Netherlands had not Íker Casillas provided an outstanding performance. Despite the talent available on the pitch, the Dutch team engaged however in a very rough game, forever symbolized by Nigel de Jong's ominous tackle on Xabi Alonso's chest[22]. To some extent the tactical objective was similar to that of Paraguay's — disrupt the smooth flow of the ball in the Spanish team — yet the execution was much more questionable (and so was the condescending attitude of the referee towards it). The game settled for the most part into a very slow pace due to the number of interruptions (Figure 13), and sequences of passes were generally short, analogously to the game against Paraguay (Figure 12). Unlike the latter game, the Dutch team succeeded in disabling Xabi Alonso but the balance of the game followed a similar pattern in that the physical style of the Netherlands became less consistent as time passed. The midfield was eventually controlled by Xavi Hernández (Figures 15 and 16) and — despite two clear fast breaks by Robben mentioned before — the ball advanced away from Spain's defensive zone. The average number of consecutive passes gained momentum at the end of the 2nd half and was rather stable during the 1st half of extra time. Not surprisingly, the pace of the game decreased steadily during the extra time (Figure 13–right side), due to both the exhaustion of the players and the team's strategy shifting to a safer mode of holding possession in order to avoid conceding a late goal and trying to exploit a more direct approach when possible. Indeed, the clustering coefficient during the 2nd part of the extra time is much lower than in remaining games (Figure 14–bottom), which is consistent with a less structured development of the game and larger distances between team lines. The whole extra time witnessed the emergence of Cesc Fàbregas (Figures 17 and 18), coming from the bench to dominate the midfield, eventually providing the assist that resulted in Andrés Iniesta's decisive goal.
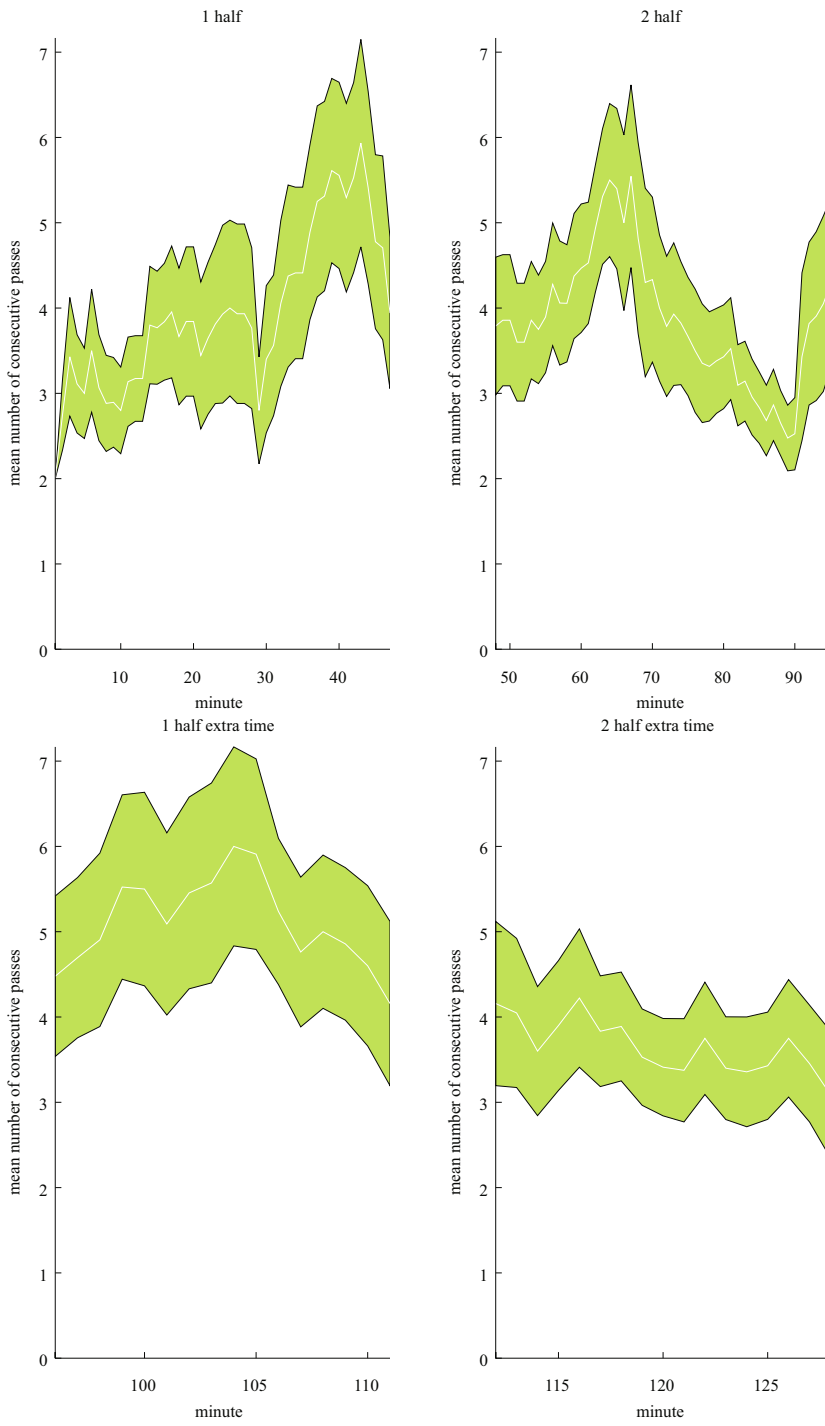
Figure 12 Netherlands-Spain. Average number of consecutive passes in the match time (top) and in the extra time (bottom)
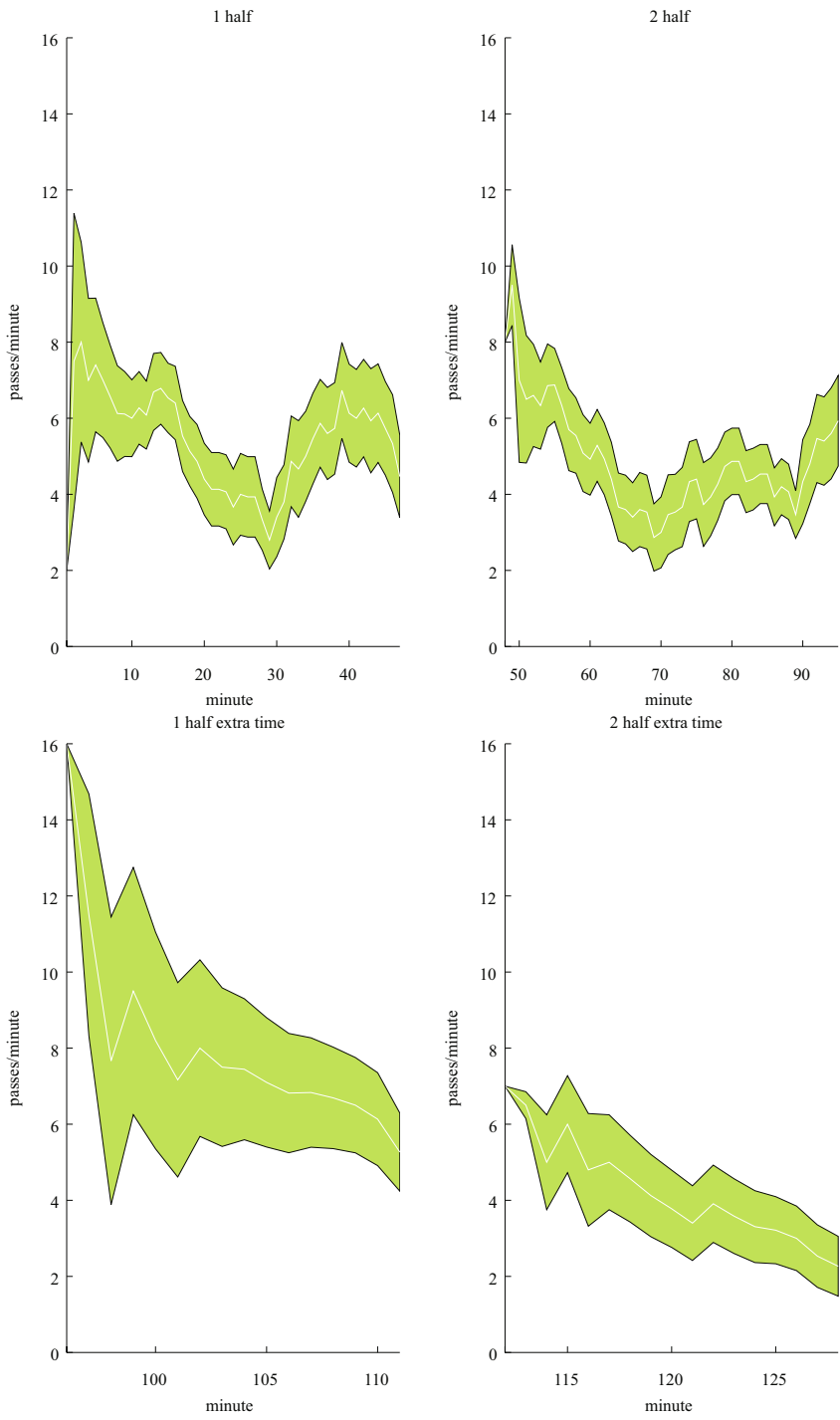
Figure 13 Netherlands-Spain. Passes per minute in the
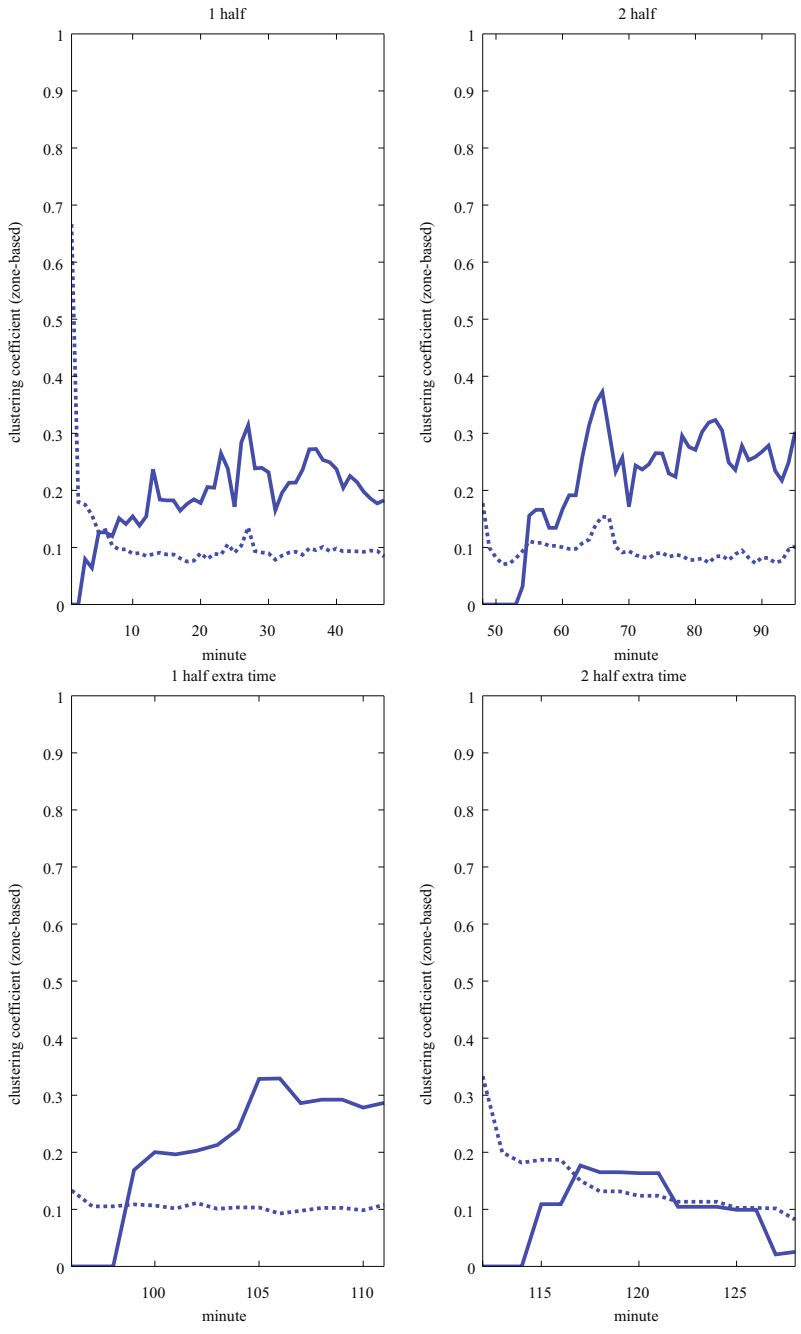match time (top) and in the extra time (bottom)

Figure 14 Netherlands-Spain. Clustering coefficient in the match time (top) and in the extra time (bottom)
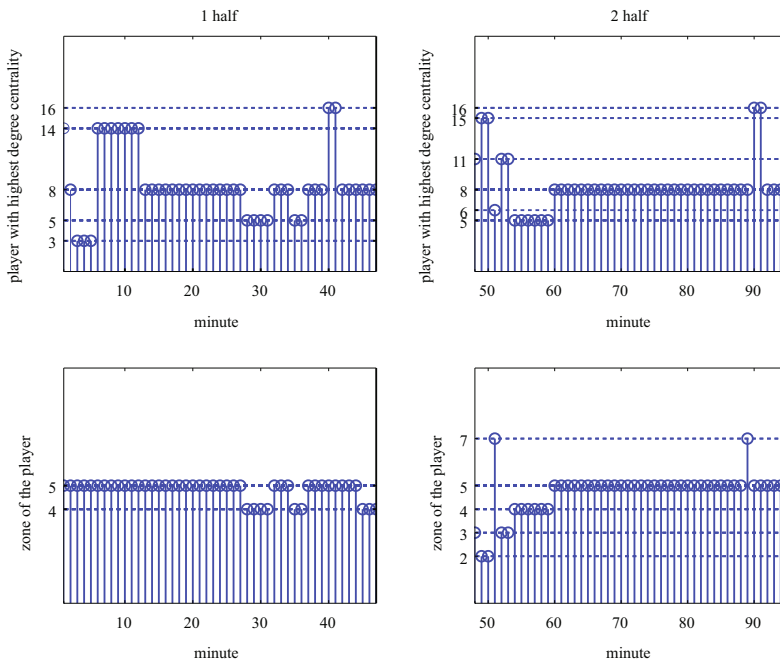
Figure 15 Netherlands-Spain. Zone-based player centrality based on number of passes
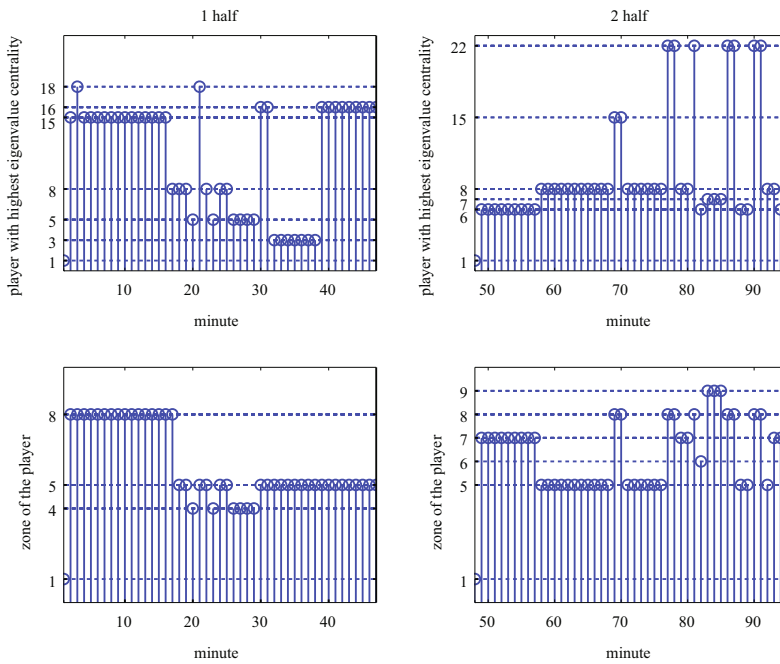

Figure 16 Netherlands-Spain. Zone-based player centrality based on random walks
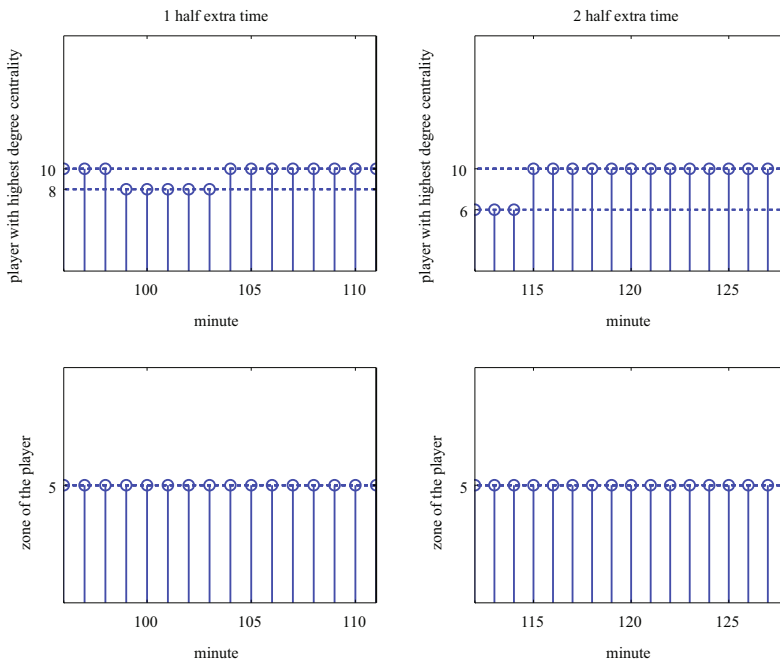
Figure 17 Netherlands-Spain. Zone-based player centrality
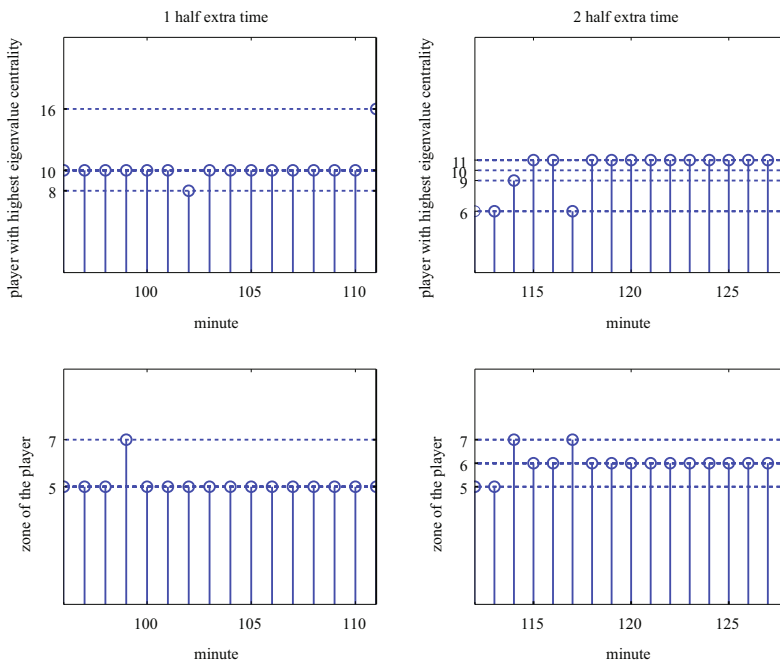in the extra time based on number of passes

Figure 18 Netherlands-Spain. Zone-based player centrality
in the extra time based on random walks

5 Conclusions and Future Work

This paper has attempted to use some graph and network metrics to analyze the performance and playing style of the Spanish national football team in the FIFA World Cup 2010. When analyzed from a temporal perspective, global measures such as the number of consecutive passes or the number of passes per minute provide a measure of the success of the team in imposing its style (or alternatively the success of the opposing team in disrupting the style of the Spanish team). A deeper insight is obtained by observing the clustering coefficient which captures the combinative nature of the "tiki-taka" style, and hints that the passing network exhibits the small-world property. While prone to a kind of baroqueness in the sense that many passes are often made in a short distance (contributing to the Spanish team's having the highest percentage of completed passes in the tournament: 87%[23]) and might be expendable, it is evident that it has prime value as a defensive strategy, by depriving the opponent of ball possession, one of the factors that is determinant in the game's outcome[24,25]. Even in the games in which the performance of the Spanish team has been deemed worse by analysts and the general public, this combinativeness imprint globally remains, with marked valleys in the aftermath of scoring a goal (and the subsequent push of the trailing team) and sporadically during the final, a special game in which emotions and passion often led to lower precision and less elaborated game-play.

There is of course much work to be done. From a purely methodological point of view, it is of crucial importance to keep analyzing other network measures such as node and edge betweenness, which naturally capture the hubs and essential associations in the distribution of the ball (let us note en passant here as an example that the role of Sergio Busquets was initially much criticized by the press and a large part of the public due to his lower creativeness; however it is clear that he was instrumental in providing balance to the midfield; this balance should not be measured in terms of outstanding possession of the ball, a task for which more creative midfielders should be responsible, but more likely for his contribution to the robustness of the network). This can have not just an explanatory value, but also a predictive value in terms of identifying the weak points of the circuit that may be targeted by the opponent. From a more global perspective, it may be interesting to analyze whether there are other high-level properties of the network — e.g., scale-freedom – preserved over time. In addition to these methodological issues, it will be very useful to deploy this kind of analysis on further data both at club level (e.g., to identify similarities and dissimilarities between team and player performance in their clubs and in the national team), and at the national team level such as in the Euro 2012 in Poland and Ukraine or — looking beyond — the World Cup 2014 in Brazil.

References

[1] Passos P, Davids K, Araújo D, Paz N, Minguéns J, and Mendes J, Networks as a novel tool for studying team ball sports as complex social systems, *Journal of Science and Medicine in Sport*, 2011, **14**(2): 170–176.

[2] Euro 2004 network predictions at Revista Redes (accessed January 2012), URL: http://revista-redes.rediris.es/webredes/eurocopa2004.htm, 2004.

[3] Mendes R S, Malacarne L C, and Anteneodo C, Statistics of football dynamics, *European Physical Journal B*, 2007, **57**(3): 357–363.

[4] Brillinger D R, A potential function approach to the flow of play in soccer, *Journal of Quantitative Analysis in Sports*, 2007, **3**(1): 3.

[5] Hughes M and Franks I M, *Notational Analysis of Sport: Systems for Better Coaching and PErformance in Sport*, Routledge, New York, 2004.

[6] Onody R N and De Castro P A, Complex network study of brazilian soccer players, *Physical Review E*, 2004, **70**(3): 037103.

[7] Yamamoto Y and Yokoyama K, Common and unique network dynamics in football games, *PLoS ONE*, 2011, **6**(12): e29638.

[8] Duch J, Waitzman J S, and Amaral L A N, Quantifying the performance of individual players in a team activity, *PLoS ONE*, 2010, **5**(6): e10937.

[9] Lee J, Borgatti S P, Molina J L, and Merelo Guervós J J, Who passes to whom: Analysis of optimal network structure in soccer matches, Poster at the Sunbelt XXV conference, URL: http://revista-redes.rediris.es/webredes/novedades/soccer_sunbelt.pdf, 2005.

[10] Heuer A and Rubner O, Fitness, chance, and myths: An objective view on soccer results, *The European Physical Journal B — Condensed Matter and Complex Systems*, 2009, **67**: 445–458.

[11] Heuer A, Mueller c, and Rubner O, Soccer: Is scoring goals a predictable poissonian process? *EPL (Europhysics Letters)*, 2010, **89**: 38007.

[12] Bundio J and Conde M, Análisis del desempeño deportivo durante la eurocopa 2004 a partir del análisis de redes sociales, *REDES-Revista hispana para el análisis de redes sociales*, 2009, **13**(0) (in Spanish).

[13] Simon H A, On a class of skew distribution functions, *Biometrika*, 1955, **42**(3–4): 425–440.

[14] Karbowski M R, Video-based social network analysis data collection in sport, *QMSS*2 2011 *Workshop: Social Network Data Collection*, URL: http://vladowiki.fmf.uni-lj.si/lib/ exe/fetch.php?media =events: qmss2:abs:mariusz_karbowski.doc, 2011.

[15] Dowie J, Why Spain should win the world cup, *New Scientist*, 1982, **94**(10): 693–695.

[16] Berrone P, Business lessons from the soccer world cup, *Management Research: Journal of the Iberoamerican Academy of Management*, 2011, **9**(1): 73–78.

[17] Holland P W and Leinhardt S, Transitivity in structural models of small groups, *Comparative Group Studies*, 1971, **2**: 107–124.

[18] Li X, Leonard D, and Loguinov D, On reshaping of clustering coefficients in degree-based topology generators, ed. by Leonardi S, *Algorithms and Models for the Web-Graph* (WAW 2004), volume 3243 of *Lecture Notes in Computer Science*, Springer-Verlag, Rome, Italy, 2004, 68–79.

[19] Bonacich P, Factoring and weighting approaches to status scores and clique identification, *Journal of Mathematical Sociology*, 1972, **2**: 113–120.

[20] Erdős P and Rényi A, On the evolution of random graphs, *Publications of the Mathematical Institute of the Hungarian Academy of Sciences*, 1960, **5**: 17–61.

[21] Hughes M and Franks I M, Analysis of passing sequences, shots and goals in soccer, *Journal of Sport Sciences*, 2005, **23**(5): 509–514.

[22] BBC news piece: "World Cup 2010: Xabi Alonso slams Nigel De Jong tackle" (accessed May 2012), URL: http://news.bbc.co.uk/sport2/hi/football/world_cup_2010/8813661.stm, July 13th, 2010.

[23] Datablog: "World Cup 2010 statistics: Every match and every player in data" (accessed May 2012), URL: http://www.guardian.co.uk/news/datablog/2010/jul/09/world-cup-2010-statistics, July 9th, 2010.

[24] Lago C and Martín R, Determinants of possession of the ball in soccer, *Journal of Sports Sciences*, 2007, **25**(9): 969–974.

[25] Bate R, Football chance: Tactics and strategy, *Science and Football*, 1988, 293–301.

Appendix: Spain's Squad Numbers

#	Name	Position
1	Íker Casillas (C)	Goalkeeper
2	Raúl Albiol	Defender
3	Gerard Piqué	Defender
4	Carlos Marchena	Defender
5	Carles Puyol	Defender
6	Andrés Iniesta	Midfielder
7	David Villa	Forward
8	Xavi Hernández	Midfielder
9	Fernando Torres	Forward
10	Cesc Fàbregas	Midfielder
11	Joan Capdevila	Defender
12	Víctor Valdés	Goalkeeper
13	Juan Mata	Forward
14	Xabi Alonso	Midfielder
15	Sergio Ramos	Defender
16	Sergio Busquets	Midfielder
17	Álvaro Arbeloa	Defender
18	Pedro Rodríguez	Forward
19	Fernando Llorente	Forward
20	Javi Martínez	Midfielder
21	David Silva	Midfielder
22	Jesús Navas	Forward
23	José Manuel Reina	Goalkeeper