What's next in Complex Networks? Capturing the concept of attacking play in invasive team sports.

João Ramos1, 2, Rui J. Lopes1, 3 & Duarte Araújo4

1 ISCTE-Instituto Universitário de Lisboa, Lisbon, Portugal

2 Universidade Europeia, Laureate International Universities, Estrada da Correia, 53, 1500-210 Lisboa, Portugal

3 Instituto de Telecomunicações, Lisbon, Portugal

4 CIPER, Faculdade de Motricidade Humana, Universidade de Lisboa, Cruz Quebrada Dafundo, Portugal

Key Points

Network theory can contribute to performance analysis in invasive team sports by describing the complex and dynamic features of interaction between players.

Complex networks, notably temporal and bipartite networks, can capture the concept of attacking play by including play actions and their temporal sequence.

Typical questions on the team interaction properties can be answered by applying, possibly at different scales, wellknown bipartite network metrics, yielding different results from usual social network analysis.

Abstract

The evolution of performance analysis (PA) within sports sciences is tied to technology development and practitioner demands. However, how individual and collective patterns self-organize and interact in invasive team sports remains elusive. Social network analysis (SNA) has been recently proposed to resolve some aspects of this problem, and has proven successful in capturing collective features resulting from the interactions between team members as well as a powerful communication tool. Despite these advances, some fundamental team sports concepts such as an attacking play have not been properly captured by the more common applications of SNA to team sports performance. In this review article, we propose a novel approach to team sports performance centered on sport concepts, namely that of an attacking play. Network theory and tools including temporal and bipartite or multilayered networks were used to capture this concept. We put forward eight questions directly related to team performance to discuss how common pitfalls in the use of network tools for capturing sports concepts can be avoided. Some answers are advanced in an attempt to be more precise in the description of team dynamics and to uncover other metrics directly applied to sport concepts, such as structure and dynamics of attacking plays. Finally, we propose that, at this stage of knowledge, it may be advantageous to build up from fundamental sport concepts toward complex network theory and tools, and not the other way around.

1. Introduction

The evolution of performance analysis (PA) as a sub-discipline of sports sciences has seen significant advances in team sports. Technological progress has had an important role in this process, which is reflected in the availability of more data and with better accuracy and higher precision (1). The data processing phase of PA has also evolved from a mostly descriptive and qualitative approach to a quantitative and complex software-based analysis. PA approaches began with *biomechanics* and *notational analysis*, which were centered on individuals and their positions, actions and time (2-4). Notational analysis is based on the quantification of critical events through frequency counting followed by qualitative and quantitative feedback analysis. First, the performance indicators for the evaluation are determined (mainly on-the-ball actions of players), then the detail level is selected manually or computationally, and finally extensive flowcharts are created for the analyzed performance indicators (4). This quantification of critical events follows a sequential formula (player → position → action → time) normally associated with success (3-5). The biomechanics approach mainly entails individual fine-grain analysis and even when it is applied to team sports, it focuses on individual technique performance (3). This method has been mostly used in the context of closed sports techniques, particularly in performance environments (3). Aspects such as the non-linear interactions between the many independent parts of emergent movement patterns, or the interactions between players or even between teams, clearly show that theoretical guidance is essential for understanding complex adaptive sociobiological systems like team sports (2). A single event such as an injury at the microscopic level in a soccer player can have a large-scale impact on his/her performance, and an even greater effect on the team's structure and team's performance (micro-macro relatedness). A comprehensive PA therefore requires interdisciplinary approaches complementing biomechanics and notational analysis. Several authors (2, 6-8) believe that the dynamical systems theory (DST) gathers the necessary theoretical framework conditions for PA because it takes into account the stability, variability and transitions among interaction states, including at different levels of analysis, and compares them with specific outcomes. Indeed, self-organization explains how order emerges from the interaction of different components; for instance, how an individual performance results from the interaction of the player's body segments, or how team performance results from the interaction of individual team players (9).

Following such innovative approaches, the next logical step in PA is to address complex issues such as the relationship between team structure and dynamics, as well as the types of team interactions and their interdependency within and between teams (1-3, 5, 10, 11). In particular, it is important to understand how players' actions can disrupt the equilibrium of the other team or create scoring opportunities. Dutt-Mazumder and colleagues (6) proposed more visual methods to overcome the complexity of these issues and to encourage practitioners (coaches, athletes) to use these approaches. In a similar fashion, sports scientists have utilized social network analysis (SNA) (5, 12-15); notably, the communicative power that network visualization offers for new insights into network structures (16).

Concerning the dynamic features and its study, in network theory, two different types of dynamics are usually considered: dynamics *on* the network and dynamics *of* the network (17). Dynamics on the network focuses on flows on the network structure, e.g. ball passes between players. On the other hand, dynamics of the network are concerned with changes in the network structure itself, e.g. how the players' positions can provide information on a possible pass or player action. In the current study, we focus on the first type of network dynamics. Successful applications of dynamic network visualization in the literature mostly represent the "ball flux" in static *flip books*, where node position remains constant but edges cumulate over time, and therefore the researcher obtains cumulative snapshots of the network as a function of time (18). Typically SNA applied to PA uses only the 'last snapshot', that is, the network resulting from the aggregate of all the interactions occurring during the entire match; focusing on the structure and not so much on the dynamics. We believe that this impairs not only the study of relevant dynamic features of the team, but also conceals import concepts such as that of an attacking play. Dynamic relations in team sports games have other basic dimensions, in addition to passes, that have not yet been fully captured in sports settings: i) relational space (i.e. interactions considered in a geographical space); ii) their time structure (i.e., rate change, order or sequence, or simultaneity of interactions); and iii) their relations with types of nodes (i.e. colleagues or adversaries), meaning cooperation or competition interactions (18), as we discuss in the following section. This is a possible approach for handling complex adaptive systems including team sports analyses in competition. Thus, network dynamics analyses may explain how and why disequilibrium situations such as scoring opportunities occur.

In this study we aim to leverage the understanding of the structure and dynamics of invasive sports teams that may influence their performance. In particular, we focus in the interactions between players of the same team (ball passes), their actions during the attacking phase (e.g. ball recovery, shot at goal), and the temporal structure of both. We do not address directly neither the players' position nor the interaction between both teams. We start, in section 2.1, by discussing how SNA has been commonly used to address team structure and dynamics in invasive team sports performance. Typically this means aggregating all the passes between players in a single directed network for which measures are taken and related to success indicators (e.g. reaching a competition stage, goals scored) (12, 14, 15, 19). In section 2.2, we describe temporal bipartite networks in depth and the way they may overcome the shortcomings often found in applying SNA to PA. We start by analyzing the attacking play concept introducing nodes that represent other player's actions than passes. The chief contribution in this proposal is that these nodes are part of a bipartite network, thus enabling to keep the key concept of attacking play during network analysis and relate it to its possible outcomes. In addition, by applying temporal networks, dynamic aspects of the team process can be captured, namely the time structure of the different attacking plays. Both concepts, bipartite and temporal networks, can be extended hierarchically enabling performance analysis at different event and time scales. In section 2.3, a set of questions that are commonly found in the literature are placed as illustrative examples of how our suggestions can be used in match analysis in a way distinct from previous studies. Although we do not address directly neither players' position nor the interaction between both teams, we provide, in sections 2.4 and 3, guidelines for future studies in this direction. Notably, in order to use spatio-temporal structure and dynamics of both teams (represented by players' positions) in the analysis of the team and players' actions, the temporal structure of such actions must be adequatelly represented. Temporal networks can provide such representation. On the other hand, the outcomes of these dynamic interactions are naturally represented by different layers in bipartite networks.

2. Network Theory and Tools in Performance Analysis

2.1 Why use Network Analysis for Performance Analysis?

Has PA succeeded in understanding the processes, such as structure and dynamics, leading to improved performance? In this article we have raised some questions and problems that currently challenge the research in the field. Specifically, we ask what is the importance of a player in the structure and dynamics of the team/network, besides providing individual performance indicators? And who is the most connected player (i.e. playmaker) regarding his/her number of colleagues and interactions? Finally, what team sub-units, such as pairs or triangles of players, have the strongest influence on team performance? These and other questions, in particular those addressing team structure and dynamics, can be investigated using networks theory and tools such as SNA. However, we do not aim and it would not be possible to comprehensively review the applications of SNA to team sports performance. Instead, we will discuss the limitations of SNA for explaining team performance and how they may be overcome. In our view, until these issues are fully resolved SNA will be insufficient for clearly describing and explaining the dynamic nature of the processes associated with team performance. We propose the adoption of multilayer networks, in particular, of bipartite and temporal networks, to overcome some of the limitations of SNA. To achieve this aim, we present solutions from netwok theory for accurately representing dynamics and suggest alternative metrics that we find more adequate for explaining sport teams performance.

The definition of *network* is a "collection of vertices joined by edges", which can represent the pattern of connections between different objects (20). Thus, social networks are structures with persons or sometimes groups of persons (actors) who are related by some form of social interaction (ties) (20, 21). In team sports settings, we can consider the actors as the players, and the ties as the interactions between players. Team goals are achieved if each individual's effort is coordinated with those of the other team-mates through dynamic interactions, i.e., a complex network is considered rather than the simple summation of the individual performances (2, 5, 19, 22, 23). To address these team sports characteristics, we focus on eight questions that relate directly to team performance and network metrics.

Technical terminology and heavy jargon from complex systems approaches can be overcome by using network visual representations, which are powerful and versatile tools widely used to describe dynamical systems (5, 6, 12, 16, 18, 24, 25). Typically, in SNA applied to sport matches, the nodes represent team players and the links between nodes correspond to the interaction between those players (19), specifically "ball passing"[1]. This relationship is characterized by a *transport action* (i.e. a token - ball - is passed between players) and a *directionality* (i.e. for each interaction there is a sender and a receiver), which combined make a directed network (i.e. the links have a direction).

[1] Analysing only at ball passing restricts the analysis of team performance to the attacking phase. In the current article, we do not attempt to directly resolve this limitation.

The two most commonly used systems for representing a network are matrices[2] (adjacency or incidence) and graphics. While matrices are particularly relevant and useful for the development of formal methods and computational processes, graphics have an extraordinary communicative power. Notably, by using adequately network graph visualization features, such as the nodes' and edges position, size and color, particular properties of the network, i.e., the importance of a player or frequency of an interaction, can be highlighted and perceived intuitively without requiring a specialized knowledge on network theory. For example, in order to identify the number of passes of each player, instead of summing matrix values the reader uses less cognitive resources by simply observing the size of the nodes (to identify the most connected player) or the weight of the edges that represents the number of passes between two players (identifying the dyadic that interacted more). This intuitive characteristic of network graphs and the non-specialized skills required from the reader are being explored by generalist newspapers in pieces devoted to soccer match analysis. Figure 1 uses one of these generalist newspapers examples to illustrate some of the features mentioned above (26).

[INSERT FIGURE 1 APROXIMATELY HERE]

2.2 From static, single-layered networks to temporal bipartite networks.

Graphs as illustrated in Figure 1 present features associated with nodes and links that are cumulative (i.e. they aggregate all the interactions that have occurred during the match) but not the relevant events in the match such as goals and gaining/losing ball possession, which are not usually represented in these analyses. However, although global metrics can be obtained in this type of network representation (e.g. which player does more passes), the fundamental concept of attacking play is not apparent. An attacking play is defined as the tactical situation when one team is in possession of the ball moving towards the opponent's goal (e.g., Lucchesi [(27)]). To visualize an attacking play it is essential to include in the network other types of nodes to illustrate the beginning and end of the attacking play, such as

[2] An adjacency matrix, *A*, is a square matrix, with rows and columns representing nodes (e.g., players) with entry a_{ij} of *A* taking value 1 if there is a link between node *i* and node *j*; and 0 otherwise. Different types of networks lead to different matrix structures: undirected graphs are represented in symmetric adjacency matrices, the fact that the link between nodes *i* and *j* has no directionality is expressed in equality $a_{ij} = a_{ji}$; in directed graphs (or digraphs) the links between nodes have a directionality; a link from node *i* to node *j* is expressed by entry a_{ij} taking value 1 independently of the value of a_{ji}. In this paper the links represent actions by the players (e.g., making a pass) and are thus directed leading to digraphs.. In what are called weighted graphs, the entries of the matrix can take other values w_{ij}, called weights, that are nor restricted to 0 or 1. The value taken by entry w_{ij} reflects the intensity or strength of that link.

In an incidence matrix, *E*, rows represent nodes and columns represent links. The entry e_{ij} takes value 1 if the link *j* is incident on nodes *i* and *j*; 0 otherwise. In directed networks values -1 and 1 are used to distinguish link origin and destination.

gaining/losing ball possession, gaining a free-kick, and scoring a goal. However, adding these nodes to graphs in a simplistic manner not only breaks the semantic homogeneity of nodes and links, which do not always correspond to passes between players, but also changes the metric values of the network. We propose that multilayer networks can bring an important contribution to the understanding of team attacking dynamics, specifically, through the combination of temporal and bipartite networks.

Typically, the formalization of a temporal network starts with the definition of M time snapshots for the entire duration of the game, T, with equal interval $\tau = T/M$, and the N nodes of interest. The next step is the aggregation of all the interactions that occur in the time interval (or time snapshot) t, between $(t-1)\tau$ and $t\tau$ where $t = 1, \ldots, M$ (28). Fixing a value for τ not only raises the problem of allocating it an appropriate value but also does not guarantee that the concept of play is accurately represented. We suggest that each interval in the temporal network should correspond to the duration of an attacking play. The definition of the beginning and end of each of these intervals is therefore defined by the beginning and end of an attacking play, typically corresponding to those instants when ball possession has been won or lost. Formally, time snapshot t_i, corresponding to the i^{th} attacking play, is defined by the time interval bound by instant t_{Gi}, where the team gained ball possession and instant t_{Li} where possession was lost.

The introduction of these new nodes to temporal networks takes us to the second element of our proposal: bipartite networks. In bipartite networks, also known as two-mode networks, two different types or classes of vertices are considered, and nodes of the same type cannot be connected directly. Consequently, the links are always incident on nodes of different types. A possible application of the bipartite network model could consider as the two node classes i) the set of players; and ii) the set of technical skills or play actions performed by the players. The application of the temporal and bipartite network concepts in combination is illustrated in Figure 2.

[INSERT FIGURE 2 ABOUT HERE]

Therefore, the use of any kind of team performance representative nodes (e.g. shot at goal) allows a direct way to relate PA metrics and network-based metrics associated with the attacking play structure. For example, the in-degree metric for the shot at goal node provides directly the number of shot at goal. On the other hand, the average length of the walks leading to this node provides the average number of passes till a shot at goal is made and thus can provide an hint on the teams attacking style of play.

An interesting feature that results from projecting the bipartite network into the two one-mode projections is illustrated in Figure 3.

[INSERT FIGURE 3 ABOUT HERE]

These two projections focus on different aspects of the team performance and uncover important limitations of traditional network match presentations. The events projection reveals the pass path between gaining and losing ball possession. The players' projection in Figure 4 shows the interactions between players (passes) in a similar way to the network in Figure 1.

It is worth highlighting the network concepts that emerge from both projections: i) in the events projection the result is always a path (no node is visited more than once); ii) in the players projection the result is a walk, where a node can be visited more than once (right defender [RD] in the example represented in figure 4). Moreover, team characteristics can be observed including a cluster of vertices that are all connected to each other and are known as a 'clique' in network jargon (e.g. cluster formed between players left defender [LD] in figure 4, RD and Midfielder [MF] in figure 4) (20). In the events and players projections one can perform an aggregation operation as shown in Figure 4 (aggregation of a play showing a shot at goal).

[INSERT FIGURE 4 ABOUT HERE]

In addition, some filtering can also be applied to the aggregation process, for instance, to consider only the attacking plays that ended in a 'shot at goal' event, as represented in Figure 4. The projections in Figure 4 clearly show that the concepts of 'path' and 'walk'(29) are of great relevance, as they uncover another distinctive aspect between team sports and other networks, specifically, how each node establishes the outbound link when a token (i.e. a ball in team sports such as soccer) is received. Typically in an invasive team sports match, the trajectory along the network nodes does not follow the shortest path nor even a path where neither nodes or links are repeated (obtained over the full aggregation graph), as is often assumed in sports sciences (15, 30-32). However, as illustrated by Figure 3 (attacking play 1), in many networks the trajectory can be a trail (where links cannot be repeated) or a walk (nodes and links can be repeated with no restrictions), and is not deterministic. This fundamental difference has a very strong impact in many networks metrics, centrality in particular, as discussed in sections 2.3.2. and 2.3.3..

2.3 Using temporal bipartite networks: some illustrative questions.

2.3.1. Question 1: who is the most interactive player?

The activity level of a player in his/her interaction with other team-mates can be captured by the *node degree*[3] of a vertex in a given graph, in which case only the adjacencies of the node/player are considered; this is therefore a local analysis of node centrality. In sports settings, the node centrality is the most widely used network metric (12-15, 19, 33) and it is typically based on ball flow. Given the directed nature of the interaction, the degree centrality is divided into two categories: *in-degree*, which measures the number of players who pass the ball to the focal player, and *out-degree*, which measures the number of players to whom the focal player passes the ball to. As centrality is focused on each individual player and his/her participation in the team ball passing activity, it can be represented for the duration of the entire game in simple graphs (Figure 1). The approach proposed in this article (section 2.4) extends the reach of this metric in two ways: i) by defining different time spans for the aggregation process, the activity of each player can be measured for time intervals other than the entire match; ii) when appropriate filtering is applied to the aggregation process, centrality can be applied to the player's activity in attacking plays of certain characteristics (e.g. plays that lead to a shot at goal event).

2.3.2. Question 2: which players have an intermediary role?

Betweenness[4] is a widely used centrality metric that could provide an answer to the important question of which players play an intermediary role. Betweenness analyzes the *global*

---

[3] *Degree of a vertex* i, hence v_i, is given by the number of nodes that are directly connected with the focal node;

$$Centrality_{degree}(v_i) = \text{degree}(v_i) = \sum_{j}^{N} a_{ij}$$

where i is the focal node, j represents all other nodes, N is the total number of nodes, and a is the adjacency matrix, in which cell a_{ij} is defined as 1 if node i is connected to node j, and 0 otherwise.

[4] *Betweenness centrality* expresses the degree in which one node lies on the shortest path between two other nodes;

structure of the network, notably the fraction of shortest paths corresponding to each node, and it shows potential for acurately representing how much each player contributes as an intermediary between other players. However, this metric is typically based on the shortest path between any two nodes computed over the graph resulting from the aggregation of all the interactions in the match (17). In these conditions the metric is not directly supported by any fundamental concept of an attacking play and this is, in our view, a strong limitation. Indeed, in team ball sports the ball flow does not necessarily follow the shortest paths over the aggregation graph; instead players projections reveal mostly walks (with relevant levels of randomness) and not paths, similarly to most flows in other networks (34). In addition, the paths that lead to specific events (e.g. shot at goal) may have a more direct impact on team performance than the connectivity between players per se (19). Moreover, how such connectivity is directed in specific events can also affect team performance (12). Nevertheless, these limitations of betweeness can be overcome by counting the fraction of walks leading to a certain event in which the focal player is involved, rather than considering the shortest paths between players (12). Freeman and colleagues (16) proposed a metric based on the idea of maximum flow, *flow centrality*, whereby different paths can be used for the same purpose and which has been applied to basketball research by Fewell and colleagues (19). Newman's *random-walk betweenness* considers all paths between nodes including those that are not optimal, although more weight is given to the shortest paths (34). Bonacich (35) refers to *power centrality*, a metric based on the assumption that centrality is related to power and as such an individual's status is a function of the statuses of his/her connections.

2.3.3. Question 3: how central is a player?

Comparing players in terms of the number of passes achieved can reveal important individual characteristics of a player's performance and also the soccer team's style of play. *Closeness*[5] is a measure of centrality that considers the length of the shortest paths between the focal node and all the other nodes. We propose that this metric should be used as follows: passing

---

$$Centrality_{betweenness}(v_i) = betweenness_i = \frac{g_{st}(i)}{g_{st}}$$

where g_{st} is the number of shortest paths between vertices *s* and *t*, and $g_{st}(i)$ is the number of those paths that pass through vertex i.

[5] *Closeness centrality* for each node, v_i, is the inverse sum of the shortest distance, $distance(i,j)$ to all other nodes, *j*, from the focal node, *i*, or how long the information takes to spread from a given node to others.

$$Closeness_{centrality}(v_i) = closeness_C(i) = \left[\sum_{j=1}^{N} distance(i,j)\right]^{-1}$$

path projections convey length (measured by the absolute frequency of passes) and duration of each play. The latter can be used directly to identify the team's style of play (see Passos et al. [1]). Moreover, by analyzing passing path projections (see Figure 3) one can capture the distance (also measured by the absolute frequency of passes) between the focus player and an event of interest, which can be utilized for example to identify which players contribute directly to shooting at goal or assisting other players. Alternatively, Noh and Rieger's (36) *random-walk centrality* metric describes the average speed at which messages are transmitted from one node to another in random walks (37). This is similar to the *closeness centrality* metric except that it considers the length of a random walk rather than the shortest path.

2.3.4. Question 4: how does each player contribute to the performance of others?

A player can contribute to the team's performance directly (e.g. goals scored) and indirectly by assisting the performance of team-mates (e.g. assistances/passes). *Eigenvector centrality*[6] (38) can be used to assess the contribution of each player to the team's performance. Similar to closeness, eigenvector centrality considers the global structure of the network (29) but assumes unrestricted walks, rather than paths, emanating from a node. Thus, this measure counts the number of walks of all lengths, weighted inversely by length, and as a result it can determine how each node affects all its neighbours at a given moment (29). Alternatively, Bonacich (35) argues that the concept of centrality should be more general due to its positive relationship with power (39). This way Bonacich (35) proposes a *power centrality* metric related to power and hence the individual or node's status is dependent on the status of its connections. The power is therefore attributed to the nodes/players in the negotiation of any single play with their team-mates, and it results from the players' efficacy to resolve previous moves. Moreover, this metric can be complemented by pre-defining weights for relevant events in the graph, for instance, ball loss with a negative weight and scoring a goal with the

---

[6] *Eigenvector centrality*, takes into consideration not only how many connections a vertex has (i.e., its degree), but also the *degree* of the vertices that it is connecting to. Each vertex i is assigned a weight $x_i > 0$, which is defined to be proportional to the sum of the weights of all vertices that point to i: $x_i = \lambda^{-1} \sum_j A_{ij} x_j$ for some $\lambda > 0$, or in matrix form

$$Ax = \lambda x,$$

where A is the (asymmetric) adjacency matrix of the graph, whose elements are A_{ij}, and x is the vector whose elements are the x_i, and λ is a constant (the eigenvalue).

highest positive value. These weights can then be propagated to other nodes/players, similarly to the eigenvector centrality measure.

2.3.5. Question 5: are there "hot" nodes in the team?

During a play situation, players tend to search for the team-mate who can typically offer more solutions to the problems of the game (40). This preference could be measured by identifying those players who connect more often with the 'powerful' ones. Passos and colleagues (13) suggested that identifying these *preferential attachments*[7] could lead to the 'decision-makers' of each team. Grund (14) has provided some innovations in this regard by assessing team ball flow with a traditional binary system (pass =1, no pass = 0) but also by considering centralities based on node strength and ties weight. The author focused on network structure (41-44) thus confirming that increased interaction intensity (density) leads to increased team performance (measured by the absolute frequency of goals scored), and increased centralization of interactions leads to decreased team performance. Given the intrinsic temporal nature of the graphs describing a match, we suggest that preferential attachments could be assessed by determining if those players with the highest node degree are more likely to be selected in the next attacking plays. By applying appropriate algorithms to the aggregation process, it is possible to obtain distinct metrics for different attacking play outcomes, and therefore to determine whether the preferential attachment process is related to those outcomes.

2.3.6. Question 6: are there clusters in the team?

Fewell and colleagues (19) assessed basketball team dynamics through degree centrality, clustering, entropy, and flow centrality, in order to uncover the play strategies of the 2010 National Basketball Association play-offs. In a group of nodes/players, it is possible to identify the players who are muttualy highly connected and those who are less so. Such

---

[7] *Preferential attachments*, also known as cumulative advantage or '*rich-get-richer paradigm'*.
This property means that every new vertex probability (p_i) to connect the existing vertices is higher for those who have already a large number of connections (connectivity k_i). For example, in a given team sports with ball, when a player attracts more interactions from the game's beginning, his/her connectivity will increase at a higher rate when compared to his/her team-mates as the game is played (network grows). Therefore, starting with a small number (m_0) of players interacting at the beginning of the game, at every time step that a new player $m(\leq m_0)$ interacts with m different team-mates already active in the game, for preferential attachment, there is a probability $p_i(k_i) = \frac{k_i}{\sum_j k_j}$ that the new player i will interact with a certain team-mate, depending on the connectivity k_i of the latter.

highly-connected groups are called *clusters*. The local *clustering coefficient*[8] can be used to measure this connectivity property (known as transitivity in social networks) by capturing the probability of cooperation between players as a function of their mutual acquaintances/interactions. It can also be said that a *triadic closure* is formed around the focal node, i.e. "if two people in a social network have a friend in common, then there is an increased likelihood that they will become friends themselves at some point in the future" (45, 46). However, when looking for the most relevant players in the formation of clusters, applying the local clustering coefficient to graphs representing an entire match (Figure 1) restricts its usefulness, as the passes (i.e. links) between players considered in the cluster formation may have occurred in different attacking plays, and possibly with a long temporal gap between them. Such is the case for the clique RD-MF-CF in Figure 4, showing an aggregation projection of a play that culminated with a shot at the goal . Given the attacking play granularity of bipartite and temporal networks (see Figure 3), only clusters formed within the same attacking play can be identified with certainty (e.g. in Figure 3, only the clique RD-LD-MF is identified).

2.4 What is still to be done? The dynamics of a network and multilayer networks.

Although the questions we have addressed thus far could be answered through *static networks*[9] or by looking at the flows between nodes, our next questions require the analysis of changes in the network structure. In typical team interactions representations (described in section 2.2), only the interactions (flows) between players are considered for building the network. There is therefore a superposition between the flows *on* the network and the definition *of* the network structure that must be taken into account when considering the applicability of SNA metrics. However, these relationships occur throughout a certain time span, and these important time changes must also be considered.

The proposal in section 2.2 of using bipartite networks to identify nodes representing players and technical actions in different layers can be extended hierarchically. Technical actions

---

[8] The local *clustering coefficient* (cc_i) for player i is defined by the proportion of actual edges/interactions (e_i) between between the n_i common neighbors of a vertex/player i and the number of possible edges between them.

$$cc_i = \frac{2e_i}{n_i(n_i-1)}$$

[9] We define as static network the static structure resulting from the aggregation over a time interval (e.g., the entire match) of all the observable edges (e.g., passes) within that interval.

(level *n* events) can be linked in another bipartite network to level *n+1* events corresponding to higher-level concepts, as illustrated in Figure 5.

[INSERT FIGURE 5 APROXIMATELY HERE]

2.4.1. Question 7: how does a player influence the team structure?

It is relevant to ask how player behaviors other than passing influence the overall network of interactions in the team. According to Barzel and Barabási (47), behavior prediction in a complex system requires a quantitative description of the system's structure and dynamics. The *dynamics of a network* considers various phenomena, including self-organization, that promote changes in the topology of the network (25), however, this metric has not yet been applied to sports sciences. The study of the dynamical interplay between the players's state and the topology of the network is recent and mostly theoretical (25, 48-51). Indeed, few theorizations have been corroborated by empirical results (25, 52). Interestingly, one of these studies (40) revealed an unexpected time dependence in network centrality (*dynamic centrality*) indicating that well-connected nodes can quickly become weakly-connected or even disconnected (25). Moreover, *dynamic centrality* expressed in adaptive networks (dynamic scale-free [DSF] networks) emerges from a reinforcement rule whereby each node considers only the importance or popularity of its neighbours (25). These surprising results reinforce the need for further studies, especially in sports settings.

Guillaume and Latapy (53) proposed another relevant approach to team sports performance showing how all complex networks may be described as bipartite structures, or alternatively via hypernetworks. The authors introduced a model that can be tested for any type of real-world complex network. Moreover, these bipartite networks can be used to represent relations that are not dyadic (i.e. they involve more than two actors). Non-dyadic relations among players, such as geographical proximity in the pitch, can describe other dimensions of the players' actions during the game and these descriptions can be used to understand the dynamics *of* the network of passes, in particular, by explaining why certain spatial team configurations lead to specific pass paths.

2.4.2. Question 8: how does the adversary team constrain the team's interactions and structure?

The manner and extent to which the opposing team constrains a team's interactions and structure is a much more complex question, as it considers the influence of the adversary team (individuals and structure) on the interactions between team individuals as well as on the team's structure. Some of the PA existing studies focus on the interplay between attackers and defenders and are therefore based on dyadic interactions between players, sub-units or teams. Typically, such studies associate these interactions with the players' spatial organization, which is computed from the surface area in so-called centroids (23, 54, 55) or Voronoi diagrams[10] (56). However, in network theory, *spatial networks* represent the nodes and edges based on their interactions in an Euclidean space. Using these metrics, research on social networks and space has identified ordering principles such as *homophily states* (57-59) and *focus constraint* (59, 60). Notably, while *homophily* depends on non-structural features such as connections fostered by status or interests (e.g. dyadic attacker-defender interactions in team ball sports), *focus constraints* are dependent on geographical proximity, enabling face-to-face interactions (59).

3. Conclusions

In this article, we reviewed how PA emerged as a sub-discipline of sports sciences by building on notational analysis and biomechanics approaches and with further contributions from DST. Additionally, we discussed what new directions, tools and potential methods network theory and complex networks can further contribute to PA. Early studies with network methodologies in team ball sports mainly considered the dynamics on networks through ball flow, which represent the interactions between players of the same team, or actions to score.

However, these studies did not consider the dynamics of networks, assuming teams to be static structures, whereby players retain the same performance level throughout the entire game, independent of the constraints imposed by team adversaries and the players' positioning. Moreover, the players' skills and technical actions as well as the evolution of the interactions between players over time were also not considered.

Finally, this static networks structure is obtained via the aggregation of all the interactions (passes) that occurred in the entire match. This leads not only to conceal important concepts such as attacking play but also to metrics that may be misleading. A notable example are metrics based on `shortest paths´ over the aggregated network, that occurs in two pitfalls: i) pass interactions form a walk and not necessarly a path; ii) interactions between players in a

[10] Voronoi diagrams are geometric constructions that represent the nearest geographical region of a player, a sub-set of a team, or even a team.

match do not follow this principle but are may be better described by geographic networks and random walks (34).

We propose that temporal and bipartite networks could be an alternative approach for representing the interactions between players during a game. Using the flexible time structure of temporal networks it is possible to capture the sequence of passes in an attacking play, which is one of the main concepts of team collective behaviour. We have highlighted how temporal bipartite network representation empowers existing metrics for capturing sports fundamental concepts (e.g. style of play) with greater adequacy. Moreover, we suggest that methods combining spatial and hypernetworks (61) with temporal networks represent a promising direction for future research, as they allow the analysis of dynamics of the network. These complex networks could integrate concepts such as how time changes the structure of the network, as well as the players' technical resources and their positioning relative to the position of other players (team-mates or adversaries).

Compliance with Ethical Standards

Funding

This work was partly supported by the Fundação para a Ciência e Tecnologia, under Grant UID/DTP/UI447/2013 to CIPER – Centro Interdisciplinar para o Estudo da Performance Humana (unit 447).

Conflicts of Interest

João Ramos, Rui J. Lopes and Duarte Araújo declare that they have no conflicts of interest relevant to the content of this review.

References

1. Passos P, Araújo D, Volossovitch A. Performance analysis in team sports. London: Routledge. Taylor & Francis Group.; 2017.

2. Glazier PS. Game, set and match? Substantive issues and future directions in performance analysis. Sports Med. 2010;40(8):625-34.

3. Hughes M, Franks IM. The essentials of performance analysis : an introduction. London ; New York: Routledge; 2008. xxxii, 312 p. p.

4. Frencken W. Soccer tactics: dynamics of small-sided games and full-sized matches. Groningen: University of Groningen - Faculty of Medical Sciences; 2012.

5. Vilar L, Araújo D, Davids K, et al. The role of ecological dynamics in analysing performance in team sports. Sports Med. 2012;42(1):1-10.

6. Dutt-Mazumder A, Button C, Robins A, et al. Neural network modelling and dynamical system theory: are they relevant to study the governing dynamics of Association Football players? Sports Med. 2011. 41: 1003. doi:10.2165/11593950-000000000-00000.

7. Bartlett R, Button C, Robins M, et al. Analysing team coordination patterns from player movement trajectories in soccer: methodological considerations. Int J Perform Anal Sport. 2012;12(2):398-424.

8. Davids K, Araújo D, Shuttleworth R. Applications of dynamical systems theory to football. Science and football V. 2005:537-50.

9. Araújo D, Davids K. Team synergies in sport: Theory and measures. Front Psychol. 2016;7.

10. Chinellato DD, de Aguiar MA, Epstein IR, et al. Dynamical response of networks under external perturbations: Exact results. arXiv preprint arXiv:07054607. 2007.

11. Silva P, Chung D, Carvalho T, Cardoso T, et al. Practice effects on intra-team synergies in football teams. Hum Mov Sci. 2016;46:39-51.

12. Duch J, Waitzman JS, Amaral LAN. Quantifying the performance of individual players in a team activity. PLoS ONE. 2010;5(6):e10937.

13. Passos P, Davids K, Araújo D, et al. Networks as a novel tool for studying team ball sports as complex social systems. J Sci Med Sport. 2011;14(2):170-6.

14. Grund TU. Network structure and team performance: The case of English Premier League soccer teams. Soc Networks. 2012;34(4):682-90.

15. Travassos B, Bourbousson J, Esteves PT, et al. Adaptive behaviours of attacking futsal teams to opposition defensive formations. Hum Mov Sci. 2016;47:98-105.

16. Freeman LC. Visualizing social networks. J Soc Struct. 2000;1(1):4.

17. Rocha L. Exploring patterns of empirical networks. UmeÃ¥: UmeÃ¥ University, Department of Physics; 2011. 66 p.

18. Moody J, McFarland D, Bender-Demoll S. Dynamic network visualization. AJS. 2005;110(4):1206-41.

19. Fewell JH, Armbruster D, Ingraham J, et al. Basketball teams as strategic networks. PLoS ONE. 2012;7(11):e47445.

20. Newman MEJ. Networks : an introduction. Oxford ; New York: Oxford University Press; 2010. xi, 772 p. p.

21. Brandes U, Freeman L, Wagner D. Social Networks. In: (ed.) RT, editor. Handbook of Graph Drawing and Visualization. CRC Press2012.

22. Eccles DW, Tenenbaum G. Why an expert team is more than a team of experts: A cognitive conceptualization of team coordination and communication in sport. J Sport Exerc Psychol. 2004;26:542-60.

23. Duarte R, Araújo D, Freire L, et al. Intra-and inter-group coordination patterns reveal collective behaviors of football players near the scoring zone. Hum Mov Sci. 2012;31(6):1639-51.

24. Bender-deMoll S, McFarland DA. The art and science of dynamic network visualization. J Soc Struct. 2006;7(2):1-38.

25. Hill SA, Braha D. Dynamic model of time-dependent complex networks. Phys Rev E Stat Nonlin Soft Matter Phys. 2010;82(4):046105.

26. Araújo D, Lopes RJ, Ramos J, et al. A final à lupa: Alemanha pressionante, Argentina recuada. Lisbon. Portugal: Jornal Público Online; 2014 [Available from: https://http://www.publico.pt/desporto/noticia/a-final-a-lupa-pressao-sustentada-da-alemanha-venceu-uma-argentina-recuada-e-a-procura-do-contraataque-1662863.

27. Lucchesi M. Attacking Soccer: A Tactical Analysis. 2001. Spring City, PA: Reedswain Publishing.

28. Rocha L, Masuda N. Random walk centrality for temporal networks. New J Phys. 2014;16(6):063023.

29. Borgatti SP. Centrality and network flow. Soc Networks. 2005;27(1):55-71.

30. Pena JL, Touchette H. A network theory analysis of football strategies. arXiv preprint arXiv:12066904. 2012.

31. Clemente FM, Martins FML, Kalamaras D, et al. General network analysis of national soccer teams in FIFA World Cup 2014. Int J Perform Anal Sport. 2015;15(1):80-96.

32. Clemente FM, Couceiro MS, Martins FML, et al. Using network metrics in soccer: A macro-analysis. J Hum Kinet. 2015;45(1):123-34.

33. Correia V, Araújo D, Duarte R, et al. Changes in practice task constraints shape decision-making behaviours of team games players. J Sci Med Sport. 2012;15(3):244-9.

34. Newman MEJ. A measure of betweenness centrality based on random walks. Soc Networks. 2005;27(1):39-54.

35. Bonacich P. Power and centrality: A family of measures. AJS. 1987:1170-82.

36. Noh JD, Rieger H. Random walks on complex networks. Phys Rev Lett. 2004;92(11):118701.

37. Stephenson K, Zelen M. Rethinking centrality: Methods and examples. Soc Networks. 1989;11(1):1-37.

38. Bonacich P. Factoring and weighting approaches to status scores and clique identification. J Math Sociol. 1972;2(1):113-20.

39. Cook KS, Emerson RM, Gillmore MR, et al. The distribution of power in exchange networks: Theory and experimental results. AJS. 1983:275-305.

40. Braha D, Bar-Yam Y. From centrality to temporary fame: Dynamic centrality in complex networks: Research Articles. Complex. 2006;12(2):59-63.

41. Reagans R, Zuckerman EW. Networks, diversity, and productivity: the social capital of corporate R&D teams. Organization Science. 2001;12(4):502-17.

42. Sparrowe RT, Liden RC, Wayne SJ, et al. Social networks and the performance of individuals and groups. Acad Manage J. 2001;44(2):316-25.

43. Balkundi P, Kilduff M. The ties that lead: A social network approach to leadership. Leadersh Q. 2006;17(4):419-39.

44. Cross R, Cummings JN. Tie and network correlates of individual performance in knowledge-intensive work. Acad Manage J. 2004;47(6):928-37.

45. Rapoport A. Spread of information through a population with socio-structural bias: I. Assumption of transitivity. Bull Math Biophys. 1953;15(4):523-33.

46. Easley D, Kleinberg JY. : Networks, crowds, and markets: reasoning about a highly connected world. Cambridge University Press. ISBN: 0521195330, 9780521195331. 2010.

47. Barzel B, Barabási A-L. Universality in network dynamics. Nat Phys. 2013;9(10):673-81.

48. Holme P, Ghoshal G. Dynamics of networking agents competing for high centrality and low degree. Phys Rev Lett. 2006;96(9):098701.

49. Pacheco JM, Traulsen A, Nowak MA. Coevolution of strategy and structure in complex networks with dynamical linking. Phys Rev Lett. 2006;97(25):258103.

50. Van Segbroeck S, Santos FC, Pacheco JM. Adaptive contact networks change effective disease infectiousness and dynamics. PLoS Comput Biol. 2010;6(8):e1000895.

51. Shaw LB, Schwartz IB. Fluctuating epidemics on adaptive networks. Phys Rev E Stat Nonlin Soft Matter Phys. 2008;77(6):066101.

52. Barabasi A-L. The origin of bursts and heavy tails in human dynamics. Nature. 2005;435(7039):207-11.

53. Guillaume J-L, Latapy M. Bipartite graphs as models of complex networks. Physica A. 2006;371(2):795-813.

54. Lames M, Erdmann J, Walter, F. Oscillations in football – order and disorder in spatial interactions between the two teams . Int J Sport Psychol. 2010;41:85-6.

55. Frencken W, Lemmink K, Delleman N, et al. Oscillations of centroid position and surface area of soccer teams in small-sided games. Eur J Sport Sci. 2011;11(4):215-23.

56. Fonseca S, Milho J, Travassos B, et al. Measuring spatial interaction behavior in team sports using superimposed Voronoi diagrams. Int J Perform Anal Sport. 2013;13(1):179-89.

57. Kossinets G, Watts DJ. Empirical analysis of an evolving social network. Science. 2006;311(5757):88-90.

58. McPherson M, Smith-Lovin L, Cook JM. Birds of a feather: homophily in social networks. Annu Rev Sociol. 2001;27(1):415-44.

59. Expert P, Evans TS, Blondel VD, et al. Uncovering space-independent communities in spatial networks. Proc Natl Acad Sci USA. 2010;108(19):7663–7668. doi:10.1073/pnas.1018962108.

60. Crandall DJ, Backstrom L, Cosley D, et al. Inferring social ties from geographic coincidences. Proc Natl Acad Sci USA. 2010;107 (52) 22436-22441; doi:10.1073/pnas.1006155107.

61. Johnson JH, Iravani P. The multilevel hypernetwork dynamics of complex systems of robot soccer agents. ACM Transactions on Autonomous and Adaptive Systems (TAAS). 2007;2(2):5.

Figure 1

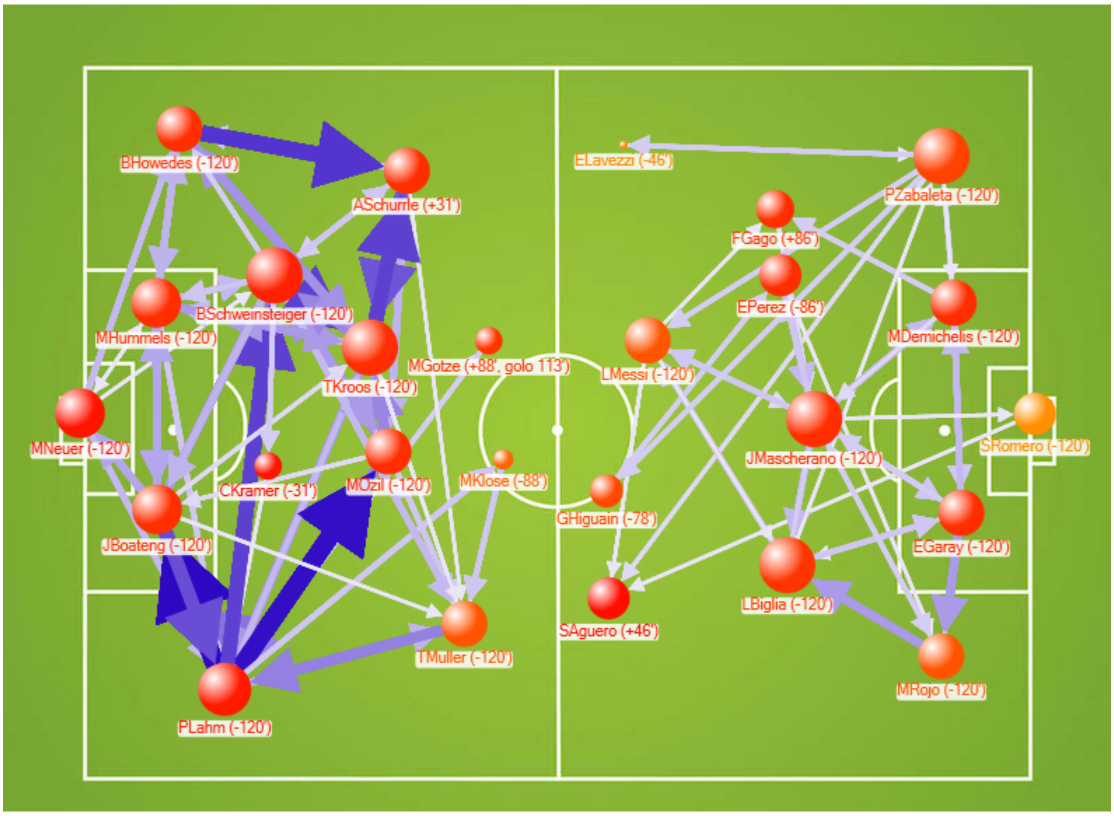


Figure 2

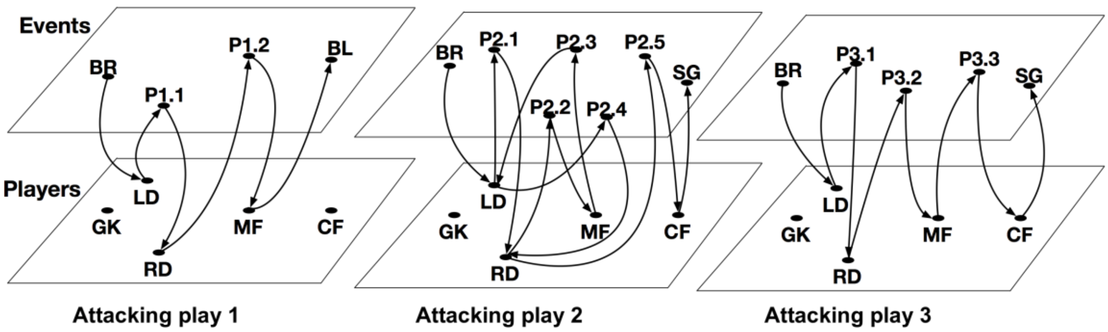


Figure 3

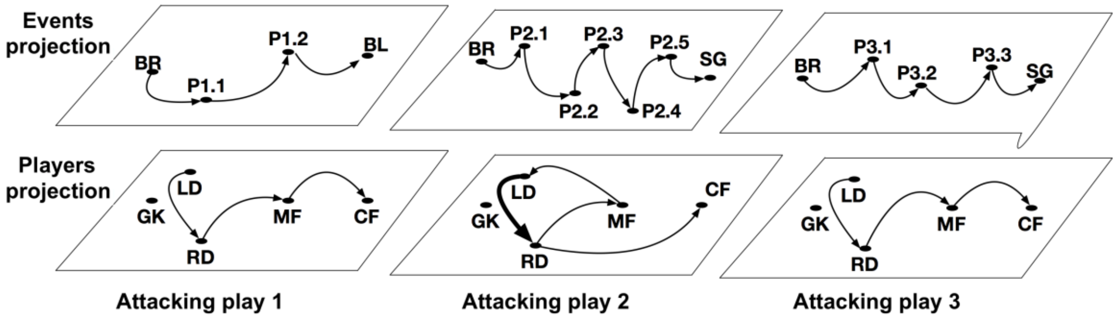


Figure 4

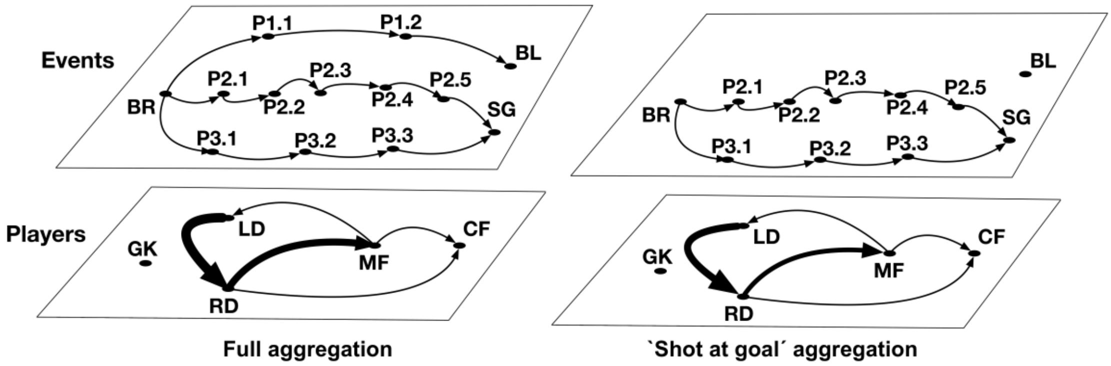
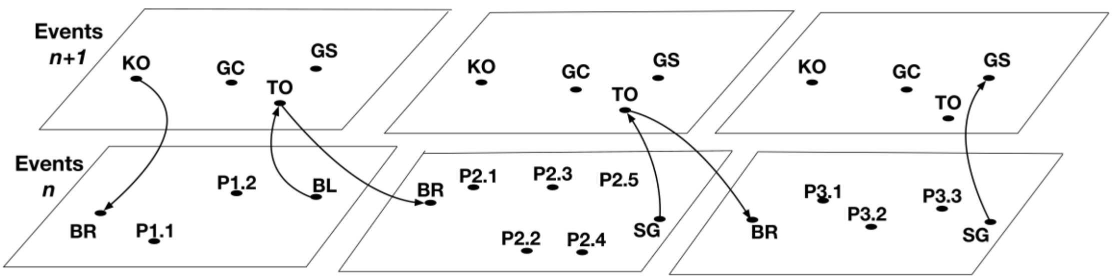


Figure 5



Captions:

Figure 1: Pass interactions in the Germany vs. Argentina match for the 2014 FIFA World Cup (26). Each circle represents a player in his relative position; the radius and colour of each circle represents the number of players that player interacts with and his pass precision (red more precision and yellow less precision), respectively; and the arrows represent the direction of the passes between players; the width and shade of each arrow represents the number of interactions (passes) between players (lighter arrows indicate less passes, darker arrows indicate more passes). The numbers in brackets represent the minutes played by the player (e.g., -120', played 120 minutes) or the moment in the match when the player started to play (e.g., +88', golo 113', entered in the match at minute 88 and scored a goal at minute 113).

Figure 2: Multilayer representation of three attacking plays in a bipartite (two-mode) network. GK goalkeeper, LD/RD left/right defender, MF midfielder, CF center forward, BR/BL ball recovery/loss, P pass, SG shot at goal. The first number represents the number of the play and the second, the order of the event.

Figure 3: Projection of the multilayer graph in two single-layer graphs. GK goalkeeper, LD/RD left/right defender, MF midfielder, CF center forward, BR/BL ball recovery/loss, P pass, SG shot at goal. The first number represents the number of the play and the second, the order of the event. The thickened arrow in attacking play 2 represents the aggregation of events/passes from LD to RD.

Figure 4: Temporal aggregation of events and player projections. GK goalkeeper, LD/RD left/right defender, MF midfielder, CF center forward, BR/BL ball recovery/loss, P pass, SG shot at goal. The first number represents the number of the play, and the second represents the order of the event. The thickened arrow represents the aggregation of events/passes from LD to RD and from RD to MF.

Figure 5: Hierarchical events represented through multilayer networks. KO kick off, TO turn over, GS goal scored, GC goal conceded, BR/BL ball recovery/loss, P pass, SG shot at goal. The first number represents the number of the play, the second number represents the order of the event.

Similar metrics to the previously proposed can still be applied, although naturally corresponding to different concepts. For example, the concept of walk length applied to the n and $n+1$ level relationships reveals the number of attacking plays till a goal is scored.