

The *Haka* Network: Evaluating Rugby Team Performance with Dynamic Graph Analysis

Paolo Cintia¹, Michele Coscia² and Luca Pappalardo¹

¹ KDDLab - ISTI CNR, Via G. Moruzzi 1 - 56124 Pisa, Italy Email: {name.surname}@isti.cnr.it

² Naxys - University of Namur, Rempart de la Vierge 8, 5000 Namur Belgium Email: michele.coscia@unamur.be

Abstract—Real world events are intrinsically dynamic and analytic techniques have to take into account this dynamism. This aspect is particularly important on complex network analysis when relations are channels for interaction events between actors. Sensing technologies open the possibility of doing so for sport networks, enabling the analysis of team performance in a standard environment and rules. Useful applications are directly related for improving playing quality, but can also shed light on all forms of team efforts that are relevant for work teams, large firms with coordination and collaboration issues and, as a consequence, economic development. In this paper, we consider dynamics over networks representing the interaction between rugby players during a match. We build a pass network and we introduce the concept of disruption network, building a multilayer structure. We perform both a global and a micro-level analysis on game sequences. When deploying our dynamic graph analysis framework on data from 18 rugby matches, we discover that structural features that make networks resilient to disruptions are a good predictor of a team’s performance, both at the global and at the local level. Using our features, we are able to predict the outcome of the match with a precision comparable to state of the art bookmaking.

I. INTRODUCTION

Mining dynamics on graphs is a challenging, complex and useful problem [4]. Many networks are representation of evolving phenomena, thus understanding graph dynamics brings us a step closer for an accurate description of reality. Sensing technologies open the possibility of performing dynamic graph analysis in an ever expanding set of contexts. One of them is competition events. Sports analysis is particularly interesting because it involves a setting where both environment and rules are standardized, thus providing us an objective measure of players’ contributions.

These possibilities power a number of potentially useful applications. The direct one is related to the improvement of a team’s performances. Once the most important factors contributing or preventing victories are identified, the team can work on strategies to regulate their collective effort toward the best practices. However, there are non-sport related externalities too. A sport team is nothing more than a group of people with different skills that is trying to achieve a goal in the most efficient way possible. This description can be applied to any other form of team [15]: a start-up enterprise, a large manufacturing firm, a group of scientists writing a scientific

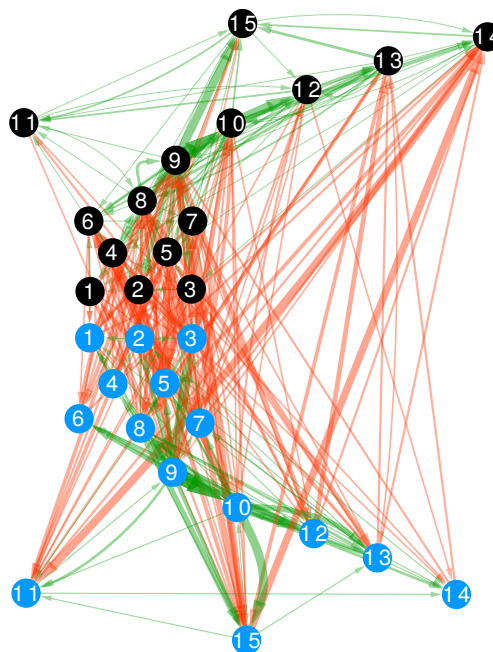


Fig. 1: The pass and tackles networks imposed on the same structure for one match in our dataset: Italy (blue nodes, home team) versus New Zealand (black nodes, away team). The number of the node refers to the role of the player. Green edges are the pass network and red edges are the tackle network. Node layout is determined by the classical rugby player positioning on the pitch during a scrum.

paper. If we are able to shed light on how group dynamics affect sport teams, we can try to universalize collaboration best practices to enhance productivity in many different scenarios.

It comes as no surprise that analyzing sport events is a fast growing literature in data science. Previous works involve the analysis of different sports: from soccer [14] [9] to basketball [27], from American football [20] to cycling [8]. In this paper we focus on a sport that was not analyzed before: rugby. The reason is that rugby has some distinctive features that makes it an ideal sport to consider. Like American football, it is a sport where there is a very clear and simple success measure: the number of meters gained in territory. Again like American football, it contains a disruption network: tackles. Unlike American football, performance is also related to how well a team can weave its own pass structure, involving all

the players in the field in the construction of uninterrupted sequences that can lead to a score.

One of the main contributions of the paper is to use both the pass and disruption networks at the same time, creating a multilayer network analysis [13] [3] of dynamics on graphs. Figure 1 depicts an example of such structure. To the best of our knowledge, there has not been such attempt elsewhere in the literature. We also do a multiscale analysis: we first focus on the multilayer network as a whole – as a result of all interactions during a match like in [9] – and we perform also a micro-level analysis on each match sequence.

Our data comes from the 2012 Tri-Nations championship, 2012 New Zealand Europe tour, 2012 Irish tour to New Zealand and 2011 Churchill Cup (only USA matches), for a total of 18 matches. The data was collected by Opta¹, using field positioning and semi-automatically annotated events.

We find that there are some features of the pass networks that make teams more successful in their quest for territorial gain. In particular, the connectivity of the network seems to play an important role. Being able avoid structurally crucial players, that would make part of the team isolated if they were removed from the network by a tackle, is associated with the highest amounts of meters made. This is also confirmed if we do not analyze only the global network as the result of the entire match, but also mining the patterns of each single sequence of the game. In the latter case, the signal is harder to disentangle from noise, and most analyzed features did not yield any significant result. This highlights how much rugby is a game dependent on a grand match strategy, rather than on just a sequence-by-sequence short term tactical one.

When applied to a real world prediction task, our framework fares well in comparison with state of the art attempts. The power in predicting the winner of a match is comparable with the one of bookmakers, who have access to the full history of teams and of the players actually performing on the pitch. In particular, our framework is less susceptible to reputation bias: the algorithm is not afraid to design New Zealand as a loser in what bookmakers saw as a great upset result – its loss to England in the December 1st, 2012 Twickenham clash.

II. RELATED WORK

In the last decade, sports analytics has increased its pervasiveness as large-scale performance data became available [23]. Researchers from different disciplines started to analyze massive datasets of players' and teams' performance collected from monitoring devices. The enormous potential of sports data is affecting both individual and team sports, providing a valid tool to verify existing sports theories and develop new ones. As an example in individual sports, Cintia et al. [8], [6] develop a first large scale data-driven study on cyclists' performance by analyzing the workouts of 30,000 amateur cyclists. The analysis reveals that cyclists' performances follow precise patterns, thus discovering an efficient training program learned from data. In tennis, Yucesoy and

Barabási develop a predictive model that relies on a tennis player's performance in tournaments to predict her popularity [28]. In NBA basketball league the performance efficiency rating introduced by Hollinger [12] is a stable and widely used measure to assess players' performance by combining the manifold type of data gathered during every game (pass completed, shots achieved, etc.). Vaz de Melo et al. [27] introduced network analysis to the mix. In baseball, Rosales and Spratt propose a new methodology to quantify the credit for whether a pitch is called a ball or strike among the catcher, the pitcher, the batter, and the umpire involved [21]. Smith et al. propose a Bayesian classifier to predict baseball awards in the US Major League Baseball, reaching an accuracy of 80% in the predictions [25]. In soccer, networks are a widely used tool to determine the interactions between players on the field, where soccer players are nodes of a network and a pass between two players represents a link between the respective nodes. For example, Cintia et al. [9], [7] exploit passing networks to detect the winner of a game based on the passing behavior of the teams. [14] shared the aim, without using networks. They discovered that, while the strategy of the majority of successful teams is based on maximizing ball possession, another successful strategy is to maximize a defense/attack efficiency score. Dynamic graph analysis has also been applied for ranking purposes [24], [20], [18]. Other examples of wider applicability of team-based success research comes from analysis of citation [22] and social [15] networks.

Methodologically, our paper is indebted with the vast field of dynamics on networks analysis [4]. This field is usually approached both from a statistical [26] and a mining perspective [5]. We adopt the latter approach. In mining network dynamics, the aim is to find regularities in the evolution of networks [2]. There are many applications for these techniques, beyond sports analytics: epidemiology [17], mobility [10] and genetics [16]. In this work, we focus on a narrower part of this field, since our networks are multilayer: nodes can be connected with edges belonging to multiple types. A good survey on multilayer networks, both modeling and analysis, can be found in [13]. The specific multilayer model we adopt is the one of multidimensional networks firstly presented in [3].

III. DATA

The data has been collected by Opta and made available in 2013 as part of the AIG Rugby Innovation Challenge². Opta performs a semi-automatic data collection that happens as the game unfolds. Sensors feed a team composed by two or more humans, who annotate the various actions of the match. Once the event ended, the data is double-checked for consistency and then serialized as an XML file.

A pass p is defined by an action that was coded as successful pass in the data. This means we drop forward, intercepted or otherwise erroneous passes, that directly result in the team losing possession. A pass is composed by a pair of players:

¹<http://www.optasports.com/>

²<http://optasports.com/news-area/aig-rugby-innovation-challenge-competition-launched.aspx>

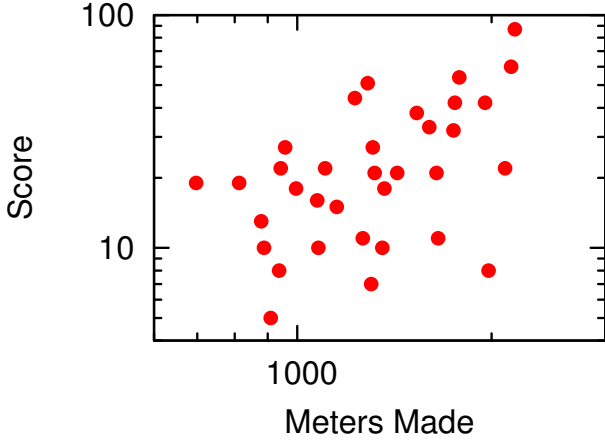


Fig. 2: The relationship between meters made on a carry and final score in our dataset. We exclude outliers for teams who did not score in a match (two occurrences in our data).

the pass originator and the pass receiver. Both players have to be part of the same team. We refer to $P_{t,g}$ as the set of passes made by team t in game g – i.e. what we call “Pass network”. $|P_{t,g}|$ is the number of total passes made, which is equivalent to the sum of the edge weights of the pass network.

A disruption d , also called tackle, is defined by an action that was coded as successful tackle in the data. This means we keep all different tackle types recorded by the data provider (chases, line, guard, etc), but we drop the ones whose result was not a clean tackle, meaning that the defender conceded a penalty or allowed the attackers to continue the offense. A disruption is composed by a pair of players: the tackler and the tackled. Players have to belong to different teams, making this a bipartite network. Similarly to the pass notation, $D_{t,g}$ and $|D_{t,g}|$ denote the set and number of tackles made, respectively. For convenience, we define an aggregate of the disruption action: $d_{u,t,g}$ is the number of all disruptions targeted at player u by team t in in game g over all disruptions made by t in g .

A sequence s is a list of passes and disruptions. Roughly, we define a sequence as a phase of the game going from the starting of a possession to its end. In rugby terminology, this means defining a sequence as a phase of game going from an interruption of the game to another. Interruptions are tries, penalties, drop goals, scrums and lineouts. Note that clean changes in possession (intercepts, ruck steals and others) are considered interruptions too. The union of the pass and tackle networks result in a structure of the type depicted in Figure 1.

In Section V, the quality measure used to distinguish between successful and unsuccessful teams is the number of meters made with a carry. Rugby is a very territory oriented game, and it can be won basically by gaining more meters ball in hand. There is a very high correlation between meters made and score/victory, and Figure 2 depicts this relationship (in the figure, the Pearson correlation is $\sim .58$). As a consequence, we use the information recorded in the `metres` attribute of each `actionrow` element in the data as our success measure.

| Symbol | Meaning |
|---------------------|--|
| u | Player |
| t | Team |
| g | Single game (match) |
| p | Pass |
| d | Disruption (tackle) |
| $d_{u,t,g}$ | Relative number of disruptions targeted to u by t in g |
| s | Sequence, a set of passes and tackles |
| $P_{t,g}$ | Passes made in game g by team t |
| $D_{t,g}$ | Disruptions made in game g by team t |
| $m_{s,t}$ | Meters made in sequence s by team t |
| $\mathcal{M}_{g,t}$ | Meters made in game g by team t |

TABLE I: The notation used in the paper.

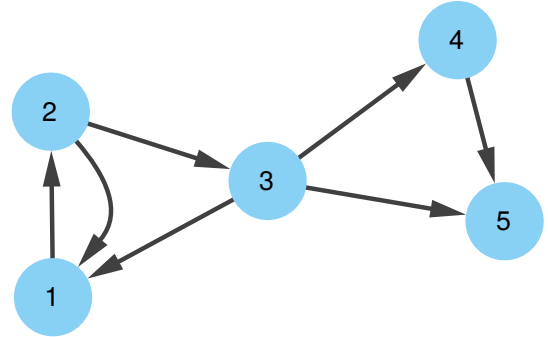


Fig. 3: A toy model of a directed graph.

We refer to meters gained in a sequence s by team t as $m_{s,t}$, and we use it in the sequence analysis (Section V-B). In the case of the structural analysis (Section V-A), we consider the game g as a whole, and we define the overall success as all the meters made, i.e. $\mathcal{M}_{g,t} = \sum_{s \in g} m_{s,t}$.

Table I sums up the notation used in the paper.

IV. FEATURES

In this section we consider the multilayer network of passes and disruptions as a whole, as it results from all interactions between and within the two teams over the entire length of the match. We define a set of features characterizing the network. We aim at predicting the number of meters made on a carry by the team over the match. The features are team-dependent, thus for each match we have two observations: the feature for the home team and for the away team. Each team has two set of features: the pass features and the disruption features.

A. Pass Features

The pass features are features calculated over the team’s pass network $P_{t,g}$. We consider the topology of $P_{t,g}$ in isolation, as a directed graph that is not interacting with other external events. The features we define are purely topological ones and they are: connectivity ($\gamma_{t,g}(P_{t,g})$), assortativity ($\rho_{t,g}(P_{t,g})$), number of strongly connected components ($\sigma_{t,g}(P_{t,g})$), and clustering ($\Delta_{t,g}(P_{t,g})$).

Connectivity is defined as follows. Given two nodes, connectivity is the number of nodes that must be removed to break all paths from the two nodes in $P_{t,g}$ [1]. For instance, in Figure 3, to separate nodes 2 and 4 you need only to remove

one node (3). With $\gamma_{t,g}(P_{t,g})$ we refer to the average of local node connectivity over all pairs of nodes of $P_{t,g}$.

Assortativity ($\rho_{t,g}(P_{t,g})$) is the Pearson correlation coefficient of the degrees of all pairs of nodes connected by an edge [19]. A positive assortativity means that in the network nodes with high degree tends to connect with other nodes with high degree and vice versa. Figure 3 represents an assortative network, as the degree correlation coefficients is ~ 0.15 .

A strongly connected component in a network is a set of nodes for which there is a path from any node of the component to any other node in the component following the directed edges of the graph. Given a directed network, there might be zero, one or more strongly connected components. Our $\sigma_{t,g}(P_{t,g})$ detects and counts the number of strongly connected components in $P_{t,g}$. Figure 3 depicts a directed graph with three strongly connected components: one composed by nodes 1, 2 and 3; and two other components composed by the single nodes 4 and 5, since they are connected by a single directed edge which does not allow to reach node 4 from 5.

The clustering of a node u is the fraction of possible triangles through that node that exist:

$$c_u = \frac{2T_u}{k_u(k_u - 1)},$$

where T_u is the number of triangles through node u , and $k_u > 1$ is its degree (for $k_u \leq 1$ the convention is to fix $c_u = 0$). The $\Delta_{t,g}(P_{t,g})$ feature is the mean clustering, or:

$$\Delta_{t,g}(P_{t,g}) = \frac{1}{n} \sum_{u \in t} c_u,$$

where n is the number of nodes in $P_{t,g}$ (for rugby usually $n = 15$, because there are 15 players in a rugby team, although in some cases a player might never receive a pass, setting $n = 14$). Note that clustering is defined for undirected graphs. Thus, in this case, $P_{t,g}$ is projected in a derived structure in which we ignore edge direction. This is the only one of the four measures for which we have to perform this projection. The graph in Figure 3 has a high clustering coefficient (~ 0.87), because nodes 1, 2, 4 and 5 all have $c_u = 1$ – they are all part of the only triangle they could be part of – while node 3 is part of only two of its six possible triangles.

B. Disruption Features

The disruption features exploit the multilayer nature of the conjunction between $P_{t,g}$ and $D_{t,g}$. We consider how different the features of $P_{t,g}$ become when one of its nodes gets disabled by a tackle. We compare these features with the version of $P_{t,g}$ where no node gets removed and we weight this difference by the relative number of times that the disruption has been made. For each disruption feature we use the same notation as the pass feature, with an overline. We define $\overline{P_{t,g}^{\sim u}}$ as the pass network $P_{t,g}$ when deprived of node u .

For connectivity, our definition is as follows:

$$\overline{\gamma_{t_1,g}}(P_{t_1,g}) = \sum_{u \in t_1} (d_{u,t_2,g} \times (\gamma_{t_1,g}(P_{t_1,g}^{\sim u}) - \gamma_{t_1,g}(P_{t_1,g}))).$$

In practice, the $\gamma_{t_1,g}(P_{t_1,g}^{\sim u}) - \gamma_{t_1,g}(P_{t_1,g})$ term calculates what happens to the connectivity of $P_{t_1,g}$ when removing u . A negative number means that the connectivity decreases, or $\gamma_{t_1,g}(P_{t_1,g}^{\sim u}) < \gamma_{t_1,g}(P_{t_1,g})$: you need to remove fewer nodes to disconnect pairs of nodes in the network. A positive value means that the connectivity increases. The $d_{u,t_2,g}$ term weighs this connectivity change with the relative number of times t_2 was able to disable player u with a successful disruption. The sum term aggregates the measure over all players representing team t_1 , the team receiving the tackles.

To have an intuition of this operation, consider the tackle features as an expression of the resilience level of the pass network: they estimate how much the network is resistant to disruptions. The higher the value, the more resilient the pass network is.

The other disruption features are defined following the same template:

$$\overline{\rho_{t_1,g}}(P_{t_1,g}) = \sum_{u \in t_1} (d_{u,t_2,g} \times (\rho_{t_1,g}(P_{t_1,g}^{\sim u}) - \rho_{t_1,g}(P_{t_1,g}))),$$

$$\overline{\sigma_{t_1,g}}(P_{t_1,g}) = \sum_{u \in t_1} (d_{u,t_2,g} \times (\sigma_{t_1,g}(P_{t_1,g}^{\sim u}) - \sigma_{t_1,g}(P_{t_1,g}))),$$

$$\overline{\Delta_{t_1,g}}(P_{t_1,g}) = \sum_{u \in t_1} (d_{u,t_2,g} \times (\Delta_{t_1,g}(P_{t_1,g}^{\sim u}) - \Delta_{t_1,g}(P_{t_1,g}))),$$

for assortativity, number of strongly connected components, and clustering, respectively.

In the case of disruptions, we have an additional feature. For each player u in a t, g pair we can calculate a centrality value, answering the question: how central was player u for team t in game g ? Thus, $\beta_{u,t,g}$ is defined as u 's closeness centrality [11] in $P_{t,g}$. The tackle centrality disruption is then defined as:

$$\overline{\beta_{t_1,g}}(P_{t_1,g}) = \sum_{u \in t_1} (d_{u,t_2,g} \times \beta_{u,t_1,g}).$$

It represent the weighted average closeness centrality of the players tackled in t_1 by t_2 .

V. ANALYSIS

In this section, we perform the network analysis to estimate the meters gain by rugby team using their network features. We start by looking at the global pass and disruption features, calculated over the entire aggregated match (Section V-A). We then focus on an analysis considering each match sequence as a single observation (Section V-B).

A. Structural Analysis

In this section we calculate the network features presented in the previous section over the pass and disruption networks. We use these features as independent variables of a simple OLS model. The dependent variable of the model is the number of meters made carrying the ball, as described in Section III. We log transform the dependent variable.

| | Dependent variable: | | | |
|---------------------|---------------------|---------------------|---------------------|----------------------|
| | $M_{g,t}$ | | | |
| | (1) | (2) | (3) | (4) |
| h | 0.058 (0.096) | 0.160* (0.084) | 0.081 (0.096) | 0.077 (0.084) |
| $\log P_{t,g} $ | 0.366* (0.201) | -0.218 (0.217) | 0.379** (0.158) | 0.159 (0.165) |
| $\Delta_{t,g}$ | 0.451 (0.758) | | | |
| $\gamma_{t,g}$ | | 0.285*** (0.075) | | |
| $\rho_{t,g}$ | | | 0.471 (0.360) | |
| $\sigma_{t,g}$ | | | | -0.073*** (0.024) |
| Constant | 5.118*** (0.724) | 7.547*** (0.891) | 5.451*** (0.764) | 6.609*** (0.816) |
| Observations | 36 | 36 | 36 | 36 |
| R^2 | 0.247 | 0.474 | 0.277 | 0.409 |
| Adjusted R^2 | 0.176 | 0.424 | 0.210 | 0.353 |
| Residual Std. Error | 0.273 | 0.229 | 0.268 | 0.242 |
| F Statistic | 3.500** | 9.602*** | 4.096** | 7.375*** |

Note: *p<0.1; **p<0.05; ***p<0.01

TABLE II: The regression results predicting meters made using pass network features.

Note that in the regression we control for the home factor with a binary variable h . This is done because home advantage is very strong in rugby³, and we want to make sure it does not affect our results. A second control we impose is the number of passes made. There is an expected correlation between how many meters a team will advance and the time it has possession of the ball. The number of passes made is a good proxy for this information.

We first check the correlations between the pass features and the number of meters made. Table II reports the results of our models. We check one feature at a time, always including our controlling factors. Two of the four features do not exhibit a significant correlation with the number of meters made: clustering ($\Delta_{t,g}$) and assortativity ($\rho_{t,g}$). This means that, when crafting their own pass network, teams are not required to encourage or discourage triadic closure – a receiver of a pass passing to the passer that originated the action – and assortativity – passing the ball to a player with a team connectivity similar to their own.

More interesting are the significant associations with connectivity ($\gamma_{t,g}$) and number of strongly connected components ($\sigma_{t,g}$). These two measures are related and it is no surprise they have opposite signs: the higher the connectivity of a network the fewer components it has. We interpret these coefficients as follows: a rugby team’s pass network should ensure a strong connectivity, likely establishing that there are multiple and reciprocal pathways for the ball to reach all players. Redundancy and structural strength are important in a rugby team.

These two factors are able to explain away the simple control on quantity of possession, represented by the number of passes made. Note that it is remarkable to obtain a significance

³As of 2016, a weak team like Italy won only 12 out of 85 matches in the Six Nations, and 11 of them were in Italy.

| | Dependent variable: | | | | |
|----------------------|---------------------|---------------------|---------------------|---------------------|----------------------|
| | $M_{g,t}$ | | | | |
| | (1) | (2) | (3) | (4) | (5) |
| h | 0.162* (0.091) | 0.195* (0.103) | 0.183** (0.085) | 0.225** (0.109) | 0.128 (0.092) |
| $\log D_{t,g} $ | -0.203 (0.129) | -0.263* (0.149) | -0.175 (0.121) | -0.190 (0.152) | -0.082 (0.131) |
| $\bar{\beta}_{t,g}$ | 1.601*** (0.372) | | | | |
| $\bar{\Delta}_{t,g}$ | | 1.379** (0.508) | | | |
| $\bar{\gamma}_{t,g}$ | | | 0.243*** (0.048) | | |
| $\bar{\rho}_{t,g}$ | | | | 0.805** (0.378) | |
| $\bar{\sigma}_{t,g}$ | | | | | -0.082*** (0.019) |
| Constant | 7.206*** (0.564) | 7.360*** (0.642) | 7.401*** (0.519) | 8.077*** (0.646) | 7.747*** (0.544) |
| Observations | 36 | 36 | 36 | 36 | 36 |
| R^2 | 0.415 | 0.250 | 0.483 | 0.191 | 0.414 |
| Adjusted R^2 | 0.360 | 0.179 | 0.434 | 0.115 | 0.359 |
| Residual Std. Error | 0.241 | 0.273 | 0.227 | 0.283 | 0.241 |
| F Statistic | 7.563*** | 3.548** | 9.956*** | 2.517* | 7.537*** |

Note: *p<0.1; **p<0.05; ***p<0.01

TABLE III: The regression results predicting meters made using tackling network features.

of $p < .01$ with a very small sample size ($N = 36$). We do not show a model containing all features at once, because the collinearity between $\gamma_{t,g}$ and $\sigma_{t,g}$ makes them both not significant.

We now turn our attention to the disruption features. In this case, instead of controlling for the number of passes, we control for the number of disruptions. The control whether a team played home or away is still in place. Table III reports the results of these models.

Differently from the pass features, all disruption features are significantly correlated with the number of meters made. The result of the average tackle centrality $\bar{\beta}_{t,g}$ seems surprising: the highest the centrality of the most targeted players the more meters the team will make. However, this is only an apparent surprise: the most central players are the scrum half and fly half (number 9 and 10 in Figure 1) which are players that do not usually cover a lot of ground. On the other hand, the wings (number 11 and 14 in Figure 1) are very peripheral but also expected to cover most of the ground. This result is indeed expected and not very interesting.

More interesting is that the more a team can maintain a high value of clustering, connectivity and assortativity after being disrupted by the average opponent tackle, the more it can advance on the pitch. As expected, the number of strong components is still negative: if after disruption the team still does not have many isolated components it is expected to be able to advance. The advice for a team would be to target their tackles to the players who are the most responsible to keep their pass network connected, but not the most central. The significance levels of $\bar{\gamma}_{t,g}$ and $\bar{\sigma}_{t,g}$ are higher also in this case.

Note that connectivity is the single most important feature discovered. Both in Table II and Table III the associated R^2 is the highest. Together with our controls, $\gamma_{t,g}$ and $\bar{\gamma}_{t,g}$ are able to explain 47–48% of the variance in number of meters made.

Is it true that connectivity is the most important feature to

Tackle Connectivity

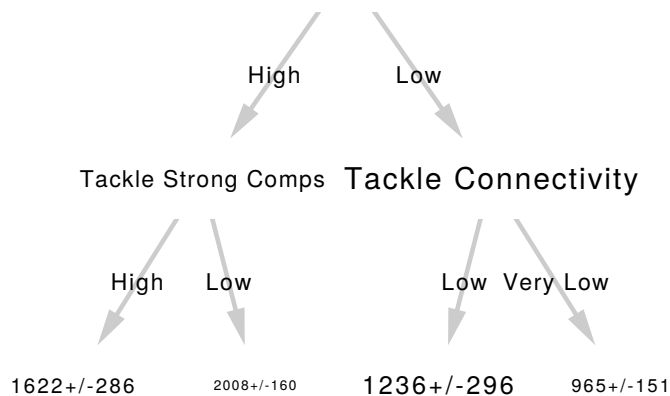


Fig. 4: The decision tree for predicting the number of meters made in a match using the pass/disruption features. The tree node’s font size is proportional to the number of observations with the given characteristics. Leaves report the expected number of meters made for the branch. Edge labels the value of the node’s feature.

predict a rugby team’s performance? We conclude this section by showing how we can use these features in a data mining framework to predict the number of meters a team will cover ball in hand. We apply a simple decision tree technique⁴, where the target variable is the number of meters made and the predictors are all the features we discussed so far.

Figure 4 depicts the result. The most important split variable the algorithm found was indeed a connectivity measure, the one resulting from tackles. It is such an important feature that it has been selected twice at two different tree levels – note that we pruned the tree to avoid overfitting. A very low tackle connectivity means that the team, as result from the opponent’s disruptions, lost most of its original connectivity. This is associated with the poorest performances in meters gained: a very low tackle connectivity resulted in less than 1,000 expected meters made. This is as little as half the expectation for a team with a high connectivity retention, plus the ability of not having its own strongly connected components broken apart. In this case, the team is expected to advance 2,000 meters.

B. Sequence Analysis

A deeper evaluation of a team performance can be obtained by the analysis of how team networks are built, action after action. To do that, we split each game in sequence of possession phases, i.e. time intervals where a team is controlling the ball. The split into possession phases is made by selecting all the events between two events that identify a start of possession. In particular, we sort all the events according to their timestamps, then we select all the sequences between possession events of the two teams playing. The possession of a team lasts until

⁴Implementation obtained from <http://www.borgelt.net/>.

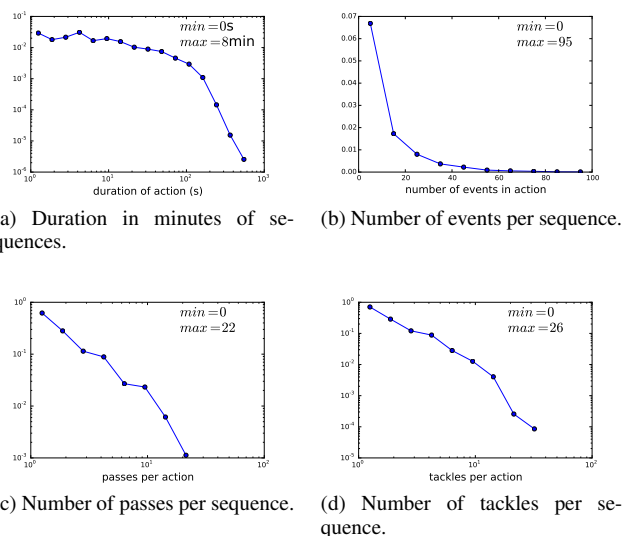


Fig. 5: The distribution of sequence statistics across all the observed matches.

the opposite team performs a possession event. The possession events we consider to split a game into possession phases are the following: *50m Restart*, *22m Restart*, *Free Kick*, *Turnover Won*, *Lineout*, *Lineout Steal*, *Scrum*, *Scrum Steal*, *Ruck Won*, *Maul Won*, *Penalty*, *Pass intercepted*.

Figure 5 depicts the distribution of various features across game sequences. The overall trend suggests that these are all broad distributions: the number of observations does not allow us to test for a power-law hypothesis, but we can conclude that there are heavy tails. In particular, Figure 5(a) shows that usually game sequences are very brief, but some can last for multiple minutes, when a team performs prolonged phases of possession – for instance attempting to score a try. This in turn implies that the number of events per action also distributes broadly (Figure 5(b)) even when broken down in passes (Figure 5(c)) and tackles (Figure 5(d)).

The goal of such an analysis is to understand the dynamics of a game relying on the network features already described in the previous sections. Here we are not exactly interested on predicting the result of a game: the focus is on the analysis of the single features in a dynamic context, in order to evaluate the additional knowledge provided by the network features we consider.

Once a game is subdivided into possession phases, we can observe how the pass network – i.e. the passing interactions between players – grow across time. As a performance measure, we use the quantity of meters gained by a team from which we subtract the meters gained by its opposition. We analyze the average value of each feature during time and we compare it to the performance of the team, for each of the 18 games we have in our dataset. Among the features we are interested in (*Connectivity*, *Assortativity*, *Strongly connected components*, *Clustering*), we observe a significant negative correlation be-

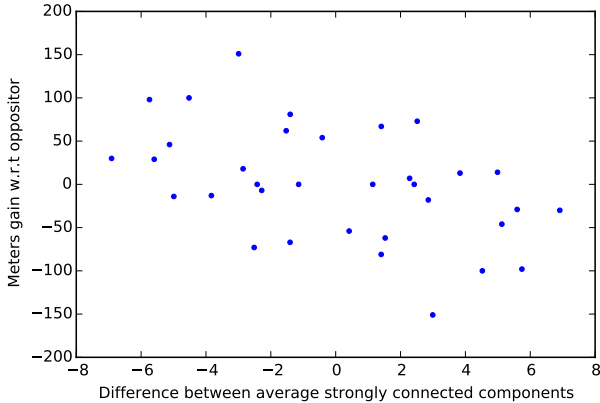


Fig. 6: The correlation between strongly connected components and meters gained w.r.t. the opposition.

tween the average number of strongly connected components, and the meters gained (minus the opposition gains). In Figure 6 such a correlation ($\rho = 0.49$) is highlighted.

This confirms the global analysis: if a team breaks down its effort in many isolated components it is unlikely to be able to gain additional meters. The fact that this is the only relevant feature – and that no tackle feature was found significant – suggests two additional insights. First, that rugby is a game fundamentally different than soccer: in the literature it has been shown that single sequence features were more relevant than here to evaluate team success [9]. Second, since these features were relevant when calculated over the entire match pass network (as Section V-A shows), it suggests that rugby has a peculiar dynamics. Our evidence points that, in rugby, each action might matter not in isolation, but as the part of a grand match strategy, that can be only appreciated by analyzing the whole pass network.

VI. PREDICTIVE MODEL

In this section, we test if our model based on pass and disruption network features is able to accurately predict the result of the game. We build a cross validation framework where we train our model on 17 matches, leaving one out, and then we predict the result of the match left out using the model trained on the other 17 matches. We repeat this procedure for all matches in the dataset. Since Section V-B showed that sequence features are not significant, we build our model using exclusively global pass network features.

We perform two prediction tasks. The two tasks differ in the target variable of interest. In the first task, we aim at predicting which one between the two teams will gain more meters during the match. With our model, we are able to obtain the correct answer for 15 out of 18 matches, i.e. with an accuracy $\sim 83\%$.

Since in rugby the number of meters gained is highly correlated with both score and odds of winning, we can use our model to predict also who is going to win the match. We say that the team predicted to gain more meters is going to win. In this case, we make the correct prediction for 14 out

of 18 matches, i.e. with an accuracy $\sim 77\%$. Note that the reduced accuracy is due to the fact that one of the matches in the dataset ended up in a draw. This is a very rare occurrence in rugby, and it was not encoded in the model⁵.

How good is our prediction? A random predictor would flip a coin and get the right answer 50% of the times. However, rugby matches tend to be predictable, given enough information about past performances of teams and players. These performances are recorded by the World Rugby organization, which publishes weekly updated national rankings of teams. It is reasonable to assume that the higher ranked team of the two playing is expected to win. If we use the World Rugby rankings to predict the outcomes of the matches, we obtain a very similar accuracy: $\sim 76\%$ ⁶.

We can do slightly better by looking at historical odds data⁷. Bookmakers are more invested in getting right a specific match prediction than World Rugby. In fact, their accuracy was higher, both of World Rugby rankings and of our model: $\sim 86\%$. However, we could find data only for 14 matches, the ones involving New Zealand, because there is no historic record for the USA rugby matches. This makes the prediction task easier: lower ranked teams are more unpredictable when playing each other, and the bookmakers always picked New Zealand for all the matches it played, being New Zealand such a dominant rugby team.

To conclude this section, it is worthwhile noting two things. First, our model was able to successfully predict the biggest upset of the 18 played matches: the victory of England over New Zealand. Neither World Rugby nor bookmakers predicted that. Second, our model is a purely structural system, that has no information about which team and which players are performing. As such, its information pool is more restricted than the one available to both World Rugby and bookmakers. The fact that the model’s performance are on par with theirs is rather encouraging. It is true that we then feed the model perfect information recorded during the match, but we detail in the conclusions how we plan to create a truly predictive model.

VII. CONCLUSION AND FUTURE WORKS

In this paper we build a multilayer network analysis framework to describe the performance of rugby teams during a match. We build two layers: a pass network and a tackle network. We extract features from these layers and we use them at two analytical levels. First, we extract features from the network as a whole, representing the entire match. Then, we divide the match in sequences of a single match action and we extract sequence features. We use these features as correlates of match performance, estimated by the number of meters a team advanced on the field. We discover that,

⁵The match was Australia v New Zealand, played on October 20th, 2012. Our model predicted a win for New Zealand.

⁶Note that this is calculated over 17 matches, not 18, because one match involved the reserve English national team, the England Saxons, which is not ranked by World Rugby.

⁷From <http://www.oddsportal.com>

using the global features, connectivity is very important for a team to successfully advance. Second, we perform single sequence mining and we find that, when considering actions in isolation, most features have no significant relation with a team's performance. This shows how much rugby is a different game than soccer, where this analysis yielded the opposite result [9]. Our interpretation is that soccer is a game of tactics, where each sequence yields results that are mostly independent from the other sequences; while rugby is a game of strategy, where sequences build on each other to obtain the intended result. This is not to say that sequence analysis is not useful: by looking at the dynamic graph one might be able to understand key moments of games – i.e. moments where the prospective winner change. Finally, we show how our predictive framework is on par with state of the art bookmaker estimates, and better suited to predict upset results.

There are a number of directions that we can explore for future works. First we can work on a network comparison of different sports, mainly soccer. This would build on top of the differences between the two sports we highlighted here. Second, we can investigate other reasons of the poor predictive performance of sequences. Sequences of sequences might give more insight to appreciate the dominance of a team not during the whole game or just one sequence, but during an intermediate period of time. Third, we are planning to use our framework to discuss how it is able to shed light on large performance shocks. In our data, we have several New Zealand versus Ireland matches, with very different outcomes: one ended 22-19 and another 60-0. The question would be: what changed in Ireland's performances across these two matches, played only a week apart? Another interesting case study would be the England versus New Zealand upset: what made the English team the only one able to beat New Zealand in the 14 matches in our data? Finally, we could test our predictive model on an actual prediction. We could build a predicted match network before the match starts and use it to pick the winner before the match happens, not after as we did here. To do so, we would need more data, as 18 matches are not enough for reliably training our system.

Acknowledgments. We thank Opta for having made available the data on which this paper is based. We thank Vittorio Romano for helping us with his knowledge of rugby rules and game dynamics. Michele Coscia has been partly supported by FNRS, grant 24927961. This work has been partially funded by the following European projects: Cimplex (Grant Agreement 641191) and SoBigData RI (Grant Agreement 654024).

REFERENCES

- [1] Lowell W Beineke, Ortrud R Oellermann, and Raymond E Pippert. The average connectivity of a graph. *Discrete mathematics*, 252(1):31–45, 2002.
- [2] Michele Berlingerio, Michele Coscia, Fosca Giannotti, Anna Monreale, and Dino Pedreschi. Evolving networks: Eras and turning points. *Intelligent Data Analysis*, 17(1):27–48, 2013.
- [3] Michele Berlingerio, Michele Coscia, Fosca Giannotti, Anna Monreale, and Dino Pedreschi. Multidimensional networks: foundations of structural analysis. *World Wide Web*, 16(5-6):567–593, 2013.
- [4] Stefano Boccaletti, Vito Latora, Yamir Moreno, Martin Chavez, and D-U Hwang. Complex networks: Structure and dynamics. *Physics reports*, 424(4):175–308, 2006.
- [5] Björn Bringmann, Michele Berlingerio, Francesco Bonchi, and Aristides Gionis. Learning and predicting the evolution of social networks. *Intelligent Systems, IEEE*, 25(4):26–35, 2010.
- [6] P Cintia, L Pappalardo, and D Pedreschi. Mining efficient training patterns of non-professional cyclists. In *22nd Italian Symposium on Advanced Database Systems (SEBD 2014), Sorrento Coast, Italy, June 16–18, 2014*, pages 1–8, 2014.
- [7] Paolo Cintia, Fosca Giannotti, Luca Pappalardo, Dino Pedreschi, and Marco Malvaldi. The harsh rule of the goals: Data-driven performance indicators for football teams. In *2015 IEEE International Conference on Data Science and Advanced Analytics, DSAA 2015, Campus des Cordeliers, Paris, France, October 19-21, 2015*, pages 1–10, 2015.
- [8] Paolo Cintia, Luca Pappalardo, and Dino Pedreschi. Engine matters: A first large scale data driven study on cyclists' performance. In *Data Mining Workshops (ICDMW), 2013 IEEE 13th International Conference on*, pages 147–153. IEEE, 2013.
- [9] Paolo Cintia, Salvatore Rinzivillo, and Luca Pappalardo. A network-based approach to evaluate the performance of football teams. In *Machine Learning and Data Mining for Sports Analytics Workshop, Porto, Portugal, 2015*.
- [10] Michele Coscia, Salvatore Rinzivillo, Fosca Giannotti, and Dino Pedreschi. Optimal spatial resolution for the analysis of human mobility. In *Advances in Social Networks Analysis and Mining (ASONAM), 2012 IEEE/ACM International Conference on*, pages 248–252. IEEE, 2012.
- [11] Linton C Freeman. Centrality in social networks conceptual clarification. *Social networks*, 1(3):215–239, 1978.
- [12] John Hollinger. The player efficiency rating, 2009.
- [13] Mikko Kivelä, Alex Arenas, Marc Barthelemy, James P Gleeson, Yamir Moreno, and Mason A Porter. Multilayer networks. *Journal of Complex Networks*, 2(3):203–271, 2014.
- [14] Joaquin Lago-Ballesteros and Carlos Lago-Peñas. Performance in team sports: Identifying the keys to success in soccer. *Journal of Human Kinetics*, 25:85–91, 2010.
- [15] David Lazer and Allan Friedman. The parable of the hare and the tortoise: Small worlds, diversity, and system performance. 2005.
- [16] Nicholas M Luscombe, M Madan Babu, Haiyuan Yu, Michael Snyder, Sarah A Teichmann, and Mark Gerstein. Genomic analysis of regulatory network dynamics reveals large topological changes. *Nature*, 431(7006):308–312, 2004.
- [17] Robert M May and Alun L Lloyd. Infection dynamics on scale-free networks. *Physical Review E*, 64(6):066112, 2001.
- [18] Shun Motegi and Naoki Masuda. A network-based dynamical ranking system for competitive sports. *Scientific reports*, 2, 2012.
- [19] Mark EJ Newman. Mixing patterns in networks. *Physical Review E*, 67(2):026126, 2003.
- [20] Juyong Park and Mark EJ Newman. A network-based ranking system for us college football. *Journal of Statistical Mechanics: Theory and Experiment*, 2005(10):P10014, 2005.
- [21] Joe Rosales and Scott Spratt. Who is responsible for a called strike? In *The 9th annual MIT Sloan Sports Analytics Conference*, 2015.
- [22] Emre Sarigöl, René Pfitzner, Ingo Scholtes, Antonios Garas, and Frank Schweitzer. Predicting scientific success based on coauthorship networks. *EPJ Data Science*, 3(1):1–16, 2014.
- [23] R. Schumaker, O Solieman, and H. Chen. *Sports Data Mining*. Springer, 2010.
- [24] Clément Sire and Sidney Redner. Understanding baseball team standings and streaks. *The European Physical Journal B*, 67(3):473–481, 2009.
- [25] Lloyd Smith, Bret Lipscomb, and Adam Simkins. Data mining in sports: Predicting cy young award winners. *Journal of Computing Sciences in Colleges*, 22(4):115–121, 2007.
- [26] Tom AB Snijders. The statistical evaluation of social network dynamics. *Sociological methodology*, 31(1):361–395, 2001.
- [27] Pedro OS Vaz de Melo, Virgilio AF Almeida, Antonio AF Loureiro, and Christos Faloutsos. Forecasting in the nba and other team sports: Network effects in action. *ACM Transactions on Knowledge Discovery from Data (TKDD)*, 6(3):13, 2012.
- [28] Burcu Yucesoy and Albert-László Barabási. Untangling performance from success. *arXiv preprint arXiv:1512.00894*, 2015.