

Research article

Evaluating Australian Football League Player Contributions Using Interactive Network Simulation

Jonathan Sargent  and Anthony Bedford

School of Mathematical & Geospatial Sciences, RMIT University, Australia

Abstract

This paper focuses on the contribution of Australian Football League (AFL) players to their team's on-field network by simulating player interactions within a chosen team list and estimating the net effect on final score margin. A Visual Basic computer program was written, firstly, to isolate the effective interactions between players from a particular team in all 2011 season matches and, secondly, to generate a symmetric interaction matrix for each match. Negative binomial distributions were fitted to each player pairing in the Geelong Football Club for the 2011 season, enabling an interactive match simulation model given the 22 chosen players. Dynamic player ratings were calculated from the simulated network using eigenvector centrality, a method that recognises and rewards interactions with more prominent players in the team network. The centrality ratings were recorded after every network simulation and then applied in final score margin predictions so that each player's match contribution—and, hence, an optimal team—could be estimated. The paper ultimately demonstrates that the presence of highly rated players, such as Geelong's Jimmy Bartel, provides the most utility within a simulated team network. It is anticipated that these findings will facilitate optimal AFL team selection and player substitutions, which are key areas of interest to coaches. Network simulations are also attractive for use within betting markets, specifically to provide information on the likelihood of a chosen AFL team list "covering the line".

Key words: Interaction matrix, negative binomial distribution, eigenvector centrality, player ratings.

Introduction

Australian Rules football, or AFL, is an invasion game played between two teams, each with 18 on-field players (and four reserves); a regular season consists of 18 teams each playing 22 matches. The dynamics of the game are similar to world football (association football or soccer), except that AFL players are permitted to use their hands to punch (handball) the ball to the advantage of a team member. The ultimate objective is to score a goal—worth six points—by kicking a ball through two upright posts at either end of the ground. Like other invasion games, scoring is the result of a series of critical events, or performance indicators, executed between the individuals involved in the contest (Nevill et al., 2002). These events are mostly discrete in nature, whether they are the number of kicks by player i or the number of times player j marks (catches) a kick from player i . In modern sports media, player performance indicators are intensively collected and published online across an ever-increasing number of

sports, both prior to and during a match. It is common for player i 's indicators from a match to be weighted and linearly combined, resulting in a numerical performance appraisal, or player rating. This methodology has become a standard for many fantasy sporting leagues—that is, to calculate players' post-match ratings then proportionally adjust their (fantasy) market value according to their rating fluctuations, as determined by a moving average from past matches. A criticism of this methodology is that it is too player-centric, ignoring an important underlying concept that a team is supposed to be more than the sum of the individual players (Gould and Gatrell, 1979/1980). Duch et al. (2010) argue that the real measure of player performance is "hidden" in the team plays, and not derived from strictly individual events associated with player i . Moreover, in their research on football-passing patterns from EURO 2004, Lee et al. (2005) measured passing between players at a group level rather than at an individual level, demonstrating how a player's passing patterns determined his location in the team network.

Discussions about network analysis commonly refer to the use of relational data or the interactions that relate one agent (player) to another and, so, preclude the properties of the individual agents themselves (Scott, 2000). The objective of this research was to move beyond such individual performance exploits, towards a measurement of each player's contribution to a dynamic system of team play. This was conceived through the identification of link plays within AFL matches (Sargent and Bedford, 2011), or sequences of play involving two or more players from team a where the ball's movement effectively increased scoring likelihood. Links were produced from data representing every "interaction" between the players; most games exceeded 2,500 cases. The interaction between pairs of players from team a within each link made it possible to generate an interaction matrix with which to observe player relations, or the number of times the ball passes from player i to player j on team a (Gould and Gatrell, 1979/80). For this research, symmetric interaction matrices were generated for each match played by the Geelong Football Club in 2011 and negative binomial distributions (nb) fitted to each player pair in the matrix so that their interaction frequency could be simulated. Pollard et al. (1977) concluded that the nb is a closer fit to events resulting from groups of players, rather than from individual performances; for example, an improved fit is observed from batting partnerships in cricket, rather than from individual batsman scores. Reep and Benjamin (1968) successfully modelled effective passes in world football

with *nbd*, while Pollard (1973) demonstrated how the number of touchdowns scored by a team in an American Football match closely followed the *nbd*. The *nbd* was considered to be a suitable fit to the AFL interaction data, able to simulate higher order interactions between pairs of superior players and lower order interactions between less prominent players.

After each match simulation, a rating for each player in the network was calculated using eigenvector centrality, a measure of the importance of a particular node (player) in a network (team)—that is, by determining the extent to which player *i* interacted with other central players. Centrality is a core concept in network analysis and has been applied in countless environments to determine patterns of flow, for example, infections, forwarded emails or money flowing through markets. Borgatti (2005) provides excellent definitions and applications of centrality in its various forms. The appeal of eigenvector centrality is its ability to measure the long-term influence of a node on the rest of the network, not just its immediate effect on adjacent nodes, as in degree centrality (Borgatti, 2005). Furthermore, a team strength index was calculated after each simulation from player centrality mean and variance, which was predictive of the team's final score margin. Through multiple iterations of the line-up and Jimmy Bartel's resulting net simulated effect on margin, the paper ultimately evaluates his contribution to a selected side. It is anticipated that the network simulation model should aid in determining the probability of a team "covering the line" in betting markets, once the team list has been released for the upcoming round of football.

Methods

Player interaction

Interaction frequency between any pair of players, $[i, j]$ from team *a* in a match is represented by the discrete random variable, r_{ij} . Three forms of interaction were recognised from our link play data:

- i) Primary interaction: efficient ball movement achieved through {Kick_{*i*}; Mark_{*j*}}, {Handball_{*i*}; Handball Received_{*j*}} or {Hit Out_{*i*}; Hit Out Received_{*j*}};
- ii) Secondary interaction: less efficient ball movement, namely player *j* gathering the ball off the ground ("Ball Get") due to an inaccurate player *i* event; team *a* retains possession of the ball;
- iii) Negative interaction: inefficient ball movement where player *i* relinquishes possession of the ball to player *k* from team *b* ("Turnover").

The interaction methodology is similar to "r-pass movement" in world football as defined by Reep and Benjamin (1968), but is enriched by recognising the combinations of players involved in the movement. Given the directional nature of the data within the link plays, the initial interaction matrices were asymmetric, where each A_{ij} was the frequency of player *i* "sending" the ball to and being "received" by player *j* (see points i) and ii) above). This research, however, required an undirected network—that is, any and all relations between players regardless of the directional flow (Scott, 2000). A directed network

would be preferred if we were interested in a player's send/receive ratio. For example, because he is mostly attempting to score, a forward would receive the ball from teammates more than he would send the ball. The undirected network required each matrix to be symmetrised, using $r_{ij} = r_{ji} = A_{ij} + A_{ji}$, ($i, j = 1, \dots, 22$). Frequency distributions could then be calculated for each $[i, j]$ in each of Geelong's 25 matches (22 regular season games and three finals matches). Geelong fielded 34 players throughout the season, so a total of $(34 \times (34-1))/2 = 561$ distributions were computed. In this calculation, the subtraction of 1 removed player *i*'s interaction with himself, and the divisor of 2 halved the distributions to be calculated because $r_{ij} = r_{ji}$. Figure 1 displays the observed interaction, $f(r)$, between Geelong's Jimmy Bartel and Andrew Mackie for all 2011 season matches. This player pair was more likely to interact between one and six times in a match than not at all. The maximum number of interactions measured in the season between any pairing from the team was eight.

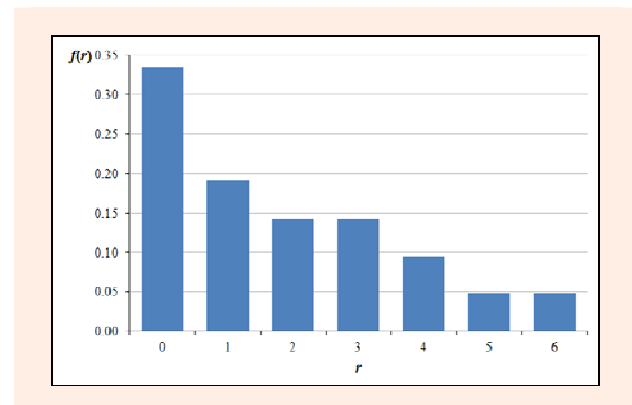


Figure 1. Frequency distribution of [Bartel, Mackie] interactions.

Interaction simulation

If the average rate of discrete events that occur between two players within an AFL match remained constant over its course, the events could be described with a Poisson distribution. However, interaction rates between any $[i, j]$ are stochastic, depending on factors such as the position of the two players, their skill levels and the defensive quality of the opposition. For this reason, the negative binomial distribution (*nbd*) was deemed more appropriate than Poisson. Although the performances of individual players do not give close fits to the *nbd*, the fit improves as more players become involved (Pollard et al, 1977). From the negative binomial distribution, the probability of *r* interactions for each $[i, j]$ is:

$$P(r) = \binom{k+r-1}{k-1} p^k q^r, r = 0, \dots, 8 \quad (1)$$

where $k > 0$, $0 < p < 1$ and $q = 1 - p$.

The parameters, *k* (the threshold number of successes) and *p* (the probability of a success) were estimated so as to minimize the Pearson's chi-squared statistic, χ^2 for each $[i, j]$, by using the observed (*O*) and

expected (E) probabilities derived from Equation (1), or:

$$\min \chi^2 = \sum_{r=0}^8 \frac{(O-E)^2}{E} \tag{2}$$

s.t. $0 < p < 1$ and $k > 0$.

where r is the number of failures (interactions).

Fitting nbd to various sports, Pollard et al (1977) estimated k and p by a method of moments, so:

$$k = m^2/(s^2 - m), p = m/s^2 \tag{3}$$

where m is the sample mean and s^2 is the sample variance. We concluded that Equation (2) was a more adequate fit to the interaction data, providing lower χ^2 values for the majority of Geelong's $[i, j]$. The [Bartel, Mackie] example is displayed in Table 1 where k and p in each $P(r)_1$ were estimated using Equation (2) and in each $P(r)_2$ using Equations (3).

Table 1. Probabilities and χ^2 values for [Bartel, Mackie]

r	$f(r)$	$P(r)_1$	$P(r)_2$
0	.3333	.3333	.2897
1	.1905	.2222	.2675
2	.1429	.1481	.1853
3	.1429	.0988	.1141
4	.0952	.0658	.0659
5	.0476	.0439	.0365
6	.0476	.0293	.0197
7	.0000	.0195	.0104
8	.0000	.0130	.0054
χ^2		.0819	.1178

A Visual Basic module was written to fit the optimized nbd to all combinations of players in the Geelong club and to simulate the players' interactions for any chosen team list in the 22 x 22 team matrix. The initial routine produced a random probability, $u \sim U(0,1)$ for each $[i, j]$ in the match, with r_{ij} determined by the cumulative distribution function:

$$F(r) = P(R \leq r) \tag{4}$$

where R represents the cumulative probability. For example, a randomly generated probability of $u = 0.3$ would produce r [Bartel, Mackie] = 0 as $u < P(R \leq 1) = 0.0000 + 0.3333$ (see Table 1). For each simulation, all $(22 \times (22 - 1))/2 = 231$ elements of the interaction matrix assumed a value for r as determined by u and Equation (4), enabling calculation of player ratings from the simulated matrix.

Player ratings

Measuring a player's net contribution to a match in any team sport is an ambiguous task, in particular for the AFL, because 36 players compete on the field at any single moment. The different positional duties performed by each player add to the complexity: defenders prevent goals; forwards kick/create goals; and midfielders obtain and retain possession of the ball to increase the chance of their team scoring. A network algorithm was introduced for rating purposes to better understand the causality of player i 's performance with respect to that of his teammates. Centrality is one of the most widely studied concepts in network analysis and allows implicit assumptions about the prominence of an individual in a network (Lusher et al., 2010). A specific type, eigenvector

centrality, was trialled as a valid player-rating model, under the assumption that the higher a player's centrality in the Geelong network, the greater his interaction with other players. The eigenvector centrality rating, e , for player i , was measured using:

$$e_i = \frac{1}{\lambda} \sum_j r_{ij} \cdot x_j \tag{5}$$

expressed in matrix form as: $Ax = \lambda x$, where x is the corresponding eigenvector from our interaction matrix, A , and the eigenvalue, λ , was solved using an automated power method. Following n multiplications of A and x , the point at which λ_{n-1} and λ_n converged prompted calculation (Equation (5)) of the ratings for all players within the actual or simulated interaction matrix.

The simulated network and corresponding ratings detailed in this paper provided a pragmatic framework for estimating player i 's utility within a selected side. An important step in this procedure was calculating team a 's network "strength", π , after each match, by:

$$\pi_a = \frac{1}{n} \sum_{i=1}^n e_i, i = 1, \dots, 22 \tag{6}$$

where each e is derived from Equation (5). We compared Geelong's 25 network indices from Equation (6) with each match's final score margin and discovered a linear regression line effectively approximates the margin ($R^2 = 0.5302$) (see Figure 2top)). In practical terms, a team increases its likelihood of winning if more players force themselves to be central in the match network. This is analogous to the finding that soccer teams, skilful enough to retain possession for longer periods than their opposition, have a greater chance of scoring (Hughes and Franks, 2005).

To validate the centrality ratings, an "individual" rating equation, Y_i , was developed, ignoring network methodology and focussing solely on player i 's post-match performance indicator totals—the same four indicators (m) as in the primary interaction data (kick; mark; handball; handball received). The equation was of the form:

$$Y_i = b_o + \sum_{m=1}^4 b_m X_m \tag{7}$$

where X_m is the frequency of performance indicator m for player i , b_m are weights and b_o is the intercept. The weights were optimized to maximize the linear relationship between the mean ratings and final score margin in each Geelong match. Substituting Y for e in Equation (6) produced a comparable measure of team strength for the individual ratings. Figure 2bottom) confirms team strength was not as accomplished at predicting score margin when each player was assessed individually ($R^2=0.2837$), rather than as an agent within a team's network. This finding was analogous to the work of Pollard et al (1977) when fitting the negative binomial distribution to groups of players.

Results

Before investigating player effects within the network, we performed a preliminary examination on our simulator, testing the hypothesis of similar means between the observed and simulated total interactions from Geelong's 22 regular season matches. Opposition effect was ignored for this stage of the research. One hundred interaction

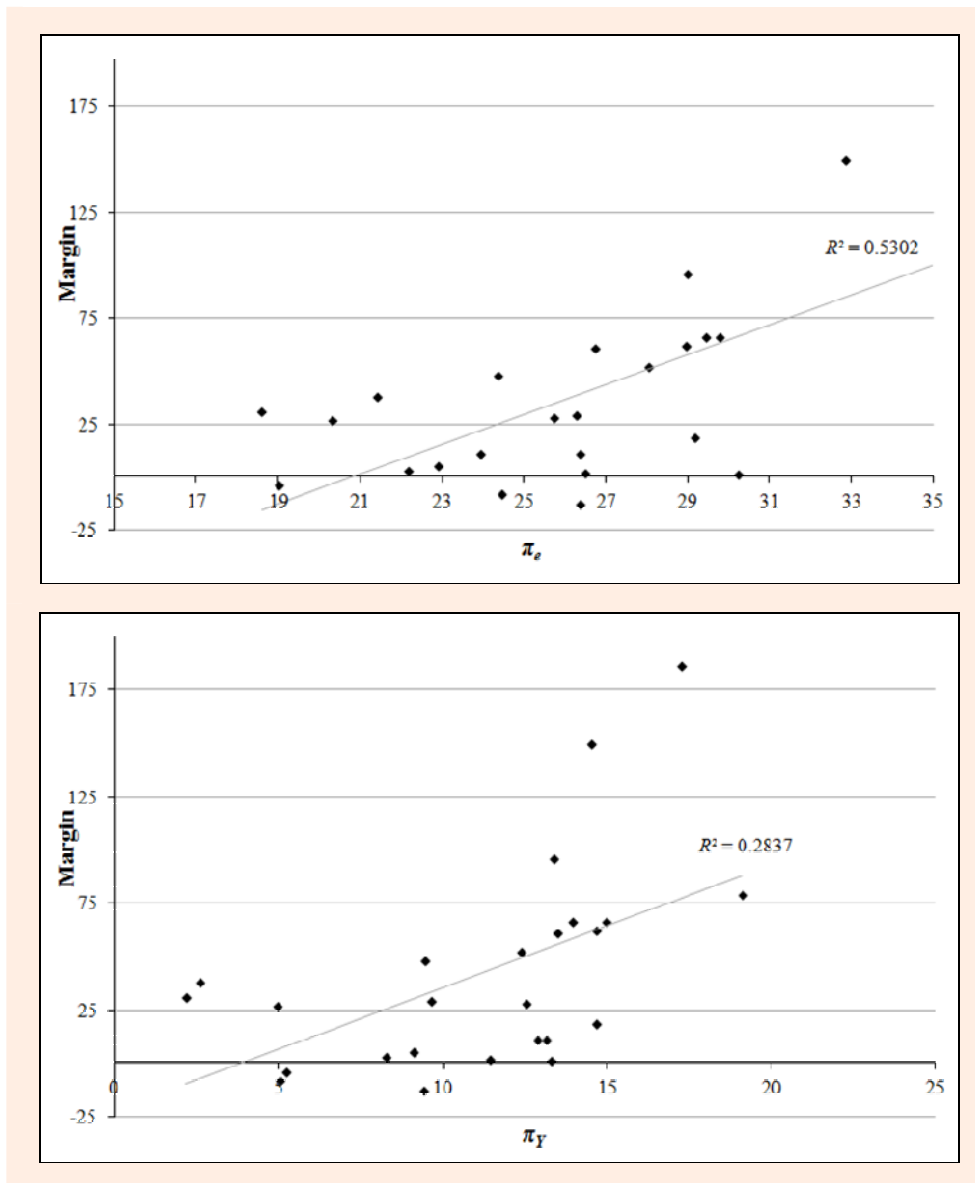


Figure 2. Top: Relationship between Geelong’s mean network rating and final score margin. **Bottom:** Relationship between Geelong’s mean individual rating and final score margin.

simulations were run on each round’s interaction totals, Σr_i , ($i = 1, \dots, 22$), and the mean and standard deviation of each distribution compared with the total observed interactions in each match. Figure 3 reveals a satisfactory fit for the model, with no significant difference between the simulated and observed series means ($p = 0.764$, $\alpha = 0.05$). Moreover, the majority of observed totals fell within 95% confidence intervals associated with each simulated match mean. Match 13 was considered an anomaly in the series—Geelong fielded their weakest side for the season, as acknowledged by the simulator, but managed to achieve almost 600 interactions and to win by 52 points, most likely due to their home-ground dominance. The outlier at Match 18 was Geelong winning by 186 points—the second-highest margin in AFL history—yet the simulator acknowledged the strength of this side, offering the largest simulated interaction mean of all matches ($\Sigma r_i = 654$). The satisfactory totals fit gave us confidence to proceed to analysis of individual player

effects.

A case study was undertaken on Geelong’s 2011 grand final team list, beginning with one thousand network simulations. Using the regression line in Figure 2(top), final score margins were predicted and logged after each simulation. The black curve in Figure 4 represents the normal distribution ($\bar{x}_1 = 47.03$, $\sigma_1 = 14.31$) of predicted margins given Geelong’s actual grand final team. Geelong won the game by 38 points, which reflects the model’s predictive properties. Another one thousand simulations were run on the same side, but we replaced Bartel with a player of lesser skill, Shannon Byrnes. The light grey curve in Figure 4 represents the normal distribution ($\bar{x}_2 = 31.96$, $\sigma_2 = 13.15$) of margins after Byrnes replaced Bartel in the side. Interpretation of this result is important; we concluded that, given his replacement (Byrnes), Bartel’s estimated net contribution to the selected team was $\bar{x}_1 - \bar{x}_2 = 15.07$ points.

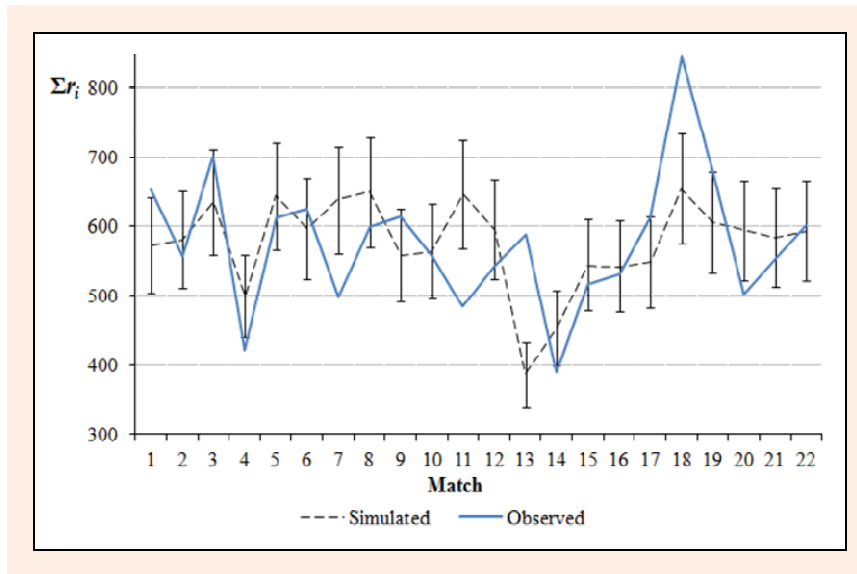


Figure 3. Simulated and observed interaction totals for Geelong matches.

Stressing the selected side was necessary as it could be hypothesised that Byrnes replacing Bartel in a stronger side may have less impact on margin due to the contribution of the other high-calibre players. To conceptualise the importance of selecting the best replacement player, we ran a third iteration in which we replaced Bartel with Darren Milburn—a highly regarded player, but not as skilful as Bartel—and again ran one thousand simulations. The normal distribution ($\bar{x}_3 = 42.85, \sigma_3 = 14.87$) is represented by the dark grey curve in Figure 4, from which we concluded that, given his replacement (Milburn), Bartel’s estimated net contribution to the selected team was $\bar{x}_1 - \bar{x}_3 = 4.18$ points. The difference between the mean of the Byrnes and Milburn distributions ($\bar{x}_2 - \bar{x}_3 = -10.89$) implied a coach would be more inclined to replace Bartel with

Milburn in that side because the negative effect on margin is reduced. It is logical that a player may be selected on grounds other than his net effect on margin; for example, Byrnes’s style of play may be more suited than Milburn’s to the game-day conditions, but this is outside the concerns of this paper.

Discussion

If a prominent player is removed from the network, remaining r_{ij} distributions are not recalculated—that is, we assume teammates do not improve their performance to cover the absence of the excluded player. This phenomenon of players exceeding expectation will be explored further in ongoing research. Furthermore, this paper has not considered the presence of covariance between any r_{ij} . The initial stages of this research

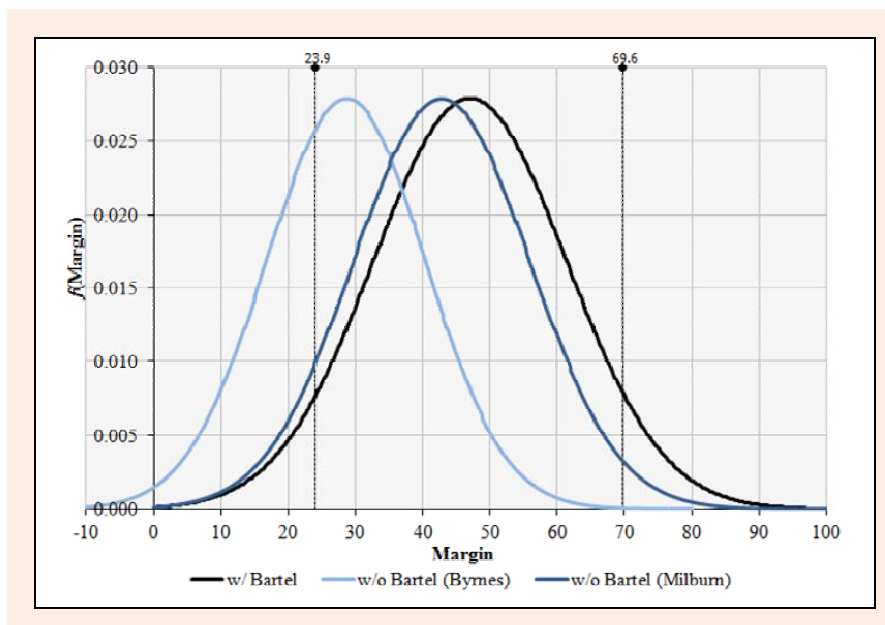


Figure 4. Margin distributions with and without Bartel.

governed that each r_{ij} is independent, even though degrees of interaction covariance between sets of $[i, j]$ are almost certain. The thousands of $[i, j]$ permutations and covariance between each would command a separate research paper. Ongoing research will also focus on improving the predictive power of the networks by weighting the three forms of player interactions in Section 2i with respect to the levels of efficiency, scoring capacity and ground and opponent effects.

Conclusion

Player-based statistical analysis is as important in today's sporting environments as ever before, with coaches continuously searching for the right mix of players to include in a team. In the AFL, the decision to include in a team one player over another can have serious repercussions on the outcome of the game. We developed a model to assist in such selection decisions by simulating different players' interactions with one another and by measuring the effect of such networks on final score margin. Negative binomial distributions were fitted to all pairs of players within a side so that interactions between players could be simulated prior to a match. It was discovered that the strength of the Geelong team's networks was predictive of its final score margin; therefore, it was possible to measure the contribution any player could make to the final margin. Hence, when a team's line-up is revealed, so too is the likelihood of the team winning. From a pre-match betting perspective, it is possible to calculate the odds of the selected team "covering the line". It is anticipated that an in-play model will add further value because coaches and punters can make informed decisions with knowledge of live match scenarios.

References

- Borgatti, S.P. (2005) Centrality and network flow. *Social Networks* **27**, 55-71.
- Duch, J., Waizman, J.S. and Amaral, L.A.N. (2010) Quantifying the performance of individual players in a team activity. *PLoS One* **5**, 1-7.
- Gould, P. & Gatrell, A. (1979/80). A structural analysis of a game: the Liverpool v Manchester United Cup final. *Social Networks*, **2**, 253-273.
- Hughes, M. and Franks, I. (2005) Analysis of passing sequences, shots and goals in soccer. *Journal of Sports Sciences* **23(5)**, 509-514.
- Lee, J., Borgatti, S.P., Molina, J.L. and Guervos, J.J.M. (2005) Who passes to whom: analysis of optimal network structure in soccer matches. In: *Proceedings of the Twenty-fifth International Sunbelt Social Network Conference, Redondo Beach, CA, USA*.
- Lusher, D., Robins, G. and Kremer, P. (2010) The application of social network analysis to team sports. *Measurement in Physical Education and Exercise Science* **14**, 211-224.
- Nevill, A.M., Atkinson, G., Hughes, M.D. and Cooper, S-M. (2002) Statistical methods for analysing discrete and categorical data recorded in performance analysis. *Journal of Sports Sciences*, **20**, 829-844.
- Pollard, R. (1973) Collegiate football scores and the negative binomial distribution. *Journal of the American Statistical Association* **68**, 342, 351-352.
- Pollard, R., Benjamin, B. and Reep, C. (1977) Sport and the negative binomial distribution. In: *Optimal strategies in sport*. Eds: Ladany, S.P. and Machol, R.E. Cleveland: North-Holland Publishing Company. 188-195.
- Reep, C. and Benjamin, B. (1968) Skill and chance in association football. *Journal of the Royal Statistical Society A*, **131**, 581-585.
- Sargent, J. and Bedford, A. (2011) Influence and prestige: rating Australian Football League player performance with network analysis. In: *Proceedings of the 3rd IMA International Conference on Mathematics in Sport. Manchester, date?? UK*. (Compact disc).
- Scott, J. (2000) *Social Network Analysis: A Handbook*. Sage Publications Ltd.

Key points

- A simulated interaction matrix for Australian Rules football players is proposed
- The simulations were carried out by fitting unique negative binomial distributions to each player pairing in a side
- Eigenvector centrality was calculated for each player in a simulated matrix, then for the team
- The team centrality measure adequately predicted the team's winning margin
- A player's net effect on margin could hence be estimated by replacing him in the simulated side with another player

AUTHORS BIOGRAPHY

Jonathan SARGENT

Employment

Stats On (Director)

Degree

PhD Candidate

Research interests

Player ratings, prediction

E-mail:

jonathan.sargent@student.rmit.edu.au

Anthony BEDFORD

Employment

RMIT University (Associate Professor)

Degree

PhD

Research interests

Simulation, match prediction

E-mail: anthony.bedford@rmit.edu.au

✉ Jonathan Sargent

RMIT University, Plenty Road, Bundoora East, VIC, Australia, 3083