Discovering patterns of play in netball with network motifs and association rules

Browne, Peter, Morgan, S, Bahnisch, J and Robertson, Samuel (2019) Discovering patterns of play in netball with network motifs and association rules. International Journal of Computer Science in Sport, 18 (1). pp. 64-79. ISSN 1684-4769

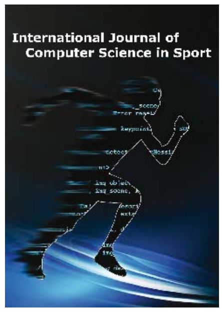
International Journal of Computer Science in Sport

Volume 18, Issue 1, 2019

Journal homepage: http://iacss.org/index.php?id=30

sciendo

DOI: 10.2478/ijcss-2019-0004



Discovering patterns of play in netball with network motifs and association rules

Browne, P.[1,2], Morgan, S.[3], Bahnisch, J.[2], Robertson, S.[1]

[1] Institute for Health and Sport (IHES), Victoria University, Melbourne, Australia

[2] Victorian Institute of Sport, Melbourne, Australia

[3] Department of Computer Science and Information Technology, La Trobe University, Melbourne, Australia

Abstract

In netball, analysis of the movement of players and the ball across different court locations can provide information about trends otherwise hidden. This study aimed to develop a method to discover latent passing patterns in women's netball. Data for both pass location and playing position were collected from centre passes during selected games in the 2016 Trans-Tasman Netball Championship season and 2017 Australian National Netball League. A motif analysis was used to characterise passing-sequence observations. This revealed that the most frequent, sequential passing style from a centre pass was the "ABCD" motif in an alphabetical system, or in a positional system "Centre–Goal Attack–Wing Attack–Goal Shooter" and rarely was the ball passed back to the player it was received from. An association rule mining was used to identify frequent ball movement sequences from a centre pass play. The most confident rule flowed down the right-hand side of the court, however seven of the ten most confident rules demonstrated a preference for ball movement down the left-hand side of the court. These results can offer objective insight into passing sequences, and potentially inform team strategy and tactics. This method can also be generalised to other invasion sports.

KEYWORDS: MATCH ANALYSIS, MACHINE LEARNING, PERFORMANCE ANALYSIS, PASSING PREFERENCES, APRIORI ALGORITHM

Introduction

Pattern recognition is a useful concept to understand the underlying features of a team sport. An objective analysis of pattern can change the way the game is played by offering tactical knowledge for players and coaches (Gyarmati, Kwak, & Rodriguez, 2014). The need for tactical insights has been recognised in sports, such as soccer and field hockey, and the quantity of research devoted to understanding passing patterns has subsequently grown (Gyarmati & Anguera, 2015; López Peña & Sánchez Navarro, 2015; Morgan, 2011; Stöckl & Morgan, 2013). However in netball, the majority of research has focused on the physical playing demands of the sport and how they vary between positions (Davidson & Trewartha, 2008; Steele & Chad, 1991; Sweeting, 2017; Sweeting, Morgan, Cormack, & Aughey, 2014).

An understanding of player interaction is vital in team sports due to its complex nature . A range of analysis techniques to discover tactical insights in team sport have been reported in the literature ranging from social network analysis (Lusher, Robins, & Kremer, 2010; Passos et al., 2011) to visualisations with Voronoi diagrams and measures of entropy to attempt to understand the complexity of performance (Fonseca, Milho, Travassos, & Araújo, 2012). Milo et al. (2002) defined network motifs as interlinking patterns which occur within complex systems at an increased rate compared to other randomised connections. These can aid in identifying player-pairs that have relatively more or less interaction within a team (Passos et al., 2011). The concept of network motifs was expanded upon with '*flow motifs*' by Gyarmati et al. (2014), in which the highly dynamic nature and strength of connections between players in various game states can better be determined.

Combining network-based approaches with spatial information has been identified as the next step forward in performance analysis, and is crucial in considering the interactions between players during a match (Clemente, Martins, Couceiro, Mendes, & Figueiredo, 2014; Fewell, Armbruster, Ingraham, Petersen, & Waters, 2012; Gyarmati & Anguera, 2015). The *Apriori* algorithm is a classical data mining technique used to find co-occurring items in large datasets, and can also be used to detect spatiotemporal trends (Agrawal & Srikant, 1994; Morgan, 2011; Spencer, Morgan, Zeleznikow, & Robertson, 2016). It is an unsupervised machine learning approach, meaning that the algorithm searches for latent patterns in data with no prior information about which rules or patterns should be considered 'right' (Gentleman & Carey, 2008). The basis of the *Apriori* algorithm can be demonstrated as follows: where $I = \{i_1, i_2, i_3 .., i_n\}$ defines a set including items. D is a database of transactions (T), where each T includes an itemset such that $T \subseteq I$. Let A be an itemset where $A \subseteq T$. An association rule is defined as the implication of the form $A \Rightarrow B$, where $A \subseteq I$, and $B \subseteq I$, and $A \cap B = \emptyset$ (Morgan, 2011). The implication symbol ($\Rightarrow$) represents that when a combination of item sets meets the conditions in A, it will also satisfy the conditions of B.

Morgan (2011) used the *Apriori* algorithm to generate rules describing frequently-occurring ball movement events in field hockey. Stöckl and Morgan (2013) extended this idea in women's field hockey with the aim of discovering the most common spatial trends, drawing probabilistic inferences about ball movement patterns and identifying which passages of play resulted in scoring opportunities.

Netball is an invasion-style team sport with a large female participation base in Commonwealth countries (Steele & Chad, 1991). Netball is played on a court measuring 30.5 metres by 15.25 metres, and is divided in equal thirds (International Federation of Netball, 2018). The objective of the game is to score a goal in a ring which is 3.05 metres above the ground. Seven players from each side take to the court at any one time, each assigned to one of seven positions.

Each player role is restricted to specific areas on the court. A centre pass is used to restart play at the beginning of each quarter and after each goal, with teams alternating in taking each centre pass (International Federation of Netball, 2018). Player movements, and the range of legal passes, are constrained by the rules of the game; players are not permitted to run when in possession of the ball, they must also dispose of the ball within three seconds of gaining possession. These constraints mean that passes and passages of play are consecutive and unfold quickly without offering time to organise complex "set plays". This contrasts with other invasion sports, where the need to move the ball quickly is a result of game style and opposition pressure – not a consequence of the rules of the sport.

These constraints allow for a clear understanding of ball movement in netball relative to many other invasion sports. In light of this, previous literature has demonstrated flow motifs and association rules can help understand player interaction spatiotemporal data, respectively (Gyarmati et al., 2014; López Peña & Sánchez Navarro, 2015; Morgan, 2011). In netball, flow motifs can be used to describe player interactions and connections. Association rules focus on ball movement and court position in match play. Whilst these concepts differ, analysis of the two techniques are complementary and provide a more comprehensive representation of passing patterns by describing both the 'who' and the 'where' respectively. This study combined motif analysis and association rules to explore passing patterns in netball by demonstrating; i) a network-based approach highlighting the frequency of passing connections between player roles and, ii) association rule mining to gain an understanding of the spatiotemporal characteristics of those passes.

Methods

Data collection

Data was divided into training and validation datasets. The training dataset included ball movement data from eight games during the 2016 Trans-Tasman Netball League season. A total of 67 players from seven teams were included in the recorded matches. Raw data values were first extracted for each centre pass (n=656). The raw data from each centre pass and subsequent pass were collected for the pass receiver as well as location (n= 4401). The validation dataset collected data from three games from the 2017 Australian National Netball League and consisted of centre pass plays (n=358) which were broken down into pass receiver (n=1781) as well as location (n= 1786). The validation dataset was used to demonstrate whether the results derived from the trained models were generalisable to netball across different seasons and leagues. Prior to data collection, the study was approved by the relevant human research ethics committee.

Observational coding was used to record pass location and receiver using a specially constructed output window in Sportscode (10.3.14, Hudl, Lincoln, Nebraska, United States of America). Two researchers coded the eight games independently and were blinded to the other's result. The centre pass and second phase pass location and receiver were compared from eight games to determine the level of agreement between researchers using the Cohen's Kappa statistic. The statistical software package, SPSS was used for this analysis (IBM Corp. Released 2013. IBM SPSS Statistics for Windows, Version 22.0. Armonk, NY: IBM Corp.). The reliability testing showed high agreement between both researchers, with a Cohen's Kappa measure of agreement reported as 0.891 across 643 valid centre pass plays, with a standard error of 0.013.

Each centre pass play was broken down into phases, where each pass formed the beginning of a new phase. For example, the first pass is the centre pass, the second pass is the second phase, the third pass the third phase, and so on.

The pass location and identity of the receiver for each pass was recorded until the possession chain was interrupted, either by i) a shot, ii) a turnover, iii) a penalty was called or, iv) the ball went out of bounds. This was regardless of whether possession was maintained or lost after the event. These conditions were incorporated to confine the data to controlled, deliberate actions which best demonstrated set or deliberate game play. To code location, the netball court was divided into 20 court segments (seven in the attacking and defensive thirds respectively, and six in the centre third) (Figure 1). These segments were based on segments used by a participating team in the Trans-Tasman Netball Competition.

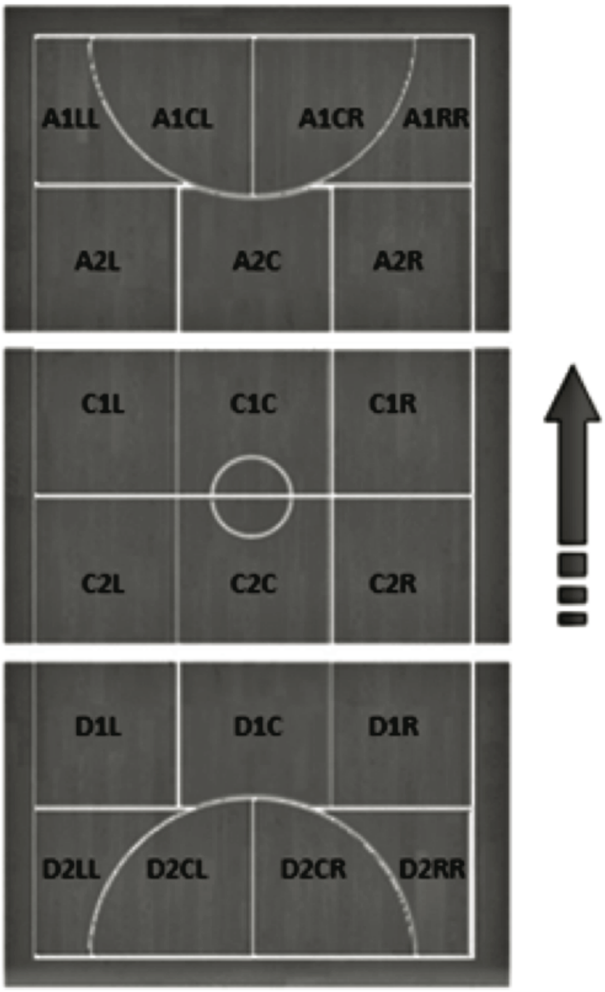


Figure 1. Netball court with segment breakdown. A netball court divided into 20 segments with unique coding labels. Arrow indicates direction of attack.

Experimental Design

The following sections outline the two different types of analysis and the four analyses performed. Firstly, flow motifs were explored looking at player passing sequences occurring from a centre pass from a general perspective (see Figure 2). Secondly, specific playing position (i.e., Centre, Wing Attack) are incorporated into the flow motifs. Thirdly, association rules are used to explore the first two passes in a centre pass play based on location. Finally, again using association rules the entire centre pass play is analysed based on the court segments outlined in Figure 1.

Flow Motifs

Flow motifs form when players frequently link with one another in uninterrupted passing sequences. The analysis considered all possible 3-pass motifs (n = 5) outlined in Figure 2. This analysis labelled passes by an alphabetical labelling system, in which players were assigned a letter depending on when they received the ball within the passing sequence (e.g. A, B, C). The alphabetical lettering system allows competition trends to be observed without the specificity of individual playing positions. As each sequence begins with a centre pass, the first letter is always 'A' and represents the player in the Centre position. Following this, an alternate labelling system was utilised using specific playing positions such as Centre (C), Wing Attack (WA), Wing Defence (WD), Goal Attack (GA), Goal Shooter (GS), Goal Defence (GD) and Goal Keeper (GK).

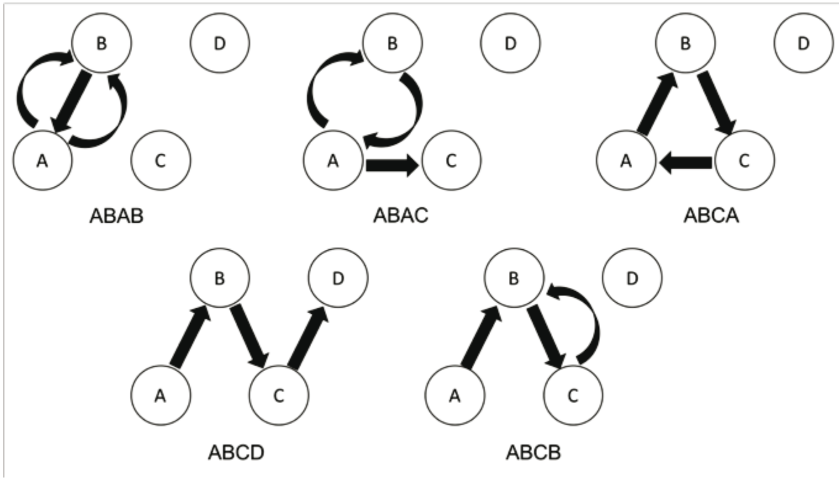


Figure 2. Five core flow motifs. Arrows represent the passing direction; each core motif always begins with the Player A.

For each match, the number of times a unique motif occurred was counted and z-scores were calculated based on the variation of the absolute count from each game. The mean was calculated as the average number of times a motif occurred per game, and the standard deviation also calculated. As centre passes alternate between teams in a game, approximately the same number of centre passes are taken by each team. Gyarmati et al. (2014) and López Peña and Sánchez Navarro (2015) demonstrated the value of z-scores, also known as a Standard Score, in the analysis of networks to compare the differences in the occurrences of individual motifs. This analysis method was incorporated in this study in two styles as follows:

Analysis 1: Passing chains consisting of three passes (i.e., without penalties or turnovers) were analysed. Five core motifs were explored, see Figure 2.

Analysis 2: Specific player positions were incorporated within the flow motifs.

Association Rules

The *Apriori* algorithm was utilised to explore frequent passing court locations. The algorithm was run in the software program R (Version 3.1.2) using the software package *'Arules'*.

The algorithm begins by scanning the database to count the support of each item, i.e., court location. This generates a full set of non-empty frequent l-item set, L_1 (Large frequent item set 1), this was then used to find L_2 (Large frequent item set 2), and L_3, until all the frequent candidates are found with L_k (see Figure 3). A large item set is a count of how many passes went to each court location with a specified pass in the passing chain.

These large item sets were then used to generate the ten strongest rules, based on measures of confidence. Rules which met the minimum support of 0.015, the equivalent of 10 instances, and the metric of confidence with a minimum level of 0.2 (20%), were included (Morgan, 2011). The training and validation dataset rules were compared with by the average confidence difference between rule sets as outlined in Dudek (2010).

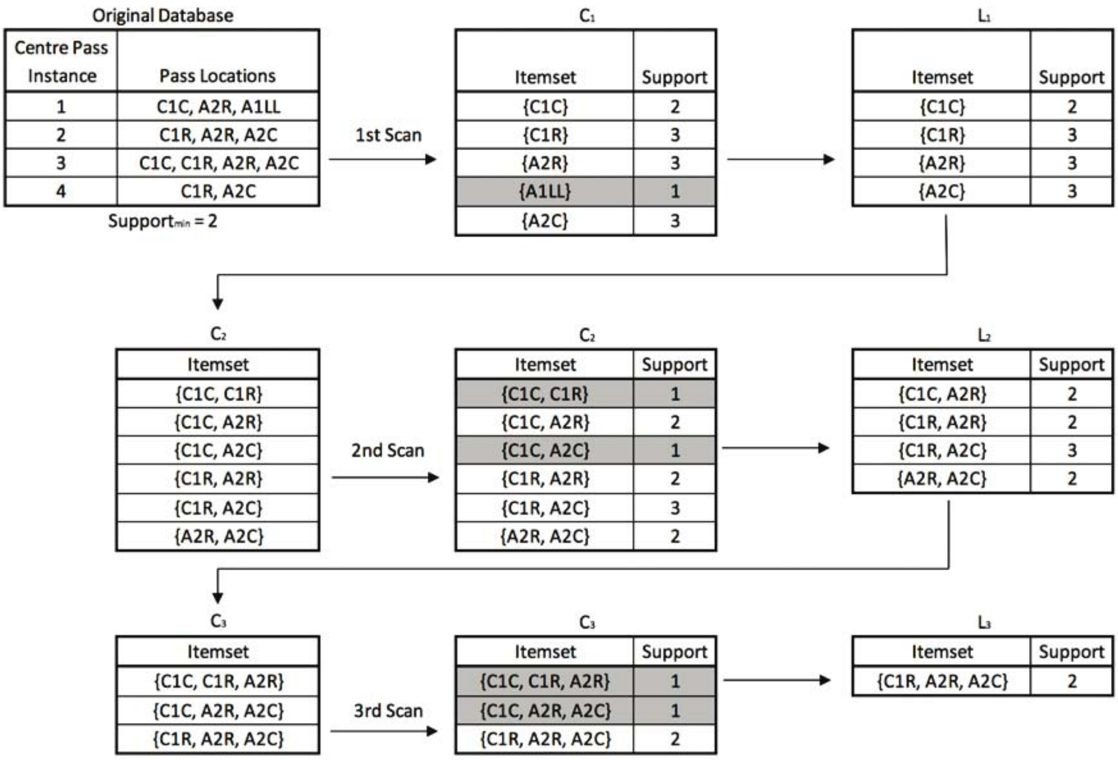


Figure 3. Tabular example of the Apriori algorithm process. Processing example of the Apriori algorithm moving from the original dataset down to the most frequent rule. Where Ck is representative of a candidate itemset of size k and Lk represents the frequent itemsets of size k. Greyed rows represent those which do not meet the minimum support and therefore do not progress through the algorithm.

$$support(A => B) = P(A \cup B) \quad (1)$$

$$confidence(A => B) = \frac{support(A \cup B)}{support(A)} = P(B|A) \quad (2)$$

The *Apriori* algorithm was run with two different criteria, initially limited to the centre pass and second phase, but then expanded to include all uninterrupted phases. Rules were ranked from highest to lowest based on their levels of confidence (equation 2). Thus, if the rule {2P - A2R, 3P A1CR} => {CP - C1R} displays a confidence of 0.1/0.2 = 0.5 in the dataset than the rule is true half of the time; 50% of centre pass plays which in the second phase and third phase were located in grid locations A2R and A1RR respectively originated from a centre pass to the C1R grid. This analysis method was incorporated in this study in two styles as follows:

Analysis 3: Two attributes were initially analysed, the centre pass and second phase pass.

Analysis 4: Following this, 14 attributes were included in the analysis, from the centre pass through to the fourteenth phase.

Results

Results from Analysis 1 can be seen in Figure 4 and demonstrate that the motif of ABCD was the most frequent motif in the training dataset. The mean score of motif occurrences was 110.2 and the standard deviation was 65.3. The *z*-score for each score motif was: ABCD: 1.54, ABCB: -0.03, ABCA: -0.08, ABAC: -0.17, ABAB: -1.26.

Analysis 2 was also conducted with the validation dataset. The ten most common motifs from the training dataset and the corresponding results from the validation dataset are shown in Figure 6.

Results from Analysis 2 are displayed in Figure 5, these demonstrate the motif style of ABCD was still the most prevalent. The position specific analysis demonstrated this as: C – GA – WA – GS. The mean of the number of times a motif occurred per match was 18.0 with a standard deviation of 15.10. The motif C – GA – WA – GS had a *z*-score of 2.12; the least preferred motif, C – WD – WA – GA had a motif score of -0.79.

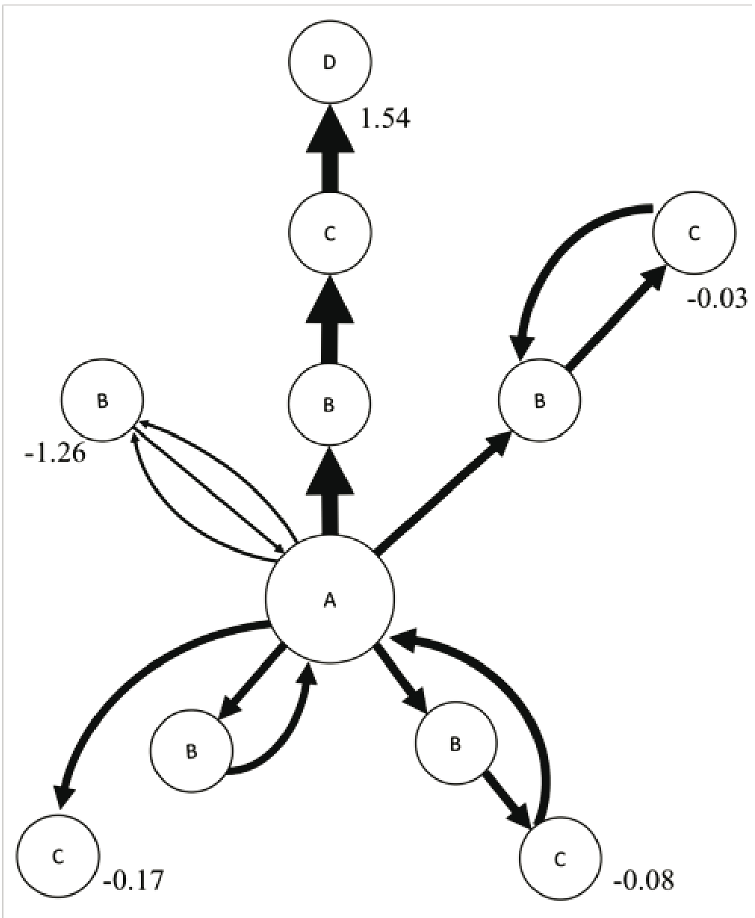


Figure 4. The first three passes of a centre pass play with the core motifs using the lettering system. The value of the *z*-score is listed and displayed by the weighting of the line connecting the players, i.e., the thicker the line the greater the *z*-score.

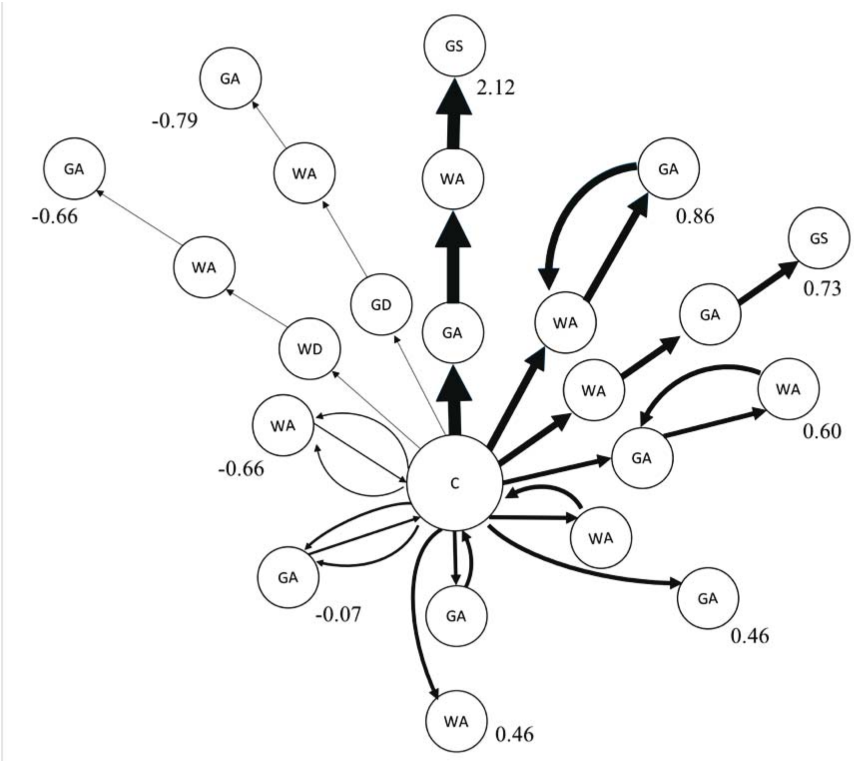


Figure 5. Ten most frequent passing motifs using playing position system generated from the first three passes of a centre pass. The value of the *z*-score is listed and displayed by the weighting of the line connecting the players, i.e., the thicker the line the greater the *z*-score. The largest *z*-score begins at the top and rotating clockwise decreasing in size.

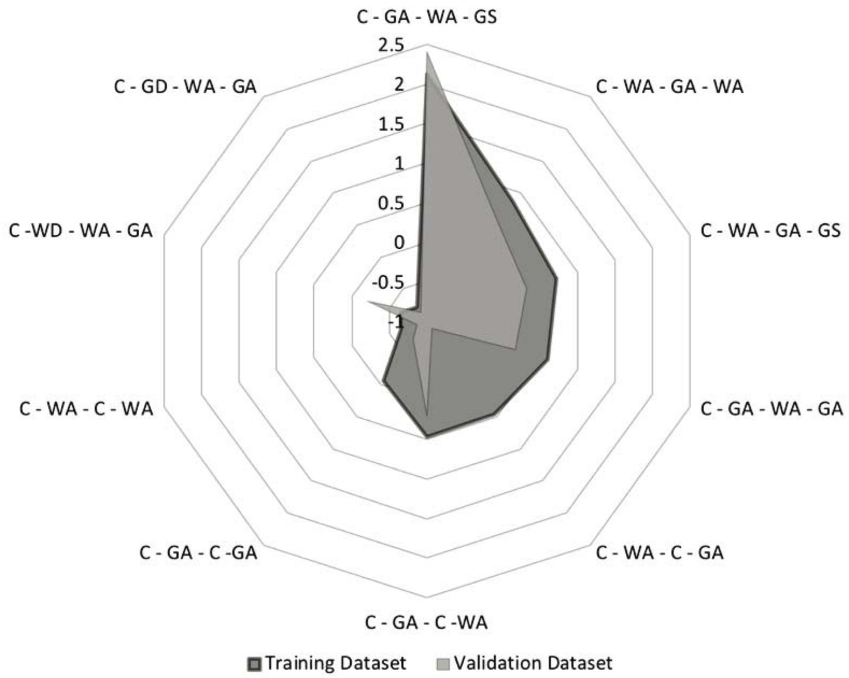


Figure 6. Comparison of z-scores for training and validation dataset. Radar chart comparison of *z*-scores for ten most frequent motifs using the lettering system in training dataset and the equivalent motifs in validation dataset.

Results from Analysis 3 are shown in Table 1 and Figure 7. For each rule, a count is displayed for ‘A’ and ‘B’. A total of nine rules met the criteria outlined above.

Table 1. List of association rules generated using only centre and second phase passes. Minimum support set at 0.015 (equating to 10 instances) and minimum confidence at 0.2 (20%).

Rule	**A**			**=>**	**B**		**Confidence**
	Pass	Location	Count	Pass	Location	Count	
1.	2P	A2R	79	CP	C1R	43	0.54
2.	2P	A2L	81	CP	C1L	44	0.54
3.	2P	A1LL	48	CP	C1L	26	0.54
4.	2P	A2R	79	CP	C1C	35	0.44
5.	2P	A1RR	42	CP	C1C	18	0.43
6.	2P	A2C	93	CP	C1C	38	0.41
7.	2P	A1RR	42	CP	C1R	17	0.40
8.	2P	C1L	68	CP	C1L	26	0.38
9.	2P	A2L	81	CP	C1C	28	0.35

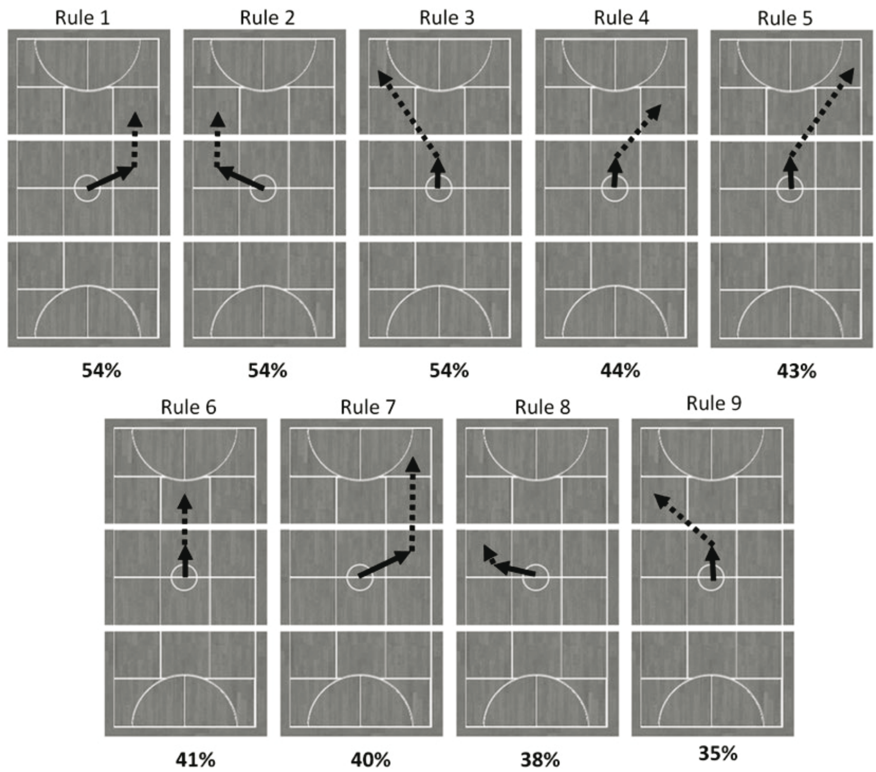


Figure 7. Association rules generated using the centre pass and second phase pass only. In relation to the association rule equation, A => B, the solid arrow represents ‘B’ and the dotted arrow represents ‘A’ of the equation. Percentages shown are level of confidence. Minimum support was 0.015 equating to 10 instances and minimum confidence was set at 0.2 (20%).

The ten most confident rules which met the support and confidence criteria are displayed in Figure 8 and Table 2. The most confident rule flows down the right-hand side of the court. However, the majority (seven out of ten) of the association rules demonstrate a preference for ball movement down the left-hand side of the court.

Analysis 4 was also conducted with the validation dataset, this consisted of the 2017 Australian National Netball League data. Four rules from the training dataset were unique from the validation dataset. The confidence difference was measured as 0.40. The confidence of the rules is displayed in Table 2.

Table 2. Ten strongest association rules by confidence, with training and validation set comparison. Ten strongest association rules based on confidence from the 2016 training dataset with the corresponding confidence from the 2017 validation dataset where the rules met the minimum levels of support and confidence.

	A				=>	B					
Rule	Pass	Location	Pass	Location	Count	Pass	Location	Count	**2016 training dataset confidence**	**2017 validation dataset confidence**	**Change in confidence**
1.	2P	A2R	3P	A1RR	21	CP	C1R	15	0.71	0.70	0.01
2.	CP	C1C	4P	A1LL	16	5P	A1CL	11	0.69	0.46	0.23
3.	5P	A1LL			37	6P	A1CL	25	0.68	0.57	0.11
4.	2P	A2L	3P	A1CL	15	CP	C1L	10	0.67	0.86	-0.19
5.	CP	C1R	3P	A1RR	23	2P	A2R	15	0.65	0.70	-0.05
6.	3P	A2C	5P	A1CL	19	4P	A2C	12	0.63	not found	-
7.	2P	A1LL	3P	A1CL	16	CP	C1L	10	0.63	not found	-
8.	2P	A2L	4P	A1CL	17	CP	C1L	10	0.59	0.72	-0.13
9.	2P	A2L	3P	A1LL	21	CP	C1L	12	0.57	not found	-
10.	CP	C1L	5P	A1CL	21	4P	A2C	12	0.57	not found	-

Discussion

Coaches often refer to match statistics to reinforce their beliefs about what took place (Hughes & Bartlett, 2002). This study paired player connections with court location through flow motifs and association rules respectively to provide a hybrid approach to understanding the tactical nature of netball (Croft, Willcox, & Lamb, 2017) . Each technique has been used in isolation in soccer and hockey (Gyarmati & Anguera, 2015; Morgan, 2011), but the ability to explore the results from the independent methodologies in combination increases the practicality and insight for performance analysts and coaches. A network approach applied in basketball has been able to demonstrate what is occurring on court (Fewell et al., 2012). The practical application of this model extends from a description of offensive ball movement to the development of defensive strategies which counter preferred ball movement patterns.

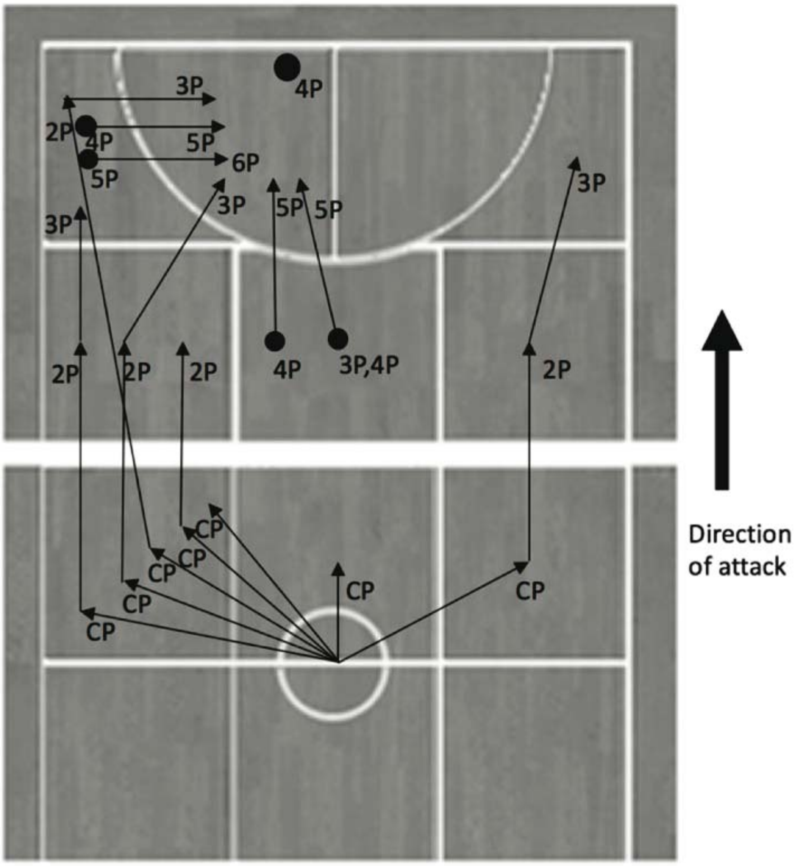


Figure 8. Ten strongest association rules based on confidence. Ten strongest association rules based on confidence displayed on attacking two thirds of court, using the centre pass up to the 14th phase where applicable in continuous centre pass plays. The rules displayed do not differentiate between the left and right side of the association rule equation, it is purely a display of the ball movement. Rule 1 and 5 display the same pass and court locations, with different sequences in their respective association rule outputs.

An external validation method comparing results from the 2016 training and 2017 validation datasets demonstrated that the study observations are relatively generalisable using this approach. This was despite the Trans-Tasman league split and forming of two independent leagues with the introduction of three new teams and the large number of player movements between teams and countries. The influx of new teams, coaches and players was also met with a change in external constraints such as reduced travel, but also a reduced playing list (12 to 10 players). Over nine seasons between 2009 to 2016 it was suggested that team profiles remained relatively similar (Bruce, Brooks, & Woods, 2018). It was due to these major changes in the competition structure that external validation was analysed between the 2016 training dataset and 2017 validation dataset. The similarity of the validity analysis demonstrates this methodology of this study does not simply describe what occurred during the 2016 Trans-Tasman season, but is beginning to demonstrate the underlying features of netball. Therefore, this study establishes the potential benefits arising from the combined use of network motifs and association rules when applied within elite sport. It also providing another form of analysis to aid in overcoming coaches historic use of statistics to confirm their beliefs (Hughes & Bartlett, 2002). It offers an insight across the league and not specific to any individual team, they provide an overview about typical ball movements which occur across both the Australian and New Zealand styles of play, and not a direct measure of an individual teams preferred ball movement.

Flow motifs were generated using three passes, the centre pass plus the subsequent two passes. A limited number of passes better represents preconceived passes, as opposed to reacting to the game. The motif ABCD was the most prevalent motif using the alphabetical lettering system. Motifs ABAC, ABCB and ABCB recorded negative z-scores, however with z-scores close to zero. Motif ABAB was by far the least utilised passing sequence across the games analysed.

The addition of player specific positions deepens the information gained by analysing player involvement, including connectivity and frequency. This gives tangible information about player involvement, such as which playing positions are connecting and how frequently, removing some of the static nature given by the alphabetic lettering system through demonstrating position specific tendencies (Rocha & Blondel, 2013). For example, the strong statistical tendency of the C – GA – WA – GS and ABCD motifs demonstrate the players' preference to attempt to move the ball in a sequential pattern without a player receiving the ball twice. Moreover, the range of z-scores between the second and tenth strongest motifs is 1.65. This relatively small range across a large number of motifs signifies the close relationship of these passing sequences and emphasises the distinction of these to the C – GA - WA – GS passing preference.

Coaches may gain practical insight through the player specific labelling system (Gyarmati et al., 2014). For instance, knowing the most preferential passing sequence was C – GA – WA – GS indicates to coaches preferred passing patterns and therefore can influence both offensive and defensive decision making. For example, if the above passing sequence was the oppositions most preferred pattern it may be due to the defence channelling the centre pass to the Goal Attack, rather than the Wing Attack. This strategy may occur more frequently because the WA is often a specialist mid-court player, and the GA may represent a less-potent scoring threat when in possession of the ball in the mid-court (compared to the shooting circle). There is an incentive for the defence to stop the WA from receiving the ball, which reduces the likelihood of that motif occurring. This strategic inference can be used by coaches offensively to understand whether the team is getting the ball to the most threatening player at the right time. For instance, were a team to have their WA receiving the centre pass they may have a better offensive structure and would be able to measure this. This indicates the practical nature for coaches in utilising the flow motif model to understand game dynamics. It also demonstrates the need to acknowledge the influence of the opposition on which passing structures can be applied in the game setting.

Interestingly, in contrast to the findings in this study, Gyarmati et al. (2014) identified that the most successful team in soccer used the motif ABCD less frequently than of any other team, and alternatively focused their game plan around a back and forth passing strategy (e.g. ABCB). Major differences exist between the two sports, but it raises the question if a netball team were able to break out of the 'norm' and use a passing structure built around back and forth passing, could they too become a more successful team? Success with such a strategy may simply be due to playing a different style to the 'norm' and the different nature of the two sports must be considered.

The *Apriori* algorithm was used to discover recurring spatiotemporal ball movement patterns. Similar to Morgan (2011), rules were generated under a minimum support constraint where at least 10 instances of each itemset were found in a sample of 656 centre pass plays. Whilst this number accounts for less than two percent of all centre pass plays, they are relatively frequent amongst seemingly random passages of play. Association rules are powerful in that they allow large-scale searches for these seemingly random frequent patterns of play and are robust to the presence of non-relevant distracting features. Association rules uncover nuanced patterns that are not easily perceivable by the human observer.

The *Apriori* algorithm was run with two sets of constraints, using first two passes and then all complete passes, the algorithm outputs showed differences in pass location preferences. For example, initially looking at the centre pass and second phase pass, five of nine rules from the algorithm described the centre pass going directly forward. Yet, when every phase was included in the algorithm for the second set of analysis the rules indicate the prevalence of centre pass direction was not straight forward but favoured the left-hand side. As these ten rules progress, all except two conclude on the left-hand side of the court. Seven frequent passing patterns conclude in the shooting circle, and surprisingly all seven are on the left-hand side. Other sports, such as field hockey, have been shown to have a right-hand side bias, however scoring plays more often progressing down the left-hand side of the field (Stöckl & Morgan, 2013). Moreover, this study indicates only that the pathways used to get to these locations may favour the left-hand side of the court, not necessarily that this is the preferred shooting position. Moreover, it does not necessarily provide the whole picture about what is occurring or why. Additionally, a limitation of the *Apriori* algorithm is that does not always find consequential passing chains. The rules generated were not all for a sequential passing sequence.

Seven playing styles have been outlined using forty performance variables in netball (Croft et al., 2017), however through the addition of a motif and rule-based approach a greater understand of team preferences may be understood within each playing style. This analysis provides an insight into the way attacking teams progress the ball forward, but also offers a perspective of tactics of the defensive team. If a team could pass straight down the centre of the court, move the ball to the top of the shooting circle and feed straight into their shooter, it is likely that they would do so. Defensive strategy is to push passes to a less threating area or to win the ball, forcing the attack to make responses such as swinging the ball and moving deep into the pockets where they have limited space, as defensive style can influence the number of successful passes in netball (Bruce, Farrow, Raynor, & May, 2009). It raises a question - are the majority of possession chains going to the left-hand side of the court because the attacking team want to go there, or because the defensive team are forcing them there? These observations reinforce the notion that the role of the defensive team cannot be ignored and where possible should be incorporated within the analysis. This area requires further research across multiple seasons which could offer an increased understanding of teams both offensive and defensive tactics.

The ability to draw inferences from both flow motifs and association rules can provide practical insight into the way plays unfold. This study demonstrated how flow motifs described a preference for a direct passing sequence, whilst the association rules demonstrated that ball movement is typically forward and left. For instance, the restriction on ball possession time increases the difficultly for a player to pass forward and then run down the court, evade the defence and receive the ball back. If the ball is getting passed directly forward and not crosscourt, it must mean that players are cutting into open space. Flow motifs help inform this. For example, the second and fourth strongest motifs using playing positions show the use of the same player twice; C – WA – GA – WA and C – GA – WA – GA (both ABCB motif styles). The ball returns to the same player on the third phase. The association rules demonstrated this third phase is typically occurs in the attacking third, where the ability to pass forward is restricted by court boundaries and court access for certain playing positions. Restricting the ability and ease of a direct passing style. However, the role of the opposition must be acknowledged in the generation of these frequent patterns. The defensive team's strategy and personal will alter and influence which patterns are frequent and successful. The combination of flow motifs and association rules begins to demonstrate *why* the play unfolds, and provides coaches with a subjective view of what has occurred. These patterns may not be apparent to the human observer (Morgan, 2011).

A larger dataset collected over multiple seasons would confirm the generalisability of the study's findings and increase their specificity. For example, more data could provide insight into the progression of tactics throughout the length of the competition. Incorporating defensive positions may allow insight into how passing patterns are influenced by defensive presence. In addition, individual teams and players could be analysed separately to determine differences in patterns of play across each team and individual (Beckkers & Dabadghao, 2017). This may offer insight into the stability and progression of both successful and losing teams. Furthermore, this helps coaches and list managers to better understand the roles and strengths of players and improve recruiting decisions when selecting players to align with the team's game style (Beckkers & Dabadghao, 2017). The analysis could be furthered with smaller court segment sizes or spatiotemporal tracking data to provide more specificity. The size of player location segments has a pronounced impact on the complexity and specificity of rule outcomes (Gudmundsson & Wolle, 2014). However, it must be noted if sector sizes were too small it may hide potential generalised insights. Additional outcome measures such as goals and turnovers could also be included.

Conclusion

This study demonstrates a new hybrid methodology for understanding recurring patterns of play from a centre pass in elite netball. The combination of flow motifs and association rules provides an unsupervised and scalable method to discover frequently occurring passing patterns in netball. This approach has the potential to offer coaches actionable intelligence into ball movement patterns in netball and other invasion-style sports. The practical application of this model extends from better understanding offensive structure, to investigating and informing defensive strategy.

Disclosure of Interest

The author(s) declared no potential conflict of interest with respect to the research, authorship, and/or publication of this article.

References

Agrawal, R., & Srikant, R. (1994). *Fast algorithms for mining association rules.* Paper presented at the Proc. 20th int. conf. very large data bases, VLDB.

Beckkers, J., & Dabadghao, S. (2017). *Flow Motifs in Soccer: What can passing behaviour tell us?* Paper presented at the In Proceedings of the 11th MIT sloan sports analytics conference, Boston, Massachusetts.

Bruce, L., Brooks, E. R., & Woods, C. T. (2018). Team and seasonal performance indicator evolution in the ANZ Championship netball league. *Journal of sports sciences, 36:24*, 2771-7777. doi:10.1080/02640414.2018.1473099

Bruce, L., Farrow, D., Raynor, A., & May, E. (2009). Notation analysis of skill expertise differences in netball. *International Journal of Performance Analysis in Sport, 9*(2), 245-254.

Clemente, F. M., Martins, F. M. L., Couceiro, M. S., Mendes, R. S., & Figueiredo, A. J. (2014). A network approach to characterize the teammates' interactions on football: A single match analysis. *Cuadernos de Psicología del Deporte, 14*(3), 141-148.

Croft, H., Willcox, B., & Lamb, P. (2017). Using performance data to identify styles of play in netball: an alternative to performance indicators. *International Journal of Performance Analysis in Sport, 17*(6), 1034-1043.

Davidson, A., & Trewartha, G. (2008). Understanding the physiological demands of netball: A time-motion investigation. *International Journal of Performance Analysis in Sport, 8*(3), 1-17.

Dudek, D. (2010). *Measures for Comparing Association Rule Sets.* Paper presented at the International Conference on Artificial Intelligence and Soft Computing, Berlin, Heidelberg.

Fewell, J. H., Armbruster, D., Ingraham, J., Petersen, A., & Waters, J. S. (2012). Basketball teams as strategic networks. *PloS one, 7*(11), e47445.

Fonseca, S., Milho, J., Travassos, B., & Araújo, D. (2012). Spatial dynamics of team sports exposed by Voronoi diagrams. *Human Movement Science, 31*(6), 1652-1659.

Gentleman, R., & Carey, V. (2008). Unsupervised machine learning *Bioconductor Case Studies* (pp. 137-157): Springer.

Gudmundsson, J., & Wolle, T. (2014). Football analysis using spatio-temporal tools. *Computers, Environment and Urban Systems, 47*, 16-27.

Gyarmati, L., & Anguera, X. (2015). Automatic Extraction of the Passing Strategies of Soccer Teams. *arXiv preprint arXiv:1508.02171.*

Gyarmati, L., Kwak, H., & Rodriguez, P. (2014). *Searching for a unique style in soccer.* Paper presented at the 2014 KDD Workshop on Large-Scale Sports Analytics, New York City.

Hughes, M. D., & Bartlett, R. M. (2002). The use of performance indicators in performance analysis. *Journal of sports sciences, 20*(10), 739-754.

International Federation of Netball. (2018). *Rules of Netball.* Retrieved from http://netball.org/wp-content/uploads/2017/12/INF-Rules-of-Netball-2018-Edition-text.pdf

López Peña, J., & Sánchez Navarro, R. (2015). Who can replace Xavi? A passing motif analysis of football players. *arXiv preprint arXiv:1506.07768*.

Lusher, D., Robins, G., & Kremer, P. (2010). The Application of Social Network Analysis to Team Sports. *Measurement in Physical Education and Exercise Science, 14*(4), 211-224. doi:10.1080/1091367x.2010.495559

Milo, R., Shen-Orr, S., Itzkovitz, S., Kashtan, N., Chklovskii, D., & Alon, U. (2002). Network motifs: simple building blocks of complex networks. *Science, 298*(5594), 824-827. doi:10.1126/science.298.5594.824

Morgan, S. (2011). *Detecting Spatial Trends in Hockey Using Frequent Item Sets.* Paper presented at the Proceedings of the 8th International Symposium on Computer Science in Sport.

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

Rocha, L. E., & Blondel, V. D. (2013). Flow motifs reveal limitations of the static framework to represent human interactions. *Physical Review E, 87*(4).

Spencer, B., Morgan, S., Zeleznikow, J., & Robertson, S. (2016). *Clustering team profiles in the Australian Football League using performance indicators.* Paper presented at the Proceedings of the 13th Australasian Conference on Mathematics and Computers in Sport, Melbourne, 11-13 July, 2016.

Steele, J. R., & Chad, K. E. (1991). An analysis of the movement patterns of netball players during match play: implications for designing training programs. *Journal of human movement studies, 20*, 249-278.

Stöckl, M., & Morgan, S. (2013). Visualization and analysis of spatial characteristics of attacks in field hockey. *International Journal of Performance Analysis in Sport, 13*(1), 160-178.

Sweeting, A. (2017). *Discovering the movement sequences of elite and junior elite netball athletes.* (Doctorate of Philosophy), Victoria University.

Sweeting, A., Morgan, S., Cormack, S., & Aughey, R. (2014). *A movement sequencing analysis of team-sport athlete match activity profile.* Paper presented at the Proceedings of the 10th Australasian Conference on Mathematics and Computers in Sport, Darwin, Australia.