

Passing Network Motifs: A Convincing Quantification of Team and Player Passing Style

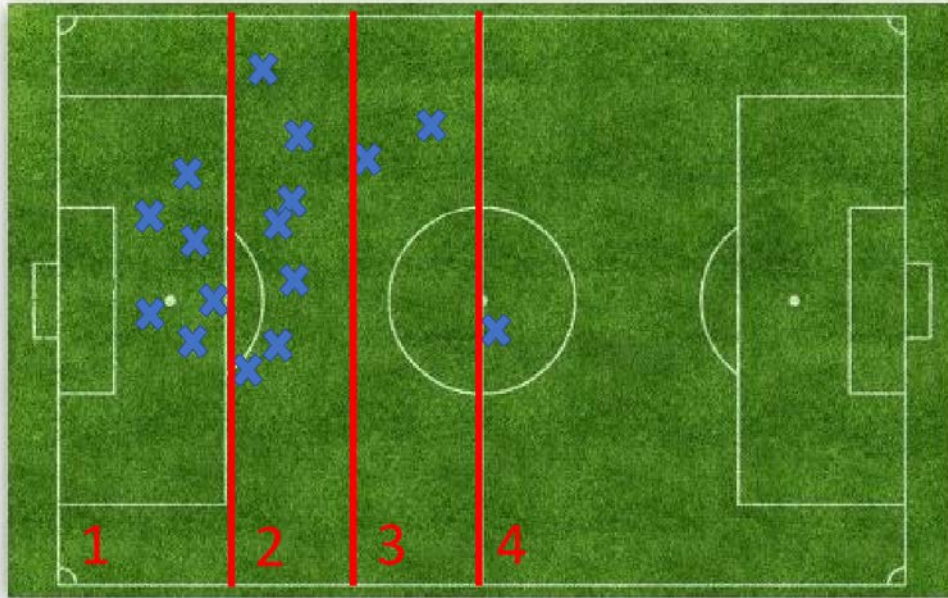
David A. Perdomo Meza

#OptaProForum

Opening Thoughts

- ‘Operationalising’ or ‘Vectorising’ raw football information in high dimensions.
 - Example:

BLOCKS

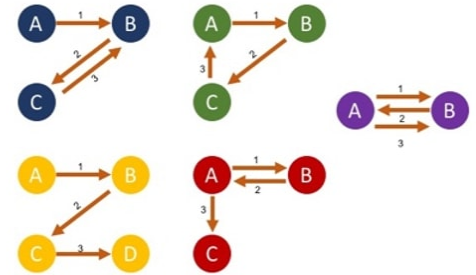
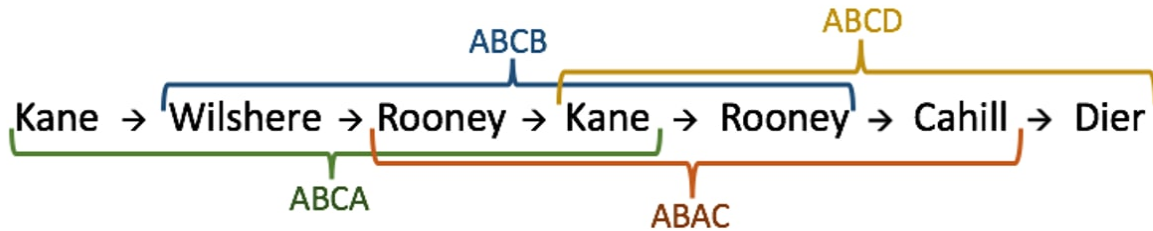


Opening Thoughts

- ‘Operationalising’ or ‘Vectorising’ raw football information in high dimensions.
- As vectors, players can be seen as ‘points’ in a ‘measurable space’. This is true even in high dimensions.
- **Question:** What do players’ positions in this space tell us about them? Do they tell us anything at all?
 - **Example:** Banks use this sort of conceptualisation to decide who to give and who not to give a loan.
- **Question:** How can we use this ‘Big Data’ approach to help and advise a club’s recruitment policy?

Passing Network Motifs

- **Original Idea:** 'Searching for a Unique Style in Soccer' by Pedro Rodriguez, Laszlo Gyarmati and Haewoon Kwak

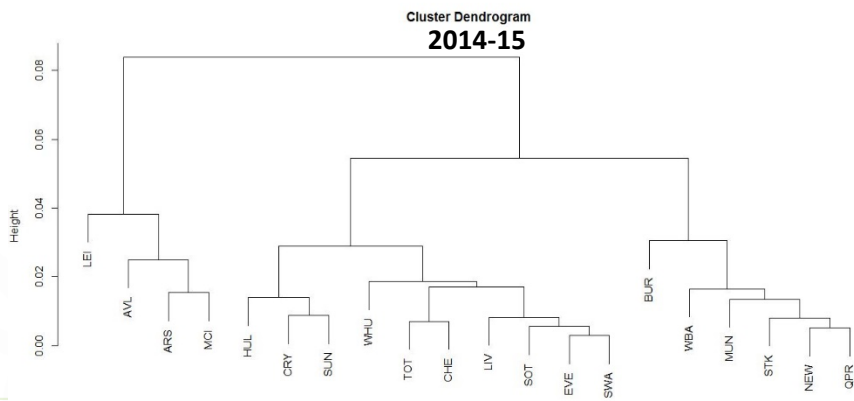
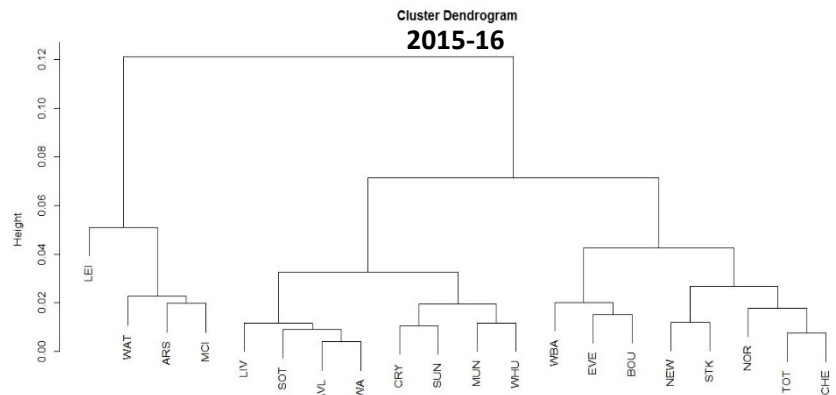


There are 5 possible 3 pass long motifs, identified by their acronyms **ABAB, ABAC, ABCA, ABCB** and **ABCD**.

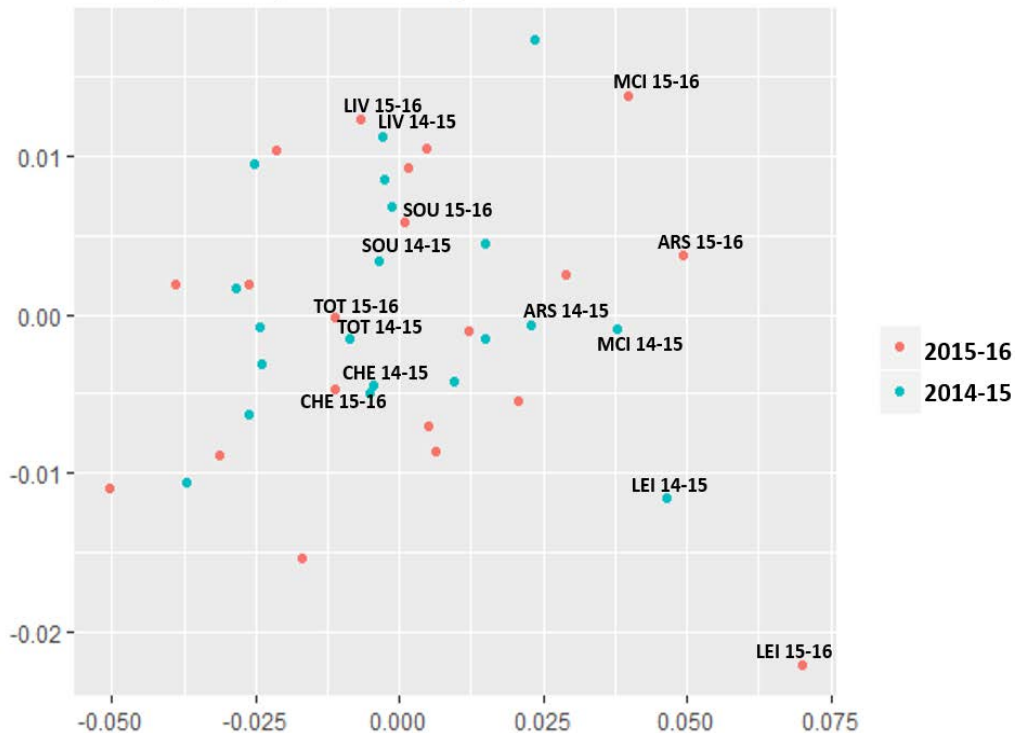
Passing Network Motifs: Team Level

- Rodriguez, Gyarmati and Kwak proposed that by looking at the *relative frequencies* with which each team uses the different motif categories, a representation of a team's "style" would emerge.
- They analysed teams from La Liga and (unsurprisingly) found that FC Barcelona have a passing motif frequency that substantially sets them apart from other teams.
- In particular, they use 'ABAB' and 'ABAC' much more frequently than other teams and 'ABCD' much less frequently.

Passing Network Motifs: Team Level



Principal Component Analysis



Passing Network Motifs: Player Level

1. **'Relative Frequencies' Idea:** What percentage of a player's participation in 3-pass sequences corresponds to each motif category.
2. **'Divergence' Idea:** If a team uses 'ABAB' 40% of the time and a player uses it 43%, he gets a +0.03 score for that motif.
3. **'Team's Proportion' Idea:** What proportion of a team's 'ABAC' instances did the player participate in.
4. **'15 Roles Idea:** A player can participate in an 'ABCA' as the 'A', the 'B' or the 'C'. Looking at all 5 motif categories there are 15 possibilities (weighted by minutes).
5. **'15 Roles Relative Frequencies' Idea**

Player Representation: How can we know if it's useful?

1. Repeatability and Stability:

A player's positions in the space should be 'close' to each other for consecutive seasons if it captures a (somewhat) stable property such as "skill", "style" or "potential".

2. Certain Output Metrics Cluster Together:

Example: If all the players who had a high number of 'key passes' are relatively close together as opposed to spread out randomly, then it provides evidence of this information being represented by their position in the 'space'.

Question: If it **is** useful, **how** can we use it?

Player Representation: Some Results

Highest Scored Players (Premier League)			
Rank	Name	Team	Key Passes
1	Alexis Sanchez	Arsenal (2)	58
2	Mesut Ozil	Arsenal (2)	125
3	Juan Mata	Man Utd (5)	48
4	Dimitri Payet	West Ham (7)	107
5	David Silva	Man City (4)	54
6	Jesus Navas	Man City (4)	47
7	Moussa Sissoko	Newcastle (18)	49
8	Daniel Drinkwater	Leicester (1)	34
9	Aaron Ramsey	Arsenal (2)	33
10	Willian	Chelsea (10)	75
11	Kevin De Bruyne	Man City (4)	72
12	Eden Hazard	Chelsea (10)	62
13	Ross Barkley	Everton (11)	47
14	Raheem Sterling	Man City (4)	33
15	Christian Eriksen	Tottenham (3)	102

Highest Scored Players (Bundesliga)			
Rank	Name	Team	Key Passes
1	Raffael	B. Monchengladbach (4)	45
2	Kingsley Coman	Bayern Munich (1)	29
3	Thomas Muller	Bayern Munich (1)	55
4	Douglas Costa	Bayern Munich (1)	40
5	Henrikh Mkhitaryan	B. Dortmund (2)	67
6	M. Risse	FC Koln (9)	61
7	Leonardo Bittencourt	FC Koln (9)	33
8	Marco Reus	B. Dortmund (2)	42
9	Daniel Caligiuri	VfL Wolfsburg (8)	28
10	Thorgen Hazard	B. Monchengladbach (4)	32
11	Ivo Ilicic	Hamburger SV (10)	23
12	Filip Kostic	VfB Stuttgart (17)	70
13	E.M. Chuopo-Moting	Schalke 04 (5)	28
14	Gonzalo Castro	B. Dortmund (2)	21
15	Lukas Rupp	VfB Stuttgart (17)	32

Player Representation: Some Results

Highest Scored Players (xA)			
Rank	Name	Team	xA (mins)
1	Alexis Sanchez	Arsenal (2)	13.1
2	Raheem Sterling	Man City (4)	4.92
3	Anthony Martial	Man Utd (5)	7.88
4	Jesus Navas	Man City (4)	7.64
5	Moussa Sissoko	Newcastle (18)	4.52
6	Dusan Tadic	Southampton (6)	10.79
7	Riyad Mahrez	Leicester (1)	9.89
8	Dimitri Payet	West Ham (7)	14.16
9	Kevin De Bruyne	Man City (4)	12.18
10	Mesut Ozil	Arsenal (2)	18.15
11	Juan Mata	Man Utd (5)	6.58
12	Memphis Depay	Man Utd (5)	4.46
13	David Silva	Man City (4)	10.56
14	Nathan Redmond	Norwich (19)	5.72
15	Marko Arnautovic	Stoke (9)	6.23

Highest Scored Players using 'Goals'			
Rank	Name	Team	Goals
1	Olivier Giroud	Arsenal (2)	16
2	Gerard De Boer	Evert	2
3	Diego Costa	Ch	12
4	Son Heung-Min	Tottenham (3)	24
5	Romelu Lukaku	Leicester (1)	18
6	Wilfried Zaha	Southampton (15)	2
7	Jefferson Moreira	Leicester (12)	0
8	Dusan Tadic	Southampton (6)	8
9	Michy Batista	Leicester (1)	8
10	Kevin De Bruyne	Man City (4)	25
11	Modeste	Swansea (12)	1
12	Kelechi Ighalo	Man City (4)	8
13	Jeremain Lens	Sunderland (17)	3
14	Wayne Routledge	Swansea (12)	2
15	Son Heung-Min	Tottenham (3)	4

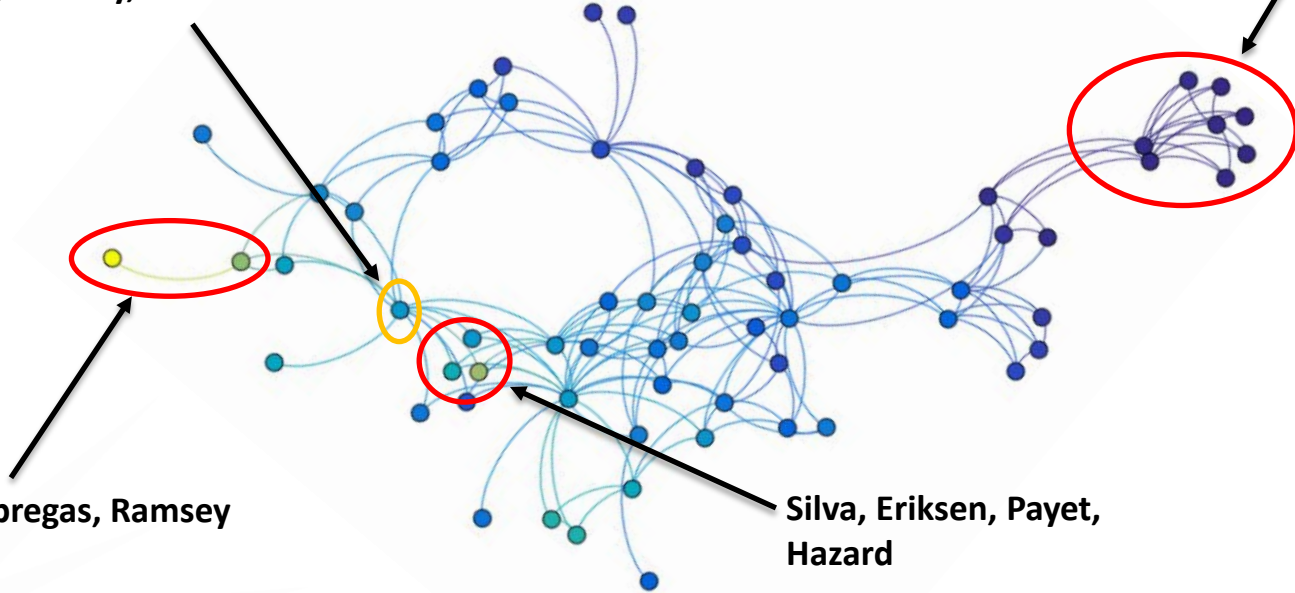
Exploring the Results: Topological Data Analysis

Mata, M'Villa, Yaya Toure, Noble,
Gueye, Barkley, Bellerin

Goalkeepers

Ozil, Fabregas, Ramsey

Silva, Eriksen, Payet,
Hazard



Closing Thoughts

- It seems that the results of analytics need to be validated by aspects that can be *directly and observably verified on the pitch* if they are to be trusted by the people inside football.
 - It is as if banks' analysts provided Rules of Thumb for loans: Women, 3 children and married – YES.
- Traditional Scouting and Analytical Scouting **can** generate cooperative workflows that benefit everybody and greatly increase the efficiency and accuracy of the scouting process.
- The passing motifs methodology **is not** the absolute and ultimate answer regarding player creativity.
- However, its “success” provides valuable insight into the operationalisation of raw data.

Thank you for listening

Questions

@dperdomomeza1

<http://theopenfootball.blogspot.co.uk/>

#OptaProForum