

# サッカーのパス網分析における中心性測定の新指標

New indicator for centrality measurement in passing network analysis of football

神島 正稔      折原 良平      清 雄一      田原 康之      大須賀 昭彦  
Masatoshi KANBATA      Ryohei ORIHARA      Yuichi SEI      Yasuyuki TAHARA      Akihiko OHSUGA

電気通信大学大学院情報理工学研究科

Graduate School of Informatics and Engineering, The University of Electro-Communications

Data analytics is used in various field including business, science and sports. The evaluation of players and teams affects tactics, training and scouting in football teams. Players and teams are often evaluated by data such as shots and goals in game results. However it is not enough to fully understand potential of the players and teams. This paper describes a new analysis method using football passing data. In order to evaluate performance of players and teams, we employ graph mining. There is an indicator called centrality that evaluates individual contribution within an organization and it is used to evaluate the players and the teams. In calculation of the centrality for a given player pair, we consider not only the shortest path of passing but also longer ones. As a result, we found our method to be more consistent with game results than conventional methods.

## 1. はじめに

近年、スポーツにおける選手の走行距離や得点数など、試合において取得されるデータの分析が行われている。チームや選手の評価、戦術や選手の育成、スカウティングなどのチーム運営において、このような分析が影響を与えている。

しかし、サッカーにおいては、選手を評価するデータ指標がゴールとアシストである場合が多く、試合においてそれらを記録した選手に高評価が偏る傾向がある。このことから、ゴールとアシストを記録していない選手も含めた全選手は、チームにおいてどのように貢献したのか、それをどのように評価すべきなのかを検証する。また、より適切な選手評価からチームを評価することで、チーム成績の要因の論理的な分析に貢献することができると考えられる。

本研究では、サッカーにおいてゴールやアシストに繋がると考えられる、パスに着目して評価を行う。具体的にはグラフ理論を用いてシュートで終了したパスの流れを活用して、選手の試合における貢献を分析をする。また、グラフ理論における中心性の概念を活用することで、選手及びチームを適切に評価するための手法を提案する。

## 2. 関連研究

スポーツに関連する研究はこれまで数多く行われており、グラフ理論を適用する研究が、サッカーのパスに対してされてきた。また、このような研究を行う際、試合におけるパスの出し手と受け手、そのパス数を記録した passing distribution が公開データとして用意されている。

サッカーにおいて、相手ゴールにボールを移動させることができる手段の一つであるパスは、非常に重要な項目であると考えられている。このことから、passing distribution を適用してグラフネットワークを作成する分析がされてきた。グラフネットワークにおいて試合における選手評価、およびチーム評価を行うための手法として中心性が測定される。中心性測定のために用いられる指標として、betweenness centrality と flow

連絡先: 神島正稔, 電気通信大学大学院情報理工学研究科, 東京都調布市調布ヶ丘 1-5-1,  
Email:kanbata.masatoshi@ohsuga.lab.uec.ac.jp

centrality がある。ただし、passing distribution は選手間のパス数を記録しているだけであり、3人以上の選手間のパスの流れは再現できない。

### 2.1 Betweenness centrality

中心性とはグラフネットワークにおいて、任意のノードが他のノードに与える影響度を示す指標である [Tsugawa 14]。サッカーの試合において、より多くパスに関わった選手が貢献を示すという考え方がある。これをグラフ理論に適用すると、passing distribution により作成したグラフから、選手ノードの中心性を測定することができる。

$\sigma_{st}$  をノード  $s$  からノード  $t$  までの最短経路の総数、 $\sigma_{st}(v)$  をノード  $s$  からノード  $t$  までの最短経路の内ノード  $v$  を通る総数として、ノードの中心性を表す betweenness centrality は以下のように定義される。ただし、 $V = \{1, 2, \dots, 11\}$  としこれを選手ノードの集合とする [Barthelemy 04]。

$$betweenness(v) = \sum_{\substack{s,t \in V \\ s \neq v \neq t}} \frac{\sigma_{st}(v)}{\sigma_{st}} \quad (1)$$

### 2.2 Flow centrality

betweenness centrality を変更し、シュートまでに至るパスの過程のみを最短経路として扱う [Duch 10]。このことから、直接得点の機会を得た場合を考慮して選手の評価を行う。新たにシュートノードを作成しノード  $t$  とする flow centrality を以下のように定義する。ただし、 $U = \{12, 13\}$  をシュートノードの集合とし、12 をゴール外、13 をゴール内へのシュートとする。

$$flow(v) = \sum_{\substack{s \in V \\ t \in U \\ s \neq v}} \frac{\sigma_{st}(v)}{\sigma_{st}} \quad (2)$$

[Duch 10] は、試合における選手の影響力はゴールに直結したパスの流れが最も大きくなると考えており、betweenness centrality より flow centrality のほうが、より選手やチーム評価を適切に行うことができると主張している。

### 3. 提案手法

flow centrality における最短経路の探索において、ノード  $t$  をシュートノードとしたためノード間のパスの経路が限定的となる。そのため、算出される経路の総数が少なくなり、測定において得点に関わった選手が高く評価されない場合がある。

そこで本研究では、シュートに至るパスに関わった選手をより考慮するため、経路の探索において最短経路+1した場合の経路を求めることで、算出される経路の総数を増やす。 $\sigma'_{st}$  をノード  $s$  からノード  $t$  までの最短経路+1の総数、 $\sigma'_{st}(v)$  をノード  $s$  からノード  $t$  までの最短経路+1の内ノード  $v$  を通る総数として、 $flow^+$  の式を以下のように定義する。

$$flow^+(v) = \sum_{\substack{s \in V \\ t \in U \\ s \neq v}} \frac{\sigma'_{st}(v)}{\sigma'_{st}} \quad (3)$$

### 4. 評価

#### 4.1 実験

提案手法が試合において、選手及びチームに対して適切な評価を行っているか検証する。UEFA EURO2008 の passing distribution<sup>\*1</sup> を用いて、任意のチーム  $A$  における各試合に先発出場した選手の *betweenness*, *flow*,  $flow^+$  を、 $P_{betweenness}^A$ ,  $P_{flow}^A$ ,  $P_{flow^+}^A$  とし降順にする。それらの選手を  $n = \{1, 2, \dots, 11\}$  とした試合におけるチームの評価を以下のように定義する。

$$\bar{P}_{betweenness}^A(n) = \frac{1}{n} \sum_{i=1}^n P_{betweenness}^A(i) \\ (P_{betweenness}^A(1) > P_{betweenness}^A(2) > \dots) \quad (4)$$

$$\bar{P}_{flow}^A(n) = \frac{1}{n} \sum_{i=1}^n P_{flow}^A(i) \\ (P_{flow}^A(1) > P_{flow}^A(2) > \dots) \quad (5)$$

$$\bar{P}_{flow^+}^A(n) = \frac{1}{n} \sum_{i=1}^n P_{flow^+}^A(i) \\ (P_{flow^+}^A(1) > P_{flow^+}^A(2) > \dots) \quad (6)$$

本研究では、参考論文に基づいて  $n=3, 11$  として、チーム  $A$  とチーム  $B$  の試合における  $\bar{P}^A(n)$  と  $\bar{P}^B(n)$  の大小が勝敗と直結しているかどうかを検証し、その精度を算出する [Duch 10].

#### 4.2 結果

データの予選リーグに対して前述した実験を行った結果を以下に示す。表 1 は  $\bar{P}^A(11)$  と  $\bar{P}^B(11)$ ,  $\bar{P}^A(3)$  と  $\bar{P}^B(3)$  の大小と勝敗の整合精度、表 2 は各試合の  $\bar{P}_{flow}^A(3)$  と  $\bar{P}_{flow}^B(3)$ ,  $\bar{P}_{flow^+}^A(3)$  と  $\bar{P}_{flow^+}^B(3)$  を示す。ただし、表 2 における国名は FIFA コードで表記し、試合結果と値が整合しない場合は F とする [FIFA 10].

表 1: 試合結果との整合精度

	<i>betweenness</i>	<i>flow</i>	$flow^+$
$n = 11$	0.43	0.33	0.43
$n = 3$	0.48	0.43	0.71

表 2: 予選リーグの各試合の  $\bar{P}_{flow}^A(3)$ ,  $\bar{P}_{flow^+}^A(3)$

Gr.A	SUI-CZE	POR-TUR	CZE-POR	SUI-TUR	SUI-POR	TUR-CZE
<i>flow</i>	4.37-7.83	4.66-3	4.86-4.89	3.19-4.67	5-5.31,F	3.9-4.05,F
$flow^+$	6.64-8.71	7.32-4.55	7.4-7.66	6.88-7.09	8.17-7.94	6.6-6.49
Gr.B	AUT-CRO	GER-POL	CRO-GER	POL-CRO	AUT-GER	
<i>flow</i>	4.83-5.12	5.67-3.87	3.23-4.22,F	4.07-3.51,F	4.72-2.71,F	
$flow^+$	7.23-7.7	8.86-6.98	6.38-6.87,F	6.45-6.87	6.92-5.24,F	
Gr.C	NED-ITA	NED-FRA	NED-ROU	FRA-ITA		
<i>flow</i>	3.27-3.33,F	4.58-3.26	3.68-5,F	5.33-3.57,F		
$flow^+$	6.04-5.19	7.24-6.24	6.69-7.18,F	8.33-5.69,F		
Gr.D	ESP-RUS	GRE-SWE	SWE-ESP	GRE-RUS	GRE-ESP	RUS-SWE
<i>flow</i>	3.57-4.61,F	3.72-3.12,F	4.33-4.5	4.97-4.43,F	4.58-4.73	2.86-3.2,F
$flow^+$	6.79-7.13,F	6.07-6.25	5.91-7.26	6.89-7.53	6.26-8.29	5.95-6.34,F

### 5. 考察

表 1 から  $flow^+$  が、*betweenness*, *flow* よりも高い値を示していることがわかる。 $flow^+$  では最短経路+1の経路を算出して flow centrality を行っているため、通常の flow centrality と比較して、シュートまでのパスの過程がより多く算出されている。このことから、シュートに至るパスに関わった選手が従来手法より評価され、精度を上回ったと考えられる。尚、本研究では経路を最短経路+1として検証を行ったが、今後は経路の適切な長さを推測し検証を行う必要があると考えられる。

また、表 2 からチームの評価が試合結果に整合しなかった場合があった。この理由を調査した結果、*flow* と  $flow^+$  の相対的な選手評価の傾向は変わらなかったが、特定の選手の評価が極端に高くなる傾向にあり、それによってチーム評価が高くなることが判明した。これは、シュートを打った特定の選手にのみパスを集めることにより発生する事象であると考えられる。このような事象は、試合においてシュートの本数が極端に少なくなる場合や、特定の選手にボールを集める戦術を遂行する場合に発生すると推測され、今後はパスの総数を考慮して調査を行う必要がある。

### 6. おわりに

本研究では、サッカーのパスからグラフ理論の中心性の概念を活用して選手の試合における貢献を分析した。分析の結果、提案手法が従来手法を上回り、より適切な選手評価及びチーム評価を行うことができた。

### 参考文献

- [Tsugawa 14] S. Tsugawa and H. ohsaki :A Study on the Robustness of Centrality Measures in Network Analysis(2014), 第 57 回自動制御連合講演会論文集 pp.1072-1077
- [Barthelemy 04] M. Barthelemy:Betweenness centrality in large complex networks(2004), The European Physical Journal B pp.163168
- [Duch 10] Jordi Duch, Joshua S. Waitzman, Luis A. Nunes Amaral: Quantifying the Performance of Individual Players in a Team Activity(2010), PLoS ONE 5(6): e10937. <https://doi.org/10.1371/journal.pone.0010937>
- [FIFA 10] FIFA's 208 Member Associations(2010):[http://www.fifa.com/mm/document/fifafacts/organisation/52/00/10/fs-120\\_01a\\_ma.pdf](http://www.fifa.com/mm/document/fifafacts/organisation/52/00/10/fs-120_01a_ma.pdf)

\*1 試合数: 31 試合, チーム数: 14 チーム, 全パス数: 21624, 1 試合の平均パス数: 698