

特 集 「スポーツ競技とAI」

サッカーにおける集団プレイの検出とその指標化に向けて

Towards Detection of Collective Plays and Their Quantification in Soccer

新井 優太

Yuta Arai

東京大学空間情報科学研究センター

Center for Spacial Information Science, the University of Tokyo.

yuu.arai@csis.u-tokyo.ac.jp

家富 洋

Hiroshi Iyetomi

新潟大学理学部

Faculty of Science, Niigata University.

hiyetomi@sc.niigata-u.ac.jp

清水 千弘

Chihiro Shimizu

日本大学スポーツ科学部, 東京大学空間情報科学研究センター

College of Sports Sciences, Nihon University./ Center for Spacial Information Science, the University of Tokyo.

shimizu.chihiro@nihon-u.ac.jp, <https://www.shmzlab.jp/>

Keywords: soccer, tracking data, collective play, triangle, network.

1. はじめに

サッカーと野球を対比してみると、野球は基本的に分業制であり、個人の能力が前面に出るのに対して、サッカーでは個人の能力はもちろんのこと、集団としての機動性や協調性が合わせて問われる。その例として、サッカーでのボールの動きに着目する。もしサッカーが2～3名のプレーヤで行われたとすると、個人プレイの特徴が現れ、ボールの動きは直線的であろう。逆に100名規模で行われいたら個人の特性は全く失われ、ボールはブラウン運動する粒子のごとく単にマクロの拡散運動を行うに過ぎない。実際のボールの運動は直線的でもなく、ブラウン運動的でもなく、それらの中間で複雑な振舞いを示すことが知られている [Kijima 14, Luzuriaga 10]。このように、サッカーは個人（ミクロ）と集団行動（マクロ）の両側面が同時に重要なスポーツである。特に、サッカーの戦術を分析するにあたって、集団行動に着眼する必要性は、「オフ・ザ・ボールでの動き」として以前から指摘されている [加藤 99]。また、サッカーは学問的にも複雑系として興味深い研究対象である。

サッカーを科学的に分析するためには、ゲームの客観化をまず最初に行わなければならない。サッカーは野球に比べて非常にアナログ的なスポーツである。野球のスコアブックの様式は非常に整備され、スコアブックから試合に関するさまざまな情報を後からでも容易に引き出すことができる。残念ながら、それに対応するものはサッカーには存在しない。サッカーの試合のテキスト速報はあるが、読んでも試合の様子を思い描くことはほとんど

不可能である。サッカーのラジオ中継の難しさの要因がそこにある^{*1}。

1980年代からコンピュータを活用した記述分析法（プレイの分類・記号化）の研究が開始された [HUGHES 88]。我が国においても先駆的研究 [河合 92, 河合 94, 城戸 02, 大橋 87, 大橋 88] がある。さらに最近のIT技術の進展によってサッカーに関するデータ、特にフィールド上の選手やボールの位置が時々刻々集められつつある。そのようなトラッキングデータの分析がサッカーの質的改善にとって重要であることは、衆目の一一致するところである。特にデータ利用ではドイツが先行し、2014ワールドカップ（ブラジル大会）での優勝の形でその成果が現れた。我が国のプロサッカーリーグ（Jリーグ）においても2015年から各試合のトラッキングデータの収集が開始され、同年8月にはデータを活用したJリーグのテレビ番組「必見！ データがサッカー中継を変える」がNHKによって実験的に放送された。

サッカーの分析に用いられるデータとしては、個人のミクロな情報（パス成功率、ボール奪取数など）やチームのマクロな情報（ボールポゼッション、オフサイド、反則やセットプレイの数など）が古くから集計されてきた。最近ではトラッキングデータと称して走行距離、ス

*1 著者の1人(HI)は、1993年10月28日にカタールの首都ドーハで行われた「ドーハの悲劇」と呼ばれている日本代表対イラク代表のサッカーの試合を当時ラジオ中継で聞いていた。試合の様子が全然わからず、フラストレーションがたまっていったことを思い出す。終了間際、アナウンサーがしばらく「ア」と声を上げたままの状態となり、とっさに何が起ったのかわからなかった。

プリント回数、ボール保持率、プレイエリアなどのデータも利用可能となっている。しかし、それらのデータは集団行動を直接計測するものではない。集団行動を測る方法も研究され始め、すでに数多くの論文（総説 [Rein 16] を参照）が出版されているが、研究のフェーズとしてはまだまだ初期の段階である。もちろん機械学習の応用も始まっている^{*2}。

現代サッカー戦術の代表的潮流は、パスの出し手とサポート役の二人の受け手がつくる三角形を基軸としたパスサッカーである。そのため、集団運動を選手ではなく三角形を基本構成要素として捉える分析方法の開発が必須である。本稿では、攻撃側選手、特にゲームメイクを行うMF（Midfielder）がつくる三角形の周長（選手間の距離感を計測）および形状（正三角形度を計測）に対して統計的な分析を行うとともに、攻撃側選手の三角形の有効性（防御側選手との攻防の状況を定量化）の観点からチームの運動性を評価・指標化する試みについて紹介する^{*3}。

2. トラッキングデータ

本研究では、データスタジアム株式会社より提供を受けた、2015年度J1柏レイソルの開幕10試合分のトラッキングデータを使用する。このトラッキングデータは1/25秒単位でフレーム番号、選手（審判も含む）、ボールに関する情報が記録された高頻度データとなっている。まずフレーム番号は、フレームを特定するためのユニークな番号を表している。次に選手情報については、フィールドにいるプレーヤー11人×2チーム、控え3人×2チーム、審判1人の計29人分の情報が記録されている。審判を含むプレーヤー情報とボール情報について、トラッキングデータに記録されている項目を表1、表2にまとめる。

表1および表2を見るとわかるように、選手の位置やポゼッション、ボールステータスについてはトラッキングデータから取得できる。しかしながら、選手のポジション情報などはトラッキングデータに記録されていない。そのためJリーグの公式サイト（<https://www.jleague.jp/>）より、ポジションなどのトラッキングデータに記録されていない情報を抽出して分析に利用した。

また、試合途中で選手の入替えが生じると分析が煩雑になってしまう。10試合のすべてで前半に選手の入替えが生じていなかったため、今回の解析では前半のデータのみを用いた。

*2 本特集における井上寛康氏の寄稿（pp. 517-524）を参照。

*3 本稿についての予備的解析結果は、日本行動計量学会第44回大会（CAI-5、札幌学院大学、2016年9月1日）および日本物理学会第72回年次大会（20aK-PS-11、大阪大学、2017年3月20日）で報告された。

3. チーム配位の可視化

3.1 攻撃側選手間のネットワーク

攻撃側の選手がつくる三角形の構築にはドロネー（Delaunay）による三角形分割アルゴリズム（三角形の最小の内角を最大化する）を利用する。そのような三角形ネットワークの例を図1に示す。ここで、ドロネー三角形を計算する際には、ボール保持チームのゴールキーパーを含めた選手11人だけをネットワークの頂点として用いるのではなく、相手ゴールの位置も含めて三角形分割を行った。パスコースが最終的に相手ゴールへつながっているかどうかも、パスにゴールへ向けての戦術的位置が込められているかどうかを評価するうえで重要なと考えたからである。ボールが三角形ネットワーク上を動く場合は、パスサッカーの戦術をとっていると判断できる。他方、三角形ネットワークを逸脱してボールが動いた場合は、サイドチェンジやクロスなどの戦略が採用されたと判断できる。

ドロネーの三角形分割はボロノイ（Voronoi）領域分割と数理的に相補的な関係にある。ボロノイ領域分割は各選手の支配領域を可視化するためにしばしば用いられる。図1と同じ選手配置にボロノイ領域分割を適用した結果が図2である。

3.2 防御側選手の守備範囲

攻撃側がつくる三角形配位がどれも同じように有効というわけではない。防御側選手によってパスコースが消されている場合も多々ある。図1における円が守備側選手の防御領域を表す。攻撃側選手がつくる三角形の有効度を評価することにより、より現実的な解析結果を得る

表1 選手に関するデータ項目一覧

フィールド名	説明
チームID	1:ホーム、0:アウェイ、3:審判
ターゲットID	各選手に振られたID、控え選手は-1
背番号	各選手の背番号
x座標	各選手のx座標 [cm]
y座標	各選手のy座標 [cm]
速度	各選手の走行速度 [cm/s]

表2 ボールに関するデータ項目一覧

フィールド名	説明
x座標	ボールのx座標 [cm]
y座標	ボールのy座標 [cm]
z座標	ボールのz座標 [cm]
速度	ボールの速度 [cm/s]
ポゼッション	Home or Away
ボールステータス	Alive or Dead

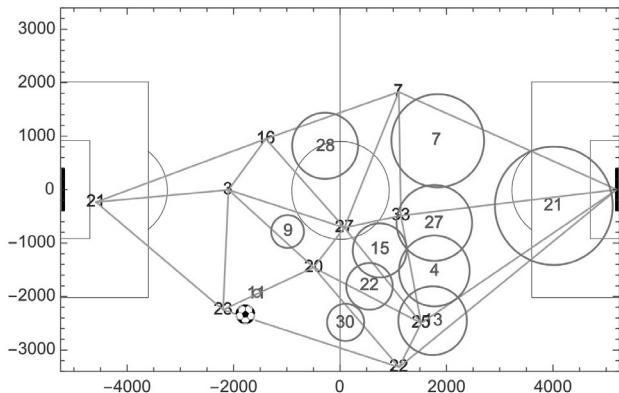


図1 攻撃側選手間のドロネー三角形ネットワークと守備側選手の防御円

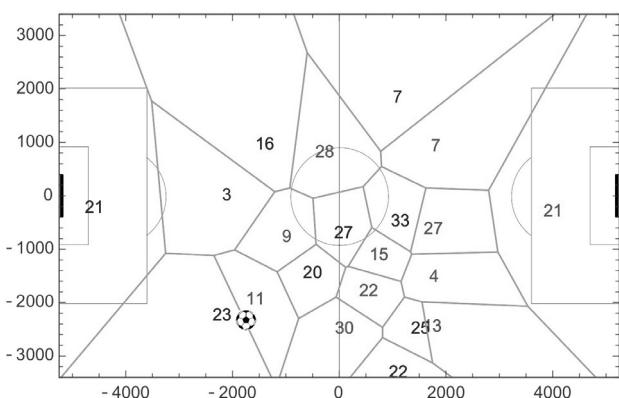


図2 選手によるピッチのボロノイ領域分割（図1の選手配置に対忾）

ことができると期待される。守備側選手の防御円の半径は、論文 [城戸 02] にならない、ボールが現在の位置からその選手まで仮想的に高速のパススピード (22 m/s) で移動する時間内にその選手が中速のスピード (4 m/s) で移動できる距離とした。

図1の例では、攻撃側 23 番の選手 (センターバック) がボールを所持し、22 番 (サイドバック), 3 番 (センターバック) および 21 番 (ゴールキーパー) へのパスコースが空いている。他方、20 番の選手 (ボランチ) へのパスコースは塞がれている。そのため、ボールをもっている 23 番の選手を中心とする三角形の内で有効な三角形は 3 番と 21 番の選手とつくる三角形に限られてしまっている。また、ボールから離れているものの、3 番、16 番 (サイドバック) および 27 番 (ボランチ) のつくる三角形は守られておらず、有効であることもわかる。

なお、ここで説明したチーム配位の可視化手法をデモンストレーションするための動画を作成している^{*4}。

4. ドロネー三角形の数理解析結果

本稿では選手がつくる三角形について分析を行っていく。特に中盤でゲームメイクを行う MF に注目し、ドロネー三角形分割のうち、MF だけによってつくられた三角形のみに着目する。

2章で説明したように、本研究で扱うデータは 1/25 秒単位という非常に高頻度なデータである。そのためすべてのフレームについてドロネー三角形分割を行って分析を行うと、ほぼ重複した状況を分析することになってしまふ。そこで本分析では、ボール保持チームの最近接選手とボールの距離が 1 m 以内であり、かつボールの速度が 1 m/s の条件を満たす場合、ある特定の選手がボールを保持している状況だとして、そのフレームの情報を抽出して分析に使用した^{*5}。連続して同じドロネーの三角形分割が続く場合には、同一の状況だとして最初のフレームのみを使用し、それ以降のフレームは除外した。

4・1 三角形の出現数

まず手始めに、MF だけによってつくられるドロネー三角形の出現数についての統計を調べた。三角形の出現数は、チームのボール保持に関する指標として使うことができる。図3は、柏レイソルと対戦相手の三角形の出現数を全 10 試合について示している。柏レイソルは全 10 試合中半数の試合で対戦相手を上回り、2 試合では拮抗、残り 3 試合では下回っている。実際の得点経過と比較すると、従来のボール保持率と同様に、三角形の出現数との間で顕著な関係性はないようである。

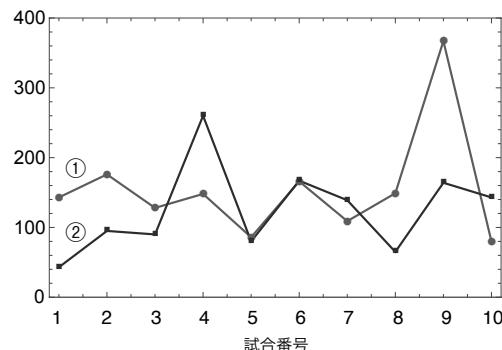


図3 MF がつくるドロネー三角形の総数
（①は柏レイソル、②は対戦チーム）

いずれにしても有意義な統計が取れるほどのサンプルがないことが大きな問題である。以降、実際の試合結果との比較には深く追求せず、得られたドロネー三角形の数理解析結果がチームの特徴付けに有用であるかどうかに着眼する。

*4 <https://www.sc.niigata-u.ac.jp/sc/res/intro/IyetomiLab.html>

*5 ワンタッチでパスが行われた場合には、検出できない可能性がある。

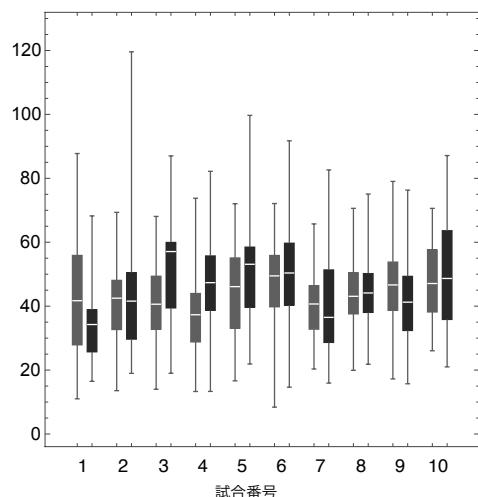


図4 MFがつくるドロネー三角形の周長の分布を表す箱ひげ図(左は柏レイソル、右は対戦チーム)

4・2 三角形の周長

選手がつくる三角形の周長は選手間の距離感を表す。周長が長ければ長いほど、選手間の距離が間延びしていることを意味する。図4で、MFがつくるドロネー三角形の周長がどのように分布しているかを箱ひげ図で表示した。第4試合の柏レイソル、第1試合の柏レイソルの対戦相手チームは、特にコンパクトなサッカーを行っていたことが見て取れる。

現代サッカーにおいて適度な選手間の距離は10 mから15 m程度[加藤99]といわれている。もし選手どうしが13 mの距離で正三角形をつくれば、その周長は約40 mとなる。そこで周長40 m以下のドロネー三角形の割合をチームのコンパクト性を表す指標とすることができる。その計算結果を図5に示す。柏レイソルは第3, 4, 5, 8試合において対戦相手に比べてよりコンパクトに試合を進めていたことがわかる。

4・3 三角形の形状

次にMFがつくるドロネー三角形の形状に目を向けてみる。ここでは三角形の形状としては正三角形が最も理想的なものと仮定する^{*6}。ドロネー三角形の形状は、選手間の距離 a, b, c を用いて次のように定義されるハーフィンダール・ハーシュマンインデックス HHIで特徴付けることができる。

$$HHI = \frac{a^2 + b^2 + c^2}{l^2} \quad (1)$$

ここで、 $l = a + b + c$ は周長である。ハーフィンダール・ハーシュマンインデックスは、もともとは経済学においてある産業の市場における企業の競争状態を表す指標の一つとして導入された。MFが形成する三角形が正三

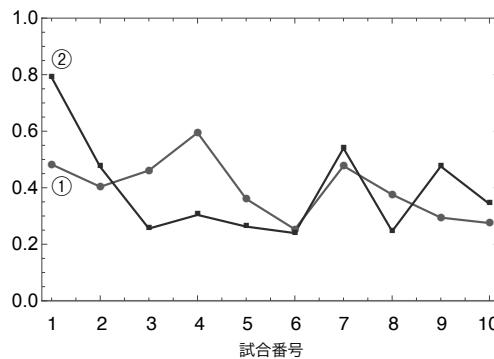


図5 周長40 m以下のMFがつくるドロネー三角形の割合
(①は柏レイソル、②は対戦チーム)

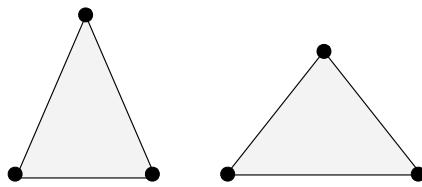


図6 $HHI = 0.34$ に対応する二つの二等辺三角形

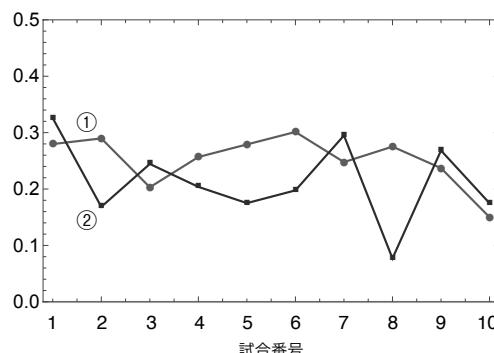


図7 $HHI \leq 0.34$ 以下のMFがつくるドロネー三角形の割合
(①は柏レイソル、②は対戦チーム)

角形の場合には、 HHI は最小値 $1/3$ をとり、三角形の形状が正三角形からはずれればずれるほど、 HHI の値は大きくなる。もし正三角形からの変形が二等辺三角形的と仮定すれば、 $HHI = 0.34$ に対応する二等辺三角形は二通りある。図6は、二つの二等辺三角形を図示したものであり^{*7}、それら二等辺三角形の正三角形からの変形度は約30%（辺の長さで評価）である。

HHI が0.34以下の三角形の割合を形成された三角形の質を特徴付けるもう一つの指標（正三角形度）とし、その結果を図7に示す。この指標の観点からは、柏レイソルは全10試合中ちょうど過半数の試合（2, 4, 5, 6, 8）で対戦相手に比べて良いパフォーマンスを示したことがわかる。

*6 代表例として、ドイツ代表のプラネットサークル戦術がある。

*7 これら二つの二等辺三角形を識別するためにはもう一つのパラメータを導入する必要がある。

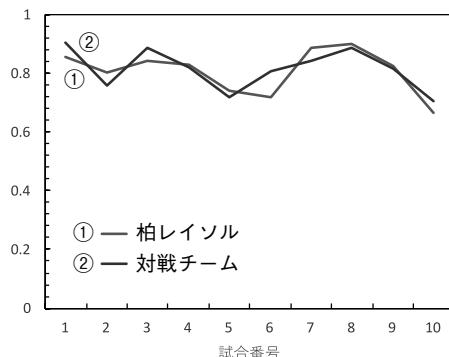


図8 各試合におけるパスコース数が一つ以上存在する状況の割合

4・4 パスコースの分析

最後に、選手間のパスコースに関する分析結果を述べる。ある1人の選手がボールをもった場合、パスする相手としては10人存在する。しかしながら、実際には、その選手の近くにいる味方の選手にパスすることが多いことは明らかであろう。

そこでボールを保持している選手が関わるドロネーの三角形分割で得られた辺をパスコースとして捉え、パスコースの数について調べた。その際に、そのパスコースが有効なパスコースかどうかが一つの問題となる。例えば、パスコースの数が多くとも、そのパスコース付近に相手選手がおり、コースのほとんどを封じられているのではあまり意味がない。そこで、3・2節で紹介した守備範囲を用いて、守備範囲にかかるパスコースは無効とした。実際には、選手は常に移動しており、選手一人一人の身体能力も異なるが、本研究では簡単のために上記のような簡略化を行った。また、得点につなげるためには相手チームゴール方向へパスコースが伸びていることが重要だと考え、自陣ゴール方向へのパスコースは除外した。ボールを保持している選手として、これまでと同様にMFに注目した。ただし、パス相手の選手については、MFに限定しない。

有効なパスコースの本数として、少なくとも1本はないとパスができないことから、1本以上のパスコースが存在する割合についてまずは調べた。結果を図8に示す。図を見るとわかるように、MFがボールを保持している状況において、パスコースが1本以上存在する状況はどの試合においてもおよそ7割以上存在する。逆にいえば、MFがボールを保持している状況において、3割程度の状況においては攻め方向へのすべてのパスコースが塞がれており、(パスコースが存在すれば)バックパスしかできない。そしてその割合に、柏レイソルと対戦チームで大きな差は見られない。

次に、有効なパスコースが1本の場合、そこへパスする可能性が高いことから、容易にコースを読まれてしまう。そこで、選択肢を増やすために3本以上の有効なパスコースが存在する状況の数について、その割合を調べた。その結果を図9に示す。

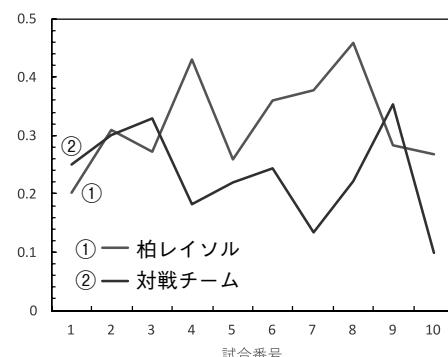


図9 各試合におけるパスコース数が三つ以上存在する状況の割合

多くの試合(2, 4, 5, 6, 7, 8, 10)で柏レイソルのほうが対戦チームよりもそのような状況が多いことが見て取れる。このことから、柏レイソルは比較的ほかの対戦チームよりも多くのパスコースを生かしている可能性が高い。しかしながら、柏レイソルが得点をあげて前半をリードして終了した試合は、第4, 5, 10試合だけであり、第2, 6試合では前半をリードされ、第1, 3, 7, 8, 9試合では前半を同点で終了している。そのため、有効なパスコース数が多いからといって、必ずしも得点につながっているわけではないことがある^{*8}。

本指標にとっての改良すべき点として、現在の分析ではある時点の静的な分析しかできていないことがある。実際のプレイでは、たとえパスコースがふさがっていたとしても、他選手が相手を引き付けるなどしてスペースをつくったり、ロングパスを出したりということも考えられる。今後はそのような動的な分析も行っていきたい。

5. おわりに

現代サッカー戦術の代表的潮流は、パスの出し手とサポート役の2人の受け手がつくる三角形を基軸としたパスサッカーである。そのため、攻撃選手、特にゲームメイクを行うMFがつくる三角形に注目した分析を行い、チームの運動性を明らかにすることを試みた。MFがつくるドロネー三角形の形状に注目した分析では、HHIを用いて形状の評価を行ったが、柏レイソルはちょうど過半数の試合において対戦相手に比べて良いパフォーマンスを示したことがわかった。

次に、守備側選手の防御の可能性を加味したパスコースの分析を行った。MFがボールを保持している際に、ドロネー三角形の辺がパスコースになると仮定し、有効

*8 もし図9の複数パスコース度の結果に図4のコンパクト度の結果および図7の正三角形度の結果も加味すると、柏レイソルにとって三拍子そろった試合は、第4, 5, 8試合である。少なくともそれらの試合の前半では柏レイソルは対戦相手にリードを許していない。もちろん、サンプル数が10試合とごく少数であり、確固たる結論を出すことは早計である。

(守備側選手によってブロックされていないという意味で) かつ対戦相手ゴール方向に向かっているパスの本数を調べた。その結果、パスコースが1本以上ある状況の数は、いずれの試合においても柏レイソルと対戦チームで大きな差は見られなかった。しかし、パスコースが3本以上ある状況の数というように、集団的行動によってプレイの幅がより広がる状況について調べてみると、柏レイソルは対戦チームよりもその割合が多いことがわかった。

以上のことから、MFどうしがつくる三角形の周長・形状の割合やパスコースの数はチームとしての連動性を特徴付けるための有力な指標となる可能性がある。しかし、サッカーの得点には偶然性が大きく関与し^{*9}、それらの指標が試合の勝敗に直接結び付くことはないであろう。さまざまな指標と組み合わせ、総合的にゲームを評価することが必要であると考えられる。本稿では、MFがボールを保持している状況の始点に着目して分析を行ったが、今後は一連の連動するプレイについて分析することで、状況の時間的变化を調べ、よりふさわしい指標を考案していく予定である。

ご覧のように、我々の研究は機械学習を駆使するまでには至っていない。サッカーではプレイ、特に集団的プレイの記述がまだ未発達であり、機械学習に分析を効果的に落とし込むことができないからである。機械学習さらに人工知能がサッカー戦術研究の分野でも重要な役割を果たすであろうことはいうまでもない。ここで紹介した研究が、そのような方向に向けての研究推進の一助になれば幸いである。

謝 辞

本研究を進めるにあたり、貴重なデータを提供していただいたデータスタジアム株式会社に感謝の意を表したい。また、井上寛康准教授（兵庫県立大学大学院シミュレーション学研究科）、河合一武教授（日本大学スポーツ科学部）、山本大准教授（日本大学スポーツ科学部）と日頃から議論していただいた。この場を借りてお礼申し上げる。

◇ 参 考 文 献 ◇

- [Hughes 88] Hughes, M.: Computerized notation analysis in field games, *Ergonomics*, Vol. 31, No. 11, pp. 1585-1592 (1988)
- [加藤 99] 加藤久: サッカーの戦術とコンピュータ分析の現状と課題、オペレーションズ・リサーチ, Vol. 44, No. 3, pp. 125-131 (1999)
- [河合 92] 河合一武、永嶋正俊、磯川正教、鈴木滋、大橋二郎、松原裕、福井真司: サッカーのゲーム分析—コンピュータを利用した即時の分析システムの実用化—、サッカー医・科学研究, Vol. 12, pp. 71-75 (1992)
- [河合 94] 河合一武、永嶋正俊、磯川正教、鈴木滋、大橋二郎、松原裕、木幡日出男、福井真司: サッカーのゲーム分析システムの実用化—NAS-3—、サッカー医・科学研究, Vol. 13, pp. 27-32 (1994)
- [城戸 02] 城戸圭介、柳田正彦、森川達也、草刈毅司、池上敦子、大倉元宏、福井真司、鈴木滋: サッカーゲームにおける新しい記述分析の提案—組み作業分析とワークサンプリング法の応用、サッカー医・科学研究, Vol. 22, pp. 203-208 (2002)
- [Kijima 14] Kijima, A., Yokoyama, K., Shima, H. and Yamamoto, Y.: Emergence of self-similarity in football dynamics, *The European Physical Journal B*, Vol. 87, No. 2, p. 41 (2014)
- [Luzuriaga 10] Luzuriaga, J.: Mean free path in soccer and gases, *European Journal of Physics*, Vol. 31, No. 5, pp. 1071-1076 (2010)
- [大橋 87] 大橋二郎: サッカーのゲーム分析用リアルタイムデータ入力プログラムの試作、ボールゲームの分析法に関する研究, Vol. 1, pp. 17-23 (1987)
- [大橋 88] 大橋二郎: サッカーのリアルタイムパス分析システムの実用化、ボールゲームの分析法に関する研究, Vol. 2, pp. 4-9 (1988)
- [Rein 16] Rein, R. and Memmert, D.: Big data and tactical analysis in elite soccer: future challenges and opportunities for sports science, *SpringerPlus*, Vol. 5, No. 1, p. 1410 (2016)

2019年6月4日 受理

著 者 紹 介



新井 優太

2015年新潟大学大学院自然科学研究科博士後期課程修了。博士（理学）。東京大学空間情報科学研究センター研究員として不動産市場に関わるビッグデータと空間データの解析を行う。特に、経済物理学を専門とし、空間データのみならず時系列データの研究も実施している。



家富 洋

1984年東京大学大学院理学系研究科博士課程修了。理学博士。東京大学理学部助手、米国アルゴンヌ国立研究所研究員、株式会社日立製作所エネルギー研究所研究員などを経て、新潟大学自然科学系教授。経済や社会の複雑現象を統計物理学の概念と手法を用いて数理的に解明。特に、景気や物価の変動ダイナミクスをはじめとして、個々の要素が織りなす集団運動に着目している。

清水 千弘は、前掲（Vol. 34, No. 4, p. 496）参照。

*9 この事実は、サッカーの得点分布がポアソン分布でよく近似できること、サッカーにはジャイアントキリングが付き物であることなどから明らかである。