

サッカーの攻撃におけるプレーの 最適化アルゴリズムの開発

徐 広孝^{1,2}・大澤 啓亮³・見汐 翔太²・
安藤 梢⁴・鈴木 宏哉⁵・西嶋 尚彦⁶

(受付 2017 年 1 月 10 日；改訂 3 月 15 日；採択 3 月 30 日)

要 旨

近年、スポーツパフォーマンスデータの分析が盛んに行われているが、ビッグデータを用いた研究は少ない。本研究の目的は、サッカーのビッグデータを用いて、攻撃においてシュートまでたどり着くためのプレーを最適化するアルゴリズムを考案することであった。データスタジアム株式会社から提供された 2013 年の J1 全 306 試合の攻撃データを、先行研究の測定項目に従って達成データセットに変換した。測定項目間のオッズ比から連動確率行列を作成し、次の手順による最適化アルゴリズムを作成した。(1) 攻撃プレーから達成項目を保存する。(2) シュートに対する連動確率に基づいて達成項目を降順でソートする。(3) 未達成項目を達成項目間に挿入した場合の確率を計算する。(4) その確率が達成項目間の確率よりも高い場合は未達成項目を挿入する。(5) 二重ループで挿入を行う。このアルゴリズムで最適化を適用し、ハーフタイムなどの短期的な場面や、数カ月単位の長期的な場面で活用する方法が提案された。

キーワード：サッカー、Jリーグ、攻撃プレー、最適化アルゴリズム、ビッグデータ。

1. はじめに

1993 年に Jリーグが開幕して以来、日本のサッカーの競技水準は著しく向上し、1998 年の初出場から 5 大会連続で FIFA (Fédération Internationale de Football Association) ワールドカップへの出場を果たした。海外リーグでプレーする日本人選手も年々増加し、日本サッカーの競技水準は世界に近づいているといえる。しかし、FIFA ワールドカップの結果は 2010 年大会のベスト 16 が最高成績であり、2014 年大会では一勝もあげることなくグループリーグで敗退した。FIFA (2016) の男子世界ランクは 2016 年 12 月時点で 45 位であり、世界の上位には及んでいない。

日本サッカー協会技術委員会 (2010) は、世界をスタンダードとして世界大会を分析し、様々

¹ 筑波大学附属駒場中・高等学校：〒154-0001 東京都世田谷区池尻 4-7-1

² 筑波大学大学院 人間総合科学研究科：〒305-8574 茨城県つくば市天王台 1-1-1

³ 日本スポーツ振興センター：〒107-0061 東京都港区北青山 2-8-35

⁴ 新渡戸文化学園：〒164-0012 東京都中野区本町 6-38-1

⁵ 順天堂大学 スポーツ健康科学部：〒270-1695 千葉県印西市平賀学園台 1-1

⁶ 筑波大学 体育系：〒305-8574 茨城県つくば市天王台 1-1-1

な強化策を推進している。さらに、指導者の仕事としてゲーム分析・評価、計画立案等が挙げられている(日本サッカー協会技術委員会, 2001)。鈴木・西嶋(2002)は、ゲームパフォーマンス分析は、一般的にゲーム分析(game analysis)とゲーム統計(game statistics)に大別されるとしている。ゲーム分析では専門家の視認的方法によって、技術、戦術、技能、チーム力などが質的に評価、記述され、ゲーム統計ではゲームパフォーマンスの分析手法(Hughes, 1996)を用いて得点数、失点数、シュート数などの計数データを扱うとしている。鈴木・西嶋(2002)は、ゲーム分析では分析者の主観性および恣意性を排除することはできないと指摘しており、Frank and Miller(1986)もまた、コーチのゲーム観察の正確性は3割程度であると述べていることから、客観的な測定値に基づいたゲームパフォーマンス分析が重要であるといえる。

ゲームパフォーマンスの分析には、データの取得が必要である。サッカーにおけるデータの取得法の原点は、競技知識をもった人間が試合を見て判断したものを記録する方法であり(加藤, 2016)、1960年代は手作業が主流であった。太田 他(1969)はグラウンド上の線を目印にして選手の移動を視覚的にとらえ、5分間に1枚のペースで手記する方法を用い、選手の移動距離や移動パターンを分析して戦術面に活用した。しかし、この方法ではデータの取得に多くの人数が必要であった。1980年代には、ビデオに録画された試合映像からゲームパフォーマンスを中心とした統計を取って分析する方法が主流となり(田中, 1984; 難波・清, 1988)、ハードウェアやソフトウェアの進化に伴って、データの取得方法も変化した。内山 他(1989)はビデオ撮影者1名、ゲーム展開(場所、プレイヤー、プレイの結果)を口頭で言う者1名、コンピュータに入力する者1名の計3名で役割を分担し、試合中にリアルタイムでデータを取得し、ハーフタイムなどに活用する方法を考案した。Hughes(1993)やErdmann(1991)は、ビデオとコンピュータを使用した解析の重要性を指摘した。

移動距離などのCGS単位系や、シュートの回数や決定率などの頻度または割合で測定できるパフォーマンスは比較的主観に依存する程度が低く、サッカー経験のある者であれば、少ない誤差で測定することができると考えられる。しかし、これらの測定可能なパフォーマンスはサッカーの一部を表現しているに過ぎず、複合的な技能やプレー構造などの測定できない領域については、選手や指導者がパフォーマンスデータに基づいて推測するしかない。そこで、鈴木 他(2000)はシュート技能の因果構造を明らかにし、山田 他(2000)はディフェンスプレッシング技能の因果構造を明らかにした。これらの研究では、複数の専門家によるデルファイ法(Linstone and Turoff, 1975)を伴う特性要因分析を適用して技能の因果構造を定性的に分析し、内容的妥当性を確認した。さらに、構造方程式モデリングを適用して技能領域間の因果構造モデルを検証し、技能評価尺度を構成した。

近年では、ハードウェア、ソフトウェアのさらなる進化によってデータ収集が飛躍的に向上し、スポーツにおいてもビッグデータを活用した研究がなされるようになった(徐 他, 2014; 丸山 他, 2015)。ビッグデータとは、「その厳密な定義はないが、小規模では成し得ないことを大きな規模で実行し、新たな知の抽出や価値の創出によって市場、組織、さらには市民と政府の関係などを変えること」と説明されている(Viktor and Kenneth, 2013)。スポーツにおけるデータ分析の根源的なテーマは試合に勝つための情報収集であることから(加藤, 2016)、サッカーのビッグデータによって、従来にはない有益な情報を得ることが期待されている。しかしながら、サッカーのビッグデータを用いた研究はまだ歴史が浅く、研究の蓄積が必要である。

本研究は、徐 他(2014)が開発したJリーグの攻撃力の測定項目を利用し、ビッグデータ解析による攻撃プレーの最適化アルゴリズムを考案した。攻撃プレーの最適化とは、「シュートにたどり着く確率を最大化するためのプレーを導き出す」ことである。サッカーの試合の目的は相手に勝つことであり、そのためには得点の獲得が必要であることから、攻撃局面はボールを奪ってからゴールを決める(あるいはボールを失う)までとなる。しかし、攻撃プレーを最適

化することを考えると、分析の対象はボールを奪ってからシュートにたどり着くまでとしたほうがよい。なぜなら、仮に理想的なプレーでシュートにたどり着いたとしても、シュート技能が低ければゴールを決める可能性が低くなるからである。つまり、ゴールを決めるかどうかはシュート技能への依存度が高いため、シュートにたどり着くまでのプレーと、その先のシュートを分けたほうが望ましいといえる。シュートにたどり着くためには、サッカー特有の技術と戦術を駆使しなければならない。技術と戦術の関係について、Jan (1989)は、「技術は戦術の重要な基礎である」と述べ、Deniau (1977)は、「技術は、戦術に役立てるために存在している」と述べている。すなわち、攻撃力は、技術と戦術の双方の側面から評価されるべきである。

これらのことを踏まえ、続く第2章では徐 他 (2014)の攻撃力測定項目について述べ、3章で本研究のデータへの適用結果を述べる。第4章では攻撃プレーの最適化アルゴリズムを説明し、第5章で活用方法の例を提示する。

2. 攻撃力の測定項目

徐 他 (2014)は、データスタジアム株式会社が測定した2011年のJリーグ Division1 (J1)と Division2 (J2)の全686試合のデータを使用して、選手とチームの攻撃力を評価する指標を作成した。データの変数は200を超え、一試合で測定されるレコード数は2,000程度であり、一年間で測定されるデータ行列はおよそ200列×137万行のビッグデータであった。このデータのレコードは、ボールを保持したプレイヤーが、ドリブルをする、パスを出す、パスを受けるなどの何かしらのアクションを起こすごとに、位置情報と共に1行記録される。1行が1アクションで構成されるため、このデータセットを「アクションデータセット」とした。測定項目は、鈴木・西嶋 (2002)や山田 他 (2000)の手続きに準じたデルファイ法を伴う特性要因分析によって、7因子39項目が作成された(図1)。ボールを奪ってからシュートにたどり着く(またはボールを失う)までのプレーは複数のアクションの集合であるため、測定項目の規準に基づいてアクションデータセットをプレーデータセットに変換した。プレーデータセットは1行が1プレーとなっており、レコード数はおよそ189万行であった。

徐 他 (2014)は、サッカーの「攻撃力」は、「技術力」と「シュート生産力」の因子から成ると仮定した。シュートを放つためには、相手よりも優位な状態をつくる必要があり、これが戦術の

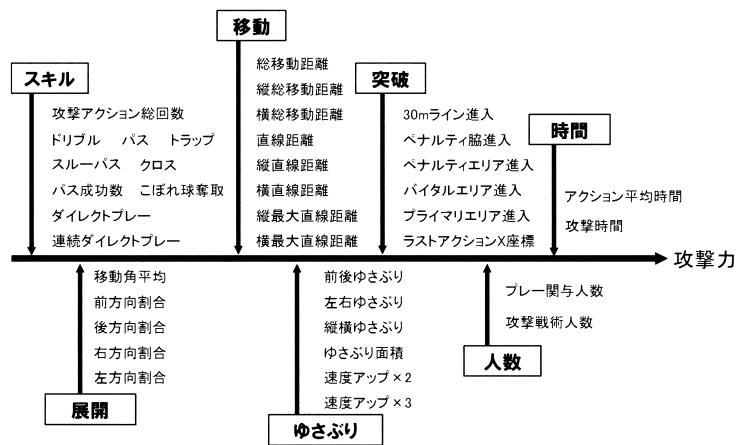


図1. 攻撃力の測定項目(徐 他, 2014).

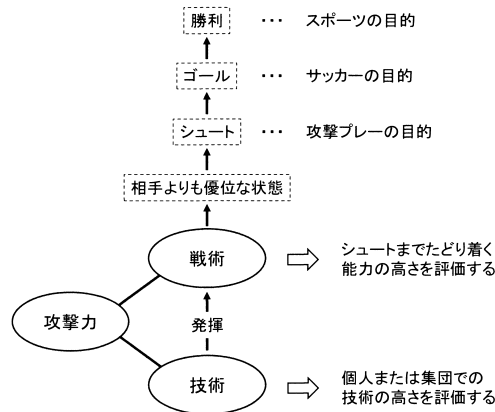


図 2. サッカーの目的と攻撃力の構造.

目的である。「シュート生産力」は、シュートにたどり着く力の大きさを表すことから、戦術技能の高さであると言い換えられる。ゆえに、技術、戦術の順に技能を発揮し、相手よりも優位に立った結果、シュートを放つことができるという因果関係が成り立つ(図 2)。これらのことから、測定項目は技術力と戦術力の両方を評価すると考えられる。

測定項目とシュートの関係を明らかにするために、CART アルゴリズム (Breiman et al., 1984) による決定木分析(分類二進木)を適用し、測定項目において、シュートした/しなかったの割合の差が最大になるような分割点を求め、これを測定項目の達成基準とした。具体的には、まず、シュートした/しなかったの 2 水準カテゴリカルデータ(以下、変数「シュート」とする)をルートノード(基準変数)とし、次式に従ってジニ係数を算出した。

$$(2.1) \quad I(S) = 1 - \sum_{j=1}^J p_j^2$$

ここで、 S は変数「シュート」、 J は水準数 2 である。 p はシュートした/しなかったの確率であり、ダミー変数であることから水準 j の分散は $p_j(1-p_j)$ となる。この分散が 0.25 であれば、シュートした/しなかったの割合が半々となり最も判別しにくい状態にあるが、サッカーではシュートにたどり着いたプレー数よりも、シュートにたどり着かなかったプレー数の方が少ないため、現実的にそうなることはない。

続いて、1つの予測変数(測定項目)を二つに分岐させる。すべての予測変数は量的尺度であるため、それぞれの分割点を探す必要がある。予測変数のレコードをソートして重複のない測定値をカウントし、その数を M とする。測定値の配列を V としたとき、分割点の候補は、 V_1 と V_2 の間から V_{M-1} と V_M の間までの $M-1$ 個である。分割点よりも大きい/小さいの二値をとるダミー変数を仮定して、次式に従い $M-1$ 個の分岐基準を計算した。

$$(2.2) \quad \Delta I = I(S) - \{P(S_L)I(S_L) + P(S_R)I(S_R)\}$$

ここで、 $P(\cdot)$ は分岐確率、 $I(\cdot)$ はジニ係数、 L は分割点よりも大きい、 R は分割点よりも小さいことを表す。 $M-1$ 個の分岐基準の中から最大のものを選び、その分割点を採用した。分割点で分岐した二つのノードにおいて、シュートした確率が高い方のノードに注目し、その割合をシュート貢献度とした。シュート貢献度は、ある測定項目の達成基準を満たすプレーをした場合に、シュートにたどり着ける確率がどの程度であるかを意味する。決定木分析は、本来で

あれば複数の予測変数で樹木を成長させるが、今回は分割点を見つけるという用途で用いた。

続いて、プレーデータセットのレコードから、シュートしたプレーのみを抽出し、さらにすべての変数を達成基準に従って達成した／しなかったの二値に変換し、これを達成データセット ($n=9,327$) とした。技術的に難しいプレーをすることは、スポーツにおいて重要であることから、2パラメータ・ロジスティック・モデル(2PLM)の項目反応理論 (Load, 1952) を適用し、項目困難度 (b) を算出した。なお、項目反応理論の前提条件となる一次元性は、テトラコリック相関行列から主因子解法の因子分析を行って、第一因子寄与率が 48.0% であり、第二因子以下に格段の差があることを確認した。

ここまで、すべての測定項目に対して、達成基準、シュート貢献度、困難度を算出した。これらのパラメータを使用して、プレーに対してシュート生産力、技術力を算出した。シュート生産力は、そのプレーにシュートまでたどり着く力がどの程度あったか、すなわち戦術技能の高さを意味する。技術力は、戦術の遂行に必要な技術技能の高さを意味する。シュート生産力 (P) は、

$$(2.3) \quad P = \frac{\sum_{i=1}^m (p_i a_i)}{A}$$

によって求められ、技術力 (B) は、

$$(2.4) \quad B = \frac{\sum_{i=1}^m (b_i a_i)}{A}$$

によって求められる。ここで、 A は達成基準を満たした測定項目数、 m は測定項目数、 p はシュート貢献度の配列、 b は困難度の配列、 a は達成した／しなかったの 1 または 0 をとる配列である。すなわち、プレーが始まってから終わるまでに達成基準を満たした項目のシュート貢献度と困難度のそれぞれの平均値を求めている。徐 他 (2014) は、 A で除する理由について、「仮に m で除した場合、無駄に多くのアクションをして達成基準を満たす項目を増やせばスコアを高めることが可能であるが、無駄なアクションをせずにシュートにたどり着いたプレーにも価値がある」と述べている。攻撃力は、因果構造に基づいてシュート生産力と技術力の積の平方根とし、因子間の相関関係を確認した。その結果、攻撃力とシュート生産力、攻撃力と技術力の相関関係が高値、シュート生産力と技術力の相関関係が低値であったことから、攻撃力を二次元で測定するモデルであることが確認された。この攻撃力指標によって、選手、チームそれぞれの技能を評価することが可能となった。しかし、どのようなプレーをすればシュートにたどり着けるかという課題は解決されていない。

3. 本研究で使用するデータの攻撃力指標への適用

第 2 章の測定項目と手続きに従い、データスタジアム株式会社から提供された 2013 年の J1 のパフォーマンスデータ (全 306 試合) を用いて達成基準を再計算し (表 1)、達成基準を満たした場合を 1、満たさなかった場合を 0 とした達成データセットを作成した。サッカーのプレーは動的であり、プレー内のある動きがその後の動きに影響すると考えられる。つまり、ある測定項目の達成は、別の項目の達成に影響するという関係が成り立つ。そこで、変数「シュート」を加えた 40 個の測定項目間の関係性をオッズ比によって求め、確率に変換した。

$$(3.1) \quad P = \frac{\text{odds}}{1 + \text{odds}} \quad \text{odds} = \frac{p(1-p)}{q(1-q)}$$

ここで、 p は片方の測定項目の達成率、 q はもう一方の測定項目の達成率である。

ある測定項目が別の測定項目に影響を及ぼすという仮定から、この確率を連動確率と命名し、すべての測定項目間の連動確率行列 (PE) を作成した (表 2)。

表 1. 測定項目の達成基準.

要因	測定項目	達成基準	要因	測定項目	達成基準	
スキル	攻撃アクション総回数	>=6	突破	30mライン進入	>=1	
	ドリブル	>=1		ペナルティ脇進入	>=1	
	パス	>=4		ペナルティエリア進入	>=1	
	トラップ	>=2		バイタルエリア進入	>=1	
	スルーパス	>=1		プライマリエリア進入	>=1	
	クロス	>=1		ラストアクションX座標	>=22	
	パス成功数	>=3		時間	アクション平均時間	>=1.05
	こぼれ球奪取	>=2			攻撃時間	>=9.5
	ダイレクトプレー	>=3		展開	移動角平均	>=31.3
	連続ダイレクトプレー	>=1			前方向移動割合	>=3.2
移動	総移動距離	>=59	後方向移動割合		>=3.3	
	縦総移動距離	>=53	右方向移動割合		>=4.8	
	横総移動距離	>=25	左方向移動割合	>=5.4		
	移動距離	>=50	ゆさぶり	前後ゆさぶり	>=2	
	縦移動距離	>=46		左右ゆさぶり	>=1	
	横移動距離	>=5		縦横ゆさぶり	>=2	
	縦最大移動距離	>=53		ゆさぶり面積	>=525	
	横最大移動距離	>=33		速度アップx2	>=2	
				速度アップx3	>=1	
				人数	プレー関与人数	>=4
				攻撃戦術人数	>=3	

4. 攻撃プレーの最適化アルゴリズムの開発

4.1 ツリー方式

連動確率行列をもとに、攻撃プレーの最適化を試みる。最も簡素的なアルゴリズムは、変数「シュート」をルートに置き、すべての測定項目の連動確率を掛け合わせて総合連動確率を求めておき、実際に行われたプレーの中から、総合連動確率が最も高いものを探すことである。しかし、この方法の総合連動確率の組み合わせは 39 の階乗である。処理に要する時間の多さと、膨大な情報の中から人が有益な情報を探すことのコストの観点から、この方法は実用的ではなかった。

4.2 挿入方式

この方法は、実際のプレーにおいて達成された測定項目をソートし、連動確率が最も高くなるように未達成の測定項目を挿入していく方法である。その手順を以下に示す。

- ① すべての測定項目の配列 V の中から、実際に行われたプレーの達成項目を保存し、配列 A とする。
- ② A を「シュート」に対する連動確率について降順でソートし、「シュート」を要素の先頭とする。
- ③ A の要素間に、 V の要素を挿入した場合の確率 p を計算する。

$$(4.1) \quad p = PE(V_i, A_j)PE(V_i, A_{j+1})$$

ここで、 PE は連動確率行列である。次の条件を満たす場合に、 V の要素を A の要素間に挿入する。

$$(4.2) \quad p > PE(A_j, A_{j+1})$$

処理①～③を、 V と A に対して二重ループさせ、 V から A に挿入された測定項目を「改善項目」、その連動確率を「改善確率」とする。

具体例を挙げて説明する。プレーの達成項目が「パス」、「縦移動距離」、「ペナ協進入」だとする(図3のA)。 V の要素数は39から3を引いた36となり、 A の要素数は3となる。 A の要素を「シュート」に対する連動確率について降順でソートすると、先頭の要素から「ペナ協進入」(0.72)、「縦移動距離」(0.71)、「パス」(0.58)となる(図3のB)。 A の先頭の要素に「シュート」を挿入すると、 A の要素間の連動確率は0.72, 0.62, 0.51となり(図3のC)、これよりも連動確率が高くなるように未達成の測定項目を挿入していく。まず、 V_1 「ドリブル」と A_1 「シュート」が0.88、 V_1 「ドリブル」と A_2 「ペナ協進入」が0.79なので、その積は0.70であり、 A_1 「シュート」と A_2 「ペナ協進入」の0.72よりも小さい。この場合、測定項目の挿入に該当しない(図3のD)。続いて、 V_2 「クロス」と A_1 「シュート」が0.92、 V_2 「クロス」と A_2 「ペナ協進入」が0.89なので、その積は0.82であり、 A_1 「シュート」と A_2 「ペナ協進入」の0.72よりも大きい(図3のE)。この場合、 A_1 「シュート」と A_2 「ペナ協進入」の間に V_2 「クロス」を挿入する(図3のF)。この処理を、 V と A のすべての要素に対して行って最適化した結果、「ペナエリ進入」、「クロス」、「ドリブル」が改善項目として挿入された(図3のG)。「シュート」を除く項目数が最適化の前後で3増加し、連動確率の積(最適化前が $0.72 \times 0.62 \times 0.51 = 0.23$ 、最適化後が $0.97 \times 0.95 \times 0.89 \times 0.84 \times 0.79 \times 0.51 = 0.28$)と項目間の連動確率(「ペナ協進入」と「シュート」は0.72から0.82、「縦移動距離」と「ペナ協進入」は0.62から0.66)が高まった。

4.3 アルゴリズムの比較

ツリー方式と挿入方式のアルゴリズムの処理速度を比較するために、両アルゴリズムのプログラムをVisual Basic for Applicationで作成し、Microsoft Office Excel 2010のアドインを実装した。一般的なノートパソコン(Sony VAIO, Intel® Core™ i7-2670QM CPU 2.20 GHz, 8 GB RAM, Windows 10 64bit)を使用して、1試合1チームの全プレーを最適化し、結果を表示するまでに要した時間を計測した。その結果、ツリー方式は76.9時間、挿入方式は78.2秒であった。

ツリー方式のアルゴリズムはシンプルであるが、すべてのパターンを計算するため処理時間が長くなってしまふ。情報量が多いことは利点のようであるが、膨大な情報の中に有益な情報が埋もれてしまい、選手やコーチが有益な情報を見落としてしまうリスクが生じる。一方、挿入方式は処理に要する時間が短く、情報量がツリー方式に比べて少ない。項目間に別の項目を挿入した場合にシュート確率が高くなるという前提に基づいたアルゴリズムであるため、関係性の低い項目が提案されることがない。そのため、プレイヤーやコーチが戦略を考えるうえで、有益な情報を見つけやすいと考えられる。しかし、挿入方式では、ある項目が達成されなかった場合にどうなるかという情報が含まれていない。これは欠点のように思えるが、この点を考慮したアルゴリズムを考えた場合、一つの問題が生じる。それは、項目を減らして計算していくと、最終的にはシュートとの連動確率が最も高い項目だけが残ってしまい、プレーとして成立しないということである。どの項目を除き、どの項目を残すかについては、プレイヤーやコーチが試合状況を総合的に判断して考えるべきであり、コンピュータに判断させることはできない。

5. チームの攻撃特性の把握とシュート率の改善

攻撃プレーの最適化を適用して、2013年のJ1上位3チーム(サンフレッチェ広島、横浜Fマリノス、川崎フロンターレ)の特性を比較した。全34節を対象にして測定項目の達成率と、改善項目として提案される確率である改善率を算出した(表3)。改善率が高い測定項目は、「エ

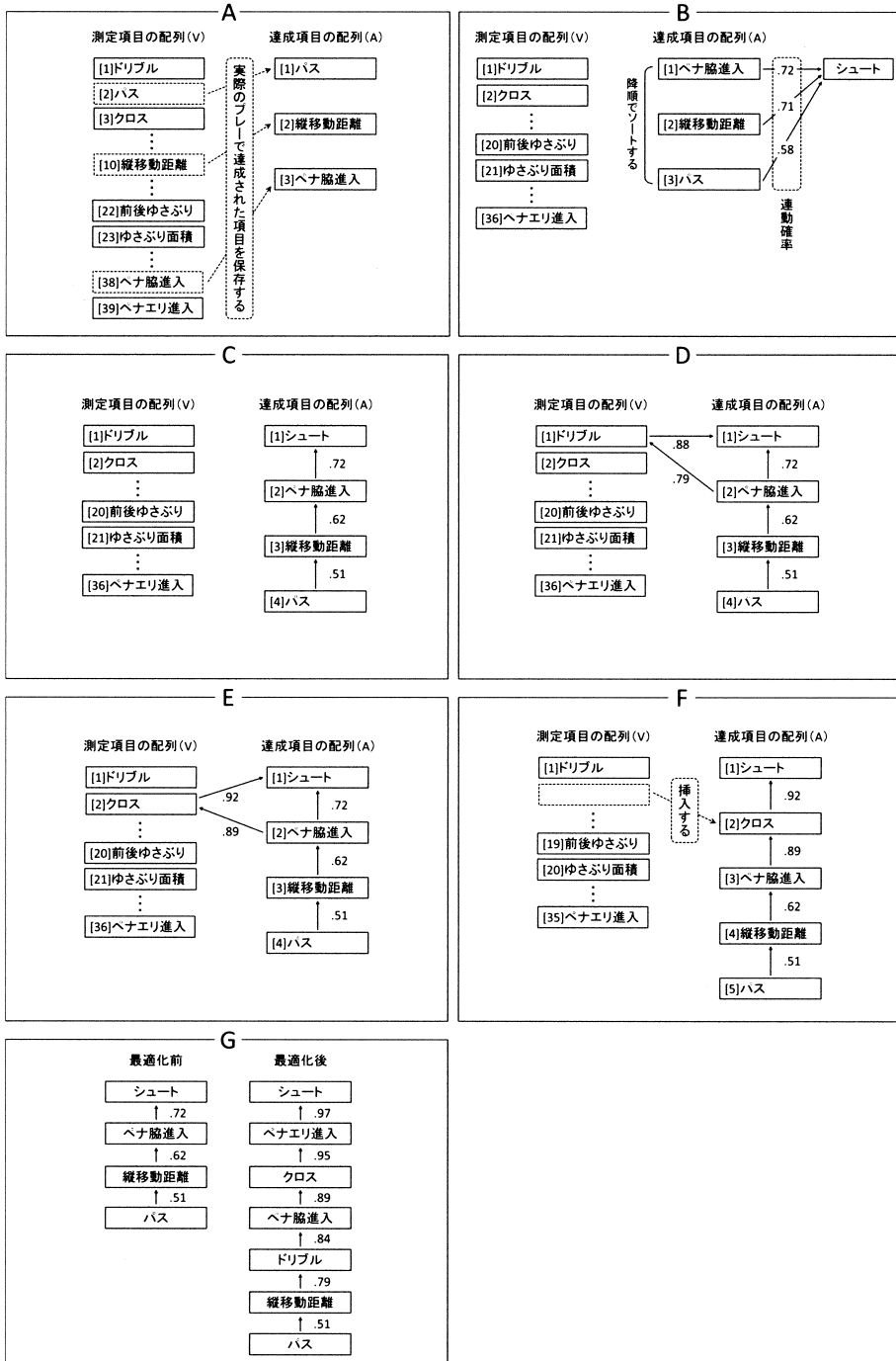


図 3. 攻撃プレーの最適化アルゴリズムの手順。

リア」要因に集中している。例えば「ペナルティエリア進入」は、ピッチにおけるペナルティエリアに進入できたかどうかを示す測定項目である。ペナルティエリアはゴールに近く守備が厚

表 3. 2013 年 J1 上位 3 チームの項目達成率と改善率、

要因	項目	広島				横浜F・マリノス				川崎					
		達成回数	改善回数	プレー回数	達成率	達成回数	改善回数	プレー回数	達成率	達成回数	改善回数	プレー回数	達成率		
スキル	攻撃アクション総回数	2,079	465	4,013	51.8	1,889	743	4,683	40.3	1,599	2,033	554	4,245	47.9	
	ドリブル	516	0	4,013	12.9	0	462	0	4,683	9.9	0	519	0	4,245	
	パス	1,817	1,038	4,013	45.3	1,635	1,478	4,683	34.9	31.6	1,693	1,237	4,245	39.9	
	スルーパス	417	160	4,013	10.4	429	111	4,683	9.2	2.4	548	142	4,245	12.9	
	トラップ	2,175	1,29	4,013	54.2	3.2	2,098	157	4,683	44.8	3.4	2,183	126	4,245	51.4
移動	クロス	579	1,446	4,013	14.4	36.0	445	1,380	4,683	9.5	29.5	352	1,513	4,245	8.3
	パス成功数	1,797	796	4,013	44.8	19.8	1,581	1,078	4,683	33.8	23.0	1,738	873	4,245	40.9
	こぼれ球奪取	87	1,406	4,013	2.2	35.0	92	1,082	4,683	2.0	23.1	91	1,142	4,245	2.1
	ダイレクトプレー	872	1,844	4,013	21.7	46.0	640	2,318	4,683	13.7	49.5	701	2,067	4,245	16.5
	連続ダイレクトプレー	2,695	82	4,013	67.2	2.0	2,993	85	4,683	63.9	1.8	2,731	105	4,245	64.3
突破	総移動距離	1,912	616	4,013	47.6	15.4	1,523	964	4,683	32.5	20.6	1,673	809	4,245	39.4
	縦総移動距離	1,553	751	4,013	38.7	18.7	1,202	940	4,683	25.7	20.1	1,318	858	4,245	31.0
	横総移動距離	2,214	311	4,013	55.2	7.7	1,902	506	4,683	40.6	10.8	2,078	383	4,245	49.0
	移動距離	1,095	837	4,013	27.3	20.9	988	648	4,683	21.1	13.8	1,134	609	4,245	26.7
	縦移動距離	835	1,014	4,013	20.8	25.3	658	838	4,683	14.1	17.9	826	815	4,245	19.5
時間	横最大移動距離	2,828	64	4,013	70.5	1.6	3,149	29	4,683	67.2	0.6	2,938	54	4,245	69.2
	縦最大移動距離	1,808	259	4,013	45.1	6.5	1,740	259	4,683	37.2	5.5	1,640	282	4,245	38.6
	30mライン進入	1,150	1,559	4,013	28.7	38.8	1,170	1,724	4,683	25.0	36.8	1,213	1,523	4,245	28.6
	ペナルティ脱走	470	1,309	4,013	11.7	32.6	514	1,127	4,683	11.0	24.1	411	1,289	4,245	9.7
	ペナルティエリア進入	397	3,593	4,013	9.9	89.5	440	4,204	4,683	9.4	89.8	513	3,702	4,245	12.1
展開	ハイレベルエリア進入	920	1,606	4,013	22.9	40.0	903	1,914	4,683	19.3	40.9	1,088	1,535	4,245	25.6
	ブライエリア進入	216	3,772	4,013	5.4	94.0	254	4,393	4,683	5.4	93.8	248	3,967	4,245	5.8
	ラストアクション保護	1,425	818	4,013	35.5	20.4	1,604	1,101	4,683	34.3	23.5	1,578	947	4,245	37.2
	アクション平均時間	2,992	99	4,013	74.6	2.5	3,255	156	4,683	69.5	3.3	2,959	182	4,245	69.7
	攻撃時間	2,037	630	4,013	50.8	15.7	1,842	914	4,683	39.3	19.5	1,865	788	4,245	43.9
ゆさぶり	移動角平均	3,375	0	4,013	84.1	0.0	3,837	0	4,683	81.9	0.0	3,496	0	4,245	82.4
	前方向移動割合	2,490	146	4,013	62.0	3.6	2,609	122	4,683	55.7	2.6	2,427	155	4,245	57.2
	後方向移動割合	2,429	166	4,013	60.5	4.1	2,559	139	4,683	54.6	3.0	2,444	168	4,245	57.6
	右方向移動割合	2,421	122	4,013	60.3	3.0	2,357	131	4,683	50.3	2.8	2,357	115	4,245	55.5
	左方向移動割合	2,397	109	4,013	59.7	2.7	2,310	122	4,683	49.3	2.6	2,438	105	4,245	57.4
人数	前後ゆさぶり	506	1,354	4,013	12.6	33.7	429	1,151	4,683	9.2	24.6	472	1,243	4,245	11.1
	左右ゆさぶり	1,220	740	4,013	30.4	18.4	1,009	692	4,683	21.5	14.8	1,193	677	4,245	28.1
	縦横ゆさぶり	1,999	198	4,013	49.8	4.9	1,763	281	4,683	37.6	6.0	1,930	228	4,245	45.5
	ゆさぶり面積	1,641	698	4,013	40.9	17.4	1,134	1,048	4,683	24.2	22.4	1,325	907	4,245	31.2
	速度アップx2	533	1,985	4,013	13.3	49.5	399	2,149	4,683	8.5	45.9	418	2,069	4,245	9.8
人数	速度アップx3	733	960	4,013	18.3	23.9	659	767	4,683	14.1	16.4	634	849	4,245	14.9
	プレー量と人数	1,913	375	4,013	47.7	9.3	1,752	488	4,683	37.4	10.4	1,804	476	4,245	42.5
	攻撃戦術人数	2,169	299	4,013	54.0	7.5	2,317	211	4,683	49.5	4.5	2,104	302	4,245	49.6
	最大値	3,375	3,772	4,013	84	94	3,837	4,393	4,683	82	94	3,496	3,967	4,245	82
	中央値	1,641	698	4,013	41	17	1,523	767	4,683	33	16	1,578	788	4,245	37
標準偏差	最小値	87	0	4,013	2	0	92	0	4,683	2	0	91	0	4,245	2
	平均値	1,528	846	4,013	38	21	1,465	942	4,683	31	20	1,483	873	4,245	35
	標準偏差	868.7	872.3	0.0	21.6	21.7	951.0	1,002.6	0.0	20.3	21.4	876.1	897.4	0.0	20.6

くなる場所でもあるため、達成率が低くなり、改善率が高くなるのは必然的である。したがって、これらの項目よりも、選手が意思選択可能な項目に着目して分析することが望ましい。

「クロス」は、1プレーの中でクロスボールを上げられたかどうかを示す測定項目である。達成率は広島が14.4%、横浜が9.5%、川崎が8.3%であり、改善率は広島が36.0%、横浜が29.5%、川崎が35.6%であった。広島はクロスを多く上げられるチームであり、なおかつ改善率が高いことから、クロスをさらに活用することでシュート率を高めることができる。横浜は達成率と改善率が低いことから、クロス以外のプレーをいかした攻撃を組み立てるべきである。川崎は達成率が低い、改善率が高いことから、クロスをより多く上げることで、シュートチャンスが生まれると考えられる。このように、項目達成率によってチームの攻撃特性を把握することができ、改善率に基づくプレーの選択は、シュート率を高めることに貢献する。

6. おわりに

攻撃プレーの最適化を活用する場面はさまざまである。試合中にリアルタイムでデータを取得できれば、ハーフタイムに前半のプレーの最適化を行うことによって、後半にどのような攻撃プレーを選択すべきであるかをフィードバックすることが可能である。また、数カ月に渡って試合のデータを測定し、その間の全プレーを最適化すれば、項目の達成率はチームの攻撃プレースタイルを反映し、改善率はシュート率を高めるためのプレー選択における情報源となる。

加藤(2016)は、「ほとんどのクラブの担当者は、いわゆる統計的なバックグラウンドを持っていない」と述べている。サッカーに限らず、スポーツではデータを分析する専門職であるスポーツデータサイエンティストが定着しつつあるが、多くのチームが有能な人材を欲している段階であるがゆえに、分析の過程や結果は、理解が難であるものは避けるべきであろう。また、加藤はスポーツデータ解析の方向性として、「どのように攻撃を組み立てて得点するかというチームごとに異なるフィロソフィーの部分からスタートして、データの取得→活用→フィードバック→改善というプロセスを繰り返していくことが必要である」とも述べている。攻撃プレーの最適化はこれらの指摘の実現の一助となると考えられる。しかし、サッカーの技術や戦術、攻撃スタイルは変わっていくものであり、それに応じて測定項目の精査、改善が必要となる。また、連動確率がオッズ比から算出されているため、測定項目間の短方向の因果関係を考慮したアルゴリズムの開発等が次の課題である。

参 考 文 献

- Breiman, L., Friedman, J. H., Olshen, R. A. and Stone, C. J. (1984). *Classification and Regression Trees*, Chapman & Hall, London.
- Deniau, G. (1977). *Tennistaktik*, Quickmedia, Gunzenhausen.
- Erdmann, W. S. (1991). Quantification of games — preliminary kinematic investigations in soccer, *Science and Football II* (eds. T. Reilly, J. Clarys and A. Stibbe), 174–179, E & FNSPON, London.
- FIFA (2016). FIFA Men's World Ranking, <http://www.fifa.com/fifa-world-ranking/>.
- Frank, I. M. and Miller, G. (1986). Eye witness testimony in sport, *Journal of Sports Behavior*, **9**, 38–45.
- Hughes, M. D. (1993). Notation analysis in football, *Science and Football II* (eds. T. Reilly, J. Clarys and A. Stibbe), 151–159, E & FNSPON, London.

- Hughes, M. D. (1996). Notational analysis, *Science and Soccer* (ed. T. Reilly), 343-361, E & FN SPON, London.
- Jan, K. (1989). *Taktik im Sport*, Verlag Karl Hofmann, Schorndorf.
- 徐広孝, 横尾智治, 安藤梢, 西嶋尚彦, 熊谷紗希, 猶本光, 鈴木宏哉, 山田庸, 中野貴博, 齋藤健治 (2014). Jリーグにおける選手とチームの攻撃力指標, スポーツデータ解析における理論と事例に関する研究集会, **1**, 21-26.
- 加藤健太 (2016). サッカーにおけるデータ分析とチーム強化, 通信ソサイエティマガジン, **37**, 29-34.
- Linstone, A. H. and Turoff, M. (1975). *The Delphi Method*, Addison-Wesley Publishing Company, Massachusetts.
- Load, F. M. (1952). A Theory of Test Scores, *Psychometric Monograph*, **7**, Psychometric Society, Richmond.
- 丸山智久, 鈴木惇一, 田畑耕司 (2015). ゾーンとパスから見える J1 チームの特徴, スポーツデータ解析における理論と事例に関する研究集会, **2**, 17-22.
- 難波邦雄, 清剛裕 (1988). 発達段階別に見たサッカーの基礎的技術の比較検討, サッカー医・科学研究, **8**, 125-148.
- 日本サッカー協会技術委員会 (2001). サッカー指導教本・2000 年度版, 日本サッカー協会, 東京.
- 日本サッカー協会技術委員会 (2010). U-12 指導指針, 日本サッカー協会, 東京.
- 太田哲夫, 浅見俊雄, 小宮喜久, 戸苅晴彦, 菊池武道, 森忠保 (1969). サッカーのゲーム分析, サッカー, **92**, 31-43.
- 鈴木宏哉, 西嶋尚彦 (2002). サッカーゲームにおける攻撃技能の因果構造, 体育学研究, **47**, 547-567.
- 鈴木宏哉, 山田庸, 大迫剛, 高橋信二, 西嶋尚彦 (2000). フォワード選手におけるゲームパフォーマンスからのシュート技能の計量, サッカー医・科学研究, **20**, 37-41.
- 田中和久 (1984). サッカー競技におけるスタイルの研究 最終ディフェンスラインの突破方法, サッカー医・科学研究, **5**, 49-56.
- 内山秀一, 今川正浩, 西野仁, 宇野勝 (1989). コンピュータを導入したサッカーのゲーム分析法, サッカー医・科学研究, **9**, 109-117.
- Viktor, M. S. and Kenneth, C. (2013). Letting the data speak, *Big Data: A Revolution That Will Transform How We Live, Work, and Think*, John Murray, London.
- 山田庸, 鈴木宏哉, 大迫剛, 高橋信二, 西嶋尚彦 (2000). ゲームパフォーマンスからのディフェンスプレッシャーの計量, サッカー医・科学研究, **20**, 32-36.

Development of Optimization Algorithm for Attack Play in Football

HirotaKa Jo^{1,2}, Keisuke Oosawa³, Syota Mishio²,
Kozue Ando⁴, Koya Suzuki⁵ and Takahiko Nishijima⁶

¹Junior and Senior High School at Komaba, University of Tsukuba

²Doctoral Program in Physical Education, Health and Sport Sciences, University of Tsukuba

³Japan Sports Council

⁴NitobeBunkagakuen

⁵Department of Health and Sports Science, Juntendo University

⁶Faculty of Health and Sport Sciences, University of Tsukuba

Although many analyses of sports performance data have been performed, few studies have worked with big data. The purpose of this study was to invent an optimization algorithm to increase player shot probability using big data. Using attacking data in all 306 matches in the J. League division 1 in 2013, supplied by DataStudiam Inc., we converted raw data to a binary dataset in accordance with the measurement items in a prior study. To create a cooperation probability matrix from the odds ratio between measurement items, we invented the “insertion algorithm,” which has the following procedure: (1) Store the ‘success’ items from attacking play; (2) sort the success items in descending order based on cooperation probability for “shoot”; (3) calculate probability in case inserting a ‘failure’ item between the success items; (4) if the probability was higher than the probability between the success items, insert the failure item; and (5) continue the insertion by double-loop. Team attack characteristics were compared by calculating the success rate and improvement rate adapted by the algorithm.