

サッカーにおけるポジショニングの修正を行う 遺伝的アルゴリズムモデルの提案

神宮司 祐哉^{1,a)} 清 雄一¹ 田原 康之¹ 大須賀 昭彦¹

受付日 2023年3月14日, 採録日 2023年10月3日

概要: サッカーのポジショニングは、少しの違いで失点のリスクが大きく変わるため解釈が難しい。そこで本研究では、ポジショニングを数値的に評価する手法 (Pitch Risk) を導入した、ポジショニング修正モデルを提案する。アマチュアの選手の試合後の振り返りでの利用などを想定している。Pitch Risk は、ピッチの支配状況を示す Pitch Control, ゴールの確率を示す xG, アシストの確率を示す SxA を用いて、確率計算を基に算出した。サッカー経験者 7 人が局面において失点のリスクが高いと回答したエリアとの JS-divergence は、Pitch Control の 118.18 に対して、Pitch Risk は 0.14 であることから、Pitch Risk は失点のリスクをより表現できている。ポジショニング修正モデルには GA を用いた。守備側 11 選手の距離と方向の計 22 変数を遺伝子とし、評価関数には Pitch Risk を用いた。GA で生成した個体数分ランダム生成した結果と比較すると、約 3% の数値的改善が生じた。また、11 選手分の学習を同時に進めたことで選手が連動した改善が見られた。モデル改善による精度向上により、ポジショニングの解釈を幅広い人々に提供できると考える。

キーワード: サッカー, ポジショニング, Pitch Control, 遺伝的アルゴリズム

Proposal of Model for Modifying Positioning of Football by Genetic Algorithm

YUYA JINGUSHI^{1,a)} YUICHI SEI¹ YASUYUKI TAHARA¹ AKIHIKO OHSUGA¹

Received: March 14, 2023, Accepted: October 3, 2023

Abstract: The positioning of soccer is difficult to interpret because a little difference can significantly change the risk of conceding a goal. Therefore, we introduce a method for quantifying the risk of conceding a goal (Pitch Risk) and propose a model to output the modification of positioning in a phase based on the method. The model is intended for use by amateur athletes in their post-game review. Pitch Risk is calculated based on a fundamental rule of probability using Pitch Control, which indicates pitch dominance, xG, which indicates the probability of a goal, and SxA, which indicates the probability of an assist. Pitch Risk is a better representation of the risk of conceding a goal since the JS-divergence with the areas where the seven experienced soccer players reported a higher risk of conceding a goal in the phase is 0.14 for Pitch Risk compared to 118.18 for Pitch Control. We use GA for the Positioning Modification model. A total of 22 variables of distance and direction for the 11 defending players were used as chromosomes, and Pitch Risk was used as the evaluation function. Improvement of about 3% occurred when compared to the best of randomly generated. In addition, there was a player-linked improvement due to the simultaneous learning of 11 players. We believe that the improvement of accuracy for the model provides positioning interpretations to a wider number of people.

Keywords: soccer, football, positioning, pitch control, genetic algorithm

¹ 電気通信大学
The University of Electro-Communications, Chofu, Tokyo
182-8585, Japan

^{a)} jingushi.yuya@ohsuga.lab.uec.ac.jp

1. はじめに

近年、計算機や分析技術が発達したことにより、スポー

ツの分野においても分析が行われるようになった。サッカーも例外ではなく、各クラブに分析チームが存在するほど戦術分析が浸透している。戦術分析には、試合から得られる非常に多岐にわたる情報が用いられ、様々な事象が分析の対象となっている。過去の試合のデータを用いた戦術の推定 [1] や、プレーのデータを用いてプレーの評価 [2]、トラッキングデータを利用した選手の動きの予測 [3]、さらにはそれらを組み合わせることでさらにサッカーにおける分析の発展 [4] を試みるなど、その例は枚挙に暇がない。

選手のポジショニングも戦術分析の情報になる。ポジショニングは、近年注目されている「ポジショナルプレー」に代表されるように重要性が増している。一方で、味方や相手との位置関係、ピッチ上の位置などが少し違うことで失点のリスクが大きく変化するため、考慮することが多く解釈が非常に難しい要素である。

ポジショニングに関する研究として、Pitch Control [5] があげられる。これは、選手の位置などの情報からピッチ上における選手の優位性を可視化できるモデルであり、一目でピッチ全体の有利不利が分かる点で有用である。Pitch Control に基づいて守備側のポジショニングについて考える場合、攻撃側が支配しているエリアを減らすようにすることで失点のリスクを軽減できる。しかし、攻撃側が有利なエリアでもゴールやアシストにつながらないエリアでは失点のリスクが低いため、守備側はそのようなエリアを無理に減らす必要はない。すなわち、失点するリスクを考慮に加えることで、守備側のポジショニングについてより正確に考えることが可能になる。

そこで、本研究ではゴールやアシストなどの試合を決定づけるプレーに関するデータと選手の位置のデータを組み合わせることで、守備側の選手のポジショニングの修正を行うモデルの提案を行う。これを実現するうえで本研究では、ポジショニング評価システムとポジショニング修正システムの作成を行った。ポジショニング評価システムは、ピッチ上のゴールの期待度を示す xG [6]、ピッチ上のアシストの期待度を示す SxA、先述の Pitch Control を用いたピッチにおける失点のリスクを数値化・可視化する。ポジショニング修正システムは、守備側の選手における移動すべき距離と方向を個体の遺伝子として用いた遺伝的アルゴリズムによって実現した。

本研究の利用例として、局面の振り返りがあげられる。主にアマチュアのチームを対象と考えており、試合の翌日のミーティングで失点シーンなどの振り返りを行う際に用いることで、具体的なポジショニングの修正案がプレー改善の一助になると考えられる。

本研究を発展させることにより、より多くの人々がポジショニングについて理解を深めるきっかけになる。これは、プレーヤーにとってはサッカーの理解やプレーの向上につながり、サポーターにとっては観戦をより面白くすること

につながると考える。

2 章では本研究に関連する研究の紹介、3 章では本研究の提案手法についての説明を行う。4 章では提案手法の実験詳細とその結果およびそれに対する評価、5 章では本研究に関する考察、6 章では本研究のまとめについて述べる。

2. 関連研究

2.1 Pitch Control

Pitch Control は、ピッチの支配状況を可視化するモデルである。文献 [5] では Pitch Control を 2 つの手法で実現している。1 つ目は「Pass Probabilities model」を拡張したものである。Pass Probabilities model とは、ボールや選手の地点への到達やボールを相手に奪われる可能性を考慮して、パスが通る可能性を算出するモデルである。パスが通るということは、その地点は味方が占有していると言い換えることができることから、Pass Probabilities model を拡張することで Pitch Control の実現をしている。もう 1 つは「player influence areas」を用いたものである。player influence areas は、選手の位置と速度によって決定づけられるある時刻でのある地点への選手の影響力を基に算出される領域である。図 1 は 2 つの手法によって算出された Pitch Control の例である。いずれのアプローチも信頼性の高い結果を得られるが、前者はすべての地点のすべてのプレーヤーの到着時刻を計算するため計算量が多いこと、後者は player influence areas には限界があるためいずれの選手にも支配されない白い部分があるという点で差が生じている。

2.2 xG

xG は、expected Goals の略であり、シュートがゴールになる確率を表す指標である。シュート自体の評価はもちろん、1 試合中や 1 シーズン中、チームや選手などを対象に和をとることでチームや選手の評価をすることも可能である。近年は戦況を考察する際に用いられることも多く、広く知られるようになっていく。

確率を表すことから、最大値は 1、最小値は 0 である。xG は目的や変数に応じて様々な算出方法が存在する。一般的に用いられている方法では、シュートした位置からのゴールまでの距離やゴールへの角度に加えて、シュートした部位（右足、左足、頭など）、直前のパスの種類（スルーパス、クロスなど）、シュートしたときの状況（キーパと 1 対 1、セットプレーなど）などが考慮されている。Stats Perform 社が提供する xG のモデル [7] は、過去の試合データから得られる何十万にのぼるシュートを基に先述した要素を変数として組み込んだロジスティック回帰モデルによって構成されている。

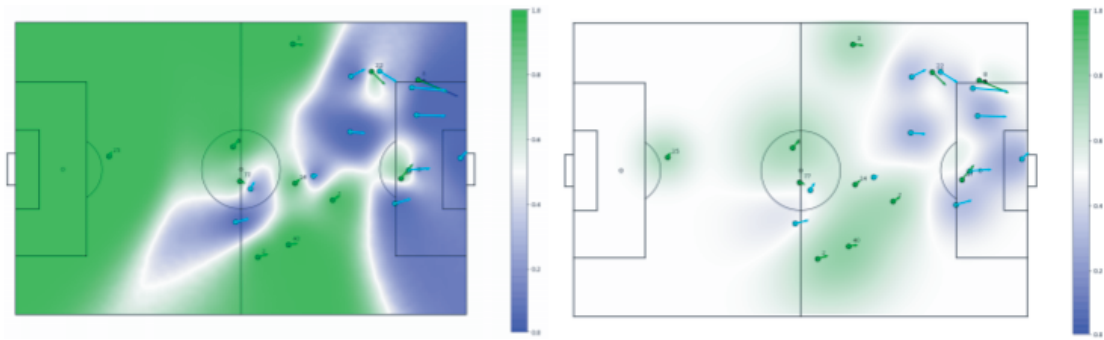


図 1 Pitch Control の例 (文献 [5] より). 左: Pass Probabilities model を拡張したもの
右: player influence areas を用いたもの

Fig. 1 Example of Pitch Control (cited from Ref. [5]). Left: Extension of Pass Probabilities model Right: using player influence areas.

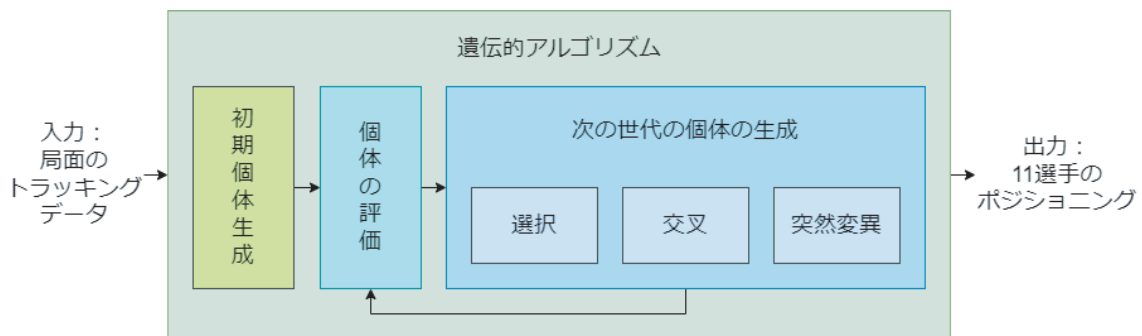


図 2 ポジショニング修正システムの概形

Fig. 2 Outline of a system for modifying defender positioning.

2.3 xA

xA は, expected Assists の略であり, パスがアシストになる確率を示す指標である. xG と同様に, 和をとることで選手やチームのパフォーマンスを評価することが可能である.

確率を表すことから, 最大値は 1, 最小値は 0 である. 一般的な xA はパスの種類 (スルーパス, クロスなど), パスの距離, パスを出した位置, パスを受ける位置, パスしたときの状況 (オープンプレー, コーナーなど) といった要素が変数として用いられている. Stats Perform 社が提供する xA のモデル [8] は, 過去の試合データから得られる何十万にのぼるパスを基に先述した要素を変数として組み込んだロジスティック回帰モデルによって構成されている.

3. 提案手法

図 2 は提案するポジショニング修正システムの概形である. 遺伝的アルゴリズムを用いるうえで, ポジショニングを数値的に評価する必要があることから, ポジショニング評価システムの作成も行った. 本章では, ポジショニング評価システムとポジショニング修正システムの説明を行う. 以降の説明における, 選手の位置や速度などの表現は特に断りがない限り, LaurieOnTracking [9] に準ずるものとする.

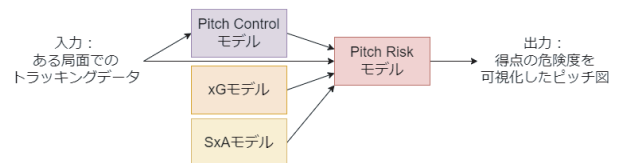


図 3 ポジショニング評価システムの概形

Fig. 3 Outline of a system for quantifying the risk of conceding a goal.

3.1 ポジショニング評価システム

図 3 は提案するポジショニング評価システムの概形である. このモデルは Pitch Control モデル, xG モデル, SxA モデル, Pitch Risk モデルの 4 つから構成されている. Pitch Control モデルと xG モデルは既存のモデルを利用し, SxA モデルと Pitch Risk モデルについては新しいモデルを提案する. 本モデルの入力はある局面におけるトラッキングデータ, 出力は入力された局面から各地点にボールが移動したときに 2 プレー以内にゴールする確率を可視化した図である. 入力であるトラッキングデータは, 指定した局面における選手とボールの位置を示す 46 の数値 $((11(\text{選手}) \times 2(\text{チーム}) + 1(\text{ボール})) \times 2(\text{次元}))$, 選手の速度と加速度を示す 88 の数値 $(11(\text{選手}) \times 2(\text{チーム}) \times 2(\text{次元}) \times 2(\text{種類}))$, 計 134 の数値からなるデータである.

以下では, 提案するモデルを構成する各モデルの説明を

行う。

3.1.1 Pitch Control モデル

Pitch Control は主に選手の位置の情報を用いて、ピッチの支配状況を可視化するモデルである。このモデルの出力はピッチの支配状況を表す。そのエリアを支配できているということは選手はプレーの選択肢に自由があることになり、よりゴールになりやすいプレーを選択することができる。つまり、ピッチの支配度はゴールのリスクに関係するため、提案するモデルに欠かせない要素といえる。

入力はある局面におけるトラッキングデータ、出力は 106×68 の配列である。出力される配列の要素は、 $106\text{m} \times 68\text{m}$ のピッチを $1\text{m} \times 1\text{m}$ に区切った各エリアに対応している。すなわち、出力される配列はピッチ上の各エリアにおける支配状況である。モデルには、LaurieOnTracking [9] で提供されている Pitch Control の算出に関連する関数を用いた。この関数は 2.1 節で述べた Pass Probabilities model の考え方を用いたものである。以降、この関数の説明である。まず、攻撃側と守備側のそれぞれについて、その地点から最も近いプレーヤの到達時間を計算する。この到達時間とボールがその地点に到達する時間、ボールをキープするために要する時間を基に場合分けを行う。いずれかのチームのプレーヤがもう一方のチームのプレーヤが到達するまでにボールをキープするのに十分な時間がある、すなわちいずれかのチームがボールをキープする可能性がない場合、ボールをキープできるほうのチームは 1、ボールをキープできないほうのチームは 0 を Pitch Control の値として返す。それ以外の場合、ボールをキープしうる距離にいるすべてのプレーヤを対象に各チームがボールをキープする確率を決定し、その合計を Pitch Control の値として返す。これらの操作を各エリアに対して行うことで、出力となる配列が完成する。

3.1.2 xG モデル

xG は、expected Goals の略であり、ゴール期待値、すなわちゴールになる確率を示す。ゴールになりやすいエリアは守備側が意識的にカバーする必要があるエリアであり、守備側のポジショニングを考えるうえでは欠かせないため、xG モデルの導入をした。しかし、ポジショニング評価システムにおいては、一般的に用いられている xG のモデルを適用することは難しい。一般的なモデルではパスの種類やシュートした部位なども変数として用いられているが、ポジショニング評価システムでは与えられた局面以降にシュートが生じることから算出時点ではシュートが行われていない。加えて、試合の映像から選手の位置の情報を得る研究はさかに行われていることから [10], [11], [12], ピッチの位置情報はアマチュアの利用者が得られる可能性が高い情報だといえる。よって本研究では、位置に関連する変数を入力とした xG モデルを用いる。このモデルの出力はその地点から撃たれたシュートがゴールになる確率を

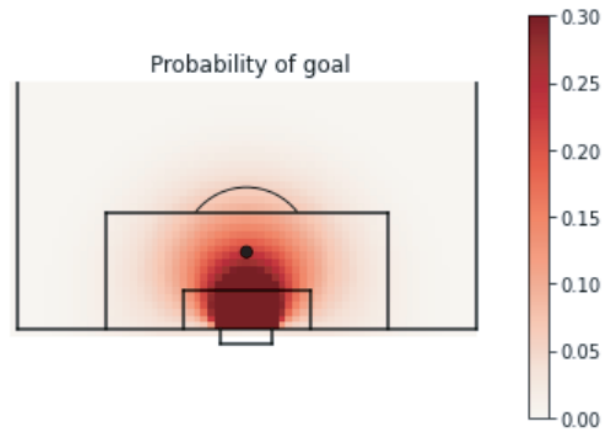


図 4 xG モデルの出力を可視化した図

Fig. 4 Figure of visualizing output of xG model.

示す。

Lu ら [6] の論文では xG のモデルとして、以下の式 (1) が紹介されている。

$$P(x, \theta) = \frac{1}{1 + \exp(Z)} \quad (1)$$

$$\text{where } Z = 4.03 - 2.53\theta + 0.12x + 0.11x\theta - 0.0069x^2$$

ここで、 $P(x, \theta)$ はゴールの確率すなわち xG、 x はゴールラインまでの距離、 θ は自身と両相手ゴールポストを結んだ線のなす角（ゴールとの角度）を表す。この式はロジスティック回帰を用いて算出されたものであり、変数についてはゴールの決めやすさに関する一般的な特徴から距離と角度が設定された。ゴールの決めやすさはゴールから近いほど上がり、サイドよりも中央、言い換えると自身と両相手ゴールポストを結んだ線のなす角が大きいくほど上がるというのは、直感的にも理解が可能だと考える。式 (1) を基にゴール付近の xG を算出し可視化した図は図 4 のとおりである。

本研究では xG モデルとして式 (1) を用いることとした。

3.1.3 SxA モデル

SxA は、Simplified expected Assists の略であり、アシスト期待値、すなわちアシストになる可能性を表す。アシストになりやすいエリアは、失点につながりやすいといえる。ゴールになりやすいエリアと同様に、守備側が意識的にカバーする必要がある。守備側のポジショニングを考えるうえでは欠かせない。しかし、先述した xG モデルと同様の理由で、ポジショニング評価システムにおいて一般的に用いられている xA のモデルを適用することは難しい。一方で、xA は xG よりも用いられる機会が少ないことや表現が複雑であることが原因と考えられるが、先述のとおりアマチュアの利用者が得やすいと考えられる。位置に関連する変数のみを用いたモデル作成が行われていない。そこで、本研究では位置に関連する変数とそこから得られる情報を入力とした SxA モデルの作成を行った。

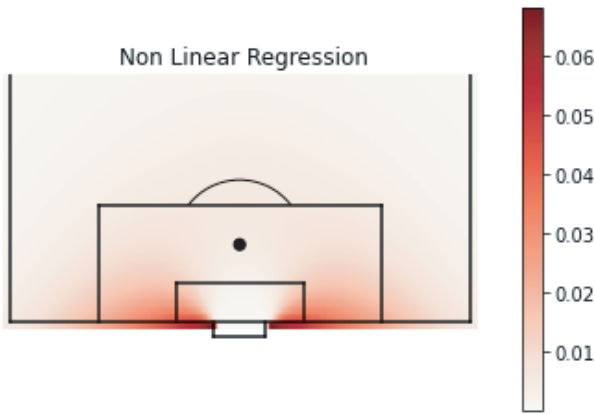


図 5 SxA モデルの出力を可視化した図

Fig. 5 Figure of visualizing output of SxA model.

本研究のアシストのモデル化は、非線形ロジスティック回帰によって行った。定義した関数を以下の式 (2) に示す。

$$P(X, Y, G) = \frac{1}{1 + \exp(Z)}$$

$$\text{where } Z = \alpha_0 + \alpha_1 X + \alpha_2 Y + \alpha_3 G + \alpha_4 XY + \alpha_5 X^2 + \alpha_6 Y^2 \quad (2)$$

この式における可変パラメータは各項の係数 α_i ($i = 0, \dots, 6$)、損失関数は交差クロスエントロピー、学習方法は L-BFGS 法を用いた。Pappalardo ら [13] によって作成されたデータのうち攻撃に関係のあるラベル (Duel, Others on the ball, Pass, Shot) を有する 2,844,909 のプレーデータを実験データとして用いた。説明変数はピッチ上の位置の 2 変数 X, Y 、その位置から算出される xG を示す G 、非線形対応のための XY, X^2, Y^2 の 6 つとした。次数を 2 次までとしたのは、表現の自由度を上げつつ、過学習にもならないようにするためである。また、 G に関する 2 次の項がないのは、 G が X, Y を用いて算出されているためである。1 次の項である X, Y, G には標準化を施した。

学習により得られた式を以下の式 (3) に示す。

$$P(X, Y, G) = \frac{1}{1 + \exp(Z)}$$

$$\text{where } Z = -8.6711 - 1.8549X - 0.43116Y - 0.25692G + 0.19006XY + 0.25621X^2 - 0.14462Y^2 \quad (3)$$

このとき、 $P(X, Y, G)$ がアシストの可能性、 X と Y が先述したピッチ上の位置、 G がその位置の xG を表す。また、学習結果をピッチに反映した画像は図 5 のとおりである。

本研究では SxA モデルとして式 (3) を用いることとした。

3.1.4 Pitch Risk モデル

Pitch Risk モデルは、その局面の直後にその地点にボールがたどり着いたときに 2 プレー以内にゴールになる確率 (Pitch Risk) を可視化するモデルである。よって、出力は Pitch Risk を可視化した図である。入力 Pitch Control

モデル、xG モデル、SxA モデルの出力と局面のトラッキングデータであり、前者は Pitch Risk の算出に必要な変数、後者は局面の出力に用いられる変数である。

Pitch Risk の算出には確率論の考え方をを用いている。Pitch Control モデル、xG モデル、SxA モデルの出力はそれぞれ確率を表していることが理由である。Pitch Risk の算出の式は以下の式 (4) のとおりである。

$$\text{Pitch Risk} = \text{Pitch Control} \times (xG + SxA \times \alpha) \quad (4)$$

ここで、 α は SxA の重みである。右辺の第 1 項である Pitch Control は該当のエリアでボールを保有できる可能性を表しており、第 2 項は該当のエリアで行われたプレーがゴールにつながる確率を表している。右辺の項は独立の関係にあると考え、積をとることで Pitch Risk を算出することとした。SxA の重みは、xG と SxA で確率空間が違うことが理由で採用した。xG がすべてのシュートが対象であるのに対して、SxA はすべてのプレーが対象である。アシストは意図せず遂行されることが多いため、対象をすべてのプレーとした。

算出した Pitch Risk の可視化には、LaurieOnTracking [9] で提供されている Pitch Control の図示に関連する関数を用いた。入力の Pitch Control の値の部分 Pitch Risk の値を用いることで Pitch Risk の可視化を行った。

3.2 ポジショニング修正システム

このモデルは、指定された局面に対して守備側の選手を対象に失点のリスクが低くなるようなポジショニングを提案するモデルである。前提として指定された局面に対するポジショニングの修正であることから、選手の位置を大幅に変えることは考えていない。このモデルは遺伝的アルゴリズムを用いて実現した。

遺伝的アルゴリズムのメリットとして、汎用性があげられる。サッカーのポジションは最適解を定めることが難しいため、最適解を明示的に示すことができない。このような問題にも対応可能な最適化手法として、遺伝的アルゴリズムが適しているといえる。加えて、仮にポジションに最適解がある場合、その周辺も解としてベターといえる。遺伝的アルゴリズムは最適解に近い解を導くことに長けており、この点においても本研究に適した手法だといえる。また、遺伝的アルゴリズムと同じ母集団ベースの最適化アルゴリズムの手法として有名なアルゴリズムとして粒子群最適化 (PSO) があるが、サッカーの文脈では PSO よりも遺伝的アルゴリズムを利用した文献が多く見つかったことから [14], [15]、今回は遺伝的アルゴリズムを採用する。

入力は局面のトラッキングデータ、出力は提案する守備側の選手 11 人における各選手の動く距離 (m) と方向 (rad) である。入力であるトラッキングデータは、3.1 節での説明と同様に、指定した局面における選手とボールの位

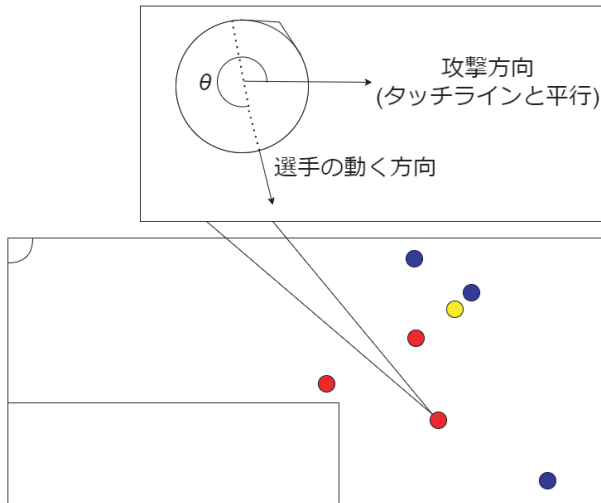


図 6 ポジショニング修正システムの出力における選手の向き. 青い点: 攻撃側選手, 赤い点: 守備側選手, 黄色い点: ボール

Fig. 6 Direction of the player in the output of a model to output the modification of positioning in a phase. blue dots: attacking players, red dots: defending players, yellow dots: ball.

置を示す 46 の数値 ((11(選手) × 2(チーム) + 1(ボール)) × 2(次元)), 選手の速度と加速度を示す 88 の数値 (11(選手) × 2(チーム) × 2(次元) × 2(種類)), 計 134 の数値からなるデータである. また, 出力における各選手の方向の値は, 図 6 の示すとおり, タッチラインと水平な攻撃方向の線を基準にして選手の動く方向のなす角とした.

遺伝的アルゴリズムによって求める対象は, 守備側の選手 11 人における各選手の動く距離 (m) と方向 (rad), 計 22 の変数である. 先述のとおり, 守備側の選手の位置は大幅に変えることを考慮していないため, 距離 r の範囲は $0 \leq r < 2$ とした. また, 方向 θ については $0 \leq \theta < 2\pi$ とした.

以下では, モデルの詳細について説明する.

3.2.1 評価関数

評価関数には, ポジショニング評価システムを利用する. このモデルの出力は, 入力された局面から, ピッチを $1\text{m} \times 1\text{m}$ に分割した各グリッドにボールが移動したときに 2 プレー以内にゴールする確率を可視化した図となっているが, その過程で各位置における Pitch Risk の算出を行っている. よって, ポジショニング評価システムを用いることで, その局面におけるピッチ上の各地点における失点のリスクを数値化できている. 各地点における Pitch Risk の合計値は局面における失点のリスクを示すことから, ポジショニング修正システムを評価値に用いた. 評価値の算出は, 以下の 2 段階で行われる. まず, 局面における守備側の選手の位置に評価対象である 22 変数を反映することで選手の移動を行う. 続いて, 選手の位置から Pitch Risk の算出を行い, 各セルの合計値を算出する. この評価値は失点のリスクを示していることから, 守備側のポジショニ

ング改善を目指す本研究では低いほうが優れていることになる.

3.2.2 選択

選択にはルーレット選択を採用した. ルーレット選択は, 評価値に応じて個体が選択される確率を設定する手法である. ある個体の評価値をその世代におけるすべての個体の評価値の和で割った値がその個体が選択される確率であり, 評価値が高いほどその個体は選択される可能性が高くなるため, 次の世代に残りやすくなる. 本研究における評価関数は値が低いほうが優れているため, そのままルーレット選択に適用すると正反対の結果が生じることになる. そのため, 選択の際に用いる場合は逆数をとることで最大化問題を扱う形で対応する. また, 評価値の差分が個体の選択される確率に及ぼす影響を大きくするために, 評価値の逆数に対するその世代の最小値との差を確率の設定に用いた. 以上をふまえた個体 i が選択される確率 p_i の算出式を以下の式 (5) に示す.

$$p_i = \frac{\frac{1}{e_i - e_{min}}}{\sum \frac{1}{e_n - e_{min}}} \quad (5)$$

このとき, e_i は個体 i の評価値, e_{min} は世代の評価値で最小の値である.

3.2.3 交叉

個体は 22 変数で構成されているが, 各選手の動きを示す 2 つの変数 (距離と方向) を 1 つのセットとして, 11 セットからなることを考える. まず, 選択により 2 つの個体を選ぶ. 続いて, 11 セットのうちランダムに選ばれた 5 セットを該当するセットごとに個体間で入れ替える. これにより, 2 つの新しい個体が生成され, 交叉が完了する.

3.2.4 突然変異

個体は 22 変数で構成されているが, 各選手の動きを示す 2 つの変数 (距離と方向) を 1 つのセットとして, 11 セットからなることを考える. 突然変異は 1 つの個体に対して行われる. まず, 11 セットのうちランダムに 15 セット選ぶ. そして, 選ばれたセットについて, 3.2 節で述べた制約 ($0 \leq r < 2, 0 \leq \theta < 2\pi$) を満たすようにランダムに生成されたセットに置き換える. これにより, 1 つの新しい個体が生成され, 突然変異が完了する.

4. 評価

実験データには, Metrica Sports Sample Data [16] で提供されているデータを用いた. このデータは, 過去のサッカーの戦術分析に関する研究 [3] などでも用いられているオープンデータであり, 実際に行われたプロの試合, 3 試合分についてのトラッキングデータが含まれている. このうちの 1 試合から適当な時間帯のシュートにつながる局面を 5 つ抽出して, 実験データとして用いた.

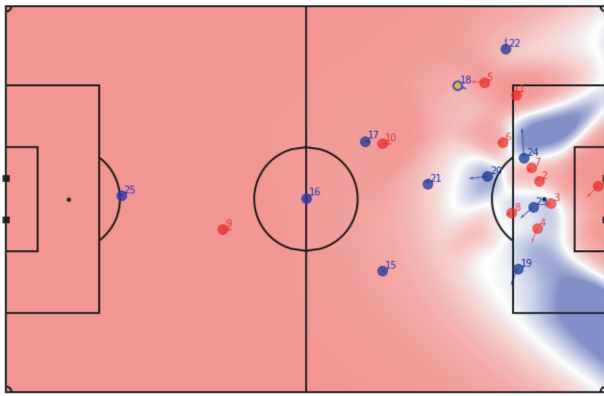


図 7 ポジショニング評価システムの出力例. 青い点：攻撃側選手, 赤い点：守備側選手, 黄色い点：ボール. 青いエリアほど失点のリスクが高いことを示す

Fig. 7 Example of a model for quantifying the risk of conceding a goal. blue dots: attacking players, red dots: defending players, yellow dots: ball; blue areas indicate high Pitch Risk.

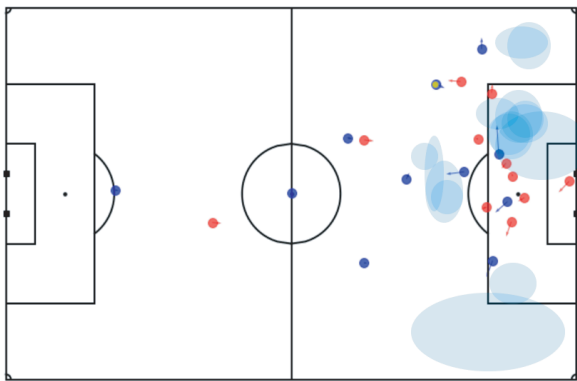


図 8 主観評価実験の結果の例

Fig. 8 The result of the subject evaluation experiment.

4.1 ポジショニング評価システム

式 (4) における変数 α について, アシストの影響が最も適当に反映されている値として, $\alpha = 10$ と設定した.

このモデルの出力例を図 7 に示す.

また, ポジショニング評価システムの評価を行うために, 主観評価実験を実施した. 本章の最初に先述したとおり, Metrica Sports Sample Data [16] で提供されているトラッキングデータから 5 つのシュートにつながる局面を選び, サッカー経験者 7 人に対して, それぞれの局面について, 2, 3 プレー以内に得点につながりそうだと思うエリアを選択してもらった. エリアは複数回答可能とした.

アンケートの結果の一部を図 8 に示す. これは, 図 7 と同じ局面に対する回答結果である.

ポジショニング評価システムの出力の妥当性を評価するために, 主観評価実験の結果との類似度を確認する. 主観評価実験の結果との類似度が高いということは, サッカー経験者が考える危険なエリアを表現できているということになり, ポジショニングの数値化・可視化の実現に近づけ

表 1 各モデルの出力と主観評価実験の JS divergence

Table 1 JS divergence between the result of the subject evaluation experiment and the output of each model.

モデル	ポジショニング 評価システム	Pitch Control
JS-divergence (\downarrow)	0.14	118.18

ているといえる.

評価の基準には, JS divergence を採用した. JS divergence は, 確率分布の類似度を示す指標である. それぞれのモデルは尺度が異なることから, 平均二乗誤差や平均絶対誤差のように単純に値の差を比べることは失点リスクの類似度を測ることにつながらない. そこで, 分布の類似度を求めることで, リスクが高いと判断される部分の一致度を求めることができ, ポジショニング評価システムの評価につながると考えた.

評価対象はポジショニング評価システム, Pitch Control である. 結果を表 1 に示す. この値は, 主観評価実験で扱った 5 つの局面における各モデルの出力と主観評価実験の結果におけるボールがあるほうのピッチ半分を対象とした JS divergence の平均値である. ポジショニング評価システムは, Pitch Control よりも主観評価実験との類似度が高いことから, ポジショニング評価システムは失点のリスクを再現できているといえる.

4.2 ポジショニング修正システム

実験データには, 本章の最初に先述したとおり, Metrica Sports Sample Data [16] で提供されているトラッキングデータから選ばれた 5 つのシュートにつながる局面を用いた. これは, 主観評価実験に採用した 5 局面と同様のものである.

モデルの詳細について改めて説明する. 3.2 節で述べたとおり, 個体は守備側の選手 11 人における各選手の動く距離 r (m, $0 \leq r < 2$) と方向 θ (rad, $0 \leq \theta < 2\pi$), 計 22 の変数の組である. また, 初期個体生成は一様乱数によるランダム生成で行われ, 1 世代の個体数は 100, 世代数は 100, 選択についてはルーレット選択, 交叉の確率は 0.3, 突然変異の確率は 0.1 とした. 加えて, 評価関数については, 3.1 節におけるポジショニング評価システムを採用した.

また, 比較実験としてランダム生成を行う. ポジショニング修正システムで生成した個体数と同じ数, すなわち $100 \times 100 = 10000$ の個体を生成し, 最も評価値が低い個体をランダム生成の出力として採用する.

結果を表 2 に示す. 表 2 は, 用意した各局面における, 元の局面の評価値, ポジショニング修正システムの出力を反映した場合の評価値の 5 回分の平均, ランダム生成の出力を反映した場合の評価値の 5 回分の平均を表している.

表 2 各モデルの出力を反映した局面ごとの評価値 (低いほうが優性)
Table 2 Evaluation value for each phase reflecting the output of each model (lower value is superior).

	局面 1	局面 2	局面 3	局面 4	局面 5
元の局面	89.94	101.79	52.68	63.72	104.01
提案システム	78.65	80.62	44.20	31.79	85.99
ランダム生成 (最良/1 万個体)	80.86	82.48	46.08	32.38	90.97

表 3 t 検定の結果

Table 3 The result of t-test.

自由度	t 値	境界値
4	3.19	2.13

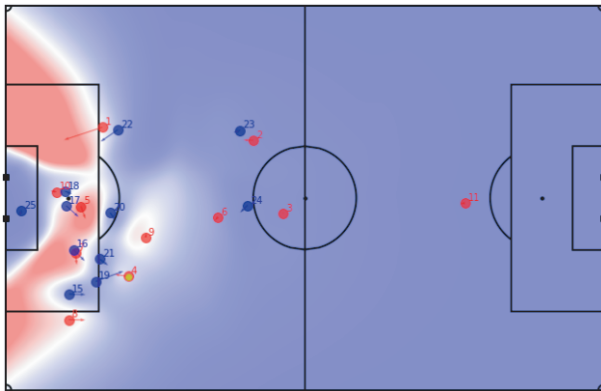


図 9 局面 2 における元の局面の評価を反映した図

Fig. 9 Figure of original phase.

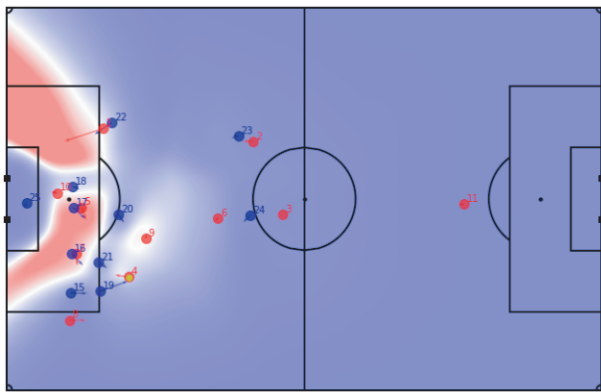


図 10 局面 2 におけるポジショニング修正システムの出力を反映した場合の評価を反映した図

Fig. 10 Figure of the phase reflecting the output of the proposed model.

また、ポジショニング修正システムがランダム生成より優れているかを確認するために表 2 の値を用いた t 検定を行った。帰無仮説を「ポジショニング修正システムの評価値とランダム生成の評価値に差がない」とし、有意水準 5% として行った片側検定の結果を表 3 に示す。t 値が境界値を上回っていることから帰無仮説は棄却され、ポジショニング修正システムの評価値とランダム生成の評価値に差が

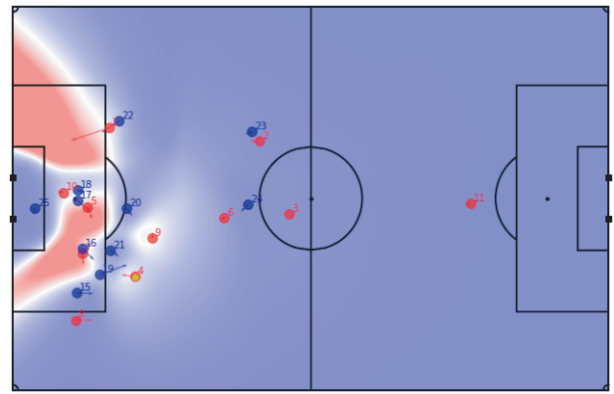


図 11 局面 2 におけるランダム生成の出力を反映した場合の評価を反映した図

Fig. 11 Figure of the phase reflecting the output of random generation.

ある、すなわちポジショニング修正システムのほうが優れていることが示された。

加えて、結果を可視化した図として、図 9、図 10、図 11 を示す。図 9、図 10、図 11 はそれぞれ、局面 2 における、元の局面の評価を反映した図、ポジショニング修正システムの出力を反映した場合の評価を反映した図、ランダム生成の出力を反映した場合の評価を反映した図である。

5. 考察

5.1 実験の妥当性について

5.1.1 実験データ

各選手のプレーを画像から特定する (シュートをしているのかパスをしているのかなど) のように正解データが容易に取得できるタスクでは、機械的に多くのシーンについて評価を行い、平均的な精度を出すことが可能である。しかし、ゾーンごとの失点リスクを特定して可視化するタスクや各選手のより良い行動を提案するタスクのように、正解データを用意することが困難なタスクの場合、人が時間をかけてその判断を行う必要がある。そのため、評価に用いるシーンの数を限定する必要が生じる。サッカーに関する研究分野では、5 個程度のシーンに対して評価を行うことも多く行われている。

たとえば、Seebacher らは試合中のチーム状況を検索・比較するための新しいビジュアル分析デザインを提案している。彼らは、専門家たちに 3 つのシーンを見せて提案手法の評価を行った [17]。Abe らは、強化学習を用いてサッカーエージェントの協調行動を促す研究を行い、2 つまたは 3 つのシーンについてエージェントの振舞いを解析して評価した [18], [19]。Yao らは、選手の動きの軌跡を可視化する方法を提案した。彼らの研究はサッカーに限定したものではないが、評価は 7 つのシーンについて行われた [20]。Allegre らは、サッカーにおける空間所有モデルの可視化手法を提案している。評価は 2 つのゴールシーンに対して

行われた [21].

このように、人が判断を行う必要があるサッカータスクにおいては、2-7 個程度のシーンに対して評価が行われることがある。本論文ではこれらの既存研究と同様に、5 個のシーンに対する評価を行った。これら 5 個のシーンはまったく別の時間帯から取得されたものであり、状況として独立している。また、4.2 節に記述したとおり、5 個のシーンに対する評価結果では統計的に優位な差が得られている。しかし、より多くのシーンで評価を行うことが望ましいため、多くのシーンに対する評価については課題である。

5.1.2 評価方法

本研究の評価は、与えられた状況からの改善を指標としたが、理論値との比較という方法も考えられる。しかし、サッカーというスポーツの性質上、真の理想状況すなわち理論値を定義することは大変困難なことである。実際、過去の論文でも状況改善による評価が行われており、理論値との比較という形での評価は行われていない。

たとえば、Marutens らは論文内で選手の影響力を表現する手法を提案している。この手法の評価は既存手法との比較と設定した評価値によって行われており、理論値との比較は成されていない [22].

このように、理想的な状況の定義が難しいサッカーという研究分野においては、理論値との比較ではない形での評価が行われている。本論文ではこれらの既存研究と同様に、理論値を用いない形での評価を行った。

5.2 ポジショニング評価システムに関して

ポジショニング評価システムは、Pitch Control から得られるピッチの支配状況に関する情報と、xG と SxA から得られる得点につながる確率を基に局面における失点の可能性を算出する仕様である。前者は局面の状況を表し、後者は局面によらずゴールにつながる可能性を表している。失点の可能性という単純に数値化することが難しい情報を、局面の情報と一般的な情報という形でタスクを分割することによって表現の質を向上させることができたと考える。

ポジショニング評価システムの出力と主観評価実験の結果を比較すると、ポジショニング評価システムで高い値が出力されているエリアと主観評価実験における回答が集中しているエリアが一致する部分があった。ゴール前の攻撃チームが走りこんでいるエリア、アシストにつながりそうなゴール脇のエリアでその傾向があることが分かる。一方で、図 7 と図 8 においてはペナルティスポット付近が該当する。選手の動きや向きとは逆方向のエリアの出力は、アンケートでは失点のリスクが高いとされていなかった。

ポジショニング評価システムを構成する各モデルの精度については改善の余地があると考えられる。Pitch Control モデルは、運動学を基に算出されるその地点への到達可能性によって算出されることから、選手が行わないであろう非

現実的な行動も考慮されている。たとえば、進行方向を急激に変更する動きがあげられる。特に、背後への動きは明確な意図をもって行わないと不自然な動きであることから、予測の時点では非現実的になると考えられる。モデルの出力にも、選手の動きや向きとは逆方向のエリアの出力が主観評価実験の結果と一致していないという点でこの傾向が見受けられる。これに対しては、実データから得られる選手の動きの傾向などを含めることで、より現実的な表現になり精度を上げることが可能だと考える。文献 [22] では、データドリブンな選手の動きのモデル化が紹介されている。これは、選手の過去の動きから可動範囲を学習することでモデル化している。このようなモデルを Pitch Control モデルとして導入することで、より現実的なピッチの支配状況の表現が可能になる。SxA モデルは、過去の試合データを基にアシストの可能性を算出するモデルだが、データの偏りが原因として考えられる。精度が不十分なエリアが存在する。その一例がゴールライン付近で、SxA の出力が高いエリアが集中している。ゴールライン付近はその周辺と比べてプレーやアシストの数が少ない。この差を埋めるためにモデルが予測した結果、それまでの傾向でゴールに近いほうが SxA の値が高いことからゴールライン付近もその値が高くなったと考えられる。サッカーにおけるメタ知識を用いて新たなパラメータを追加することで、SxA モデルの精度をさらに高めることが可能だと考える。ボールホルダがゴールライン際でプレーする際、ゴール方向に向いていれば攻撃側はゴール前にクロスを入れられる可能性が高く、必然的にアシストの可能性も高くなる。一方、ボールホルダがゴールに背を向けるような状況の場合、守備側はプレッシャをかけることで外側に追い出しやすくなるため、一転してアシストの可能性は低くなる。このように、サッカーで前提となるような知識から必要なパラメータを導き出し、それをモデルに加えることが SxA のモデルの改善への近道だと考える。

また、ポジショニング評価システムで前提としているプレーヤーは、一般的なチームの一般的なプレーヤーである。本研究の設定では選手の最大スピードは選手によらず一定、「ドリブルをしがち」や「リスクの低いパスを選択しやすい」といった選手のプレー選択の傾向などは反映されていない。すなわち、チームやプレーヤーの個性をいっさい考慮していない。実際の試合では、先にあげた要素以外のチームのプレースタイルや体力の状態なども局面に影響を及ぼしうる。スピードや体力などの比較的数値的に表しやすい要素については、選手のパラメータとして設定できるようにすることで、モデルに反映することが可能である。数値化については、実際の試合で記録されるものを用いることが可能だと考える。スピードについては、試合で記録された最高速度、体力についてはスプリント数や走行距離の割合から概算することなどが考えられる。選手のプレーの選

択傾向やチームのプレースタイルといった明確に数値化することが難しい要素については、新しい指標を作ることによって解決することを提案する。実際の試合データから傾向などをディープラーニングなどを用いた特徴抽出によってデータとして扱える状態にすることで新しい指標として用いることが可能だと考える。

5.3 ポジショニング修正システムに関して

表 2 を見ると、ポジショニング修正システムの出力は元の局面から状況を改善する結果となっており、ランダム生成と比較しても優れていることが分かる。22 変数の最適化という複雑なタスクを遺伝的アルゴリズムを用いることによって実現できているといえる。

各パラメータを変化させた場合の効果について述べる。世代数は、増やしたとしても 100 世代前後で生じた個体が最良であることが多かったため、今回の設定で問題ないと考えられる。交叉、突然変異の起こる確率について、増やすことによって学習の進みが鈍化する傾向があった。これは、その世代の最良個体が残りにくくなるのが原因だと考えられる。交叉、突然変異は既存の個体に変化を加える操作であるため、次の世代にその世代の最良個体が残る確率が低くなる。また、一般的に遺伝的アルゴリズムで用いられる個体の遺伝子は 2 値で表現されることが多いのに対して、本研究における個体の遺伝子は連続値から構成されており、交叉と突然変異によるランダムな値の変換による影響が顕著に表れる。そのため、もともと適応度の高い個体が適応度の低い個体になることが比較的起こりやすいため最良個体が残りにくくなると考えられる。

図 10 と図 11 を比較すると、提案手法-修正の出力の傾向として、攻撃側の選手へのアプローチと守備側の選手の連動が見受けられる。前者は、攻撃側の選手に対して守備側の選手が近づくことで、攻撃側の選手のプレーの幅を狭めることができ、結果として失点のリスクを低減させることにつながっている。後者は、守備側のある選手が動くことで生じうるスペースを別の守備側の選手が埋めるといった動き、ディフェンスラインの選手がラインを揃えることでカバーすべきスペースを効率的に埋めることなどが該当する。いずれについても、サッカーをするうえで常識的な内容であるが、正解を探すのが難しい要素である。ピッチの状況を反映させた評価関数に基づいて学習を行うことで、そのようなことも表現することができたと考える。

改善すべき点として、連続値に対する対応をあげる。先ほども述べたとおり、連続値に対するランダムな値の変換は影響が大きく、学習の効率に影響を与えている可能性が考えられる。CMA-ES [23] のように分布に基づいて値の生成を行うことで対応できると考える。

6. おわりに

本研究では、ポジショニングを定量化・可視化することを目的として、Pitch Control, xG, SxA を用いたモデルを提案した。さらに、このモデルを評価関数として用いた遺伝的アルゴリズムによって、アマチュアの選手の試合後の振り返りなどを想定した、ポジショニングの修正を行うモデルの作成を行った。その結果、ピッチの支配状況とゴールやアシストの情報を組み合わせることで、より適切に危険なエリアを表現することが可能なことが分かった。また、それを用いてポジショニングの修正を実現することができた。一方で、本研究で提案したモデルから考えられる失点のリスクが高いエリアとサッカー経験者が考える失点のリスクが高いエリアには差異があることが分かった。モデルの改善や精度の向上、実用的な実装を行うことで、プレーヤやマネージャのポジショニングの理解を深める一助になることやポジショニングのシミュレーションがエンターテインメントとして観戦を盛り上げることにつながる可能性がある。ひいては、サッカーというスポーツの魅力を上向させることにつながると考えられる。

謝辞 本研究は JSPS 科研費 JP21H03496, JP22K12157 の助成を受けたものです。

参考文献

- [1] Suzuki, G., Takahashi, S., Ogawa, T. and Haseyama, M.: Team tactics estimation in soccer videos based on a deep extreme learning machine and characteristics of the tactics, *Ieee Access*, Vol.7, pp.153238–153248 (2019).
- [2] Decroos, T., Davis, J., et al.: Valuing on-the-ball actions in soccer: A critical comparison of xt and vaep, *Proc. AAAI-20 Workshop on Artificial Intelligence in Team Sports*, pp.1–8, AI in Team Sports Organising Committee (2020).
- [3] Ötting, M. and Karlis, D.: Football tracking data: a copula-based hidden markov model for classification of tactics in football, *Annals of Operations Research*, pp.1–17 (2022).
- [4] Tuyls, K., Omidshafiei, S., Muller, P., Wang, Z., Connor, J., Hennes, D., Graham, I., Spearman, W., Waskett, T., Steel, D., et al.: Game plan: What ai can do for football, and what football can do for ai, *Journal of Artificial Intelligence Research*, Vol.71, pp.41–88 (2021).
- [5] Alguacil, F.J.P.: *Modelling the collective movement of football players* (2019).
- [6] Lu, Z., Yang, J. and Sumpter, D.: *Statistical and visualization methods on evaluating players in football* (2020).
- [7] The Analyst: What are expected goals (xg)? (2021), available from (<https://theanalyst.com/eu/2021/07/what-are-expected-goals-xg/>).
- [8] The Analyst: What are expected assists (xa)? (2021), available from (<https://theanalyst.com/eu/2021/03/what-are-expected-assists-xa/>).
- [9] Friends of Tracking-Data-FoTD: Laurieontracking (2021), available from (<https://github.com/Friends-of-Tracking-Data-FoTD/LaurieOnTracking>).
- [10] Theiner, J., Gritz, W., Müller-Budack, E., Rein, R.,

- Memmert, D. and Ewerth, R.: Extraction of positional player data from broadcast soccer videos, *Proc. IEEE/CVF Winter Conference on Applications of Computer Vision*, pp.823-833 (2022).
- [11] Zheng, B.: Soccer player video target tracking based on deep learning, *Mobile Information Systems*, Vol.2022 (2022).
- [12] Akan, S. and Varlı, S.: Use of deep learning in soccer videos analysis: Survey, *Multimedia Systems*, pp.1-19 (2022).
- [13] Pappalardo, L., Cintia, P., Rossi, A., Massucco, E., Ferragina, P., Pedreschi, D. and Giannotti, F.: A public data set of spatio-temporal match events in soccer competitions, *Scientific Data*, Vol.6, No.1, pp.1-15 (2019).
- [14] Chen, X. and Gao, P.: Path planning and control of soccer robot based on genetic algorithm, *Journal of Ambient Intelligence and Humanized Computing*, Vol.11, pp.6177-6186 (2020).
- [15] Xu, X., Xie, J. and Xie, K.: Path planning and obstacle-avoidance for soccer robot based on artificial potential field and genetic algorithm, *2006 6th World Congress on Intelligent Control and Automation*, Vol.1, pp.3494-3498, IEEE (2006).
- [16] Metrica Sports: sample-data (2021), available from (<https://github.com/metrica-sports/sample-data>).
- [17] Seebacher, D., Polk, T., Janetzko, H., Keim, D., Schreck, T. and Stein, M.: Investigating the sketchplan: A novel way of identifying tactical behavior in massive soccer datasets, *IEEE Trans. Visualization and Computer Graphics* (2021).
- [18] Abe, T., Orihara, R., Sei, Y., Tahara, Y. and Ohsuga, A.: Acquisition of cooperative behavior in a soccer task using reward shaping, *2021 the 5th International Conference on Innovation in Artificial Intelligence*, pp.145-150 (2021).
- [19] Abe, T., Orihara, R., Sei, Y., Tahara, Y. and Ohsuga, A.: Step-by-step acquisition of cooperative behavior in soccer task, *Journal of Advances in Information Technology*, Vol.13, No.2, pp.147-154 (2022).
- [20] Yao, L., Bezerianos, A., Vuillemot, R. and Isenberg, P.: Visualization in motion: A research agenda and two evaluations, *IEEE Trans. Visualization and Computer Graphics* (2022).
- [21] Allegre, J. and Vuillemot, R.: Visualizing and analyzing disputed areas in soccer, *Visualization in Data Science (VDS)*, pp.52-60, IEEE (2020).
- [22] Martens, F., Dick, U. and Brefeld, U.: Space and control in soccer, *Frontiers in Sports and Active Living*, p.175 (2021).
- [23] CyberAgentAILab: cmaes (2023), available from (<https://github.com/CyberAgentAILab/cmaes>).



神宮司 祐哉

1998年生。2023年電気通信大学大学院情報理工学研究科情報学専攻修士課程修了。サッカーに関する研究に従事。



清雄一

1981年生。2009年東京大学大学院情報理工学系研究科博士後期課程修了。同年(株)三菱総合研究所入社。2013年より電気通信大学。現在、同大学大学院情報理工学研究科教授。博士(情報理工学)。エージェント、プライバシー保護技術等の研究に従事。2016年度土木学会水工学論文賞、情報処理学会論文賞受賞。電子情報通信学会、日本ソフトウェア科学会、IEEE Computer Society 各会員。



田原 康之

1966年生。1991年東京大学大学院理学系研究科数学専攻修士課程修了。同年(株)東芝入社。1993~1996年情報処理振興事業協会に外向。1996~1997年英国 City 大学客員研究員。1997~1998年英国 Imperial College 客員研究員。2003年国立情報学研究所着任。2008年より電気通信大学准教授。博士(情報科学)(早稲田大学)。エージェント技術、およびソフトウェア工学等の研究に従事。日本ソフトウェア科学会会員。



大須賀 昭彦 (正会員)

1958年生。1981年上智大学理工学部数学科卒業。同年(株)東芝入社。同社研究開発センター、ソフトウェア技術センター等に所属。1985~1989年(財)新世代コンピュータ技術開発機構(ICOT)外向。2007年より電気通信大学。現在、同大学大学院情報理工学研究科教授。2017年より同大学大学院情報システム学研究科研究科長兼任。2012年より国立情報学研究所客員教授兼任。工学博士(早稲田大学)。ソフトウェア工学、エージェント、人工知能の研究に従事。1986年度および2016年度情報処理学会論文賞、2013年度人工知能学会研究会優秀賞、2014年度同学会功労賞、2018年度電子情報通信学会 ISS 活動功労賞受賞。IEEE Computer Society Japan Chapter Chair、人工知能学会理事、日本ソフトウェア科学会理事、同学会監事等を歴任。電子情報通信学会、人工知能学会、日本ソフトウェア科学会、電気学会、IEEE Computer Society 各会員。本会フェロー。