

誰でも手軽に使える DEA ウェブアプリケーションの構築

小畑 経史

事業体の経営効率性を分析するために考案されたデータ包絡分析法 (DEA) を実行するには、線形計画問題を何度も解く必要がある。そのため、線形計画問題を解く手段をもたないものにとっては必ずしも手軽に利用できるとは言いがたい。そこで、インターネットを通じてウェブブラウザ上で動作するウェブアプリケーションとして DEA アプリケーションを構築することにより、誰もが手軽に利用できるようになると考えられる。本稿では DEA ウェブアプリケーションの構築と、それを利用したサッカーチームのパフォーマンス評価の例を紹介する。

キーワード：DEA, ウェブアプリケーション, HTML5, サッカーデータ, パフォーマンス評価

1. はじめに

Charnes et al. [1] によって考案されたデータ包絡分析法 (data envelopment analysis; DEA) は、同種の事業体の効率性を評価するための優れた手法である。DEA では、評価対象である事業体 (decision making unit; DMU) の効率性を図る指標である効率値や、改善の指標である入出力スラックを求めるために、線形計画問題を DMU の個数の倍の回数 (たとえば後述のサッカーデータ分析の例では DMU が 24 あるため 48 回) 解く必要がある。

ここで、DEA に触れたばかりの初学者が実際に DEA を試してみたいと考えたケースを想定してみよう。このようなものにとって、線形計画問題を解くためのツールが身近にあるとは限らない。一つの方法として、Microsoft 社の表計算ソフト Excel のアドインツールであるソルバー機能がある。これを利用することで線形計画問題を解くことができる。Excel は一般のユーザーに広く普及しており、比較的誰もが利用することができるといえよう。Cook and Zhu [2] では、ソルバー機能を利用した DEA の計算方法が詳しく紹介されている。しかし、ソルバー機能の使い方の理解に加えて、DEA ではそれを何度も実行する必要があることから、誰もが手軽に利用できる方法とは言いがたい。

ところで、近年、誰でも手軽に利用できるアプリケーションの形態として、ウェブアプリケーションが急速に広がってきた。ウェブアプリケーションとは、大ま

かに言うと、ウェブブラウザ上で動作するアプリケーションのことである。現在では、パーソナルコンピュータにはネットワーク通信機能とウェブブラウザが必須となっており、ウェブアプリケーションの動作環境が自ずと整っている。さらに、ネットワーク通信機能とウェブブラウザを備えたスマートフォンやタブレットが急速に普及しており、これらの機器でもウェブアプリケーションを実行することができる。すなわち、今やコンピュータをもたないものでもウェブアプリケーションの動作環境を手にしているといえる。そこで、DEA を実行するウェブアプリケーションを構築することにより、誰でも手軽に DEA を利用できるようになると考えられる。

2. ウェブアプリケーションと HTML5

2.1 ウェブアプリケーション

WWW (world wide web) とは、HTML (hypertext markup language) 形式で記述された情報 (ウェブページ) をインターネットを通じてやりとりする仕組みである。情報の提供側である装置やソフトウェアをウェブサーバー、受け取った情報を整形して表示するためのソフトウェアをウェブブラウザと呼ぶ。情報の提供側 (サーバーサイド) と受け取り側 (クライアントサイド) からなるシステムであるため、サーバー・クライアント型のシステムと呼ばれる。WWW では、ユーザーからの働きかけに反応する機能として、関連したほかの情報へ推移するリンク機能を当初から備えていたが、基本的にはあらかじめ用意された情報を静的に伝えるものであった。

しかし、次第にユーザーからのより自由な働きかけに動的に返答する要求が生まれてきた。そのような要求

おばた つねし
大分大学理工学部
〒 870-1192 大分県大分市大字旦野原 700
t-obata@oita-u.ac.jp

に対応するために、まず考えられたのが CGI (common gateway interface) という仕組みである。CGI では、ブラウザからの要求を受け取ったサーバソフトウェアが、別のプログラムに処理を依頼する。依頼を受け取ったプログラムが、要求に応じた返答となる HTML データを生成してサーバソフトウェアへ渡し、サーバソフトウェアがブラウザへその HTML データを返す。当時は「ウェブアプリケーション」という用語は一般的ではなかったが、CGI がウェブアプリケーションの初期の形であったとあってよいだろう。小畑と白石 [3] は CGI を使用して、オペレーションズ・リサーチの一手法である階層化意思決定法 (analytic hierarchy process; AHP) のウェブアプリケーション化を試みている。

CGI では、処理の中心はサーバサイドで行われ、ブラウザが受け取るのはあくまでも HTML 形式のデータである。そのため、クライアントサイドには高度な処理能力が必要ない、標準的な HTML データの表示さえできればブラウザの種類やプラットフォームを選ばない、といった利点がある。その一方で、処理の度にサーバとのデータのやり取りが生じるためオフラインでは動作せず、サーバに負荷が集中するためサーバ側に処理能力が要求される。また、表現力の面ではあまり高いものとはいえない。

続いて、より高度な表現を可能にするために、HTML の枠組みにとらわれずに処理を行うという発想から生まれたのが、Java アプレットや Flash である。これらはいずれも、ウェブページの一部にアプリケーションのための領域を取り、その中で動作する。動作の主体はクライアントサイドのコンピュータである。

Java アプレットはオブジェクト指向プログラミング言語である Java により記述され、VM (virtual machine) と呼ばれる仮想環境上で動作するため、一度コンパイルしたものは VM が用意されたプラットフォームであればどこでも動作する。高度な GUI (graphical user interface) を備えたアプリケーションを実現ことができ、一時は広く利用された。しかし、動作速度や VM の互換性の問題、セキュリティに対する不安などの点から徐々に利用が狭まり、現在ではあまり使用されていない。

Flash はもともとは一部のブラウザに用意されたプラグインという仕組みを利用したものであるが、多くのブラウザに広まり、主にゲームの分野で活用されてきた。現在でもブラウザ上での動画再生に使用されている。このころのウェブ技術の発展については佐

藤 [4] に詳しい。

その後、HTML そのものの表現力を上げることや、HTML に動きを加えるためのスクリプト言語の試みの中から生まれたのが、現在のウェブアプリケーションの主流といえる HTML5 技術である。

2.2 HTML5

HTML では文書の内容を色々な意味をもつタグで囲むことで、単なるテキスト以上の意味を付加することができる。たとえば、`<a>` というタグはほかの文書へのリンクを、`<h1></h1>` というタグは最上位の見出しを、`<i></i>` というタグは斜体での表示を意味する。当初の HTML では文書構造を意味するタグと表示の際の見栄えを意味するタグが混在していた。これらを分離することを目的として生まれたのが、CSS (cascading style sheet) である。CSS の登場により、文書の構造は HTML で、見栄えは CSS で記述するという役割分担がなされるようになった。最新のバージョンである HTML5 と CSS3 により、HTML 文書はこれまで以上の表現力を実現することができる。

アプリケーションとして動作するにはユーザーからの働きかけを受け取る仕組みが必要となるが、HTML にはフォームと呼ばれるユーザーからの入力を受け取る部品も用意されている。これにより HTML は入力インタフェースの役割も果たす。これと並行して、HTML 文書に動きを付加するために、文書に書き込めるスクリプト言語が開発された。このようなものの一つが JavaScript である。これら、文書構造と入力インタフェースを担当する HTML5、見栄えを担当する CSS3、動作を担当する JavaScript を組み合わせ、動的にユーザーの働きかけを処理し、結果を表示するウェブアプリケーション技術が、広義の HTML5 と呼ばれるものである (図 1)。

HTML5 アプリケーションはクライアントサイドで動作するため、当然ブラウザ側の対応が必要となる。Windows, macOS, Linux などのパーソナルコンピュータや、iOS, Android などのスマートフォン/タブレットに、現在広く普及している主要なウェブブラウザは、ほぼ HTML5 に対応しており、一部のタグの対応への違いや、見栄えの違いがあるものの、おおむね同じように動作すると考えてよい。

HTML5 アプリケーションがクライアントサイドで動作することによる利点として、オフラインで動作することが挙げられる。最初に一度はネットワーク経由でサーバに置かれたウェブアプリケーションにアクセスする必要があるが、その際に、関連したファイル

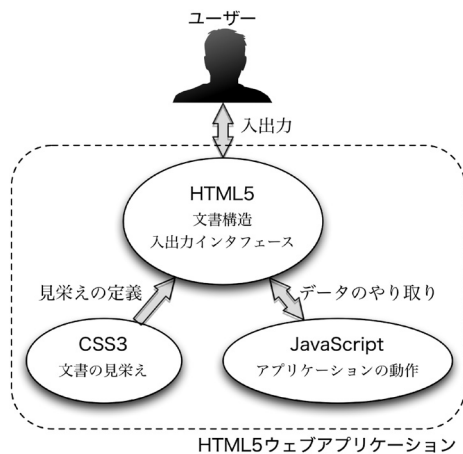


図 1 HTML5 の構成

式がキャッシュとしてローカルに保存される。その状態であれば、ネットワークから切断されていてもアプリケーションは動作する。あるいは、アプリケーションを構成するファイル一式をあらかじめ取得しておけば、ローカルでウェブブラウザから起動することで、ネットワーク接続なしにアプリケーションを使用することもできる。

3. DEA のウェブアプリケーション化

先に述べたとおり、HTML5 ではアプリケーションの動作を JavaScript で記述する。筆者が構築した DEA ウェブアプリケーションにおいても、DEA の計算部分は JavaScript が担当することになる。JavaScript はオブジェクト指向のスクリプト言語である。そのため、ほかのアプリケーションにも再利用が可能なクラスライブラリとしての利用を念頭に置いて、計算のための道具をいくつかのクラスの形で実装することにした。ここでは

- ベクトル
- 行列
- 線形計画問題
- DEA 問題

をそれぞれクラスとして定義した。線形計画問題の解法はシンプレックス法を使用している。

これらのほかに入力インタフェースと出力結果の表示を担当する HTML ファイルと CSS ファイル、HTML で入力された値を取得し前述のクラス群を使用して計算した結果を HTML へ返す動作を記述した JavaScript ファイルを作成することで、DEA ウェブアプリケーションを完成させ、「Web DE DEA」と名づけた。

Web DE DEA は次のように動作する。起動し、

表 1 入力用 CSV 形式

```

15,3,2
,住宅,家賃,パン,医師,年収
シアトル,586,581,1.45,4.49,46928
デンバー,475,558,0.97,2.79,42879
フィラデルフィア,201,600,1.5,3.64,43576
(以下省略)
  
```

Web DE DEA v.2.1
DEA (データ包絡分析法) を用いて DMU の効率性を分析する Web アプリケーションです。

データ入力
DMU の数: 15 | 入力数: 3 | 出力数: 2

	Inputs				Outputs	
	住宅	家賃	パン	医師	年収	
シアトル	586	581	1.45	4.49	46928	
デンバー	475	558	0.97	2.79	42879	
フィラデルフ	201	600	1.5	3.64	43576	
ミネアポリス	299	609	1.49	2.67	45673	
ローリー-ダ	318	613	0.99	4.94	40990	
セントルイス	265	558	0.89	3.4	39079	
シンシナティ	467	580	1.25	2.8	38455	
ワシントン	583	625	1.29	3.35	54291	
ピッツバーグ	347	535	0.99	3.66	34534	
グラス	296	650	1.5	1.96	41984	
アトランタ	600	740	1.19	2.23	43249	
ボルチモア	575	775	0.99	4.02	43291	
ボストン	351	888	1.09	5.69	46444	
ミルウォーキ	283	727	1.53	3.11	41841	
ナッシュビル	431	695	1.19	3.25	40221	

CSVファイルからデータを読み込む: (入力用 CSV ファイルのサンプルはこちら)

指向性: 入力指向 出力指向

規模の収束: CRS (CCR) モデル VRS (BCC) モデル IRS モデル DRS モデル

分析結果

図 2 データの入力

DMU の数、入力数、出力数を入力すると、入力欄が表示される。現れた入力欄に、DMU のラベル、入出力項目ラベル、各入出力値を入力する。これらはあらかじめ準備しておいた CSV (comma-separated values) ファイルから読み込むことも可能となっている。CSV ファイルの形式は、第 1 行に DMU 数、入力数、出力数を、第 2 行に入力項目ラベル、出力項目ラベルを、第 3 行以降に各 DMU のラベルと入出力値を記述するものとする。表 1 に、Cook and Zhu [2] p. 23 に示されたデータを CSV 形式に並べたものの一部を示す。このデータを入力した状態が図 2 である。

そして、入力指向/出力指向の別、モデル (CCR モデル、BCC モデル、IRS モデル、DRS モデルが選択可能) を選択し、「効率値を計算」ボタンを押すことで DEA 計算が実行される。計算の結果として、効率性、効率スコア、各入出力のスラック値、参照 DMU がページ下部に表の形で出力されるとともに、二次利用のた

分析結果	Efficiency	Score	住宅	家賃	パン	医師	年収	Ref. DMUs
	slack	slack	slack	slack	slack	slack	slack	
シアトル	**	1	0	0	0	0	0	シアトル
デンバー	**	1	0	0	0	0	0	デンバー
フィラデルフィア	**	1	0	0	0	0	0	フィラデルフィア
ミネアポリス		0.982	0	0	0.036	0.849	0	フィラデルフィア, ワシントン
ローリー・ダーラム	**	1	0	0	0	0	0	ローリー・ダーラム
セントルイス	**	1	0	0	0	0	0	セントルイス
シンシナティ		0.808	0	0	0	0	0	シアトル, フィラデルフィア, ローリー・ダーラム, ワシントン
ワシントン	**	1	0	0	0	0	0	ワシントン
ピッツバーグ		0.898	0	0	0	0	0	シアトル, フィラデルフィア, ローリー・ダーラム, ワシントン
ダラス		0.864	0	0	0	1.383	0	フィラデルフィア, セントルイス, ワシントン
アトランタ		0.822	14.197	45.583	0	0.584	0	デンバー
ボルチモア	**	1	0	0	0	0	0	ボルチモア
ポートランド	**	1	0	0	0	0	0	ポートランド
ミルウォーキー		0.812	0	5.617	0	0.444	0	フィラデルフィア, セントルイス
ナッシュビル		0.781	16.194	0	0	0	0	ローリー・ダーラム, セントルイス, ワシントン

** Efficient, * Weakly Efficient
 CSV format:
 効率, 住宅, パン, 医師, 年収
 シアトル:0.982,14.197,45.583,0,0.584
 デンバー:0.808,0,0,0,0
 フィラデルフィア:0.864,0,0,0,0
 ミネアポリス:0.898,0,0,0,0
 ローリー・ダーラム:0.812,0,5.617,0,0
 シンシナティ:0.822,14.197,45.583,0,0.584
 ワシントン:0.822,14.197,45.583,0,0.584

図3 分析結果の表示

めに CSV 形式でも出力される。先ほどのデータを入力指向 CCR モデルで分析した結果表示が図 3 である。

4. 利用例

ここからは、Web DE DEA を使用して、公開されたサッカーデータからスターティングメンバーの組み合わせごとのパフォーマンスを評価した例を示す。

DEA を使用したサッカーデータの分析は、大きく、選手を評価したもの ([5] など) とチームを評価したもの ([6] など) に分けられるが、ここではチームとしてのパフォーマンス評価を考えることにする。Hirotsu and Ueda [7] はチームとしての評価を選手の組み合わせによって行うことを提案したが、その組み合わせが実際の試合で実現したかは考慮していない。そこで、実際の試合で実現した選手の組み合わせを、その試合で発生したプレーのデータをもとに評価することを考える。

さて、サッカーに関するデータはいろいろなところで収集されており、日本のプロサッカーリーグである J リーグに関するデータが、データスタジアム社が運営するウェブサイト、フットボールラボ¹で公開されている。このサイトでは、J 1, J 2, J 3 に所属するチーム/選手の 2012 年シーズン以降のデータが閲覧できる。試合ごとのチームのデータとして、表 2 に挙げたスタッツ (プレーに関する統計数値) が掲載されている (図 4)。

あるチーム内で、どのような選手の組み合わせが高いパフォーマンスを発揮したかを評価したい、という

¹ <http://www.football-lab.jp>

表 2 公開されているチームスタッツ

シュート数	シュート成功率
枠内シュート数	—
PK によるシュート数	PK によるシュート成功率
パス数	パス成功率
クロス数	クロス成功率
直接フリーキック数	—
間接フリーキック数	—
コーナーキック数	—
スローイン数	スローイン成功率
ドリブル数	ドリブル成功率
タックル数	タックル成功率
クリア数	—
インターセプト数	—
オフサイド数	—
警告数	—
退場数	—
30 m ライン進入回数	—

HOME 進取		チーム・スタッツ		AWAY 次攻	
成功率	回数		回数	成功率	成功率
14.3%	14	シュート	19	21.1%	
-	3	枠内シュート	9	-	
-	0	PKによるシュート	0	-	
77.7%	493	クロス	328	66.5%	
0.0%	12	クロス	13	53.8%	
-	14	前後 F K	12	-	
-	2	間接 F K	0	-	
-	8	C K	2	-	
85.2%	27	スローイン	19	89.5%	
27.6%	29	ドリブル	12	58.3%	
88.5%	18	タックル	40	72.5%	
-	10	クリア	25	-	
-	2	インターセプト	2	-	
-	0	オフサイド	2	-	
-	2	警告	0	-	
-	0	退場	0	-	
-	47	30 m ライン進入	36	-	

図 4 試合ごとのチームスタッツの表示例

状況を想定しよう。そもそも何をもってパフォーマンスが高いと評価できるのだろうか。多くの先行研究では「得点」「失点」「勝ち点」そのものをパフォーマンス指標として使用している ([6] など)。しかし、結果として勝利につながらなくても「よい」プレーが多ければそれを評価したい。そこで、得点を増やす方向に働くスタッツ項目、失点を減らす方向に働くスタッツ項目を探して、それらを指標として使用することにする。また、スタッツ項目の中には意図して増やすことが難しいものもあるため、それらは指標としては使用しない。

次に選手の組み合わせについて考える。たいていの試合では、試合途中に選手交代が行われ、選手の組み合わせが移り変わる。しかし、公開されたスタッツは試合単位で集計された値であるため、交代前後での変化を読み取ることはできない。そのため、スターティングメンバーの組み合わせを、その試合を代表する組み合わせとみなす。また、キーパーについては選手の

表3 スタッツと得点、失点との相関係数

	得点	失点
シュート	0.481	0.223
パス	-0.402	-0.043
クロス	-0.274	0.442
ドリブル	0.051	0.070
タックル	0.476	0.188
クリア	0.206	-0.205
インターセプト	0.204	-0.028

組み合わせに与える影響が少ないと考え、フィールドプレーヤーの組み合わせのみを考えることにする。

分析の手順としては次のようになるだろう。

1. パフォーマンス指標として使用するスタッツ項目を選択する。
2. 同じスターティングメンバーである試合について、試合数、評価指標であるスタッツを集計する。
3. スターティングメンバーのバリエーションを、DMU、試合数を入力、評価指標となるスタッツを出力としてDEAにかけて、効率性を評価する。

それでは2016シーズンの30試合をJ3で戦った大分トリニータのデータを使用した分析例を見てみよう。

まず、得点につながるプレー、失点を防ぐプレーを見つけるために、各スタッツと得点の相関係数、各スタッツと失点の相関係数を求めた(表3)。得点と正の相関のある項目が得点につながるプレー、失点と負の相関のある項目が失点を防ぐプレーとみなすことができるだろう。クリアとインターセプトは得点と正の相関、失点と負の相関をもつことから、得点を増やし失点を減らす方向へ働くプレーといえる。意外なことに、パスとクロスは得点と負の相関を示し、得点に結びつかないことがわかる。そのため、これらは評価指標から外す。またドリブルは得点との間にわずかに正の相関があるものの、失点との正の相関がより大きいため評価指標から外す。シュートとタックルは失点と正の相関があるものの、得点との正の相関がより大きいため評価指標として使用する。以上から、シュート、タックル、クリア、インターセプトを評価指標として採用することにする。ちなみに、これらのプレーが高パフォーマンスにつながると考えられることは、チームのプレースタイルとして相手の攻撃を奪ってからの素早いカウンター攻撃が有効であったことを示唆しているといえるかもしれない。

続いて全30試合のスターティングメンバーの組み合わせを調べたところ、24のバリエーションがあった。それぞれについて、試合数、シュート、タックル、クリ

表4 DEA にかけるデータ

ID	入力		出力		
	試合数	シュート	タックル	クリア	インターセプト
1	1	15	7	7	2
2	1	9	12	15	0
3	1	11	16	29	1
4	2	26	58	35	5
5	1	13	22	19	2
6	1	20	26	21	1
7	2	29	58	50	8
8	2	24	47	55	5
9	1	19	27	26	1
10	1	14	23	18	2
11	1	19	24	22	3
12	1	18	29	18	3
13	1	17	20	9	4
14	1	7	18	13	1
15	1	16	37	11	10
16	2	39	70	42	5
17	1	18	27	28	0
18	1	10	19	49	2
19	1	8	19	15	2
20	3	32	63	50	13
21	1	12	31	29	7
22	1	17	19	17	1
23	1	12	15	15	2
24	1	14	18	10	2

ア、インターセプトを集計したデータを表4に示す。

このデータを表1の形式に整えて、Web DE DEAに渡し、入力指向CCRモデルで分析した結果、表5の効率値、参照集合が得られた。効率性の欄に丸印を付したものが効率的DMUである。本分析では、DMUの規模(すなわち同じメンバーでの試合の数)による効率性への影響があまりないと考え、CCRモデルを採用した。

この分析により得られた効率的なスターティングメンバーの組み合わせが本当に高いパフォーマンスを発揮したといえるのかどうかを見るために、得られた効率値と勝ち点の平均の間のスピアマンの順位相関係数を調べたところ、値が0.434、右片側検定の p 値が0.017となり、有意水準5%で正の関係性の統計的な有意性が確認できた。両者の散布図を図5に示す。図を見ると効率値が高いにもかかわらず勝ち点が低い組み合わせがいくつか存在するが、これらはよいプレーをしたものの相手の強さや運などのために結果に繋がらなかったケースを示しているといえるのではなからうか。

表 5 DEA 分析結果

ID	効率値	効率性	参照集合
1	0.771		15, 16
2	0.510		9, 18
3	0.766		9, 18
4	0.832		15, 16, 21
5	0.736		11, 15, 16, 18
6	1.000	○	6
7	0.929		15, 16, 18, 21
8	0.832		16, 18, 21
9	1.000	○	9
10	0.758		11, 15, 16, 18
11	1.000	○	11
12	0.939		15, 16
13	0.913		15, 16
14	0.538		16, 21
15	1.000	○	15
16	1.000	○	16
17	0.999		9, 16, 18
18	1.000	○	18
19	0.582		16, 21
20	0.704		11, 15, 16, 18
21	1.000	○	21
22	0.852		6, 16
23	0.649		11, 15, 18
24	0.723		15, 16

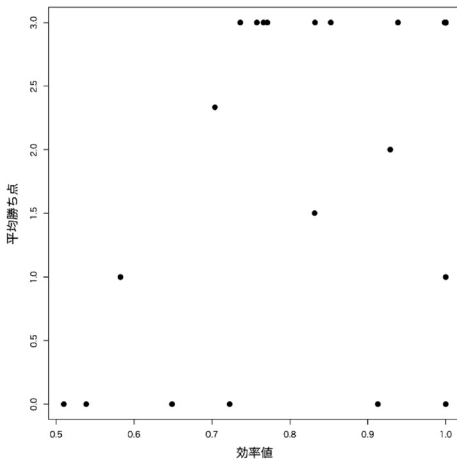


図 5 効率値と平均勝ち点の関係

5. おわりに

本稿では、HTML5 技術を利用した DEA 分析のためのウェブアプリケーションの構築について述べ、これを利用した分析例を紹介した。このような試みにより、誰もが手軽に DEA を利用できる環境が提供できたと考えている。著者はこのほかにも HTML5 による AHP のウェブアプリケーション化も試みている [8]。これらのウェブアプリケーションはいずれも

<http://lab.ms.oita-u.ac.jp/t-obata/html5/>で公開されている。

これらオペレーションズ・リサーチ手法のウェブアプリケーションが、オペレーションズ・リサーチ普及の一助になれば幸いである。

謝辞 本稿へ貴重な助言を賜った富山大学経済学部白石俊輔先生に深く感謝の意を表する。

参考文献

- [1] A. Charnes, W. W. Cooper and E. R. Rhodes, “Measuring the efficiency of decision making units,” *European Journal of Operational Research*, **2**, pp. 429–444, 1978.
- [2] W. D. Cook and J. Zhu (森田浩訳), 『データ包絡分析法 DEA』, 静岡学術出版, 2014.
- [3] 小畑経史, 白石俊輔, “WWW を利用した AHP システムの構築,” 大分大学工学部研究報告, **38**, pp. 29–34, 1998.
- [4] 佐藤和雄, “Web 技術の歴史,” *Unisys Technology Review*, **110**, pp. 175–185, 2011.
- [5] N. Hirotsu, H. Yoshii, Y. Aoba and M. Yoshimura, “An evaluation of characteristics of J-league players using data envelopment analysis,” *Football Science*, **9**, pp. 1–13, 2012.
- [6] M. Flégl, “Performance analysis during the 2014 FIFA World Cup qualification,” *The Open Sports Science Journal*, **7**, pp. 183–197, 2014.
- [7] N. Hirotsu and T. Ueda, “Measuring efficiency of a set of players of a soccer team and differentiating players’ performances by their reference frequency,” In *Proceedings of the 5th International Conference on Mathematics in Sport*, pp. 66–71, 2015.
- [8] 小畑経史, “HTML5 による AHP Web アプリケーションの構築,” 大分大学工学部研究報告, **62**, pp. 1–7, 2015.