

階層的重み付けマッチング及び Dice 関数による マッチングアルゴリズム

大 槻 明

1. はじめに

現在は、Uber や Airbnb に代表されるように、シェアリングエコノミー[1]の時代だと言われている。つまり、既存のデータや個人の持つ資源をシェアリングすることによって新たな価値を生み出せるようになった。さらに、現在は IoT や AI の時代と言われており、全自動で動く車が当たり前となり、あらゆる「モノ」がインターネットで繋がるようになった。このような状況下において、従来はオフィス出勤して仕事をするのが一般的であったが、働く場所を問わないリモートワーク（テレワーク）が行える環境が整ってきた。このような働き方は、政府が進める「働き方改革」の中でも推奨されている[2]。

場所を問わずに働いている人やモノを効率良くマッチングさせるためには、例えば、クラウドファンディング - CAMPFIRE [3] など、出資者（検索者）と技術及びアイデア提供者（情報提供者）をマッチングさせるようなオンラインマッチングサービスが極めて重要となる。しかし、既存のマッチングサービスでは、場所だけ、知識だけ、時間だけなど、何か1つだけを単純にマッチさせるようなアルゴリズムが殆どであり[3-6]、本研究が想定しているような、情報提供者及び検索者がそれぞれ重要視している項目を順位カテゴリカル変数として処理してマッチングさせるアルゴリズムは無かった。他方、本研究に近いものとして、相手同士が設定した項目間から最適なマッチングを求める二部グラフマッチング問題[7-8]があるが、これは n 対 n の中で最適なマッチングパターンを調べるものであり、本研究のコンセプトは、1 対 n において 1 に最適なマッチング相手を調べるものである。

そこで本研究では、仕事内容などの内容に関するマッチング項目については、階層的重み付けマッチングにより順位尺度に変換したうえで、ユークリッド距離により完全マッチングを行うアルゴリズムを考案した。そして、期間や金額など範囲を伴うマッチング項目については、Dice 関数を応用したマッチングのアルゴリズムを考案し、これらを線形結合することにより総合マッチング指標を算出するアルゴリズムを考案した。

2. 先行・関連研究

現在日本ではシェアリングエコノミーに関する取り組みが盛んである。2015/12 にシェアリングエコノミー協会が設立されたり[9]、経済産業省がシェアリングエコノミーの国際標準化を目指していること[10]からもうかがえる。さらに、シェアリングエコノミーに関する研究も数多く存在するが[5, 6, 11-13]、現状ではシェアリングエコノミーの萌芽的な研究が多くみられる。例示すると、寺前[11]はシェアリングエコノミーの論議の方向性について述べており、岡田[12]や中美[13]はシェアリングエコノミー

をめぐる論点について議論している。

また、エコノミーだけにとらわれないシェアリングサービスも数多く提案されている。ANYTIMES (エニタイムズ) [4]は、個人が持っているスキルや時間のシェアを仲介するサービスである。シェアするスキルにはさまざまなものがあり、掃除や洗濯、料理などの家事だけでなく、語学の指導、ペットの世話、順番待ち、引越し作業などがある。

Crowd Works (クラウドワークス) [5]は日本最大級のクラウドソーシングサービスであり、オンライン上での仕事マッチングサイトである。100種類以上の仕事を扱っており、エンジニア・ライター・デザイナーなど、様々な実績のあるプロが登録している。クラウドワークスはあらゆる個人、法人、官公庁の方々が依頼・発注・仕事をしており、他にも様々な士業(弁護士、行政書士、司法書士、公認会計士など)の方々が登録をしていて、専門的な仕事の発注も可能である。

coconala (ココナラ) [6]は自分の得意なこと(知識・スキル・経験)を、オンライン上でサービスとして出品・売買できるオンラインマーケットである。ワンコイン(500円)から出品できるので、誰でも気軽に利用できるのが魅力である

以上に概観したシェアリングサービスは、いずれもマッチングアルゴリズム自体は単純であり、マッチする項目の多さ順で表示する程度のものである。しかし、ビッグデータ時代と言われる現代においては、膨大なシェアリングデータから効率的なマッチングをするためのアルゴリズムの研究が必要となってくる。

他方、コンピューターサイエンス分野におけるマッチングアルゴリズムに関する先行研究を概観すると、古くは D.Gale ほか[14]が提案した Deferred Acceptance (DA) アルゴリズム、または Gale-Shapley (GS) アルゴリズムが存在する。GS アルゴリズムも、DA アルゴリズムも、1対1、もしくは1対nのマッチングを議論する「1集団内マッチング問題」と言われており、田中[15]の論文や宮崎[16]の解説記事の中でこのアルゴリズムの具体的な説明がなされている。

以降は DA 及び GS の応用的なマッチングモデルが数多く研究されてきた。一例を示すと、富山[17]は、受験生と大学が選好順序を持つと仮定して、大学入試制度に係るマッチング問題について論じている。鎌田ら[18]は、研修医がどの病院で研修するかという組み合わせに関する研修医マッチング制度について、DA 及び GS の応用的なマッチングモデルを考察することで、経済学の問題として定式化した。

以上に概観した、DA や GS を起源とするマッチングモデルは、相手同士の選好(希望順位)を基準にマッチングを行うモデルである。しかし、上述したシェアリングサービスでは、選考(希望順位)は無く、相手同士が設定した項目間から最適なマッチングを求める二部グラフマッチング問題に属する。

二部グラフのマッチングに関する定理には、Hall の結婚定理[7]のような完全マッチングや、Tassa[8]の最大マッチングなどが存在するが、本研究は最大マッチングの中でも「重みつき最大二部マッチング問題」に位置付けられると考えられる。重みつき最大二部マッチング問題とは、相手同士の各ペアについて、そこがペアになったら相手同士がどれだけ嬉しいかを階層的に重み付けして数値化し、そこからペアリングして各ペアの利益の総和が最大になるように最適化する問題である。しかし、重みつき最大二部マッチング問題は、 n 対 n の中で最適なマッチングパターンを調べるものであるが、本研究が取り上げている題材(ケース)は、1対 n において1に最適なマッチング相手を調べるアルゴリズムが必要である。

そこで本研究では、上述のとおり、仕事内容などの内容に関するマッチング項目については、階層的重み付けマッチングにより順位尺度に変換したうえで、ユークリッド距離により完全マッチングを行う

階層的重み付けマッチング及び Dice 関数によるマッチングアルゴリズム（大概）

アルゴリズムを考案した。そして、期間や金額など範囲を伴うマッチング項目については、Dice 関数を応用したマッチングのアルゴリズムを考案し、これらを線形結合することにより総合マッチング指標を算出するアルゴリズムを考案した。

3. 提案マッチングアルゴリズム

3.1 マッチング対象者（ロール）

付録 1 の列項目に示すように、本研究では、「アイデア提供者及びスキル提供者」と「資金等提供者」をマッチングさせることを想定している。このアイデア提供者及びスキル提供者とは、文字通り資金等提供者が実現したいプロジェクトなどの遂行に必要なアイデア及びスキルを提供する者であり、付録 1 では提供者 1 人目～6 人の 6 パターン用意した。そして、資金提供者は、個人、企業及び行政の 3 パターンを用意した。

3.2 マッチングの検証イメージ

付録 1 には、前節のロールごとにダミーデータが入力されている。これらの入力データをまとめると表 1 に示す通りとなる。このように、マッチング対象者がそれぞれ 1 つもしくは複数の内容をランダムに登録している状況下で、1 : n (1 が資金提供者で n がスキル等提供者) のネットワークを仮定して、資金提供者が希望する内容に最適なマッチングを実現するアルゴリズムを提案する。

表 1 ロールごとの入力ダミーデータ
Table 1 Input of dummy data per rolls.

	ロール	入力データ
アイデア・スキル提供者	提供者 1	ポータルサイトを 2～3 か月 20 万で作成
	提供者 2	ポータルサイトを 2 か月 30 万で作成
	提供者 3	動画コンテンツを 2～3 か月 30 万で作成
	提供者 4	ロゴを 1～2 か月 30 万で作成
	提供者 5	・ポータルサイトを 2 か月 30 万で作成 ・動画コンテンツを 2 か月以内 50 万以下で作成
	提供者 6	・動画コンテンツを 2～3 か月 100 万で作成 ・ロゴを 1～2 か月 50 万で作成
資金提供者	個人 1	動画コンテンツを 2 か月以内 30 万以下で作成できる方を探している
	行政 1	・ポータルサイトを 3 か月以内 30 万～50 万で作成できる方を探している ・動画コンテンツを 2 か月以内 50 万以下で作成できる方を探している
	企業 1	・ポータルサイトを 3 か月以内 20 万～30 万で作成できる方を探している ・ロゴを 1 か月 40 万で作成できる方を探している

3.3 行サンプルモデル及び階層的重み付けマッチング

本研究で提案する行サンプルモデルは、付録 1 に示している通り階層構造であり、内容の完全マッチで評価する階層（第 1～第 5 階層）とマッチする対象に幅がある場合を想定した階層（第 6～第 7 階層）

の2通りの階層構造に分かれている。まず、第1階層～第5階層であるが、第1階層はマッチング対象のカテゴリを表している。これらのカテゴリは、クラウドファンディング-CAMPFIRE サイト[3]の「プロジェクトをさがす」ページで示されているカテゴリを参考に作成した。第2階層～第5階層は第1階層の内容を掘り下げて完全マッチングによりマッチ度を評価するための階層であり、下位の階層になるほど重みが付くように、質的変数を順位尺度に変換する表(表2)を準備して、この表により質的変数を順位尺度に変換する。

なお、本研究では第1階層～第5階層の5階層で内容の完全一致度評価を行っているが、ケースバイケースでこの階層を自由に設定することができる。

表2 階層的重み付けマッチングのための順位尺度表

Table 2 Ordinal scale table for matching algorithm using depth weighted score matching.

階層	スコア
第1階層	1
第2階層	2
...	...
第n階層	n

3.4 ユークリッド距離による完全一致度判定

完全マッチで評価する階層(第1～第5階層)については、前節で変換した重み付きスコアを対象に、ユークリッド距離を用いて完全一致度の判定を行う。資金提供者Xとスキル等提供者Yのユークリッド距離 $D(X, Y)$ は次式のように求められる。

$$D(X, Y) = \sqrt{(X_1 - Y_1)^2 + (X_2 - Y_2)^2 + \dots + (X_n - Y_n)^2} \quad (1)$$

なお、1, 2, ..., nは、マッチング比較対象となる全階層(全項目)を表す。例えば、付録1で「個人1」と「スキル等提供者1～6」のマッチングを考える場合は、個人1が登録している次の全項目がマッチング対象となる。

1 まちづくり・地域活性化

1-1 カフェスタンド

1-1-2 イベント・広報

1-1-2-1 広報コンテンツ作成

1-1-2-1-2 動画コンテンツ作成

つまり、付録1の「個人1(X)」と「スキル等提供者1(Y)」のユークリッド距離であれば次式ように求められる。

$$\begin{aligned} D(\text{スキル等提供者1}, \text{個人1}) &= \sqrt{(1-1)^2 + (2-2)^2 + (3-3)^2 + (4-4)^2 + (0-5)^2} \\ &= 5.00 \end{aligned} \quad (2)$$

階層的重み付けマッチング及び Dice 関数によるマッチングアルゴリズム (大槻)

そして、式 3 の $MatchD(X, Y)$ に示すように、当該 X, Y 間のユークリッド距離 $D(X, Y)$ を、全ての X, Y 間のユークリッド距離の最大値 $\max(D(X, Y))$ で割り、それを 1 から引くことで、式 2 で求めたユークリッド距離を反転したうえで 0~1 の実数に変換する。

$$MatchD(X, Y) = 1 - \frac{D(X, Y)}{\max(D(X, Y))} \quad (3)$$

3.5 Dice 関数によるマッチング対象に幅がある場合の一致度判定

第 6 階層～第 7 階層は、例えば、資金提供者がホームページ作成を 3 か月以内で希望しており、スキル等提供者がホームページ作成を 3~4 か月で作成できると登録していた場合など、マッチング対象に幅がある場合を想定した階層であり、第 1~第 5 階層の重み付き階層とは別として考える階層である。

具体的には、マッチング対象に幅がある質的変数を表 3 のように 1 か月単位 (金額は 10 万単位) で分割したうえで Dice 関数を用いて一致度判定を行う。

表 3 Dice 関数のインプットデータのイメージ
Table 3 Image of the input data of the Dice similarity coefficient.

スキル等提供者 (HP を 3~4 か月で作成)	共起	資金提供者 (HP 作成を 3 か月以内で希望)
		1 か月
		2 か月
	3 か月	
4 か月		

ここで、Dice 関数[21]のアルゴリズム $Dice(X, Y)$ は次式のように表される。

$$Dice(X, Y) = \frac{2|X \cap Y|}{|X| + |Y|} \quad (4)$$

参考まで、Dice 関数により表 3 のスキル等提供者及び資金提供者の一致度を判定すると次の通りとなる。

$$\text{表 3 の Dice による一致度} = \frac{2 \cdot 1}{(2 + 3)} = 0.40$$

なお、本研究における分割の単位は、期間は 1 か月単位、金額は 10 万単位としているが、ケースバイケースで任意に設定することができる。また、表 4 に示すケースのように、スキル等提供者及び資金提供者がそれぞれ複数の項目を登録している場合は、それら全てを対象にマッチング対象項目として一致度を算出する。

表 4 Dice 関数のインプットデータのイメージ 2

Table 4 Second image of the input data of the Dice similarity coefficient.

スキル等提供者 30 万で HP を作成 50 万以下で動画コンテンツを作成	共起	資金提供者 30 万～50 万で HP 作成を希望 50 万以下で動画コンテンツ作成を希望
	10 万 20 万 30 万 30 万 40 万 50 万	40 万 50 万
Dice	0.86	

3.6 最終的な一致度判定

ユークリッド距離によって求めた第 1 階層～第 5 階層（内容）の一致度を A, Dice 関数によって求めた第 6 階層（期間）の一致度を B, 第 7 階層（金額）の一致度を C と置くと, 最終的な X, Y 間の一致度 $Match(X, Y)$ は次式にて求められる.

$$Match(X, Y) = \frac{A \cdot (B+C)}{\max(A \cdot (B+C))} \quad (5)$$

A, B, C を単純な線結合とはせずに, A を (B+C) にかけている理由は, A の内容が不一致の時に 0 を返すようにするためである.

4. 提案モデルの検証

4.1 はじめに

本章では, 付録 1 のデータを用いて, 完全マッチングの一致度判定 (4.2 節) 及びマッチング対象に範囲がある場合の一致度判定 (4.3 節) について, 関連手法との比較検証を通じて本モデル有用性について検証を行う.

4.2 完全マッチングの一致度判定

今回対象とするデータは順位尺度のカテゴリカルデータであるため, スピアマン順位相関係数 [20] との比較検証を通じて, 3.4 節で述べたユークリッド距離による一致度判定の有用性について検証した.

付録 1 の第 5 階層までのデータを元に, 最終的な式 3 ではなく, 式 2 を用いて一致度判定を行った結果を表 5 に, スピアマン順位相関係数による一致度判定を行った結果を表 6 にそれぞれ示す. 表 5 から, 例えば個人 1 と完全一致している提供者は提供者 3 だけであり, 提供者 5 と 6 は 2 個登録しているうちの 1 つが動画作成であるため一部の一致, そして, 提供者 1 と 2 はポータルサイトしか登録していないため不一致という結果であった. ユークリッド距離を用いた一致度計算では, 数値が小さいほど関係性が高いこと表しており, また, 完全一致は距離 = 0 を返していたためうまく一致度を表せていると考えられる.

対して, スピアマン順位相関係数 r_s は, 値を順位に置き換えたピアソンの積率相関係数であり, 同順位がある場合には平均順位を与える. そして, スピアマン順位相関係数は, 「 $-1 < r_s < 1$ 」の値を返し, 完全一致の場合は 1 を返す. しかし, 本モデルで採用する場合はこの順位に置き換えるという計算過程

階層的重み付けマッチング及び Dice 関数によるマッチングアルゴリズム (大槻)

に問題があると考えられる。例えば、「個人1」と「提供者5及び提供者6」は、ユークリッド距離(表5)の場合は0.00(完全一致)であったが、スピアマン順位相関係数(表6)では「 $r_s=0.99$ 」であった。この理由について表7を元に述べる。表5の第5階層までの値を順位に置き換えたものを表7の「順位」列に、置き換える前のものを「実数」列にそれぞれ示しているが、スピアマン順位相関係数は、同順位がある場合には平均順位を与えるが、アイデア等提供者の5人目、6人目には、第5階層に同順位が2つあるため、それぞれのスコアが5.5に変換されてしまう。そして、資金提供者の個人1の第5階層のスコアは5であるため、この差が原因で「 $r_s=0.99$ 」というズレが生じたものと考えられる。

表5 ユークリッド距離による一致度判定結果

Table 5 Result of the degree to which it matches in case of Euclidean distance.

	個人1	行政1	企業1
提供者1	5.00	5.00	5.00
提供者2	5.00	5.00	5.00
提供者3	0.00	5.00	7.07
提供者4	5.00	7.07	5.00
提供者5	0.00	0.00	5.00
提供者6	0.00	5.00	5.00

表6 スピアマン順位相関係数による一致度判定結果

Table 6 Result of the degree to which it matches in case of Spearman's rank correlation coefficient.

	個人1	行政1	企業1
提供者1	-0.25	0.13	0.13
提供者2	-0.25	0.13	0.13
提供者3	1.00	0.13	-0.73
提供者4	-0.25	-0.73	0.13
提供者5	<u>0.99</u>	1.00	0.14
提供者6	<u>0.99</u>	0.14	0.14

表7 表5の第5階層までの実数値及び実数値を順位に置き換えた値

Table 7 Real numbers of table 5's directory 1-5, and data values that were converted from real numbers to ranks.

	アイデア等提供者												資金提供者					
	1人目		2人目		3人目		4人目		5人目		6人目		個人1		行政1		企業1	
	実数	順位	実数	順位	実数	順位	実数	順位	実数	順位	実数	順位	実数	順位	実数	順位	実数	順位
1	1	5	1	5	1	5	1	5	1	5	1	5	1	5	1	5	1	5
1-1	2	4	2	4	2	4	2	4	2	4	2	4	2	4	2	4	2	4
1-1-2	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
1-1-2-1	4	2	4	2	4	2	4	2	4	2	4	2	4	2	4	2	4	2
1-1-2-1-1	5	1	5	1					5	<u>1.5</u>					5	<u>1.5</u>	5	<u>1.5</u>
1-1-2-1-2					5	1			5	<u>1.5</u>	5	<u>1.5</u>	5	1	5	<u>1.5</u>		
1-1-2-1-3							5	1			5	<u>1.5</u>					5	<u>1.5</u>

4.3 マッチング対象に範囲がある場合の一致度判定

表 8～25 に第 6 階層 (期間) の一致度判定を、表 26～43 に 7 階層 (金額) の一致度をそれぞれ Jaccard 係数 [22]、Dice 係数及び Simpson 係数 [22] で求めた結果について示す。なお、Jaccard 係数は式 6、Simpson 係数は式 7 にて求められる。

$$Jaccard(X, Y) = \frac{|X \cap Y|}{|X \cup Y|} \quad (6)$$

$$Simpson(X, Y) = \frac{|X \cap Y|}{\min(|X|, |Y|)} \quad (7)$$

Jaccard 係数は、2 集合における差集合の要素数が多いほど Jaccard 係数が小さくなるという欠点がある。例えば、集合 A = {1,2,3,4,5}、集合 B = {1,3} であった場合、差集合 A/B = {2,4,5} となる。例えば表 16 に示す通り、Jaccard 係数はこの差集合が多いほど分母の値が大きくなるため Jaccard 係数が小さくなる。そして、この問題を解消したのが Dice 係数及び Simpson 係数である。今回の実験結果を見ても、Jaccard 係数のスコアは、相対的にみて Dice 係数及び Simpson 係数よりも低く出ていたため、この結果からも、Jaccard 係数は差集合の要素数の多さに影響を受けていると考えられる。つまり、Dice 係数及び Simpson 係数よりも、共通要素が多いほど類似度が高いという感覚とは異なる結果となったため、本モデルでは Jaccard 係数を採用しないこととした。

また、Simpson 係数は分母 |X|、|Y| のうちの最小値を取ることから、例えば表 15 の計算式では「1/MIN (1, 3) = 1/1 = 1」となってしまう。このように、Simpson 係数では分母に極端に小さい差集合が存在する場合にはうまく一致度を求めることができないため、本モデルの第 6 階層～第 7 階層の一致度判定には Dice 係数を採用することとした。

なお、行政 1 と提供者 2 は「1-1-2-1-1 ポータルサイト作成」と「1-1-2-1-2 動画コンテンツ作成」の 2 項目で一致している。このような場合は、重複している項目も重複カウントとして計算する。また、表 8～表 43 における「一致無」の表記には次の 2 種類の意味が存在する。

項目は一致しているが、次に示す例のように条件が一致していない。

スキル等提供者	資金提供者
動画コンテンツ作成を <u>100 万</u> で行うと登録	動画コンテンツを <u>30 万以下</u> で作成する方を探している

項目すら一致していない。

スキル等提供者	資金提供者
<u>HP</u> 作成を <u>20 万</u> で行うと登録	動画コンテンツを <u>30 万以下</u> で作成する方を探している

階層的重み付けマッチング及び Dice 関数によるマッチングアルゴリズム（大槻）

【期間】の一致度一覧

表 8 提供者 1 と個人 1 Table 8 Technology provider 1 and Funder (Individual) 1.			表 9 提供者 2 と個人 1 Table 9 Technology provider 2 and Funder (Individual) 1.			表 10 提供者 3 と個人 1 Table 10 Technology provider 3 and Funder (Individual) 1.		
提供者 1	共起	個人 1	提供者 2	共起	個人 1	提供者 3 2～3 か月	共起	個人 1 2ヵ月以内
	一致無			一致無			2 か月	1 か月
						3 か月		
Simpson			Simpson			Simpson	0.50	
Dice			Dice			Dice	0.50	
Jaccard			Jaccard			Jaccard	0.33	
表 11 提供者 4 と個人 1 Table 11 Technology provider 4 and Funder (Individual) 1.			表 12 提供者 5 と個人 1 Table 12 Technology provider 5 and Funder (Individual) 1.			表 13 提供者 6 と個人 1 Table 13 Technology provider 6 and Funder (Individual) 1.		
提供者 4	共起	個人 1	提供者 5 2ヵ月以内	共起	個人 1 2ヵ月以内	提供者 6 2～3 か月	共起	個人 1 2ヵ月以内
	一致無			1 か月			2 か月	1 か月
				2 か月		3 か月		
Simpson			Simpson	1.00		Simpson	0.50	
Dice			Dice	1.00		Dice	0.50	
Jaccard			Jaccard	1.00		Jaccard	0.33	
表 14 提供者 1 と行政 1 Table 14 Technology provider 1 and Funder (Administration) 1.			表 15 提供者 2 と行政 1 Table 15 Technology provider 2 and Funder (Administration) 1.			表 16 提供者 3 と行政 1 Table 16 Technology provider 3 and Funder (Administration) 1.		
提供者 1 2～3 か月	共起	行政 1 3ヵ月以内	提供者 2 2 か月	共起	行政 1 3ヵ月以内	提供者 3 2～3 か月	共起	行政 1 2ヵ月以内
		1 か月			1 か月			1 か月
	2 か月			2 か月		3 か月	2 か月	
	3 か月				3 か月			
Simpson	1.00		Simpson	1.00		Simpson	0.50	
Dice	0.80		Dice	0.50		Dice	0.50	
Jaccard	0.67		Jaccard	0.33		Jaccard	0.33	

表 17 提供者 4 と行政 1 Table 17 Technology provider 4 and Funder (Administration) 1.			表 18 提供者 5 と行政 1 Table 18 Technology provider 5 and Funder (Administration) 1.			表 19 提供者 6 と行政 1 Table 19 Technology provider 6 and Funder (Administration) 1.		
提供者 4	共起	行政 1	提供者 5	共起	行政 1	提供者 6	共起	行政 1
			2ヵ月 2ヵ月以内		3ヵ月以内 2ヵ月以内	2~3ヵ月		2ヵ月以内
	一致無			1ヵ月 2ヵ月 2ヵ月	1ヵ月 3ヵ月		2ヵ月	1ヵ月
Simpson			Simpson	1.00		Simpson	0.50	
Dice			Dice	0.75		Dice	0.50	
Jaccard			Jaccard	0.60		Jaccard	0.33	
表 20 提供者 1 と企業 1 Table 20 Technology provider 1 and Funder (Corporation) 1.			表 21 提供者 2 と企業 1 Table 21 Technology provider 2 and Funder (Corporation) 1.			表 22 提供者 3 と企業 1 Table 22 Technology provider 3 and Funder (Corporation) 1.		
提供者 1	共起	企業 1	提供者 2	共起	企業 1	提供者 3	共起	企業 1
2~3ヵ月		3ヵ月以内	2ヵ月		3ヵ月以内			
		1ヵ月		2ヵ月	1ヵ月 3ヵ月		一致無	
	2ヵ月 3ヵ月							
Simpson	1.50		Simpson	1.00		Simpson		
Dice	0.80		Dice	0.50		Dice		
Jaccard	0.67		Jaccard	0.33		Jaccard		
表 23 提供者 4 と企業 1 Table 23 Technology provider 4 and Funder (Corporation) 1.			表 24 提供者 5 と企業 1 Table 24 Technology provider 5 and Funder (Corporation) 1.			表 25 提供者 6 と企業 1 Table 25 Technology provider 6 and Funder (Corporation) 1.		
提供者 4	共起	企業 1	提供者 5	共起	企業 1	提供者 6	共起	企業 1
1~2ヵ月		1ヵ月以内	2ヵ月		3ヵ月以内	1~2ヵ月		1ヵ月以内
	1ヵ月			2ヵ月	1ヵ月 3ヵ月		1ヵ月	
2ヵ月						2ヵ月		
Simpson	1.00		Simpson	1.00		Simpson	1.00	
Dice	0.67		Dice	0.50		Dice	0.67	
Jaccard	0.50		Jaccard	0.33		Jaccard	0.50	

階層的重み付けマッチング及び Dice 関数によるマッチングアルゴリズム (大槻)

【金額】の一致度一覧

表 26 提供者 1 と個人 1 Table 26 Technology provider 1 and Funder (Individual) 1.			表 27 提供者 2 と個人 1 Table 27 Technology provider 2 and Funder (Individual) 1.			表 28 提供者 3 と個人 1 Table 28 Technology provider 3 and Funder (Individual) 1.		
提供者 1	共起	個人 1	提供者 2	共起	個人 1	提供者 3 30 万	共起	個人 1 30 万以下
	一致無			一致無			30 万	10 万 20 万
Simpson			Simpson			Simpson	1.0	
Dice			Dice			Dice	0.50	
Jaccard			Jaccard			Jaccard	0.33	
表 29 提供者 4 と個人 1 Table 29 Technology provider 4 and Funder (Individual) 1.			表 30 提供者 5 と個人 1 Table 30 Technology provider 5 and Funder (Individual) 1.			表 31 提供者 6 と個人 1 Table 31 Technology provider 6 and Funder (Individual) 1.		
提供者 4	共起	個人 1	提供者 5 50 万以下	共起	個人 1 30 万以下	提供者 6 100 万	共起	個人 1 30 万以下
	一致無			10 万 20 万 30 万			一致無	
			40 万 50 万					
Simpson			Simpson	1.00		Simpson		
Dice			Dice	0.75		Dice		
Jaccard			Jaccard	0.60		Jaccard		
表 32 提供者 1 と行政 1 Table 32 Technology provider 1 and Funder (Administration) 1.			表 33 提供者 2 と行政 1 Table 33 Technology provider 2 and Funder (Administration) 1.			表 34 提供者 3 と行政 1 Table 34 Technology provider 3 and Funder (Administration) 1.		
提供者 1 20 万	共起	行政 1 30 万~50 万	提供者 2 30 万	共起	行政 1 30 万~50 万	提供者 3 30 万	共起	行政 1 50 万以下
	一致無			30 万			30 万	10 万 20 万
					40 万 50 万			40 万 50 万
Simpson			Simpson	1.00		Simpson	1.00	
Dice			Dice	0.50		Dice	0.33	
Jaccard			Jaccard	0.33		Jaccard	0.20	

表 35 提供者 4 と行政 1 Table 35 Technology provider 4 and Funder (Administration) 1.			表 36 提供者 5 と行政 1 Table 36 Technology provider 5 and Funder (Administration) 1.			表 37 提供者 6 と行政 1 Table 37 Technology provider 6 and Funder (Administration) 1.		
提供者 4	共起	行政 1	提供者 5	共起	行政 1	提供者 6	共起	行政 1
			30 万 50 万以下		30 万~50 万 50 万以下	100 万		50 万以下
	一致無			10 万 20 万 30 万 30 万 40 万 50 万	40 万 50 万		一致無	
Simpson			Simpson	1.00		Simpson		
Dice			Dice	0.86		Dice		
Jaccard			Jaccard	0.75		Jaccard		
表 38 提供者 1 と企業 1 Table 38 Technology provider 1 and Funder (Corporation) 1.			表 39 提供者 2 と企業 1 Table 39 Technology provider 2 and Funder (Corporation) 1.			表 40 提供者 3 と企業 1 Table 40 Technology provider 3 and Funder (Corporation) 1.		
提供者 1	共起	企業 1	提供者 2	共起	企業 1	提供者 3	共起	企業 1
20 万		20 万~30 万	30 万		20 万~30 万			
	20 万	30 万		30 万	20 万		一致無	
Simpson	1.00		Simpson	1.00		Simpson		
Dice	0.67		Dice	0.67		Dice		
Jaccard	0.50		Jaccard	0.50		Jaccard		
表 41 提供者 4 と企業 1 Table 41 Technology provider 4 and Funder (Corporation) 1.			表 42 提供者 5 と企業 1 Table 42 Technology provider 5 and Funder (Corporation) 1.			表 43 提供者 6 と企業 1 Table 43 Technology provider 6 and Funder (Corporation) 1.		
提供者 4	共起	企業 1	提供者 5	共起	企業 1	提供者 6	共起	企業 1
30 万		40 万	30 万		20 万~30 万	50 万		40 万
	一致無			30 万	20 万		一致無	
Simpson			Simpson	1.00		Simpson		
Dice			Dice	0.67		Dice		
Jaccard			Jaccard	0.50		Jaccard		

4.4 最終的な一致度判定

表 44 に 1～5 階層のユークリッド距離による一致度判定結果 ($D(X, Y)$ 及び $MatchD(X, Y)$) を, 表 45 と表 46 に期間 (第 6 階層) と金額 (第 7 階層) の Dice 関数による一致度判定結果を, 表 47 に最終的な一致度判定結果 ($Match(X, Y)$) をそれぞれ示す.

表 44 1～5 階層のユークリッド距離による一致度判定結果

Table 44 Result of the degree to which it matches by Euclidean distance in case of directory 1-5.

	個人 1		行政 1		企業 1	
	$D(X, Y)$	$MatchD(X, Y)$	$D(X, Y)$	$MatchD(X, Y)$	$D(X, Y)$	$MatchD(X, Y)$
提供者 1	5.00	0.00	5.00	0.29	5.00	0.29
提供者 2	5.00	0.00	5.00	0.29	5.00	0.29
提供者 3	0.00	1.00	5.00	0.29	7.07	0.00
提供者 4	5.00	0.00	7.07	0.00	5.00	0.29
提供者 5	0.00	1.00	0.00	1.00	5.00	0.29
提供者 6	0.00	1.00	5.00	0.29	5.00	0.29
max	5.00		7.07		7.07	

表 45 期間 (第 6 層) の Dice 係数による一致度判定結果

Table 45 Result of the degree to which it matches by Dice similarity coefficient in case of directory 6.

	個人 1	行政 1	企業 1
提供者 1	0.00	0.80	0.80
提供者 2	0.00	0.50	0.50
提供者 3	0.50	0.50	0.00
提供者 4	0.00	0.00	0.67
提供者 5	1.00	0.75	0.50
提供者 6	0.50	0.50	0.67

表 46 金額(第 7 層)の Dice 係数による一致度判定結

Table 46 Result of the degree to which it matches by Dice similarity coefficient in case of directory 7.

	個人 1	行政 1	企業 1
提供者 1	0.00	0.00	0.67
提供者 2	0.00	0.50	0.67
提供者 3	0.50	0.33	0.00
提供者 4	0.00	0.00	0.00
提供者 5	0.75	0.86	0.67
提供者 6	0.00	0.00	0.00

表 47 最終的な一致度判定結果

Table 47 Final result of the degree to which it matches.

	個人 1		行政 1		企業 1	
		Match(X, Y)		Match(X, Y)		Match(X, Y)
提供者 1	0.00	0%	1.09	42%	1.76	100%
提供者 2	0.00	0%	1.29	50%	1.46	83%
提供者 3	2.00	73%	1.13	43%	0.00	0%
提供者 4	0.00	0%	0.00	0%	0.96	55%
提供者 5	2.75	100%	2.61	100%	1.46	83%
提供者 6	1.50	55%	0.79	30%	0.96	55%
max	2.75		2.61		1.76	

提案コンセプトの有用性について、個人 1 を例に検証を行う。まず、個人 1 が望んでいた「動画コンテンツ作成」を提案していない提供者 1, 2, 4 のマッチ度は 0% であった。そして、動画コンテンツ作成を提案していた提供者 3, 5, 6 のマッチ度はそれぞれ 73%, 100% 及び 55% であった。この詳細を表 49 に示しているが、提供者 6 は、動画コンテンツ作成コスト（第 7 階層）が個人 1 とは「一致無」であるため一番低いスコアになっていた。また、提供者 5 は、動画コンテンツ作成期間（第 6 階層）が個人 1 と完全マッチしており、また、コスト（第 7 階層）も提供者 3 よりも高い一致度であったことから、結果、個人 1 と一番マッチしていたのは提供者 5 という結果になった。

表 48 個人 1 と提供者 3, 5, 6 の一致度判定結果 (第 6 階層, 第 7 階層のみ)

Table 48 Result of the degree to which it matches in case of Technology provider 3, 5, 6 and Funder (Individual) 1.

提 供 者			個人 1	
			第 6 層 動画 2 か月以内	第 7 層 動画 30 万以下
3	第 6 層	動画 2~3 か月	0.50	
	第 7 層	動画 30 万		0.50
5	第 6 層	動画 2 か月以内	1.00	
	第 7 層	動画 50 万以下		0.75
6	第 6 層	動画 2~3 か月	0.50	
	第 7 層	動画 100 万		一致無

5. おわりに

本研究では、シェアリングエコノミー時代に資するマッチングアルゴリズムについて、仕事内容などの内容に関するマッチング項目については、マッチング項目を階層構造で表し、相手同士がマッチした選択項目を重み付け順位尺度に変換したうえで、ユークリッド距離により完全マッチングを行うアルゴリズムを考案した。そして、期間や金額など範囲を伴うマッチング項目については、Dice 関数を応用したマッチングのアルゴリズムを考案し、これらを線形結合することにより総合マッチング指標を算出するアルゴリズムを考案した。今後は、関連手法との比較検証を通じて本アルゴリズムのさらなる有効性検証をしていきたい。

参考文献

- [1] 経済産業省. 「シェアリングエコノミービジネスについて」
https://www.meti.go.jp/shingikai/sankoshin/shomu_ryutsu/joho_keizai/bunsan_senryaku/pdf/004_04_00.pdf, (参照 2019-5-7).
- [2] 働き方改革実現会議. 「働き方改革実行計画 (概要)」,
<https://www.kantei.go.jp/jp/headline/pdf/20170328/05.pdf>, (参照 2019-5-7).
- [3] クラウドファンディング - CAMPFIRE サイト. <https://camp-fire.jp/>, (参照 2019-5-8).
- [4] ANYTIMES. <https://www.any-times.com/>, (参照 2019-5-8).
- [5] Crowd Works. <https://crowdworks.jp/>, (参照 2019-5-8).
- [6] coconala: <https://coconala.com/>, (参照 2019-5-8).
- [7] cut-the-knot: Marriage Theorem. <http://www.cut-the-knot.org/arithmetic/elegant.shtml>, (参照 2019/5/12).
- [8] Tassa Tamir: Finding all maximally-matchable edges in a bipartite graph, Theoretical Computer Science, 2012, 423: 50-58.
- [9] シェアリングエコノミー協会サイト:
https://sharing-economy.jp/ja/news/press_conference_20151214/, 2019/1/10 閲覧.

- [10] 経済産業省サイト「日本提案によりシェアリングエコノミーに関する国際標準化の技術委員会の設立
が承認されました」：<http://www.meti.go.jp/press/2018/01/20190107001/20190107001.html>, 2019/1/10 閲覧.
- [11] 寺前 秀一：シェアリングエコノミー論議の方向性：貸切と乗合の相対化, 横浜市立大学論叢. 人文科学
学系列 68(1), pp.113-143, 2016.
- [12] 岡田悟：シェアリングエコノミーをめぐる論点, 国立国会図書館 調査と情報—ISSUE BRIEF—, No. 983,
pp.1-13, 2017.
- [13] 中美尋：シェアリングエコノミーが日本産業に与える影響, Mizuho Industry Focus Vol. 209, pp.1-38, 2018.
- [14] Crowd Works: <https://crowdworks.jp/>.
- [15] coconala: <https://coconala.com/>.
- [16] D. Gale and L. S. Shapley: College Admissions and the Stability of Marriage, Mathematical Association of
America, The American Mathematical Monthly, Vol. 69, No. 1, pp. 9-15, 1962.
- [17] 田中嘉浩：2 集団マッチングについて：メカニズム・デザイン, 経済学研究, 62(2): pp.41-47, 2013.
- [18] 宮崎修一：安定マッチング問題, 情報処理 Vol.54, No.10, pp.1064-1701, 2013.
- [19] 富山慶典：社会的マッチング理論と大学入試制度, 理論と方法 7(2), 2_61-2_83, 1992.
- [20] 鎌田雄一郎, 小島武仁, 和光純：マッチング理論とその応用：研修医の「地域偏在」とその解決策, 医
療経済研究 23(1), pp.5-20, 2011.
- [21] Sørensen, T.: A method of establishing groups of equal amplitude in plant sociology based on similarity of
species and its application to analyses of the vegetation on Danish commons, Kongelige Danske Videnskabernes
Selskab, 1948, 5 (4): pp.1-34.
- [22] 相澤彰子：大規模テキストコーパスを用いた語の類似度計算に関する考察, 情報処理学会論文誌 49(3),
pp.1426-1436, 2008.

階層的重み付けマッチング及びDice 関数によるマッチングアルゴリズム（大槻）

付録 1 マッチング項目及びロールのイメージ

階層							アイデア・スキル提供者						資金等提供者							
1 (カテゴリー)	2 (内容1)	3 (内容2)	4 (内容3)	5 (内容4)	6 (期間, 単位:月)	7 (合計金額)	...	1 目	2 目	3 目	4 目	5 目	6 目	個人: 1	個人: 2	行政: 1	行政: 2	企業: 1	企業: 2	
1 まちづくり・地域活性化								1	1	1	1	1	1	1						
1-1	カフェスタンド							2	2	2	2	2	2	2	2	2	2	2	2	2
	1-1-1	プロダクト																		
	1-1-2	イベント・広報						3	3	3	3	3	3	3	3	3	3	3	3	3
	1-1-2-1	広報コンテンツ作成						4	4	4	4	4	4	4	4	4	4	4	4	4
	1-1-2-1-1	ポータルサイト作成					5	5	5	5	5	5	5	5	5	5	5	5	5	5
		1-1-2-1-1-1	2～3 か月				6	6	6	6	6	6	6	6	6	6	6	6	6	6
			1-1-2-1-1-1-1	200,000			7	7	7	7	7	7	7	7	7	7	7	7	7	
			1-1-2-1-1-2	2 か月			6	6	6	6	6	6	6	6	6	6	6	6	6	
			1-1-2-1-1-2-1	300,000			7	7	7	7	7	7	7	7	7	7	7	7	7	
			1-1-2-1-1-3	3 か月以内												6	6	6	6	
			1-1-2-1-1-3-1	300,000～ 500,000												7	7	7	7	
			1-1-2-1-1-3-2	200,000～ 300,000															7	
		1-1-2-1-2	動画コンテンツ作成				5	5	5	5	5	5	5	5	5	5	5	5	5	
		1-1-2-1-2-1	2～3 か月				6	6	6	6	6	6	6	6	6	6	6	6	6	
			1-1-2-1-2-1-1	300,000			7	7	7	7	7	7	7	7	7	7	7	7	7	
			1-1-2-1-2-1-2	1,000,000						7	7	7	7	7	7	7	7	7	7	
			1-1-2-1-2-2	2 か月以内										6	6	6	6	6	6	
			1-1-2-1-2-2-1	300,000 以下										7	7	7	7	7	7	
			1-1-2-1-2-2-2	500,000 以下												7	7	7	7	
		1-1-2-1-3	ロゴ作成				5	5	5	5	5	5	5	5	5	5	5	5	5	
		1-1-2-1-3-1	1～2 か月				6	6	6	6	6	6	6	6	6	6	6	6	6	
			1-1-2-1-3-1-1	300,000			7	7	7	7	7	7	7	7	7	7	7	7	7	
			1-1-2-1-3-1-2	500,000							7	7	7	7	7	7	7	7	7	
		1-1-2-1-3-2	1 月以内																6	
			1-1-2-1-3-2-1	400,000															7	

