AIによる国・地域特性を考慮した COVID-19感染拡大抑制施策の効果分析

AI-based analysis of impact of non-pharmaceutical interventions against COVID-19 with respect to country/region features

野呂 智哉[1*]　加藤 孝史[2]　福田 茂紀[1]　浅井 達哉[1]
岩下 洋哲[1]　藤重 雄大[1]　福田 貴三郎[1]　大堀 耕太郎[1]
Tomoya Noro[1]　Takashi Kato[2]　Shigeki Fukuta[1]　Tatsuya Asai[1]
Hiroaki Iwashita[1]　Yuta Fujishige[1]　Takasaburo Fukuda[1]　Kotaro Ohori[1]

[1] 株式会社富士通研究所
[1] Fujitsu Laboratories Ltd.
[2] 富士通九州ネットワークテクノロジーズ株式会社
[2] Fujitsu Kyushu Network Technologies Limited

Abstract: Each country/region has implemented some non-pharmaceutical interventions (NPIs), such as school closure, workplace closure, restriction of gatherings, and stay-at-home requirement, to control COVID-19 epidemic. It is important to analyze the association between the NPIs and COVID-19 incidence and to estimate the impact of each of the NPIs against COVID-19 for the future policy-making. Although the impact potentially differs depending on country/region features (e.g. society, economy, life style), existing works for analysis of COVID-19 transmission growth and control according to NPIs and country/region features are done separately, and no analysis of the impact of each NPIs with respect to country/region features has been done as of yet. In this paper, we extract exhaustively combination of NPIs and country/region features associated with COVID-19 incidence using "Wide Learning" developed by Fujitsu Laboratories, and propose hypotheses for impact of the NPIs against COVID-19 with respect to country/region features as well as comparing our findings with findings on the existing analysis of the impact of NPIs against COVID-19.

1　はじめに

新型コロナウイルス感染症 (COVID-19) は，2019年12月に中国の武漢で初めての感染者が報告された後，216の国・地域で2300万人以上の感染者，80万人以上の死亡者が確認され[1]，現在も感染拡大が続いている．

しかし，感染拡大の程度は国・地域により大きく異なる．ヨーロッパの国・地域では最初の感染者が確認されてから急激に感染者数が増加し，イタリアやスペイン，イギリス等では死亡者数も急増した．ヨーロッパ以外でもアメリカやブラジル，インド等で特に感染が拡大している．一方，台湾，日本，韓国等の東アジア．東南アジアの国・地域やオーストラリア，ニュージーランド等のオセアニアの国・地域では，感染拡大を比較的抑制している．

感染拡大を抑え込むため，各国・地域では様々な対策を実施している．その中には，集会制限や学校閉鎖，営業休止，外出制限等，医療や薬剤によらない対策 (non-pharmaceutical intervention, NPI) もある．感染者数や死亡者数の統計をもとにした分析 [4, 6, 7, 9] や数理モデル・シミュレーション解析 [3, 8, 14] により，これらの対策の効果を推定する研究が多く報告されている．一方，気候条件 (気温や湿度等)，経済水準 (GDP 等)，人口統計 (人口密度や年齢別人口比等)，健康や生活習慣に関する統計 (癌による死亡率，糖尿病罹患率，喫煙率等) といった各国・地域の特性がCOVID-19の感染拡大の程度に関連するとも言われており，これらの相関の有無を統計的分析 [1, 2, 10, 11] や数理モデル・

*連絡先：株式会社富士通研究所
〒 211-8588 神奈川県川崎市上小田中 4-1-1
E-mail: t.noro@fujitsu.com
ORCID: 0000-0002-5645-8497

[1]WHO 発表 (2020 年 8 月 24 日現在). https://who.int/emergencies/diseases/novel-coronavirus-2019

シミュレーション解析 [12] により推定する研究も報告されている．しかし，各国・地域の対策の効果の分析と各国・地域の特性の相関の分析は別々に行われ，統合的な分析はされていない．例えば，Hunter らは，集会制限，学校閉鎖，レストランやバーなどの一部のビジネスの休業の実施は感染拡大の抑制との関連がある一方，すべての不要不急サービスの休業や外出制限の実施に感染拡大の抑制との関連がないことを示している [4] が，効果の有無はその国・地域の特性によって異なる可能性がある．もし違いがあるとすれば，どのような特性が影響するだろうか．これが明らかとなれば，今後の COVID-19 感染拡大の第 2 波，第 3 波に対して，国・地域ごとにその特性に合った効果的な対策の実施が期待できる．

本稿では，各国・地域の COVID-19 感染拡大抑制施策の実施状況と各国・地域の特性の組み合わせと感染拡大の抑制との関連を，数理モデル・シミュレーション解析ではなく，実データをもとにした統計的分析により検出する．分析には，富士通研究所の開発技術「Wide Learning」 [5, 13] [2] を利用し，感染拡大の抑制に関連する施策と国・地域の特性の組み合わせを網羅的に抽出する．Wide Learning で抽出した組み合わせを従来研究で得られた知見と比較するとともに，従来研究で言及されていない組み合わせについて，国・地域の特性を考慮した施策効果に関する新たな仮説を提示する．

2 従来研究

2.1 各国・地域の COVID-19 関連施策の効果

Hunter らは，ヨーロッパ 30ヶ国を対象に，各国・地域が実施した集会制限 (mass gathering restrictions)，教育機関閉鎖 (educational facilitites closed)，レストランやバー等の一部の不要不急ビジネス閉鎖 (initial business closure)，すべての不要不急サービス閉鎖 (non-essential services closed)，外出制限 (stay at home order)，フェイスマスク着用 (face coverings) の 6 つの施策の効果を分析している [4]．分析にはベイズ一般化加法混合モデル (GAMM) を利用し，各施策開始日からの経過日数と感染者数および死亡者数の関係を推定している．その結果，集会制限，一部のビジネスの休業，教育機関閉鎖は感染者数，死亡者数の抑制と関連があり，すべての不要不急サービス閉鎖，外出制限，フェイスマスク着用は関連が見られないとしている[3]．

Islam らは，149 の国・地域を対象に，学校閉鎖 (school closures)，職場閉鎖 (workplace closures)，交通機関閉鎖 (public transport closure)，大規模集会制限 (restrictions on mass gathering)，ロックダウン (lockdown) の 5 つの施策の効果を分析している [6]．この分析では，個々の施策ではなく，複数の対策の同時実施や実施順序に注目し，学校閉鎖，職場閉鎖，大規模集会制限，ロックダウンの同時実施は感染者数の抑制と関連があるが，それに交通機関閉鎖を加えてもより大きな抑制との関連は見られないことと，ロックダウンを学校閉鎖や職場閉鎖より早く実施すると，遅く実施する場合に比べてより大きな抑制と関連することを示している．

これらの分析では，各国・地域の特性は考慮しておらず，特性による各施策の効果の有無の違いについては明らかとなっていない．本稿では，各国・地域の特性を考慮し，特性と対策の組み合わせから感染拡大抑制施策の効果に関連する仮説を発見する．

2.2 各国・地域の特性と COVID-19 感染拡大状況

Notari らは，各国・地域の人口，気候，経済，生活習慣，医療等に関係する特性 (統計) と感染拡大状況の相関を分析している [10]．多くの国・地域では，感染者数が 10 のオーダに到達すると，その後の数週間は指数的に感染者数が増大し，その後，ロックダウン等の施策や国民の意識変化等の影響により増加傾向は緩やかとなり，やがて増加傾向はおさまることに着目し，感染拡大の初期段階における増加ペースと各国・地域の特性の相関を判定している．その結果，気温，外国人旅行者数，高齢者・労働者比，都市部人口比，肺がん死亡率，アルコール消費量，BCG 接種率等と感染者数の増加ペースの間に相関があるとしている．

この研究では，対象を感染拡大初期の感染者数の変化に限定し，各国・地域の施策の影響を排除しながら，国・地域の特性と感染拡大の関連を分析している．本稿では，国・地域の特性だけでなく，各国・地域がとる施策も考慮し，特性と施策の組み合わせから感染拡大抑制の効果に関連する仮説を発見する．

3 分析手法

3.1 Wide Learning

富士通研究所の開発技術「Wide Learning」 [5, 13] は，分類問題を対象とし，入力データ項目 (説明変数) のあらゆる組み合わせを網羅的に探索し，分類ラベル (目的変数) の判定に適した組み合わせを重要な仮説「ナレッジチャンク (knowledge chunk, KC)」として網羅的に抽出する．Wide Learning により，各国・地域の特性や感染拡大抑制施策の実施状況に関する入力デー

[2] https://widelearning.labs.fujitsu.com
[3] 外出制限はむしろ感染者数の増加と，フェイスマスク着用は死亡者数の増加と関連があることを示している．

タ項目の組み合わせの中から抑制効果の有無に関連する KC を列挙することができる．列挙した KC は，感染拡大抑制施策の効果に関する従来研究の知見と比較するとともに，従来研究で言及されていない KC について，国・地域特性を考慮した感染拡大抑制施策の効果に関する新たな仮説として提示する．

Wide Learning を感染拡大抑制施策の効果分析に適用する場合，学校閉鎖や外出制限などの感染拡大抑制施策の実施の有無，施策開始からの経過日数が 7 日以上・未満といった施策の実施状況に関する項目や，1 人あたり GDP が高い・低い，高齢者人口比率が高い・低い，喫煙者人口比率が高い・低い，1000 人あたり病床数が多い・少ないといった国・地域の特性に関する項目を,「高齢者人口比率が高く，かつ，外出制限開始から 7 日未満」のように組み合わせることにより，その条件に合う状況では感染拡大を抑制できるか否かを予測する．一般に，項目数が増えれば組み合わせ爆発を引き起こすが，Wide Learning は,「枝刈り」「メモ化」といった効率的な組み合わせ列挙の技術を使った処理手順を最適化することにより，現実規模の問題に対し，組み合わせを高速に列挙することを可能にしている．

3.2 目的変数と問題設定

感染拡大抑制施策の効果分析のための目的変数を設定する．まず，時刻 (年月日) t でどのくらいのペースで感染者数が増加しているかを表す指標として,「感染拡大率 (spread rate, SR)」を以下のように定義する．

$$\mathrm{SR}(t) = \frac{\text{時刻 } t \text{ の直後 1 週間の新規感染者数} + \sigma}{\text{時刻 } t \text{ の直前 1 週間の新規感染者数} + \sigma} \tag{1}$$

σ は平滑化のためのパラメータである．本稿では，1 日あたり平均 1 人を加算することとし，$\sigma = 7$ とする．この感染拡大率が 1 を上回る場合は感染者数が増加傾向にあり，1 を下回る場合は減少傾向にあることを示す．また，日々の感染拡大率が 1 より大きい値でほぼ一定である場合，その期間の新規感染者数は指数関数的に増加していることを示し，感染拡大率が下がる場合はその増加ペースが緩やかになっていることを示す．

COVID-19 は，実際に感染してから発症し，PCR 検査等により感染が確認されるまでに 1，2 週間程度かかると言われている．つまり，ある国・地域が何らかの施策を開始した場合，その効果が新規感染者数の増減として現れるまでに 1，2 週間かかることが予想される．Hunter らの分析では，感染拡大抑制との関連があるとされる施策でも，その関連が現れるまでに 2 週間かかることを示している [4]．Islam らは，感染者数を予測するモデルにおいて，施策開始後の 7 日間を施策実施前として扱っている [6]．本稿では，施策の効果が現れるまでに要する期間を考慮し，時刻 t とその 14 日後 ($t + 14$days) の感染拡大率の比 (spread rate ratio, SRR) を以下のように定義する．

$$\mathrm{SRR}(t) = \frac{\mathrm{SR}(t + 14\text{days})}{\mathrm{SR}(t)} \tag{2}$$

この感染拡大率の比が 1 を下回る場合，感染者数の増加傾向が緩和している，増加傾向から減少傾向に転じている，減少傾向が加速している，のいずれかにあたる．

Wide Learning が対象とする問題は分類問題であるため，目的変数を感染拡大の抑制の成否とする．本稿では，各国・地域の感染拡大期間内で 4 日ごとに基準日を設定し，各基準日における SRR の値が 0.5 未満である場合を「感染拡大を抑制している」と判断して正例として扱い，0.5 以上である場合を負例として扱う．つまり，国・地域と基準日の組み合わせごとに 1 つの事例 (入力データ) を作る．各国・地域の感染拡大期間は，累計感染者数が 30 人を超えた日を開始日とし，感染拡大率が 1 未満の日が 7 日続く日を終了日とする．ただし，ブラジルのように，感染拡大の抑制が効かず長期化する国・地域もあるため，期間が 50 日を超える場合は，開始日から 50 日後を終了日とする[4]．また，基準日の時点で感染拡大率が 1.0 未満となる事例は，正例，負例のいずれにも含めず，除外する．

ある国・地域の感染拡大率の変動の例を図 1 に示す．感染拡大期間の初期は感染者数が指数関数的に増加するため，感染拡大率は高い値で推移する．その後，施策の実施や国民の意識，生活スタイルの変化等により，感染拡大率は低下し，やがて，1.0 を下回って新規感染者数は減少に転じる．ある程度新規感染者数が少なくなると減少のペースは緩やかになり，感染拡大率は 1.0 前後となる．図 1 において，t_1 から t_6 は基準日を表し，各基準日を始点とする矢印の終点は 2 週間後を表す．つまり，矢印の終点と始点の感染拡大率の比が，その基準日における SRR となる．この例の場合，t_2，t_3，t_5 を基準日とする事例を正例として扱い，t_1，t_4，t_6 を基準日とする事例を負例として扱う．

3.3 データ収集

目的変数の値の算出のため，各国･地域の日々の COVID-19 新規感染者数のデータを，Our World in Data が公開しているデータ[5] から収集した．

感染拡大抑制施策に関する項目を説明変数として選定するにあたり，従来研究で示されている知見をもとに仮説を立てる．従来研究には Hunter らの分析 [4] と

[4] 予備調査では，感染拡大期間が 50 日を超える国・地域は，145ヶ国中 26ヶ国であった．

[5] https://github.com/owid/covid-19-data/tree/master/public/data

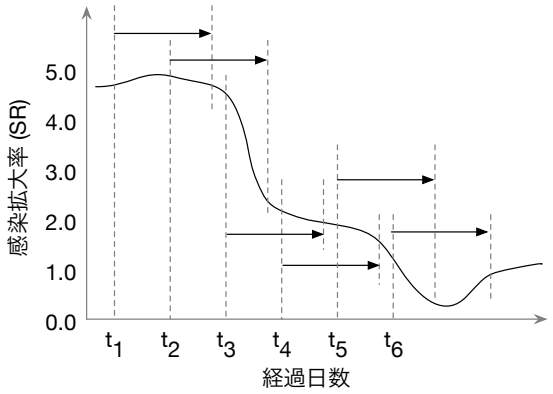


図 1: 感染拡大率の変動の例

Islam らの分析 [6] がある．Hunter らは個々の施策と感染者数および死亡者数の抑制の関連を分析している一方，Islam らは複数の施策の同時実施や実施順序と感染者数の抑制の関連を分析している．本稿では個々の施策と感染拡大抑制の関連に注目することとし，Hunter らの分析による知見をもとに下記の仮説を立てる．

仮説 1 (学校閉鎖): 初等教育から高等教育までのすべての教育機関の閉鎖は，感染拡大抑制の効果がある

仮説 2 (営業休止): 一部またはすべての不要不急ビジネスの休業は，感染拡大抑制の効果がある

仮説 3 (集会制限): 公的または私的の集会の禁止 (制限人数の違いは問わない) は，感染拡大抑制の効果がある

仮説 4 (外出制限): 不要不急の外出の制限は，感染拡大抑制の効果がない

仮説 2 について，Hunter らの分析では，レストランやバー等の一部のビジネスの休業と，すべての不要不急ビジネスの休業を区別して扱い，前者は感染拡大抑制と関連があるが，後者は関連がないとしている．しかし，この 2 つの施策は，ビジネスの休業という観点では同じ施策であり，後者の施策は前者の施策をより厳しくしたものである．本稿では，この 2 つの施策を区別せず，一部またはすべての不要不急ビジネスの休業として扱う．

これらの仮説にもとづき，Oxford COVID-19 Government Response Tracker (OxCGRT) [6] から，各施策の実施状況に関するデータを収集する．OxCGRT は，各施策の実施状況について，その厳しさ (スケール) を 3 段階から 5 段階の数値で表しているが，我々の分析では，Hunter らの分析に合わせて表 1 に示すようにスケールを設定する．Hunter らの分析の対象の感染拡大施策は，いずれも勧告 (recommendation) ではなく義務 (requirement) のレベルであるため，それ以外の感染拡大抑制施策についても，義務のレベル以上にスケールを設定する．ただし，外国からの入国制限は，Islam ら [6] が指摘するように，自国民ではなく外国人に対する施策であり，この施策の実施は他国にも影響を与えるものであるため，我々の分析においても対象外とする．

表 1: 説明変数として利用するデータ項目

感染拡大抑制施策	学校閉鎖 (スケール 3)
	営業休止 (スケール 2 以上)
	集会制限 (スケール 4)
	外出制限 (スケール 2 以上)
	公共行事中止 (スケール 2)
	交通機関閉鎖 (スケール 2)
	国内移動制限 (スケール 2)
国・地域特性	1 人あたり GDP
	高齢者 (65 歳以上) 人口比
	平均寿命
	外国からの旅行者数
	都市部人口比
	喫煙者人口比
	糖尿病罹患率
	BCG 予防接種受診率
	人口密度
	1000 人あたり病床数
	1000 人あたり医師数
	給与所得者比
	極度貧困率
	手洗い環境普及率
	飲料水設備普及率
	下水処理設備普及率
その他	国・地域のエリア分類

また，Hunter らは，ヨーロッパの 30ヶ国を対象に感染拡大抑制施策の効果の分析を行っているが，本稿ではヨーロッパに限定せずにデータを収集する．その代わりに，上述の仮説と我々の分析結果の比較のため，World Bank が設定する国・地域のエリア分類[7] を説明変数として追加する．Hunter らの分析対象のヨーロッパ 30ヶ国は，World Bank のエリア分類における「Europe and Central Asia」に相当する．

国・地域特性に関する項目は，Notari らの分析 [10] で感染拡大ペースと相関があるとされている項目を中心に，World Bank が公開するデータ[8] から選定する．我々の分析で扱う国・地域特性に関する項目を表 1 に示す．

[6] https://bsg.ox.ac.uk/research/research-projects/coronavirus-government-response-tracker

[7] https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups

[8] https://datatopics.worldbank.org/universal-health-coverage/coronavirus

3.4 説明変数の離散化および 2 値化

各感染拡大抑制施策の実施状況に関する項目について，下記の 2 種類の説明変数を用意する．

実施有無: 基準日の時点で当該の施策を実施しているか否か (True/False)

経過日数: 基準日の時点で当該の施策を開始して何日経過しているか (整数値)

施策の効果があるとすれば，それは施策開始からの経過日数が少ないうちに SRR の値の変動として現れるという仮定のもと，経過日数は 7 日と 14 日の 2ヶ所を閾値として離散化し，2 値化する．つまり，7 日以下か否か，8 日以上か否か，14 日以下か否か，15 日以上か否かの 4 つの説明変数に分けられる．

数値データである各国・地域特性に関する項目は，ある閾値でデータ集合を 2 分割した場合のそれぞれのデータ集合のエントロピーの加重平均が最小となるときの閾値で離散化する．つまり，ある閾値でデータ集合が A と B の 2 つの集合に分割されるとき，下記の式で算出する H の値が最小となる閾値で離散化する ($P_1(A)$，$P_0(A)$ はそれぞれ集合 A の中の正例，負例の割合を表す)．

$$\begin{aligned} H(A) &= -P_1(A)\log P_1(A) - P_0(A)\log P_0(A) \\ H(B) &= -P_1(B)\log P_1(B) - P_0(B)\log P_0(B) \\ H &= \frac{|A|}{|A|+|B|}H(A) + \frac{|B|}{|A|+|B|}H(B) \end{aligned} \quad (3)$$

離散化は一度だけ実行する．つまり，各項目を 2 分割し，2 値化する．

カテゴリーデータである国・地域のエリア分類は，one-hot 表現に変換することにより 2 値化する．

4 分析結果

4.1 Wide Learning 用入力データ概要

新規感染者数，各国・地域の感染拡大抑制施策実施状況，各国・地域の特性に関するデータは，2020 年 8 月 6 日に収集した．各国・地域の SRR の値や感染拡大抑制施策に関する説明変数の値を決定する際の基準日は，2020 年 1 月 1 日から 7 月 11 日までの間で 4 日ごとに設定した．データサイズは 1,002 件 (国・地域数は 145)，そのうち正例は 336 件 (正例の割合は約 33.5%) となった．国・地域のエリア分類をヨーロッパ・中央アジアに限定すると，データサイズは 306 件，そのうち正例は 142 件 (正例の割合は約 46.4%) となった．

4.2 感染拡大抑制施策単独の効果に関する仮説の確認を通じた KC 採用基準の設定

Wide Learning は，入力データ中のあらゆる説明変数の組み合わせの中から，感染拡大を抑制できるか否かと強く関連するものを重要な仮説 (KC) として列挙する．本節では，第 3.3 節で立てた，感染拡大抑制施策の効果に関する 4 つの仮説の確認を通して，組み合わせの中から KC として採用する基準を設定する．

本稿では，各組み合わせと感染拡大を抑制できるか否かの間の関連の強さを表す指標として，正規化相互情報量と確信度を利用し，ヨーロッパ・中央アジアであることと各施策の経過日数に関する説明変数の組み合わせの正規化相互情報量をもとに KC として採用するかどうかの基準を設定する．正規化相互情報量として，以下の式で計算される不確定係数を利用する．

$$\mathrm{NMI}(L|C) = \frac{I(L;C)}{H(L)} = \frac{H(L) - H(L|C)}{H(L)} \quad (4)$$

L は正例・負例を表す変数を，C は注目する組み合わせに該当するか否かを表す変数であり，$I(L;C)$ は相互情報量を，$H(L)$ はエントロピーを表す．正規化相互情報量は，注目する組み合わせに該当するか否かが，正例・負例の区別とどの程度関連があるかを表す．一方，確信度は，注目する組み合わせに該当するデータのうち，分類ラベルが組み合わせのラベルと一致するデータの割合を表す．すなわち，ラベルが肯定の組み合わせ (感染拡大を抑制することを示す組み合わせ) の場合は正例の割合を表し，ラベルが否定の組み合わせ (感染拡大を抑制しないことを示す組み合わせ) の場合は負例の割合を表す．

エリアがヨーロッパ・中央アジアであることと感染拡大抑制施策の経過日数が 7 日以内であることの組み合わせの正規化相互情報量と確信度を表 2 に示す (確信度の列の括弧内の数値は，正例，負例のデータ数を表す)．これらはすべてラベルが肯定の組み合わせであり，確信度を見ると，公共行事中止の施策が感染拡大抑制と最も関連が強く，学校閉鎖，営業休止，集会制限，外出制限，国内移動制限までは，ヨーロッパ・中央アジアの国・地域に関するすべてのデータに占める正例の割合 (約 46.4%) より高く，感染拡大抑制との関連が見られる．外出制限も抑制との関連があるという結果は，Hunter らの分析結果 (第 3.3 節の仮説 4) とは異なる結果である．一方，交通機関閉鎖の正規化相互情報量は，他の 6 つの施策と比較して極端に低く，確信度もヨーロッパ・中央アジアの国・地域に関するすべてのデータに占める正例の割合より低いことから，感染拡大抑制との関連は弱いと推定される．

この結果を踏まえ，本稿では，ヨーロッパ・中央アジアの国・地域であることと国内移動制限の経過日数が

表 2: ヨーロッパ・中央アジアの国・地域と各施策の組み合わせの正規化相互情報量と確信度

組み合わせ (ラベル: 肯定)	NMI	確信度
ヨーロッパ・中央アジア	0.0255	0.720
∧ 公共行事中止日数 ≤ 7		(36, 14)
ヨーロッパ・中央アジア	0.0163	0.594
∧ 営業休止日数 ≤ 7		(41, 28)
ヨーロッパ・中央アジア	0.0149	0.633
∧ 学校閉鎖日数 ≤ 7		(31, 18)
ヨーロッパ・中央アジア	0.0127	0.586
∧ 集会制限日数 ≤ 7		(34, 24)
ヨーロッパ・中央アジア	0.0121	0.574
∧ 外出制限日数 ≤ 7		(35, 26)
ヨーロッパ・中央アジア	0.0077	0.537
∧ 国内移動制限日数 ≤ 7		(29, 25)
ヨーロッパ・中央アジア	0.0003	0.391
∧ 交通機関閉鎖日数 ≤ 7		(9, 14)

7 日以下であることの組み合わせの正規化相互情報量 (0.0077) を，KC として採用するかどうかの基準として設定する．また，説明変数の組み合わせは最大 4 つまで (ただし，国・地域のエリア分類に関する変数を含まない場合は最大 3 つまで) とする．その結果，49,274 件の組み合わせ (ラベルが肯定の組み合わせは 19,268 件，否定の組み合わせは 30,006 件) が KC として採用された．これは，すべての組み合わせ (311,613 件) のうち，正規化相互情報量の上位約 15.8%にあたる．

4.3 国・地域特性を考慮した感染拡大抑制施策の効果に関する仮説の発見

第 4.2 節で採用された国・地域の特性に関する説明変数と感染拡大抑制施策に関する説明変数の組み合わせからなる KC の中から，極度貧困率と施策の組み合わせからなるものの例を表 3 に示す．一般に，貧困層の人々の生活環境は悪く，感染予防の徹底が難しいため，極度の貧困層の人口が多い国・地域では感染拡大を抑制することが難しいことが予想される．しかし，表 3 のとおり，極度貧困率が 1.25%を超える国・地域において，1000 人あたり病床数が 1.60 以下の場合や外国人旅行者数が 5,265,000 人以下の場合，これらの特性と集会制限の組み合わせからなる KC は感染拡大の抑制を否定するものである一方，サハラ以南アフリカで交通機関閉鎖を実施していない場合や，平均寿命が 78.7 歳より高い場合は，これらと公共行事中止の組み合わせからなる KC が感染拡大の抑制を肯定するものであり，公共行事中止の施策と感染拡大の抑制の間に関連が見られる．

我々の分析の対象としている施策の中で，日本が実施したものは学校閉鎖のみであり，その他の施策は実施していない．日本でこれらの施策を実施した場合，その効果は期待できるだろうか．また，日本で実施した学校閉鎖の施策は，効果が期待できるものだっただろうか．それを確認するため，各施策と日本の特性の組み合わせからなる KC を調査した．施策ごとに，その施策の経過日数が 7 日以内であることと日本の特性の組み合わせからなる KC の数を表 4 に示す．いずれの施策についても肯定を表す KC が存在する一方，否定を表す KC は集会制限と日本の特性の組み合わせからなるものが 2 件あるのみであった．表 5 に，各施策と日本の特性の組み合わせからなる KC のうち，正規化相互情報量が最大のものを示す．交通機関閉鎖以外の施策と日本の特性の組み合わせからなる KC の中には，正規化相互情報量と確信度の両方が高い，すなわち感染拡大抑制の成否との関連の強い KC が存在し，日本ではこれらの施策の実施による感染拡大抑制の効果が出る可能性がある．一方，交通機関閉鎖と日本の特性の組み合わせからなる KC は少ない上に，正規化相互情報量は小さく，他の施策と比較して日本での感染拡大抑制との関連があることを示す根拠が少ないことが分かる．

表 3: 極度貧困率と施策の組み合わせからなる KC

KC (ラベル: 肯定)	NMI	確信度
極度貧困率 > 1.25	0.0086	1.000
∧ サハラ以南アフリカ		(5, 0)
∧ 交通機関閉鎖実施なし		
∧ 公共行事中止日数 ≤ 7		
極度貧困率 > 1.25	0.0086	1.000
∧ 平均寿命 > 78.8		(5, 0)
∧ 公共行事中止日数 ≤ 7		
KC (ラベル: 否定)	**NMI**	**確信度**
極度貧困率 > 1.25	0.0110	1.000
∧ 病床数 ≤ 1.60		(0, 17)
∧ 集会制限日数 ≤ 7		
極度貧困率 > 1.25	0.0100	1.000
∧ 外国人旅行者数 ≤ 5265000		(0, 16)
∧ 集会制限日数 ≤ 7		

表 4: 施策と日本の特性の組み合わせからなる KC の数

施策	肯定	否定
公共行事中止日数 ≤ 7	458	0
営業休止日数 ≤ 7	342	0
学校閉鎖日数 ≤ 7	306	0
集会制限日数 ≤ 7	309	2
外出制限日数 ≤ 7	355	0
国内移動制限日数 ≤ 7	281	0
交通機関閉鎖日数 ≤ 7	37	0

5 むすび

本稿では，富士通研究所の開発技術「Wide Learning」を利用し，各国・地域の COVID-19 感染拡大抑制施策の実施状況と各国・地域の特性の組み合わせと感染拡

表 5: 施策と日本の特性の組み合わせからなる KC の例

KC (ラベル: 肯定)	NMI	確信度
高齢者人口比 $>$ 8.48	0.0560	0.898
∧ 外国人旅行者数 $>$ 5265000		(44, 5)
∧ 公共行事中止日数 $\leq$ 7		
平均寿命 $>$ 78.8	0.0343	0.78
∧ 営業休止日数 $\leq$ 7		(39, 11)
病床数 $>$ 1.60	0.0283	0.795
∧ 平均寿命 $>$ 78.8		(31, 8)
∧ 学校閉鎖日数 $\leq$ 7		
高齢者人口比 $>$ 8.48	0.0327	0.810
∧ GDP/人 $>$ 19071		(34, 8)
∧ 集会制限日数 $\leq$ 7		
人口密度 $>$ 66.9	0.0298	0.755
∧ 下水処理普及率 $>$ 64.4		(37, 12)
∧ 外出制限日数 $\leq$ 7		
喫煙者人口比 $>$ 14.1	0.0309	0.833
∧ GDP/人 $>$ 19071		(30, 6)
∧ 国内移動制限日数 $\leq$ 7		
東アジア・太平洋	0.0086	1.000
∧ 交通機関閉鎖日数 $\leq$ 7		(5, 0)
KC (ラベル: 否定)	NMI	確信度
都市部人口比 $>$ 45.0	0.0084	1.000
∧ 国内移動制限日数 $>$ 14		(0, 13)
∧ 集会制限日数 $\leq$ 7		

大の抑制との関連を検出し，従来の感染拡大抑制施策の効果分析の研究で得られた知見と比較するとともに，各国・地域の特性を考慮した感染拡大抑制施策の効果に関する新たな仮説を提示した．

国・地域特性を考慮しない，感染拡大抑制施策単独での効果については，公共行事中止の施策の実施は感染拡大抑制との関連が強く，営業休止，学校閉鎖，集会制限，外出制限，国内移動制限の施策の実施も関連が見られた一方，交通機関閉鎖の施策の実施については関連が見られなかった．国・地域特性を考慮した施策の効果の分析では，極度貧困率が高い国・地域でも，サハラ以南アフリカの国・地域や平均寿命の高い国・地域の場合，これらと公共行事中止の組み合わせからなる KC と感染拡大抑制の成否に関連が見られた．さらに，各施策と日本の特性との組み合わせからなる KC を調査すると，日本ではすべての施策について感染拡大抑制の効果が出る可能性があるが，交通機関閉鎖は他の施策に比べて感染拡大抑制の効果があることを示す根拠が少ないことが分かった．

我々の分析は最初の感染拡大時期を対象に行っているが，この分析で得られた仮説が感染拡大の第 2 波，第 3 波でも成立するか，分析する必要がある．また，今回の分析では，各施策の厳しさの変化や施策終了の影響を考慮していない．この影響を考慮した分析の検討は，今後の課題である．

参考文献

[1] Marco Tulio Pacheco Coelho, Joao Fabricio Mota Rodrigues, Anderson Matos Medina, Paulo Scalco, Levi Carina Terribile, Bruno Vilela, Jose Alexandre Felizola Diniz-Filho, Ricardo Dobrovolski: Exponential phase of covid19 expansion is driven by airport connections, *medRxiv*, https://doi.org/10.1101/2020.04.02.20050773 (2020)

[2] Serge Dolgikh: Covid-19 vs BCG: Statistical Significance Analysis, *medRxiv*, https://doi.org/10.1101/2020.06.08.20125542 (2020)

[3] Seth Flaxman, Swapnil Mishra, Axel Gandy, H. Juliette T. Unwin, Thomas A. Mellan, Helen Coupland, Charles Whittaker, Harrison Zhu, Tresnia Berah, Jeffrey W. Eaton, Mlodie Monod, Imperial College COVID-19 Response Team, Azra C. Ghani, Christl A. Donnelly, Steven M. Riley, Michaela A. C. Vollmer, Neil M. Ferguson, Lucy C. Okell, Samir Bhatt: Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe, *Nature*, https://doi.org/10.1038/s41586-020-2405-7 (2020)

[4] Paul Raymond Hunter, Felipe Colon-Gonzalez, Julii Suzanne Brainard, Steve Rushton: Impact of non-pharmaceutical interventions against COVID-19 in Europe: a quasi-experimental study, *medRxiv*, https://doi.org/10.1101/2020.05.01.20088260 (2020)

[5] Hiroaki Iwashita, Takuya Takagi, Hirofumi Suzuki, Keisuke Goto, Kotaro Ohori, Hiroki Arimura: Efficient Constrained Pattern Mining Using Dynamic Item Ordering for Explainable Classification, *arXiv:2004.08015*, https://arxiv.org/abs/2004.08015 (2020)

[6] Nazrul Islam, Stephen J Sharp, Gerardo Chowell, Sharmin Shabnam, Ichiro Kawachi, Ben Lacey, Joseph M Massaro, Ralph B D'Agostino Sr, Martin White: Physical distancing interventions and incidence of coronavirus disease 2019: natural experiment in 149 countries, *BMJ*, Vol. 370, No. m2743, https://dio.org/10.1136/bmj.m2743 (2020)

[7] Jasper S Johnston, Eloise S Johston, Sebastian L Johnston: The importance of timing of a population level intervention on COVID-19 mortality,

medRxiv, https://doi.org/10.1101/2020.04.19.20071845 (2020)

[8] Kevin Linka, Mathias Peirlinck, Francisco Sahli Costabal, Ellen Kuhl: Outbreak dynamics of COVID-19 in Europe and the effect of travel restrictions, *medRxiv*, https://doi.org/10.1101/2020.04.18.20071035 (2020)

[9] Thomas A. J. Meunier: Full lockdown policies in Western Europe countries have no evident impacts on the COVID-19 epidemic, *medRxiv*, https://doi.org/10.1101/2020.04.24.20078717 (2020)

[10] Alessio Notari, Giorgio Torrieri: COVID-19 transmission risk factors, *medRxiv*, https://doi.org/10.1101/2020.05.08.200905083 (2020)

[11] ProfileAdebayo A Otitoloju, ProfileIfeoma P Okafor, ProfileMayowa Fasona, ProfileKafilat Adebola Bawa-Allah, ProfileChukwuemeka Isanbor, Chukwudozie Solomon Onyeka, ProfileOlawale S Folarin, Taiwo O Adubi, ProfileTemitope O Sogbanmu, ProfileAnthony E Ogbeibu: COVID-19 pandemic: examining the faces of spatial differences in the morbidity and mortality in sub-Saharan Africa, Europe and USA, *medRxiv*, https://doi.org/10.1101/2020.04.20.20072322 (2020)

[12] Sen Pei, Sasikiran Kanadula, Jeffrey Shaman: Differential Effects of Intervention Timming on COVID-19 Spread in the United States, *medRxiv*, https://doi.org/10.1101/2020.05.15.20103655 (2020)

[13] 大堀 耕太郎, 浅井 達哉, 岩下 洋哲, 後藤 啓介, 重住 淳一, 高木 拓也, 中尾 悠里, 穴井 宏和: 知識発見によって信頼をつなぐ Wide Learning 技術: FUJITSU, Vol. 70, No. 4, pp. 48-54, https://www.fujitsu.com/jp/documents/about/resources/publications/magazine/backnumber/vol70-4/paper08.pdf (2019)

[14] 倉橋 節也: 新型コロナウイルス (COVID-19) における感染予防策の推定: 人工知能学会論文誌, Vol. 35, No. 3, https://doi.org/10.1527/tjsai.D-K28 (2020)

特許データを用いた製薬企業の戦略に関する研究

Research on strategies of pharmaceutical companies using patent data

米村崇[1]　松本裕介[1]　菅愛子[1]　高橋大志[1]

Takashi YONEMURA[1], Yusuke MATSUMOTO[1], Aiko SUGE[1], and Hiroshi TAKAHASHI[1]

[1]慶應義塾大学大学院経営管理研究科

[1]Graduate School of Business Administration, Keio University

Abstract: 製薬企業にとって特許は宝であり、新薬開発には欠かせない。近年、少子高齢化による社会保障費の上昇、それに伴い既存医薬品の薬価改定の速度が増している。新薬を矢継ぎ早に開発・販売しなければ製薬企業はたちまち衰退の道を歩むだろう。だが、新薬の開発には不確実性が高く、かつ莫大な予算が必要となってくる。その成果ともいえる特許データには製薬企業の戦略が詰まっているともいえる。本研究では、特許データベースである DWPI を用い、特許データを解析することで、製薬企業の戦略の分析を試みた。

1. はじめに

1-1. 国民医療費の概況

日本政府は社会保障と税の一体改革を進めているが、現実として社会保障費の増大等の問題を短期的に解決することは困難であるのは周知の事実である。厚生労働省(2020)によると、日本の 65 歳以上の人口の割合が 30%を超える 2025 年問題が騒がれ、その先の 2040 年には医療費は 65 兆円を超え GDP 比では 8%以上になると予想している。この国民的な課題を解決する策の一つに、薬剤費を下げる政策は常に騒がれてきた。だが近年、バイオ医薬品等の新薬の高額な医薬品の発売が相次ぎ、薬剤費は膨れ上がっているのが現状である。政府としても、画期的な医薬品の発売は国民の福祉や医療の質の向上、イノベーションの推進といった観点からも望んでいることではある。薬価改定の制度もそのような考えで仕組み付けられている。

1-2. 医薬品市場の現状

製薬企業は従来の生活習慣病を中心とした長期収載品から収益を得る構造から、より高い創薬力を持つ産業構造への転換に迫られている。この大きな外部環境の変化の中ではあるが、製薬企業の新薬の研究開発には従前同様、時間と費用の多くを投入する必要がある。日本製薬工業協会によると基礎研究から販売まで 9～17 年、金額も数百億円～数千億円を計上、かつ成功確率も極めて低いことから、ポートフォリオ形成の為、複数のパイプラインが重要になってくる。新薬創出、つまりイノベーションを産み出す鍵の一つに特許がある。知的財産権である特許は製薬企業にとっての生命線であり、将来の研究開発の指針を決める重要な要素である。本研究では特許データである DPWI に SCDV を用いて特許のベクトル化をすることで、企業が保有する特許の重心を測定する。合併等の当事者間の重心の距離を測定することで、特許における企業戦略を分析する。

2. 関連研究

2-1. 特許に関する価値

製薬企業にとって特許はイノベーション創出になくてならないものであり、コアコンピタンスに成り得る重要な資産である。製薬企業の研究開発力を測る指標としては特許があり、特許といった無形資産が企業価値に影響を与える先行研究は多くなされている（Griliches(1981)）。伊地知・小田切(2006)は、特許はイノベーションの利益を専有する手段としても有効であることを示している。パートナーシップと特許の関係による先行研究として、植西・伊佐田(2015)は特許の所有者又は譲受者が複数の場合、パートナーを形成していると仮定したネットワーク分析結果と財務関連データの相関関係を求めた結果、密接な関係を築くことが業績の向上に寄与することが確認されており、パートナーシップのあり方が重要になることが示唆されている。このように非財務指標である特許が企業価値や財務データに影響する

ことは先行研究で示されているのも含め、製薬企業の特許戦略は企業戦略を考える上での KSF（Key Success Factor）であり、企業が成功するか否かを大きく左右する。

2-2. 技術的類似度

研究開発と特許が密接な関係があることは前項で触れた。Jaffe(1986)は企業との技術的類似度を計測するために技術距離を活用する手法を提案した。井田(2011)は医薬品産業における合併について、企業間の製品ポートフォリオの同質的および異質的の判別にあたり、薬効別の医薬品売上高ベースを用いて、Jaffe(1986)の技術距離を参考に計算している。大西(2010)は企業の研究開発費は、同質的な企業同士の合併に限っては増加する一方で、特許出願件数や保有特許を削減する方向に作用することが明らかにし、競争促進的な効果を持つと示唆している。

3. データ

本研究では、Derwent World Patents Index（以下 DWPI）より取得した特許データを活用する。松谷・岡・小林・加藤(2013)によれば、DPWI は世界の特許を収録している特許調査において重要なデータベースであり、特徴は標題、抄録、索引にある。各技術者の専門家が標題および第三者抄録を作成していることから、高精度かつ効率的な特許検索が行える。本研究では特許データのうち、各特許の IPC 分類と抄録を活用する。特許データの分析対象は公報発行日が 2004 年、2006 年、2011~2015 年の日本の特許データ 2,225,901 件である。対象企業は、合併企業を対象にすることから 2004 年は三共、第一製薬、大日本製薬、住友製薬、山之内製薬、藤沢薬品工業、2011~2015 年は医療用医薬品を製造している会社の中において、2020 年 6 月時点で時価総額上位 20 社を対象にしている。なお各特許の抄録はすべて英語で記されている。

4. 分析手法

4-1. SCDV の作成方法

本研究では、企業の技術的類似度を比較するために Sparse Composite Document Vector(以下 SCDV)の手法を用い、特許の文書ベクトル化を行う。対象データは DWPI 抄録内にある該当特許の新規性、詳細な説明、用途、優位性を記した抄録 4 項目にあるテキスト情報を統合し、ステミング処理を行い、以下の方法で算出する（Fujiwara(2020)、藤原・松本・菅・高橋(2020)）。データに 200 次元の Skip-Gram model を用いることで、d次元の単語ベクトルを取得する。次に 60 クラスター、スパース域値 3%の混合分布モデルより、得られた単語ベクトルに確率を付与することでウェイトを持たせる($w\overrightarrow{cv}_{ik}$)。得られた$w\overrightarrow{cv}_{ik}$をクラスター数($K$）の数だけ結合($\oplus_{(1-k)}$)し、逆文書頻度 IDF(N が全文書を、dfₜ がある単語tの出現数を表す)でウェイトを付与することで$\overrightarrow{wtv}_i$を取得する。SCDV による式は以下の式(1)、(2)、(3)に示す。

$$w\overrightarrow{cv}_{ik} = wv_i \times P(C_k|w_i) \quad (1)$$

$$IDF_t = \log\frac{N}{df_t} + 1 \quad (2)$$

$$\overrightarrow{wtv}_i = IDF_t \times \oplus_{(1-k)} w\overrightarrow{cv}_{ik} \quad (3)$$

次元数等の値設定は、Dheeraj ・Vivek ・Bhargavi ・Harish(2017)や松本・菅・高橋(2019)に従って行う。SCDV を用いて特許のベクトル化をした後、検索式は津谷(2013)に準じ、国際特許分類 IPC を利用した。各 IPC は「A61K」(医薬用、歯科用又は化粧用製剤]) 又は「A61P」(化合物または医薬製剤の特殊な治療活性)のいずれかを含むものとした

4-2. 重心からの距離

次に得られた各企業の 12000 次元の特許文書ベクトルを平均することで各企業の重心（cv）を算出する。各企業における 12000 次元空間内での位置関係を把握する。単語ベクトル（p）をn特許数保有している企業iの重心は以下の（4）式で表す。次に得られた各企業の重心間をベクトルの距離を算出することと同様に計測する。式（5）に重心間距離の計算式を記す。ユークリッド距離を使用する。

$$cv_i = \begin{bmatrix} \left(\frac{p_1+p_2+\cdots+p_n}{n}\right)_1, & \left(\frac{p_1+p_2+\cdots+p_n}{n}\right)_2, \cdots, \\ \left(\frac{p_1+p_2+\cdots+p_n}{n}\right)_{12000} \end{bmatrix} \quad (4)$$

企業$_i$・企業$_{i+t}$間距離 ＝

$$\sqrt{\begin{matrix}(p_{i1}-p_{i2+1})^2+(p_{i2}-p_{i2+1})^2\cdots \\ +(p_{i12000}-p_{i12000+1})^2\end{matrix}} \quad (5)$$

ここで得た企業間の重心との距離を技術距離とする。技術距離が現実と乖離していないかを示すため

井田(2011)と比較する。同条件で比較するため、同時期で同企業の統合前の特許データの企業間の技術距離を算出する。対象データは三共、第一製薬、大日本製薬、住友製薬、山之内製薬、藤沢薬品工業の 2004 年の特許データを用いる。技術距離は数字が 0 に近いほど（値の小ささ）、該当企業の距離は小さい、すなわち技術的に類似していることを表す。市場距離は数字が 1 に近いほど（値の大きさ）、製品ポートフォリオの類似性の高さを表す。

次いで 2011~2015 年の特許データを用い、医療用医薬品を研究している会社の中で、2020 年 6 月時点で時価総額上位 20 社を対象にして分析を試みた。5～9 年後の時価総額を対象としている理由として、新薬は基礎研究から販売まで 9～17 年かかるといわれていることから、そのタイムラグを考慮している。対象企業は中外製薬、第一三共、武田薬品工業、アステラス製薬、大塚ホールディングス、エーザイ、塩野義製薬、小野薬品工業、協和キリン、参天製薬、ペプチドリーム、日本新薬、大日本住友製薬、大正製薬ホールディングス、久光製薬、ロート製薬、JCR ファーマ、アンジェス、科研製薬、沢井製薬である。分析には前述の技術距離ならびに階層的クラスター分析をおこなった。階層的クラスター分析の方法はウォード法を用いた。

5. 分析結果

5-1. 技術距離と市場距離の関係

表 1 は合併前企業と新会社名、合併時期、本研究で算出した技術距離、井田(2011)が算出した市場距離を示したものである。

表 1 国内医薬品産業の合併当事者間の技術距離と市場距離

合併前の会社名	合併後新会社名	合併時期	本研究で算出した合併前の技術距離	井田(2011)が算出した合併前の市場距離
三共	第一三共	2007年4月	0.192	0.567
第一製薬				
大日本製薬	大日本住友製薬	2005年10月	0.321	0.357
住友製薬				
山之内製薬	アステラス製薬	2005年4月	0.373	0.323
藤沢薬品工業				

井田(2011)では第一三共を同質的な製品ポートフォリオを有する 2 社の合併事例として、アステラス製薬を異質的な製品ポートフォリオを有する 2 社の合併事例としていたが、技術距離で算出した場合でも同様の結果となった。

5-2. 2011～2015 年の企業間の技術距離

表 2 は 2020 年 6 月時点での時価総額ランキング上位 5 社の企業間の技術距離を示したものである。

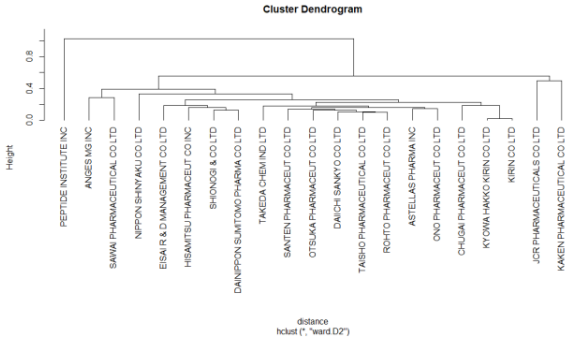
表 2 時価総額上位 5 社の企業間の技術距離

会社名	中外製薬	第一三共	武田薬品工業	アステラス製薬	大塚製薬
中外製薬	0	0.168	0.215	0.181	0.197
第一三共		0	0.175	0.147	0.124
武田薬品工業			0	0.171	0.137
アステラス製薬				0	0.149
大塚製薬					0

中外製薬、武田薬品工業、第一三共、アステラス製薬、大塚製薬間の技術距離は近いことを確認できる。

図 1 は 2020 年 6 月時点での時価総額ランキング上位 20 社の企業間の技術距離を階層的クラスター分析を行い、樹形図に示したものである。

図 1 時価総額上位 20 社の企業間の技術距離 階層的クラスター分析



Height0.4 では 17 社とは別に、ペプチドリーム、JCR ファーマ、科研製薬がクラスターとして分類された。

表 3 4 社の企業間の技術距離

会社名	ペプチドリーム	JCRファーマ	科研製薬
ペプチドリーム	0	0.876	0.891
JCRファーマ		0	0.493
科研製薬			0

該当の 3 社と他の製薬 17 社とは一定の技術距離

があることが示唆されたが、3 社間の距離も近くはないことが示されている。JCR ファーマは希少疾病領域、ペプチドリームはバイオ医薬品の、ともに販売機能を多く持たない創薬特化型のスペシャリティーファーマであることが、科研製薬は整形外科領域ならびに皮膚科領域を中心に経営資源を集中させている会社であるが農薬も扱っていることが、それぞれ技術距離が遠い要因の可能性がある。

6. まとめ

本研究では、企業が保有する特許データを用い、SCDV により特許文書のベクトルを獲得することで、当該企業の技術的情報の数値化に成功した。次に企業の技術の重心位置を、企業間で比較することで技術距離を算出し、医薬品業界の企業分析をおこなった。技術距離と市場距離は一定の関係があることが示唆された。次いで、5 年間の技術距離の算出の結果、製薬企業の研究戦略あるいは特許戦略の分析の可能性を見出せた。詳細な分析が今後の課題である。

参考文献

[1] Dheeraj Mekala., Vivek Gupta., Bhargavi Paranjape., Harish Karnick.: SCDV: Sparse Composite Document Vectors using soft clustering over distributional representations, Association for Computational Linguistics, Vol. Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing, pp. 659–669, (2017)

[2] Jaffe, A.B. : Technological Opportunity and Spillovers of R&D: Evidence from Firms'Patents, Profits and Market Value, American Economic Review. 76(5): 984-1001.(1986)

[3] Griliches Zvi. :Market value, R&D and patents , Economics Letters, 7, 183-187,(1981)

[4] Shohei Fujiwara, Yusuke Matsumoto, Aiko Suge, Hiroshi Takahashi: Constructing a Valuation System Through Patent Document Analysis, In: Jezic G., Chen-Burger J., Kusek M., Šperka R., Howlett R., Jain L. (eds) Agents and Multi-agent Systems: Technologies and Applications 2020. Smart Innovation, Systems and Technologies, vol 186. pp.355-366, Springer, (2020)

[5] 井田聡子,永田晃也,隅藏康一: 医薬品産業における企業境界の変化がイノベーションに及ぼす影響に関する分析, 文部科学省科学技術政策研究所 DISCUSSION PAPER No.75,(2011)

[6] 伊地知寛博,小田切宏之: 全国イノベーション調査による医薬品産業の比較分析,文部科学省科学技術政策研究所 Discussion Paper No.43,(2006)

[7] 植西祐子,伊佐田文彦: 国内製薬企業の特許共同出願に見るパートナーシップのネットワーク分析,研究・イノベーション学会 年次学術大会,（2015）

[8] 大西宏一郎,永田晃也: 医薬品産業における M&A が研究開発・知的財産活動に与える影響, 日本知財学会誌 Vol. 7 No. 1―2010 : 37-44 ,(2010)

[9] 津谷薫子: 癌製剤に関する医薬企業の用途発明の研究, 平成 25 年度プロジェクトレポート,(2013)

[10] 藤原匠平, 松本裕介、菅愛子, 高橋大志: 特許文書ベクトルを用いた企業価値評価, 第 34 回人工知能学会全国大会, (2020)

[11] 松谷貴己, 岡紀子, 小林伸行, 加藤久仁政: Derwent World Patents Index (DWPI)抄録の評価の試み -日本語特許公報を例に-情報管理, Vol. 56, No. 4, pp. 208-216, (2013)

[12] 松本裕介,菅愛子,高橋大志: 企業の多角化とシナジー効果の関連性-特許データを用いた分析-, 日本ファイナンス学会第 27 回大会,(2019)

[13] 厚生労働省: 平成 29 年度 国民医療費の概況 https://www.mhlw.go.jp/toukei/saikin/hw/k-iryohi/17/dl/kekka.pdf (2020 年 6 月 4 日)

[14] 日本製薬工業協会ホームページ http://www.jpma.or.jp/(2020 年 6 月 4 日)

[15] 日本製薬工業協会:産業ビジョン 2025 世界に届ける創薬イノベーション,(2016)

交渉エージェントを用いた 工場内無人搬送車システムの検討

Study of Factory Automated Guided Vehicles Systems using Negotiation Agent.

加藤　大望　矢田　昇平　倉橋　節也

Daimotsu Kato, Shohei Yada and Setsuya Kurahashi

筑波大学
University of Tsukuba

Abstract: 近年、工場では原材料、部品、完成品等の輸送に関し、無人搬送車 (AGV) を用いた自動搬送システムが利用されている。ジョブショップ型の工場では、同じ製造設備を用いて同様の工程で繰返すことから搬送車の渋滞が生じ、適切なタイミングで製品を製造装置に供給できず、納期遅延や品質トラブルなどを引き起こす可能性がある。本研究においては、ジョブショップ型の工場である半導体製造工場を想定し、交渉エージェントを用いた工場内無人搬送車システムの検討を行ったので報告を行う。

1　研究背景、および目的

近年、生産性の向上等を目的とし、工場の IoT 化、スマート化が検討されている[1, 2]。このような工場では原材料、部品、完成品等の輸送において、従来のベルトコンベア等を用いた搬送システムとは異なり、ロボット等を用いた効率的、かつフレキシブルな搬送システムが検討されている[3, 4]。工場の生産方式は大別すると、製造装置を工程順に従って同順序で処理するフローショップ型、同種の製造装置を複数配置し、処理順序が可変であるジョブショップ型の生産方式に分かれる。ジョブショップ型の生産方式は、同じ製造装置を共有して製品を生産する工場でよく用いられる生産方式であることから、生産工程間における加工中の製品搬送において、無人搬送車 (Automated Guided Vehicle: AGV)を用いた自動搬送システムの導入が進められている。使用される AGV は自動制御され、加工中の製品を運搬するが、この際、AGV の渋滞や加工前後での適切な時機での製品の受渡しができず、納期遅れが発生する等の課題が存在する。このため、搬送車の交通量を最適化し，渋滞発生を抑えることが、ジョブショップ型工場の生産性向上に重要となる。ジョブショップ型工場は様々なものがあるが、半導体製造工場では基板洗浄やフォトリソグラフィ工程など同様の装置を複数回繰り返して使用することで製造を行うことから典型的なジョブショップ型の工場であり、特に 12 インチの半導体ウェハを用いた工場では、AGV を用いた大規模な自動搬送システムが適用され、その搬送効率が工場の生産性に大きな影響を与えることから様々な研究が行われている[5-7]。

本研究では、ジョブショップ型工場である半導体製造工場を基に、交渉エージェントを用いた工場内 AGV システムの検討を行ったので報告を行う。

2　研究方法

2.1　NetLogo を用いた AGV 流量解析

工場内 AGV システムでは、図 1 に示すような全長 L [m]である AGV が速度 v [m/s] で走行し、前方車との車間距離 L_h [m] を検知し、L_h を一定に保つように v を調整しながら走行する。この AGV の交通流量を最大化し、搬送路での渋滞発生を抑えるため、本研究においてはマルチエージェントシミュレーションを用いた解析を行った。マルチエージェントシミュレーションにおいては、各 AGV をエージェントとして扱い、エージェントを生産ラインに単位時間あたり n [台/s] で流入した場合に流出する交通流量が AGV の車間距離 L_h に対し、どのように変化するかを解析した。なお、マルチエージェントシミュレーションでは、エージェント型プログラミング言語 NetLogo を用いて解析を行った[8]。

解析では、図 2 のような工場の製造ラインの一部を想定し、単位時間あたり n で搬送経路から加工装置の存在するイントラベイに AGV が流入し、反時計回りに搬送路を走行しているとした[5]。イントラベイの搬送路の両側には、加工装置を配置し、流入した AGV が加工装置に製品の受渡しを行い、再び搬送路から流出するとした。本解析では、イントラベイから流出する際の単位時間あたりの AGV の台数を交通流量[台/s]とし、Lh と n を変化させ、交通流量の解析を行った。

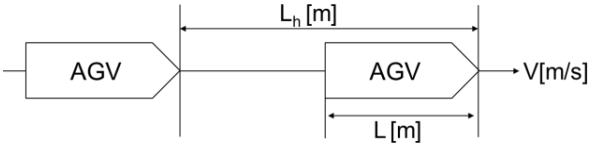


図 1　AGV 間の関係

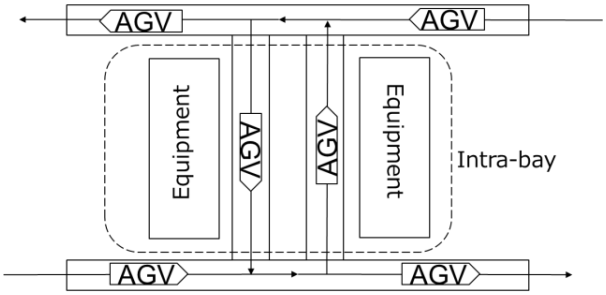


図 2　工場内レイアウト

車間距離 L_h を 1～9 m で変化させた場合における交通流量の時間変化を図 3 に示す。ここで、単位時間あたりに流入する AGV 台数 n を n = 0.2 台/s としている。図に示すように横軸が時間に対し、縦軸の交通流量は約 100 sec の時点でピークを持つことがわかる。これは、イントラベイに流入または流出する AGV が搬送路の交差点において、車間距離 L_h を一定に保つために、速度 v を変化させるためであると考えられる。一方、交通流量は一度ピークを付けた後は一定の値を示すか、徐々に低下することが確認できる。図 3(a) では、時間 800 sec における交通流量の値を比較すると、図 3(a)では、車間距離 L_h に関しては、L_h = 1, 3, 5 m において、L_h = 5 m で約 0.3 台/s 示すが、L_h = 7, 9 m では、交通流量の低下が明確に確認できる。

上記の原因を確認するため、n = 0.3 台/s における搬送路上に存在する AGV 総数の時間変化を図 4 に示す。図 4 に示すように Lh = 7, 9 m では AGV 総数が時間 500 sec を超えると 0.5 台/m で飽和していることが確認できる。この値は AGV が流入する n = 0.3 台/s よりも高い値となり、車間距離のマージンを広くとりすぎてしまい、搬送路上で渋滞が発生した結果、交通流量が低下したと考えられる。

以上の結果から単位時間あたりに流入する AGV 台数 n を増加させることで交通流量を増加させることが可能であるが、車間距離には最適値が存在し、車間距離 Lh を広く設定しすぎると渋滞が発生し、交通流量が低下する可能性があることが分かった。このような渋滞発生の条件は、装置の配置や加工処理時間に変化すると考えられるが、イントラベイに流入または流出する AGV がイントラベイ外に存在する搬送路が交差する点において、車間距離 Lh を一定に保つために、速度 v を変化させるために生じている点は同様であると考えられる。なお、上記で示した車間距離 Lh、単位時間あたりに流入する AGV 台数 n と交通流量の関係は参考文献[9]に示している。

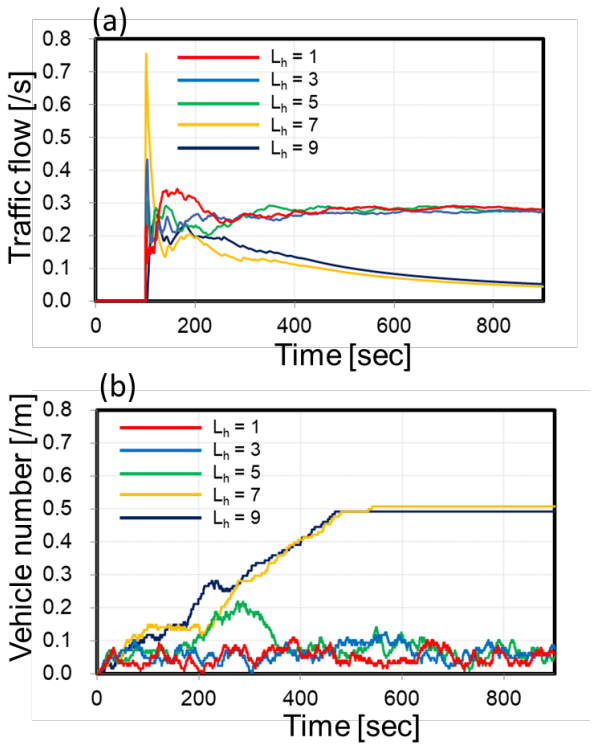


図3 (a) 交通流量の時間変化
(b)搬送路上に存在するAGV総数の時間変化
(n = 0.3 の場合)

2.2　交渉エージェントの検討

上記の研究では AGV の流量のみに着目したが、実際には各 AGV に割当てるタスクを考慮する必要がある。タスク割当においては、交渉エージェント等を用いたタスク割当システムが検討されており、このシステムでは本研究で検討した渋滞情報（図 3

に示したような搬送路上に存在する AGV 総数の時間変化等）を各 AGV に与えることで、AGV が自律的に行動し、効率的なタスク処理を行う事が重要になると考えられる[10,11]。そこで、本研究では、AGV の自律的な行動を与える検討を行った。具体的には、各 AGV に交渉エージェントの機能を与え、契約ネットプロトコル(Contract Net Protocol: CNP)に従い、タスクをオークション形式で各 AGV 与えることを検討した[12]。本研究の場合、契約交渉の際には、装置タスクを与えるエージェントを管理者、入札するエージェントを契約者とし、AGV は契約者として扱い、動的に契約交渉を行うとした。また、交渉の際には 契約ネットプロトコルに基づいたメッセージがマネージャと 契約者の間でやり取りされる。タスク管理者と契約者の間の交渉の流れは図 4 のようになる。図 4 に示すように、まず、タスク管理者は AGV にタスクを送信しい、次に AGV はそのタスクを入札するかの判断を行い、入札が可能である場合は入札を行う。つづいて、タスク管理者は複数の入札に対し、落札者を選定する。落札者は、各 AGV の異質性（例えば、目的とする装置と AGV の距離など）を考慮し、落札者の AGV に結果を報告する。最後にタスクを落札した AGV は、自己の状況を考慮し、タスクを実行するかの判断を行う。なお、落札されたタスクが実行されなかった場合は、タスク管理者は別の契約者にタスクを割り当てる。

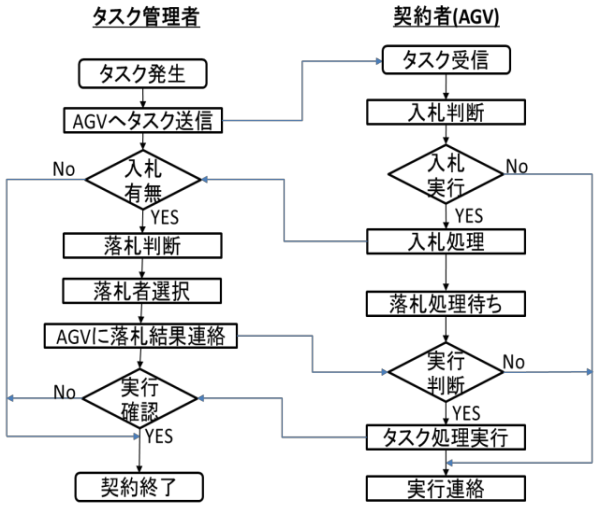


図4 タスク管理者と契約者の間の交渉の流れ

2.3 BDI スキームを用いた CNP の検討

2.2 節で上げた CNP を用いたシステムを NetLogo 上で実装する際に、BDI アーキテクチャ[13]を用いて検討を行った。BDI アーキテクチャは、人間が合理的に行動する際には、信念（Belief）、欲求（Desire）、意図（Intention）という三つの心的状態が関与しているとし、その分析を基本として自律エージェントを構築しようとするアーキテクチャである。具体的には、以下の 3 つの状態で実装されている。

・Belief：各エージェントの環境や状態
・Desire：各エージェントの達成する目標
・Intention：各エージェントが Desire を達成するために実行する行動や計画

NetLogo においては、BDI extension という形で BDI アーキテクチャが提供されており[14]、本研究においても BDI extension を用いて検討を行った。この BDI アーキテクチャをベースに、CNP を用いたシステム検討することでより、AGV の自律性を考慮したシステムの設計が可能となると考えられる。図 4 に示すような交渉の流れは、Intention を基本として実装されるが、エージェントの環境が大きく変わった場合（Belief が変化した場合）には、エージェントは Intention を変更する必要がある。本研究においては、図 3(b)に示す搬送路上に存在する AGV 総数の時間変化をモニタリングし、その値を各 AGV にフィードバックすることにより、搬送路上の AGV に渋滞が生じると類推される場合には、Belief が変化させることで、より環境の変化に対応するシステム設計が可能になると考えられる。

3 期待される成果

本研究では、マルチエージェントシステムでの解析結果をもとに、より各エージェントが自律的に状況を判断し、タスク実行を行えるよう、BDIアーキテクチャを組み込んだCNPのシステムを検討した。BDIアーキテクチャを組み込むことで、より外部環境に順応できるシステム設計が可能になると考えられる。特に、本研究で検討した渋滞情報（図3に示したような搬送路上に存在するAGV総数の時間変化等）を各AGVに与えることで、自律的に経路探索を行い、効率的な経路発見・およびタスク割当を行う事が重要であり、さらには、過去の工場内の渋滞情報、加工装置の故障予測情報、各AGVから位置情報を収集し、推論を行う事で、Beliefを変化させ、より効率的な工場内無人搬送車システムの検討を進めていく。

参考文献

[1] 内閣府『第 5 期科学技術基本計画』平成 28 年 1 月 22 日

[2] M. Hermann et al., “Design Principles for Industrie 4.0 Scenarios”49th Hawaii International Conference on System Sciences (HICSS), 2018

[3] S. Wang et al., “Implementing Smart Factory of Industrie 4.0: An Outlook” International Journal of Distributed Sensor Networks, **12**, 1 (2016).

[4] A. Kusiak et al., “Smart manufacturing” International Journal of Production Research, **56:1-2**, 508 (2018).

[5] J. Tung et al., “Optimization of AMHS design for a semiconductor foundry fab by using simulation modeling” Proceedings of the 2013 Winter Simulation Conference, 3829.

[6] Rank S, Hammel C, Schmidt T, et al., “Reducing simulation model complexity by using an adjustable base model for path-based automated material handling systems case study in the semiconductor industry” Proceedings of the winter simulation conference, Huntington Beach, CA, 6–9 December 2015.

[7] K. Kumagai et al., “Maximizing traffic flow of automated guided vehicles based on Jamology and applications” Advances in Mechanical Engineering, **10** (12) , 1 (2018)

[8] Netlogo Homepage: https://ccl.northwestern.edu/netlogo/

[9] 加藤 大望他『マルチエージェントシミュレーションを用いた 工場内無人搬送車システムの解析』, SIG-BI #14(2019)

[10] R. V. Parysi et al., “Flexible Multi-Agent System for Distributed Coordination, Transportation & Localisation” AAMAS 2018, p1832.

[11] Sebastian Mayer et al., “Adaptive Production Control with Negotiating Agents in Modular Assembly Systems” 2019 IEEE International Conference on Systems, Man and Cybernetics (SMC), p120.

[12] R.G.Smith, ”The Contract Net Protocol: High-Level Communication and Control in a Distributed Problem Solver,” IEEE Trans. on Computers, Vol.C-29, pp.1104-1113 (1980)

[13] M. Bratman et. al., “Intention, Plans, and Practical Reason,” Harvard University Press (1987)

[14] Ilias Sakellariou et. al., “Enhancing NetLogo to Simulate BDI Communicating Agents,” J. Darzentas et al. (Eds.): SETN 2008, LNAI 5138, pp. 263–275 (2008)

経営者の容姿が企業行動に与える影響に関する研究

A Study on the Influence of Managerial Appearance on Corporate Behavior

深澤あゆみ　高橋大志

Ayumi Fukasawa, Hiroshi Takahashi

慶應義塾大学大学院経営管理研究科
Graduate School of Business and Administration, Keio University

Abstract: In Europe and the United States, appearance is considered to be an important factor in business, and not a few studies have examined the effects of managerial appearance on corporate management and individual success and performance. In this study, we attempt to analyze quantitatively the faces of managers in Japan, which is one of the most important components of appearance, in order to elucidate whether a similar phenomenon is found in Japan. Photographs of managers' faces are collected from annual reports and other sources, and feature vectors are extracted. This will be clustered using the K-means method using FaceNet's trained model VGGFace2 to analyze whether there are differences in the behavior of the companies in each cluster, i.e., how the face of the manager affects the company's behavior.

1. はじめに

「容姿が良いと仕事もできる」、「見た目が良いと得をする」というイメージがあるが、これは日本に限ったことではなく、むしろ海外の方がそのイメージは強い傾向にある。特にアメリカのビジネス社会では容姿重視の風潮が強いと言われており、例えば歯並びや歯の白さなどが重要視されている。日本ではあまりメジャーではないものの、欧米ではイメージコンサルタントというビジネスパーソン向けの外見コンサルタントのような職種も存在する。ビジネスにおいて外見は非常に重要視されており、外見に投資する人間は多いという。海外では容姿が企業の経営や個人の出世・業績に及ぼす影響が研究されているものも少なからず存在している。

一方、日本を対象としたこのような研究は我々の知る限り存在していない。日本でも「見た目が良いと得をする」と感じている人は少なくはないであろうものの、「人を見た目で判断してはいけない」「見た目より中身」という風潮があるのも確かであり、本当に容姿が個人及び企業の業績に何らかの影響を及ぼしているのかは明らかではない。ただし、初対面の人間との出会いや一歩的に見られることの多いビジネスにおいては、実際の中身よりも外見的な要素の方が相手に大きな印象を残すことが考えられる。もし本当に何らかの影響があるのであれば、個人や企業が自らの評価や業績を高めるために外見に対する何らかの投資を行うことが有効な手段となる可能性がある。そのため、実際に日本で容姿と個人の評価等の関係あるのかといったことを明らかにすることにより、一つのスキルとして外見を有効的に活用できる可能性がある。

次節において、先行研究について触れたのち、分析手法、分析結果について触れる。5. はまとめである。

2. 先行研究

個人の容姿の美しさが、本人の評価や業績に及ぼす影響を分析したものとしては、海外の研究がほとんどである。見た目の美しさと収入の関係を調査した Hamermesh et al.（1994）は、カナダとアメリカの労働者を対象に調査を行なった結果、男女ともに容姿が美しい方が、収入が高くなる傾向にあり、美しさによる影響は女性よりも男性の方がやや大きいという結果を示している。

容姿の美しさが本人の評価や業績に及ぼす影響を、職業を絞って調査した研究も見られる。Biddle et al.（1995）では弁護士の容姿の美しさと卒業後の収入の関係を、アメリカのあるロースクールの入学時の

写真を用いて個人の容姿レベルを判定することによって調査している。その結果として、民間部門の弁護士は公的部門の弁護士より容姿が美しく、また容姿が美しい弁護士の方が高収入であるという傾向が判明している。また、Hamermesh et al.（2003）では、アメリカの大学教授に対する学生からの評価を分析し、大学教授の容姿への評価が高いほど、学生からの指導評価も高くなっているという事実を表している。ただし、これらの先行研究では、見た目が美しい弁護士の方が高収入、見た目が美しい大学教授の方が学生からの指導評価が高くなることが示唆されてはいるものの、これらの原因が、見た目が美しい弁護士や教授の方がクライアントもしくは学生に外見的に支持されるためなのか、実際に仕事ができるためなのかは判断することができないとしている。

また、単に個人の美しさが周囲からの評価を高めている可能性だけでなく、個人が自らの美しさを、その他のスキルと同様に生産性向上のために使用した結果、個人の業績が向上している可能性を示唆する先行研究も見られる。Salter et al.（2012）では、アメリカの不動産ブローカーの美しさと賃金の関係を調査し、美しい不動産ブローカーは賃金が高くなる傾向にあることを示している。さらに、美しい不動産ブローカーは、仕事を取るために必要な努力、知性、組織的スキルを補完するために美しさを使用していることが示唆されている。また、Cao et al.（2019）は、セルサイド・ファイナンシャル・アナリストについての調査を行い、見た目が美しいアナリストは、そうでないアナリストに比べて、より正確な業績予測を行うという結果を得ている。その理由として、見た目が美しいアナリストは雇用主からの支援を受けやすく、内部の情報にアクセスしやすいこと、メディアへの露出の機会が増え顔を売ることができること、機関投資家等との繋がりを持つ機会が増えること、が挙げられており、その結果としてより多くの情報にアクセスし、業績の向上に役立てることができる可能性があるとしている。

また、容姿が本人の属する組織の評価、業績に及ぼす影響に関する先行研究も多々存在する。Pfann et al. (2000) では、オランダの広告会社数十社を対象とした調査で、容姿の美しい幹部がいる企業ほど収益性が高く、成長が早いという結果が報告されており、その結果として個人の収入も増加する。また 1996 年にスイスのビジネス誌が実施した「スイスで一番美しい CEO」を投票するアンケートでも、上位に入った企業は比較的規模の大きい企業であったという結果が報告されている。

直接容姿との関わりがあるかは不明なものの、経営者の表情等を基に分析した経営者のナルシシズム度が、企業行動に何らかの影響を及ぼすことも示唆されている。Saisho et al.（2018）では、日本企業の経営者を対象に、アニュアルレポート内の写真の数、大きさ、サインの大きさ、表情を用いて個人のナルシシズム度を測定し、企業行動との関係を調査したところ、ナルシシズム度が高い経営者は、より積極的な投資を行うという結果を得ている。ただし、その行動が企業の業績にとって正の影響をもたらすとは限らない。

以上の通り、容姿が個人や企業の業績に正の相関をもたらしている可能性を示唆する先行研究は少なくはないものの、日本を対象としたこのような研究は限定的であり、日本でも同様のことが言えるのかは、明らかではない。

先行研究の多くは、個人の容姿と本人の評価や業績との関係を分析したものであるが、容姿の美しさと個人の評価や業績が正の相関関係にあること関しては多くの研究でも示されている。一方、個人の容姿がその人物の所属する組織自体の評価や業績にどのような影響を与えているかということについての先行研究は少なく、Pfann et al. (2000)のオランダの広告会社の調査はあるものの、必ずしも明確な結論が得られているわけではない[1]。

また、容姿の測定に関し、先行研究では容姿のレベルを、アンケート回答者や分析者の主観かつ目視で段階分けしているものがほとんどであり、対象者との面接でレベルを定めているものも存在するため、美しさの指標として顔以外に服装や印象なども含まれるものもある。

本研究では、経営者の容姿が本人の属する企業行動もしくはその他の指標に与える影響を分析することとする。まず、本研究では、容姿を構成する重要な要素の一つである「顔」を対象とし、顔を定量的に分析することを試みる。その上で、経営者の容姿（顔）が企業の業績と行動にどのような影響を及ぼしているのかを分析する。

3. 分析手法

顔の分析には FaceNet と顔写真のデータセット VGGFace2 を用いた学習済みモデル、対象とする企業の代表取締役社長の顔写真を使用する（Schroff et al. (2015), Qiong et al. (2018) ）。

企業の代表取締役社長の顔写真については、対象企業のアニュアルレポート等から抽出し、サイズを

[1] 個人の容姿とその人物の所属する組織自体の評価や業績との関係は、業種や企業の規模によっても変わってくる可能性がある（Pfann/Biddle/Hamermesh/Bosman（2000））。

揃えて分析に用いる。企業行動を示す指標は,「設備投資額 売上高比率」,「研究開発費売上高比率」,「販管費率」,「負債比率」,「子会社・関係会社の株式取得額」,「子会社・関係会社の買収件数」を用いる。また、企業業績を表す指標に「ROA (総資産利益率)」を用いる[2]。

分析では、はじめに、対象となる人物の顔写真を用いて、それぞれの容姿レベルを設定する。対象者の顔写真各 2 枚を収集し、顔認識のニューラルネットワークである FaceNet の学習済みモデル VGGFace2 を用いて、それぞれの顔をベクトル化する。次いで、主成分分析により次元圧縮した後、K-means 法を用いてクラスタリングする。これにより対象者の顔をいくつかのクラスタに分類することが可能である[3]。

4. 分析結果

日経 225 銘柄採用企業のうち、113 社の代表取締役社長の顔写真各 2 枚を収集し、上記の手法を通じ顔写真のベクトル化および主成分分析を行った。図 1 は、各主成分の寄与率を示したものである。

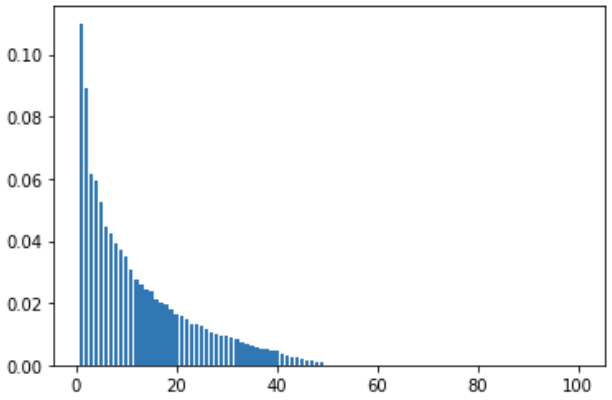


図 1：各主成分の寄与率（顔写真）

図 2 は、クラスタリングの結果を示したものである。 図中の、横軸は第一主成分、縦軸は第二主成分を示している。

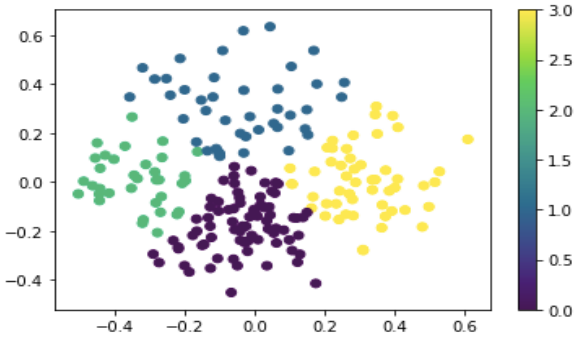


図 2　クラスタリングの結果（顔写真）

5. おわりに

本研究では日本において、経営者の外見が企業の行動に影響を与えるのかを解明することを目的とし、外見を構成する要素の一つである顔を対象として、対象企業の経営者の顔を定量的に分析することを試みた。本分析では、企業のアニュアルレポート等から抽出した顔写真を基に VGGFace2 などのモデルおよび K-means 法などを用いた分析を行った。

顔が企業行動に何らかの影響を及ぼしているとしても、企業行動はその企業の属性や財務状況によっても影響を受けるため、顔の影響は限定的である可能性がある。詳細な分析は、今後の課題である。

参考文献

[1] Daniel S. Hamermesh, Jeff E. Biddle., “Beauty and the Labor Market,” American Economic Review, vol 84, (1994), pp. 1174-1194.

[2] Jeff E. Biddle, Daniel S. Hamermesh., “Beauty, Productivity and Discrimination: Lawyers' Looks and Lucre,” Journal of Labor Economics, Vol. 16, No.1, (1995), pp. 172-201.

[3] Gerard A. Pfann, Jeff E. Biddle, Daniel S. Hamermesh, Ciska M. Bosman., "Business success and businesses' beauty capital," Economics Letters, Elsevier, vol. 67(2), (2000), pp. 201-207.

[4] Daniel S. Hamermesh, Amy M. Parker., “Beauty in the Classroom: Professors' Pulchritude and Putative Pedagogical Productivity,” Economics of Education Review, (2005), v. 24, issue. 4, pp. 369-76.

[5] Sean P. Salter, Franklin G. Mixon Jr, Ernest W. King., “Broker beauty and boon: A study of physical attractiveness and its effect on real estate brokers' income and productivity,” Applied Financial Economics 22(10), (2012), pp.811-825.

[6] Qiong Cao, Li Shen, Weidi Xie, Omkar M. Parkhi,

[2] これらの指標は Saisho et al. (2018)を基に設定した。

[3] 学習済みモデル VGGFace2 の精度の検証として、芸能人やスポーツ選手等の有名人の顔写真（８～１５枚×１２人分）を用いて上記の方法でクラスタリングを行った結果、90%程度の精度で同一人物が同一のクラスタに分類された。

Andrew Zisserman., “VGGFace2: A Dataset for Recognising Faces across Pose and Age,” 2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018), (2018).

[７] Atsutaka Saisho, Aiko Suge, Hiroshi Takahashi., “Analyzing the Relationship Between the Characteristics of Company Executives and Their Companies: Observation Through Facial Expressions,” International Conference on Innovation & Management (ICIM2018), (2018).

[８] Ying Cao, Feng Guan, Zengquan Li, Yong George Yang., “Analysts’ Beauty and Performance,” Management Science, (2019).

[９] Florian Schroff, Dmitry Kalenichenko, and James Philbin., "Facenet: A unified embedding for face recognition and clustering." Proceedings of the IEEE conference on computer vision and pattern recognition. (2015).

要約文章生成と意味類似度による株式市場におけるニュース記事の影響度の測定

Analysis the Effect of News Articles on the Stock Market by Summarizing and Semantic Similarity

野矢淳[1]　高橋大志[1]

Jun Noya[1], and Hiroshi Takahashi[1]

[1] 慶應義塾大学大学院経営管理研究科
[1]Graduate School of Business Administration, Keio University

Abstract: Stock market fluctuations are caused by various factors. This study focuses on news articles to improve the accuracy of stock market predictions. The effect of news articles is bigger when information that was not predicted is public, earning surprise. We have summarized past news articles by GPT-2 to discover news trends. We also converted the news articles into vectors and link them to stock prices by machine learning. As a result of this analysis, we were able to improve the accuracy of stock market predictions.

1 はじめに

株式市場の株価はさまざまな要因によって変動する．株価の挙動を説明するために，これまで数多くの研究が行われてきた．しかし複雑な株式市場をモデル化することは困難であり，現在でも株価の動きを完璧に説明することはできていない．

株価の変動は，一般的にこれまでに市場では認識されておらず，予測しづらい情報が発表されたときに特に大きく反応する．これをアーニングサプライズという．アーニングサプライズに関連した研究として，アナリスト数や企業業績によって検証したもの[1]や，アーニングサプライズが将来どのように株価と関係するか検証したもの[2]などが存在する．

また，株価は景気や企業業績，市場の動きなどに影響を受けて変動する．これらの情報を受信する媒体として代表的なものに有価証券報告書や四季報，ニュースなどがある．とりわけニュースは誰でもアクセスできる情報源であり，個人のような小さい投資家から機関投資家までが活用している．ニュースと株価の関係性を扱ったものは数多く存在している[3]．また，近年は機械学習モデルを使った研究も盛んに行われている[4][5].

そのため本研究では，ニュース記事を要約することで算出されるアーニングサプライズをモデルに組み込むことで，株価の説明モデルの精度の向上を目的としている．その手法として，近年の機械学習の発展により自然言語処理で文章の意味を理解することが可能となったため，言語生成モデルでニュース記事の要約文章を生成し，文章のベクトル化によって類似度を求めて分析する．

2 データ

本研究では，要約文章生成のためにニュース記事と株価の結びつけに株式市場データを使用する．それぞれ以下通りである．

2.1 ニュース記事データ

ニュース記事データは，トムソン・ロイター社が提供するロイターニュースを用いた．トムソン・ロイター社は世界最大級のマルチメディア通信社であり，日本においても幅広いニュースを提供している．特にトムソン・ロイター社の報道スピート，正確性，信頼性は高い評価を得ており数多くの投資家が活用する情報源である．本研究では，2016 年のシャープ株式会社に関連するニュース記事の発信日時，本文を使用した．

2.2 株式市場データ

株式市場データは，東京証券取引所におけるティックデータを用いた．ティックデータとは，株式の約定価格や約定数量の推移を示した，時系列の取引データのことである．本研究では，2016 年の取引時間と約定価格を使用した．

3 分析手法

本研究の分析手法の構造を，図 1 に示す．まず一定期間のニュース記事から要約文章を生成し，次の期間に発表されたニュース記事との類似度を比較した．その後，算出された類似度より株価との関連性を分析した．要約記事の作成には GPT-2 を用いて，比較のための文章ベクトルの生成は Bert を用いた．また，株価との関連性は，ニュースが発表される前後の株価の変動率を算出し，ラベル付を行った上で，NN を用いて分析を行った．

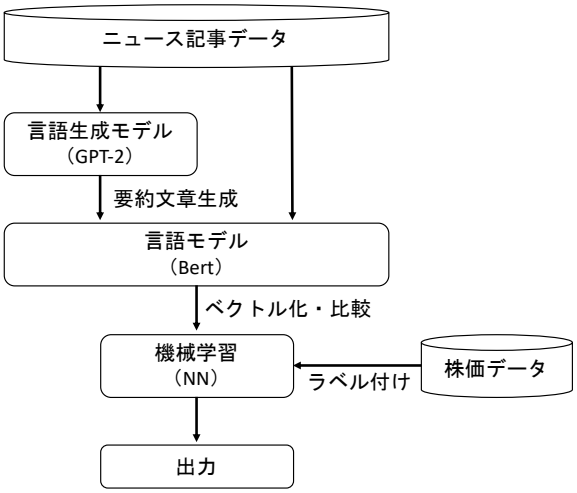


図 1　分析手法の構造

3.1 GPT-2 を用いたニュース記事の要約

Generative Pretrained Transformer 2 (GPT-2)[6]は，汎用人工知能の研究をするアメリカの非営利団体である Open AI が開発した言語生成モデルである．世界中のインターネット上にある 800 万の Web サイトの文章を学習しており，読解や翻訳，要約などのタスクを得意とする．

本研究では関連銘柄のニュース記事を読み込み，要約文章を生成するのに用いた．要約する期間は一ヶ月毎として，その月のニュース記事の要約文章を生成した．

3.2 Bert を用いた文章のベクトル化

Bidirectional Encoder Representations from Transformers(Bert)[7]は, Google が開発した言語モデルである．事前学習した後に転移学習によって様々なタスクを解くことができる．このモデルの最大の特徴は文章の文脈を理解することであり，文章の意味比較や．他の文章が続く可能性などを計算できる．

本研究ではニュース記事を比較するために，文章を多次元ベクトル化するのに用いた．

3.3 NN による分析

Neural Network(NN)は脳のように動作する機械学習モデルであり，複数のノード（ニューロン）が結合して構成されている．単一のノードにおいて，複数の入力（$x_0, x_1 \ldots x_n$）をすると，一つの出力(Y)をする（図２）．これらのノードを複数組み合わせて最適な値を算出する．本研究では文章の類似度を比較したベクトルと株価との結びつけを行うのに用いた．

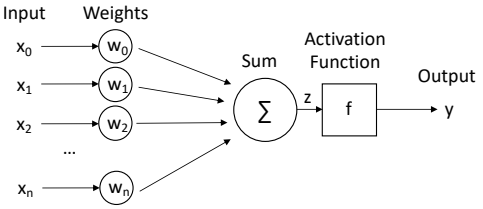


図 2　NN のノードに対する入力と出力

4 分析結果

本研究では，文書をベクトル化した後に機械学習で分析した Model 1，要約文章との類似度のみを機械学習で分析した Model 2，要約文章との類似度及び文書をベクトル化した両方を機械学習で分析した Model 3 を用いて，正答率を比較した．

4.1 要約文章の生成

GPT-2 を用いて，一年間のニュース記事を一ヶ月毎に要約した．その結果の例を，図 3 に示す．これより可読性の高い要約文章を生成できたことが分かる．

> The plan to raise Sharp's capital, saying it could raise 2 trillion yen ($18.85 billion) with investors. Mr. Murakami, former manager of the fund, was also on the call. The government is unlikely to help Sharp's plans, some experts say that is a surprise, but the government should be helpful for Sharp. The government should offer Sharp and take Sharp Sharp and Hon Hai and SoftBank.

図 3　生成した要約文章の例

4.2 文章のベクトル化と分析

Bert を用いて文章をベクトル化して要約文章との比較を行い NN で株価との関連性の分析を行った．それぞれのモデルの結果は図 4 の通りである．正答率は，Model 1 が 55%，Model 2 が 35%，Model 3 が 57%となった．これより，要約文章との類似度を組み込んだモデルである，Model 3 の結果が一番高い正

答率であることが確認できた．これらの結果は，類似度の情報が資産価格評価に貢献できる可能性を示すものである．

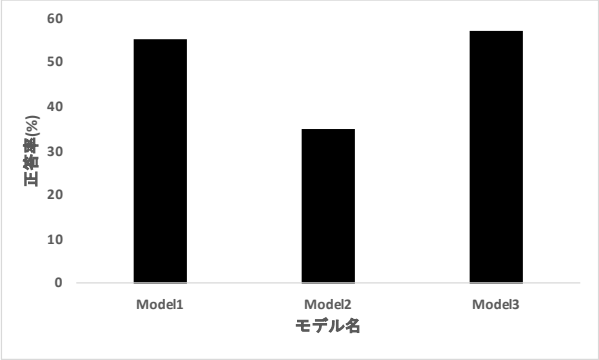


図 4　NN による分析結果

5 まとめ

本研究は, ニュース記事の要約文章生成によって, 株価の説明モデルを構築した．GPT-2 で要約文章を生成した後，Bert でベクトル化して NN で分析を行った結果，類似度と文章ベクトルを組み込んだモデルの精度が最も高いことを見出した．これらの結果は，ニュース記事と要約文章の類似度を通じサプライズを組み込むことで，株価の説明モデルの精度の向上の可能性を示すものである．

今後はデータ数を増やして結果をさらに検証する必要がある．また要約文章を生成するにあたって GPT-2 の後継モデルである GPT-3 を使うことを検討している．

参考文献

[１] Abraham, Rebecca, Harrington, Charles: Predictors of the Degree of Positive Earnings Surprises, Open Journal of Accounting, Vol. 05, pp. 25-34, (2016)

[２] Panos Patatoukas, Hongjun Yan: "The Impact of Earnings Surprises on Stock Returns: Theory and Evidence,", Yale School of Management, (2009)

[３] Gidófalvi, Győző: Using news articles to predict stock price movements, University of California, San Diego, (2001)

[４] Joshi, Kalyani, N, Bharathi, Rao, Jyothi: Stock Trend Prediction Using News Sentiment Analysis, International Journal of Computer Science and Information Technology, Vol. 8, pp. 67-76, (2016)

[５] M. G. Sousa, K. Sakiyama, L. d. S. Rodrigues, P. H. Moraes, E. R. Fernandes, E. T. Matsubara: BERT for Stock Market Sentiment Analysis, 2019 IEEE 31st International Conference on Tools with Artificial Intelligence (ICTAI),, pp. 1597-1601, (2019)

[６] Radford Alec, Wu Jeffrey, Child Rewon, Luan David, Amodei Dario, Sutskever Ilya: Language models are unsupervised multitask learners, OpenAI Blog, Vol. 1(8), pp. 9, (2019)

[７] Jacob Devlin, Ming-Wei Chang, Kenton Lee, Kristina Toutanova: BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding, NAACL-HLT 2019, (2018)

[８] Yoshihiro Nishi, Aiko Suge, Hiroshi Takahashi: Text Analysis on the Stock Market through “Fake”News Generated by GPT-2, INFORMS Annual Meeting, (2019)

[９] Yoshihiro Nishi, Aiko Suge, Hiroshi Takahashi: Construction of News Article Evaluation System Using Language Generation Model., In: Jezic G., Chen-Burger J., Kusek M., Šperka R., Howlett R., Jain L. (eds) Agents and Multi-agent Systems: Technologies and Applications 2020. Smart Innovation, Systems and Technologies, vol 186. pp.355-366, Springer, (2020)

働き方と生産性の相互作用分析のためのミーティング検知手法

Meeting Detection Method for Interaction Analysis between Work Style and Productivity

矢田 昇平[1]　加藤 大望[1]　倉橋 節也[1]

Shohei Yada[1]　Daimotsu Kato[1]　Setsuya Kurahashi[1]

[1]筑波大学大学院 ビジネス科学研究科 経営システム科学専攻
[1]University of Tsukuba

Abstract: Labor shortage due to population decline is a social issue in Japan.
On the other hand, in April 2019, the “Work Style Reform Laws” was enacted, and diversification of work styles and improvement of productivity can be seen as a management issue for Japanese companies.
In this paper, we propose a method to detect various work situations and meeting occurrences of employees, in order to find out the relationship of interaction between "work style" and "productivity", using location information data obtained from in-house Wi-Fi connection.　In addition, by analyzing the network based on the information of the detected meeting, the network in the organization can be visualized.
In this way, we aim to identify a model that maximizes organizational productivity by visualizing the way employees work styles.

1. 研究の背景

日本は少子高齢化や人口減少に伴う労働力不足が深刻化している. [1] 加えて就業時間あたりの労働生産性も著しく低く, OECD の発表によると日本は主要先進 7 カ国の中で 1970 年以降常に最下位となっている. [2] こうしたことから日本経済は「量」「質」共に危機的状況にあり, このことは社会課題ともいえるであろう.

一方で, 2019 年 4 月には『働き方改革関連法』が施行され, 生産性向上とともに就業機会の拡大や意欲・能力を存分に発揮できる環境をつくることが求められており, 企業側からみるとこのことは経営課題とも捉えられる.

このように, 働き方の多様性を受け入れながらも組織レベルでの生産性を高めていかなければならないなかで, 組織経営は未だに経験・勘・慣習などといった非科学的なアプローチが中心ではないだろうか.

2. 研究の目的

本論文の目的は「個人の働き方」と「組織の生産性」との相互作用の関係を見出すために, オフィス内での従業員間のさまざまなミーティングの発生状況を検出することにある. そして, その結果を用いて組織の生産性を最大化するモデルを特定することを目指す.

ミーティングの発生状況を検知するために, オフィス内における従業員の位置情報を利用する. 本研究対象の環境は, 東京都内にあるオフィスビルの執務フロア内であり, 基本的には全てのフロアでフリーアドレス制が導入されている. 従業員はノート PC とスマートフォンが支給されており, フロア内においては Wi-Fi 接続によって業務を行うのが標準である. この Wi-Fi 接続情報を用いてフロア内における位置情報を取得し, さまざまなタイプの異なる執務状況やミーティングの発生状況を検出する.

3. 先行研究との比較

執務エリア内での位置情報や対面情報を用いた先行研究としては, 名刺型のウェアラブルセンサーデバイスを用いて日立研究所が実施した研究[3]や, Wi-Fi の接続情報を用いてヤフー株式会社の山下らが実施した研究[4]があるが, 本研究は後者の研究を発展させる形で行うものである. 山下らの研究では, 執務フロア内の位置情報から離合集散モデルに基づくミーティング発生の検知手法を提案しているが, 本研究における発展性は主に以下 2 点である.

一つ目は, ミーティングがオフィス内のどのようなエリアで発生しているのかを特定している点である. これは, 座標系の位置情報データに, フロアレイアウトのエリア区分を重ね合わせて計測することで実現する.

二つ目は，検知したミーティングを従業員間のネットワークと捉え「ネットワークの中心性」を特徴量抽出している点である．

4. 研究手法

本研究の対象とする環境は，東京都内にあるオフィスビルの執務フロア内であり，基本的には全てのフロアでフリーアドレス制が導入されている前提である．従業員の位置情報の取得には，オフィス内の位置情報システム「pozzy」のログを用いる．[5]

現在この位置情報データは，フリーアドレス環境下で従業員の居場所を検索したり，各フロアの混雑状況を確認したりといった執務のうえで必要な形で利用されているが，本研究においてはこの位置情報データから個人を特定できる情報を一切排除することで匿名性を担保しつつ，ミーティングの検出とネットワークの中心性を抽出する．以降で，研究の具体的な手法を記す．

4.1. 執務場所の特定

まずはオフィスのレイアウトデータを分析環境に読み込む．全てのフロアは233 Feet（約71 メートル）四方の間取りである．このフロアレイアウトを二次元の XY 軸座標平面として扱い，左上の座標軸$(0, 0)$地点から最大$(233, 233)$まで正の値を持たせる．

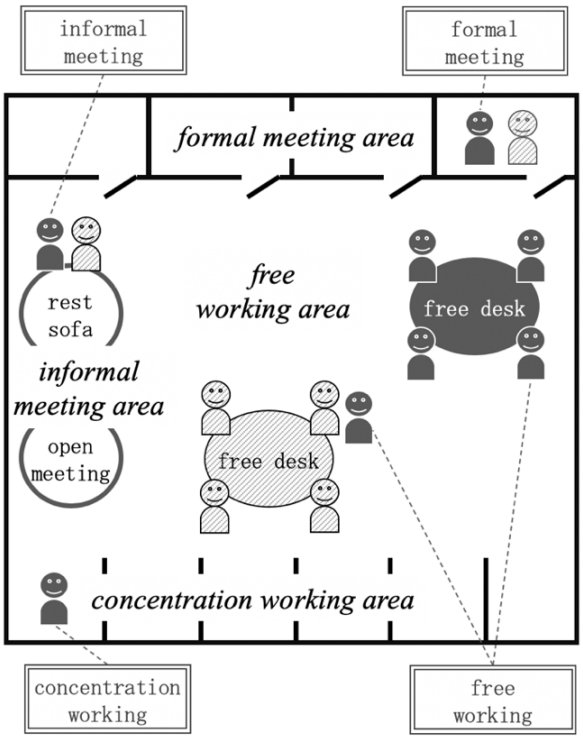
このフロアレイアウトに対し，エリア区分を付加することでその地点の属性を得ることができる．エリア区分は，実際のオフィスの利用状況に応じて「*formal meeting area*」「*informal meeting area*」「*concentration working area*」それ以外を「*free working area*」とする 4 区分であり，その概念を図 1 に記す．

こうして特定したエリア区分付きのフロアレイアウトに対して，各従業員の位置情報データを重ねることで位置情報データにエリア区分を付加し，フリーアドレス環境下において「どいうった場所で業務を行なっているか」といった働き方を表現する特徴量を抽出する．

4.2. ミーティングの検出

山下らの先行研究におけるミーティング検知の定義設定は，人数 2 人以上で上限なし（1 対 1 の立ち話等も考慮），場所はフロア内であればどこでも，時間は 5 分以上 60 分以内で，複数デバイスの離合集散の動きからミーティングの発生を検知しネットワークグラフ化することで，コミュニケーションの可視化を試みたが，本研究ではこの先行研究のミーティング検知手法を踏襲しつつ，4.1 で用いたエリア区分を，ミーティングが発生した位置情報においても適応することで，「どういった場所でミーティングが行なわれているか」といった働き方に関する特徴量を抽出した．ここまでで説明したエリア特定の概念を示したのが図 1 である．

図 1：エリア区分の概念図



4.3. ネットワークの中心性の抽出

先の通り 4.2 で検出したミーティングデータを用いて，組織内における中心的な役割を定量的に算出するために，2 種類のネットワークの中心性を抽出する．1 つ目は「次数中心性(dgc)」である．これは，それぞれの従業員をネットワークのノードと捉えたときに，多くのノードと隣接しているノードを中心的とする指標である．次数とは各ノードに隣接しているエッジ数のことであり，ノードvの次数を$deg\ (v)$としたとき，次数が高いノードほど重要とする．

$$dgc(v) = deg\ (v)$$

もう 1 つは，ノードの 2 点間を結ぶ経路上にしばしば現れるノードを中心的とする「媒介中心性(bwc)」である．任意のノードペア間の最短パスのうち，媒介しているパスの割合によりノードをランキ

ングするものであり，$\sigma_{s,t}$はノードs,t間の最短パス数，$\sigma_{s,t}(v)$はノードvを通るノードs,t間の最短パス数としたとき以下の式で表すことができる．

$$bwc(v) = \sum_{s \in V} \sum_{t \in V} \frac{\sigma_{s,t}(v)}{\sigma_{s,t}}$$

なお，これらの中心性については，ミーティングが発生した「*formal meeting area*」「*informal meeting area*」「*free working area*」別に合計で 6 パターン算出することとした．その理由は，会議室で行われるミーティングだけではなく，開放的な場で行われるオープンなミーティングや，突発的に発生する立ち話のようなコミュニケーションがもたらす効果も検証するためである．

5. 分析結果

今回の分析に利用したデータの期間は，2019 年 11 月 1 日から 11 月 7 日までの一週間の位置情報データである．この期間で抽出された全従業員の 5 分おきの位置情報のデータ数は 5,060,524 レコード，ミーティングの発生回数は，ミーティングであると検知するデバイス間の距離を 10Feet（約 3 メートル）と設定した場合で 24,835 レコードであった．

5.1. 執務場所特定の精度

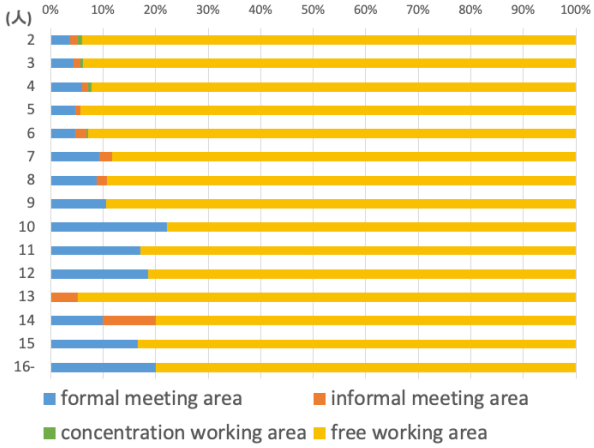
サンプルとして，あるひとりの従業員の執務場所検知結果と，実際の当時のスケジュールを比較してみたところ，検知された 49 のミーティング実施エリアでのデータ数に対して 83.7%にあたる 41 のデータが当時のスケジュール内に収まる結果となった．ただしこれは，あくまで検知されたデータに対する検証であるため，本来であれば当時の正確なスケジュールすべてに対する検証が必要である．これは検証数の向上と合わせてさらに分析を進める予定である．

5.2. ミーティング発生状況

該当期間における集合人数別のミーティング発生回数をみると，全体の 66%にあたる 16,315 回が 2 名によるミーティング（1 対 1 のミーティング）であった．研究対象組織においては上司と部下による定期的な「1on1」が組織文化的に推奨されており，この結果にはある程度納得ができる．

また，この集合人数別のミーティング発生回数を，エリア区分別にみてみると図 2 のようになった．6 名程度までの比較的少人数のミーティングは，ほとんどが「*free working area*」で発生しており，10 名を超えたあたりから，「*formal meeting area*」での発生比率が高まっていることが見て取れる．

図 2：集合人数別/エリア区分別 MTG 発生比率

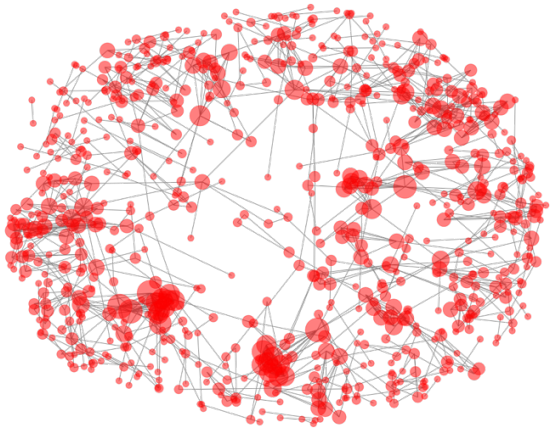


5.3. ミーティングネットワークの中心性

該当期間における「*informal meeting area*」で発生したミーティングを次数中心性によってネットワークグラフにしたものが図 3 である．それぞれの従業員を赤い円（ノード），コミュニケーションのパスを灰色の線（エッジ）で表しており，ノードの大きさはここでは次数中心性に応じて変化させている．

これをみるとネットワークの所々で次数中心性の高いノードが点在していることがわかる．

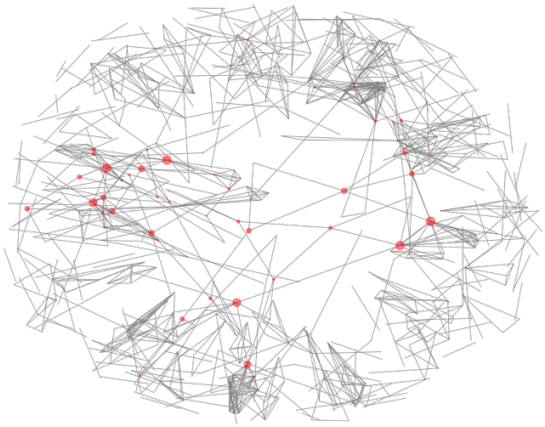
図 3：次数中心性によるネットワークグラフ



次に，同じ「*informal meeting area*」で発生したミーティングを媒介中心性によってネットワークグラフにしたものが図 4 である．媒介中心性の値が高いノードがまばらに発生しており，次数中心性とは異なる分布があるように思われる．

ただし，定義が異なる中心性間で，中心性の強さを単純比較することはできないため，この差が意味するところを，今後分析を進める予定である．

図 4：媒介中心性によるネットワークグラフ



6. 今後の展望

この結果を踏まえて本研究は以下の通り発展させていく計画をしている．

6.1. 目的変数及び媒介変数の考慮

これまでに抽出した「働き方」のデータを説明変数として扱い，これに対する目的変数として「生産性」を加えることで「働き方」と「生産性」の関係性を明らかにできると考えている．この「生産性」には半期ごとに実施する個人および組織の 11 段階業績評価を用いることを検討している．業績評価を生産性として扱う理由としては 2 点ある．

1 点目は全従業員を一律で相対比較できる評価指標が存在しないからである．営業職から技術職まで幅広く在籍する企業においては致し方ない側面である．2 点目の理由は，業績評価の設定思想が，全社の定量的な業績目標を組織・個人までブレイクダウンするという設計思想になっているためである．

続いて，個人の「働き方」が例えば職位や年齢，性別，またはモチベーションといった外的要因からの影響を受けている可能性も無視できない．そのため「個人の属性」や「組織満足度」といった媒介変数を加えることで，組織として生産性を最大化する組織形成モデルを明らかにしたいと考えている．

6.2. 働き方と生産性の因果推論

ここまでの研究過程においては，変数間に有意な相関がみてとれたとしても，その因果関係に言及することは難しい．そこでベイジアンネットワークを用いて「組織の生産性」を向上する「働き方」を導出することを試みる．

6.3. シミュレーションモデル化

ここまでで検証できるのは，あくまで個人の働き方が，当人の生産性にどう寄与しているかであるが，本来は働き方の多様性が組織の生産性をどのように高めるかを明らかにしたいと考えている．

またこれらが相互作用の関係にあると仮定すると，個人の働き方を「ミクロの行動パターン」として，組織の生産性を「マクロの社会現象」として階層的に捉え，今回の分析結果をパラメータとして用いることで，エージェントベースモデルによるシミュレーションが設計可能では無いかと考えている．

そうすることで，ミクロ・マクロループが引き起こす構造変化のメカニズムを創発的に発見する発展的な研究を計画している．

7. 参考文献

[1] 内閣府: 平成 30 年版 高齢社会白書, 第 1 章 第 1 節 生産年齢人口の将来推計
[2] 公益財団法人 日本生産性本部: 労働生産性の国際比較 2019
[3] Satomi Tsuji, Nobuo Sato, Kazuo Yao, et al (2019). “Employees’ Wearable Measure of Face-to-Face Communication Relates to Their Positive Psychological Capital, Well-Being”
[4] 山下 達雄, 寺岡 照彦, 田口 拓明 (2020). Wi-Fi 接続による屋内位置情報を用いた人間関係の抽出.
[5] ヤフーの社内システムを紹介します, Yahoo! JAPAN Tech Blog, 2016 年 12 月 5 日 . https://techblog.yahoo.co.jp/advent-calendar-2016/pozzy/
[6] 村田 剛志（2019 年）. Python で学ネットワーク分析, オーム社
[7] 和泉 潔 他（2017 年）. マルチエージェントのためのデータ解析, コロナ社
[8] 出口 弘 (2004). エージェントベースモデリングによる問題解決ーエージェントベース社会システム科学としての ABM 一, オペレーションズ・リサーチ, 2004 年 3 月号, p161-167
[9] 伏見 卓恭 他（2012 年）. ノード集合に対する媒介中心性の提案.

日本のガス市場における ダイナミック・プライシングモデルに関する研究

A Study on the Potential of Dynamic Pricing Models in the Japanese Gas Market

許 愼碩 [1]　村上 英治 [1,2]　高橋 大志 [1]

Myeongseok HEO [1] Eiji MURAKAMI [1,2] and Hiroshi TAKAHASHI [1]

[1] 慶應義塾大学大学院経営管理研究科　[2] アズビル金門(株)
[1] Graduate School of Business Administration, Keio University [2] Azbil Kimmon Corporation

Abstract: Gas and electric power are one of the essential social infrastructure for social and economic activities. The purpose of this research is to construct a Real Time Pricing (RTP) model suitable for the distribution and sales market of LP gas in Japan by utilizing the Dynamic Pricing Model. As a method of predicting demand, we would like to first create a model of power consumption and then apply it to LP gas. Analysis of customers' electric power and gas consumption can infer from what standpoint they were consuming public goods. Through this, we would like to offer users' reasonable prices to expand their choices and contribute to the efficient use of social resources.

1. 研究背景

ガス、電気は、社会、経済活動において必要不可欠な社会インフラの一つである。

本研究はダイナミック・プライシングモデルを活用し、日本の LP ガスの流通及び販売市場に適した Real Time Pricing (RTP)モデルを構築することを目的とする。

電力市場に関しては、ダイナミック・プライシングモデルに関する報告が、数多く行われているのに対し、ガス市場に関しては、相対的に分析報告の数は限定的である。本研究では、ガス市場を対象としたダイナミック・プライシングモデルについて検討を行う。

次節において関連研究について触れたのち、基礎分析の結果を示す。4. は、まとめである。

2. 関連研究

2.1 ガス消費量と電力消費量

近年、電力業界では、消費量の多い時間帯の電力消費を少なくするというピークカットが行われている。これに関連して様々なプライシングモデルが提案されており、特にダイナミック・プライシングモデルが注目を集めている。

ダイナミック・プライシングモデルは、リアルタイムに需要と供給を把握し、最適化された価格を算出するものである[1]。このようなモデルを使用することで個々の消費者の特性や個別状況が反映された価格を算出することが可能である。更に、ダイナミック・プライシングを通じ消費者が感じている価値をより正確に価格に反映することで、既存のステイティック・プライシングに比べて効率性の向上に貢献することが期待される。

2.2 日本のガス市場の現状

日本の LP ガス市場は自由料金制である[2]。自由料金制であるので、原油価格の上昇によって LP ガス料金も変動的に動くことが考えられるが、事業者によっては原油価格が下落してもガス料金を下げない場合も存在する。これらの要因の一つとして LP ガスの価格には公表義務がなく、市場の閉鎖性のため適正価格より高い価格が設定されたとしても消費者は気付きづらいことが挙げられる。

1996 年から消費者はガス会社を自由に選ぶことができるようになっている。しかし、現実の選択には、様々な要因が存在する。例えば、住宅を新築す

る際にプロパンガス設備の導入費用をガス事業者が立て替えるため長期契約を締結する場合が多い。そのような場合、適正価格より高い価格の LP ガスの利用を促される可能性がある[1]。

日本のガス市場におけるダイナミック・プライシングモデルの開発においては、こうした日本の LP ガス市場の特殊性と現実性を勘案する必要性がある [3]。

本研究では、電力市場などにおける先行研究の成果を取り込み、日本のガス市場に関する研究を行う [3-7]。

3. 基礎分析

LP ガス消費量の季節ごとの消費量を分析したところ、2015~2019 年の 5 年間の LP ガス消費量から、僅差で夏（7~8 月）の消費量が冬（12~1 月）の消費量より高いことを確認した[2]。

また、使用料データを基に顧客をグループ化することで、顧客群別のカスタマイズ対応サービスが可能になる。本分析では、電気使用量データを基に顧客のクラスター分析を実施し、可視化を行った [5]。同様の分析はガス使用量を用いても可能であり、今後の課題である。

4. まとめおよび今後の課題

本分析では、日本のガス市場におけるダイナミック・プライシングモデル構築に向けた検討を行った。現実のデータを用いた詳細な分析は今後の課題である。また、本分析においては、事業者、消費者等、複数のステークホルダーが存在する。これらを考慮したエージェントベースモデルおよび人間参加型アプローチによる分析やファイナンス分野の知見を取り込んだ分析も、今後の課題の一つに挙げられる [9-15]。

参考文献

[1] Walter R. Paczkowski: Pricing Analytics, Models and Advanced Quantitative Techniques for Product Pricing, (2018)

[2] 江田健二: かんたん解説!! 1時間でわかるガス自由化入門, https://pps-net.org/tainavi-interview (2017 年 6 月)

[3] 村上英治: 価値創出を指向するメーターデータプラットフォームガスミエール TM, Azbil Technical Review, pp. 20-23, (2020 年 4 月)

[4] Yasirli Amri, Amanda Lailatul Fadhilah, Fatmawati, Novi Setiani, Septia Rani: Analysis Clustering of Electricity Usage Profile Using K-Means Algorithm, IOP Publishing, (2016)

[5] https://archive.ics.uci.edu/ml/datasets/ElectricityLoadDiagrams20112014

[6] Sebastián de la Torre, José M. Arroyo, Antonio J. Conejo, and Javier Contreras: Price Maker Self-Scheduling in a Pool-Based Electricity Market: A Mixed-Integer LP Approach, IEEE TRANSACTIONS ON POWER SYSTEMS, Vol. 17, No. 4, pp. 1037-1042, (NOVEMBER 2002)

[7] Alireza Rahimi Vahed, Teodor Gabriel Crainic, Michel Gendreau, Walter Rei: A Path Relinking Algorithm for a Multi-Depot Periodic Vehicle Routing Problem, CIRRELT, pp.1-33, (2013)

[8] 経済産業省資源エネルギー庁, 石油等消費動態統計年報, (2015, 2016, 2017, 2018, 2019)

[9] Axelrod,R: The Complexity of Cooperation Agent-Based Model of Competition and Collaboration, Princeton University Press (1997)

[10] Epstein, J.M., Axtell, R.: Growing Artificial Societies Social Science From The Bottom Up. MIT Press (1996)

[11] Hiroshi Takahashi, and Takao Terano: Agent-Based Approach to Investors' Behavior and Asset Price Fluctuations in Financial Markets, Journal of Artificial Societies and Social Simulation, no.3, Vol.6, (2003)

[12] 山下泰央, 高橋大志, 寺野隆雄：ビジネスゲームによるファイナンスへの接近：金融資産への投資の意思決定の学習, コンピュータソフトウェア, pp.30-40, No.4, Vol.25 (2008)

[13] Hiroshi Takahashi: An Analysis of the Influence of dispersion of valuations on Financial Markets through agent-based modeling, International Journal of Information Technology & Decision Making, vol.11, issue 1, pp. 143-166, 2012.

[14] 菊地剛正, 國上真章, 高橋大志, 鳥山正博, 寺野 隆雄 : 経営意思決定表現モデルを用いたビジネスケースとエージェントモデルの意思決定過程の形式的記述, 情報処理学会論文誌, 10, 60, pp.1704 -1718 (2019)

[15] 井上光太郎、高橋大志、池田直史: ファイナンス, 中央経済社 (2020)

[1] 異なる LP ガス事業者を選ぼうとすると、多額の違約金が請求される可能性もある。

[2] 本稿では、経済産業省資源エネルギー庁より公表されている石油等消費動態統計を用いた。

アンケートデータを用いた 退職前後世代の資産状況・家計収支の分析

Analysis of Financial Situation of Retirement Generation using Questionnaire Data

菊地 剛正 [1,2]　高橋 大志 [1]

Takamasa Kikuchi [1,2] and Hiroshi Takahashi [1]

[1] 慶應義塾大学
[1] Keio University
[2]MUFG 資産形成研究所
[2] MUFG Financial Education Institute

Abstract: With the aging of society, the problem of asset formation before and after retirement is attracting attention. The issues related to asset depletion are often analyzed based on macro statistical data such as financial assets and disposable income. Whereas such analysis does not necessarily reflect the attributes of each individual so there is room for improvement. In this paper, as a first step to analyze the characteristics of the assets and household income before and after retirement, we conducted a basic analysis using individual questionnaire data. Specifically, after showing the summary statistics of the above data, cross tabulation was performed according to the attributes of the questionnaire respondents. We analyzed the tendency of various asset statuses and stances toward investment due to differences in residential area and household composition.

1　はじめに

金融庁の審議会報告書[1]に端を発し，いわゆる「老後資金 2000 万円問題」，退職前後世代の資産形成と取り崩しの問題に注目が集まっている[2]. しかし，例えば，資産の取り崩し・引き出しについての議論がほとんどなられていないとの指摘があるなど[3], 基礎となる調査･分析を高度化する余地がある.

資産の枯渇に係る論点については，金融資産額や可処分所得などのマクロ統計データを基にした分析が多い[1; 4; 5]. また, マーケットデータを基にして，保有資産の価格変動が資産の枯渇に与える影響を分析したものもある[6]. これらの先行研究については，個々人の資産状況やライフスタイルなどの属性を十分に反映するという点で，改善の余地がある.

本稿では，退職前後世代の資産や家計収支の状況の特徴を分析するための第一歩として，個票のアンケートデータ[7; 8]を用いた基礎的な分析を行う．具体的には，当該データの要約統計量を示した上で，アンケート回答者の諸属性に応じたクロス集計を行う．居住地域や世帯構成の違いにより，各種資産状況や投資に対するスタンス・老後の見通しがどのような傾向となるか分析を行う.

2　データ概要

本稿では，MUFG 資産形成研究所が行った「退職前後世代の老後の生活に関する意識調査」[7; 8]の個票アンケートデータを使用する. 当該アンケートは，株式会社マクロミルの WEB アンケートにより収集された．調査期間は 2019 年 1 月 22 日から同 25 日，調査対象は 50 歳以上の男女, 調査地域は全国であり，有効回答数は 6,192 サンプルであった．アンケート項目と選択肢については，Appendix に付す通りである．当該アンケートは，個々人の資産状況（現在の資産残高や想定される老後の収入・支出）のみならず, 予定される資産承継額や投資に対するスタンス，老後の見通しなどを包括的に調査している点に特徴を有する.

なお，世帯金融資産ごとの構成比は，総務省家計調査報告（貯蓄・負債編）「高齢者世帯の貯蓄現在高階級別世帯分布(二人以上の世帯, 2017 年)」[9]を参考に割り付けられている．また，世帯金融資産のレンジ内の男女・年代別の構成比は，均等割り付けとなっている.

3　分析結果

3.1　要約統計量

アンケート項目に対する要約統計量は下表の通りである(表 1). 以降の分析では，全項目が利用可能な 4,601 サンプルを使用する.

3.2　クロス集計

(1) 居住地域毎の特徴

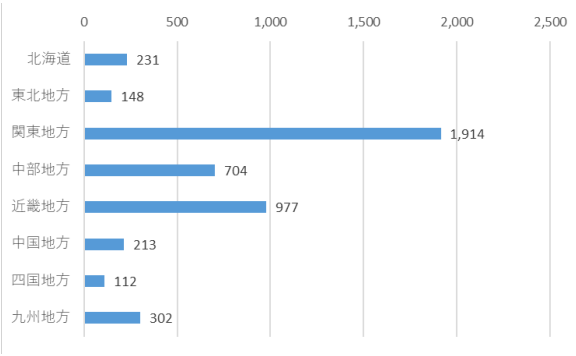


図 1a 居住地域毎のサンプル数

居住地域毎のサンプル数を示したのが図 1a である．関東・近畿・中部地方が多いことがわかる.

＜資産状況＞

居住地域毎に現在の金融資産残高の階層を比較したものが図 1b である．600 万円以上の層の割合は，関東・中部・近畿地方が多い．他方，600 万円未満の層の割合は，東北・北海道・中国地方が多い.

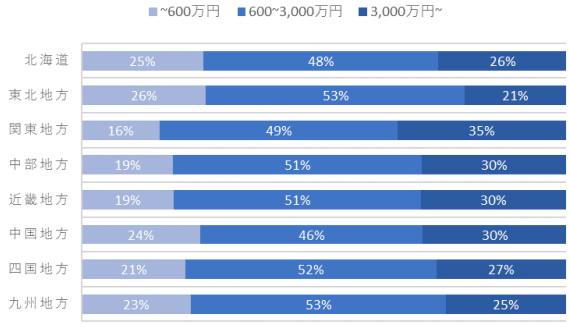


図 1b 居住地域毎の現在の金融資産残高階層

次に，居住地域毎に承継予定の金融資産残高の階層を比較したものが図 1c である.

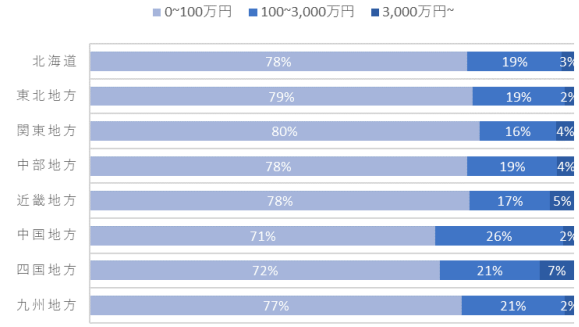


図 1c 居住地域毎の承継予定の金融資産残高階層

いずれの地域も，承継予定なしを含む 100 万円未満の階層が全体の 7～8 割を占めている．他方，100 万

表 1 アンケート項目の要約統計量（抜粋）

項目分類		項目名	要約統計量						
大	中		最頻値	中央値	最大値	最小値	第1四分位	第3四分位	サンプル数
基本属性	－	性別	男,女	－	－	－	－	－	6,192
		年齢	70	64.5	91	50	57	71	6,192
		地域	関東地方	－	－	－	－	－	6,192
		職業	無職	－	－	－	－	－	6,192
		世帯構成	夫婦のみ	－	－	－	－	－	6,192
		住宅種別	持家戸建	－	－	－	－	－	6,192
		婚姻状況	既婚	－	－	－	－	－	6,192
		子の人数	2人	2人	5人~	0人	1人	2人	6,192
資産状況	ストック	金融資産額(現状)	3~5,000万	1.5~2,000万	1億~	~100万	7~800万	3~5,000万	6,192
		金融資産額(承継予測)	0万	0万	1億~	0万	0万	0万	5,342
		退職金額(予測/実績)	0万	~1,000万	3,000万~	0万	0万	1.5~2,000万	5,743
	フロー	定期収入(現状)	1~300万	3~500万	1億~	~100万	1~300万	5~1,000万	6,090
		定期収入(老後 予測/実績)	~20万	2~30万	80万円~	~20万	~20万	2~30万	5,862
		定期支出(老後 予測/実績)	2~30万	2~30万	80万円~	~20万	~20万	2~30万	6,006
	その他	資産寿命	75~80歳	75~80歳	100歳~	50~55歳	70~75歳	85~90歳	6,192
		資産取り崩しの見通し	賄える	－	－	－	－	－	6,192
健康状況	－	生命寿命	80~85歳	80~85歳	100歳~	50~55歳	75~80歳	85~90歳	6,192
		健康寿命	75~80歳	75~80歳	100歳~	50~55歳	70~75歳	80~85歳	6,192
投資状況	－	投資経験	現在有	－	－	－	－	－	6,179
		リスク量	0%	0%	90%~	0%	0%	20~30%	6,068

円以上の資産承継を見込む層は，相対的に中国・四国地方が多い．

＜投資スタンス・老後見通し＞

投資に対する意識について比較したものが図 2a である．「現在投資をしている」または「将来投資をする予定である」層，つまり，投資にポジティブな層は，相対的に関東・中部・近畿・東北地方が多い．一方で，「現在も将来も投資するつもりがない」ネガティブな層は，北海道において過半数を占めている．

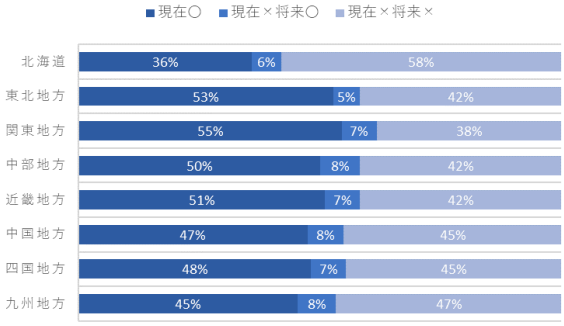


図 2a 居住地域別の投資に対する意識（現在・将来）

次に，老後の見通しについて比較したものが図 2b である．どの地方も，老後の定期的収入にて支出を賄える･賄えていると回答した層が最も多い．なお，「分からない」と答えた層は，相対的に四国が最も少ない．

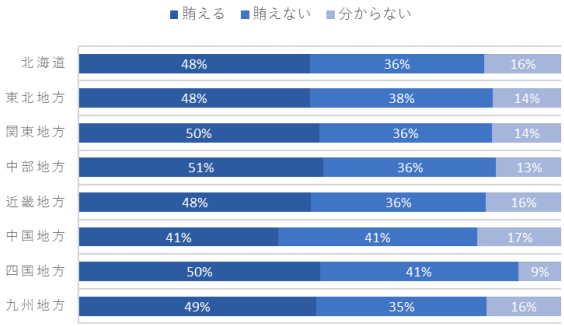


図 2b 居住地域別の老後の見通し(収入で支出を賄えるか)

(2) 世帯構成毎の特徴

＜資産状況＞

世帯構成については，退職前後で場合分けを行い，相対的にサンプル数の多い，退職前後の「単独世帯」「夫婦のみの世帯」「子どもと同居の世帯」を対象とする(図 3a).

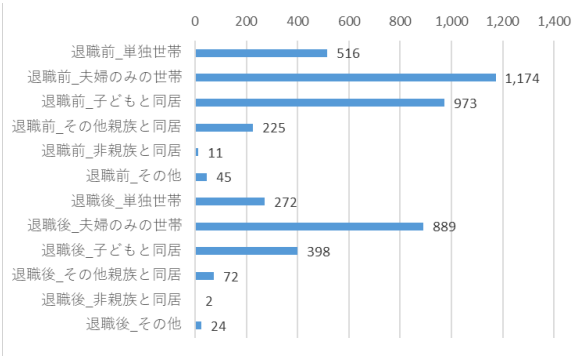


図 3a 退職前後・世帯構成毎のサンプル数

退職前後・世帯構成毎に現在の金融資産残高の階層を比較したものが図 3b である．区分によらず，600 万円～3,000 万円の層が 8 割前後と最も多い．

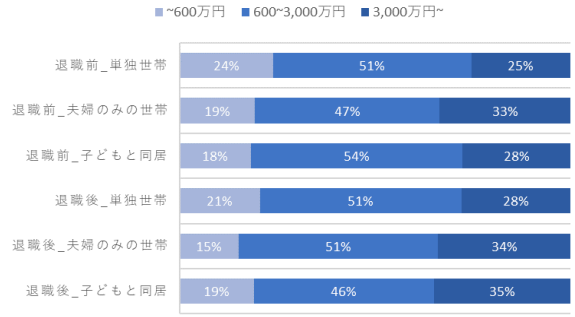


図 3b 世帯構成毎の現在の金融資産残高階層

次に，退職前後・世帯構成毎に承継予定の金融資産残高の階層を比較したものが図 3b である．こちらも区分によらず，100 万円未満の層が太宗を占める．一方，100 万円以上の承継を見込む層は，退職前の「子どもと同居の世帯」，「夫婦のみの世帯」，「単独世帯」の順で割合が多い．

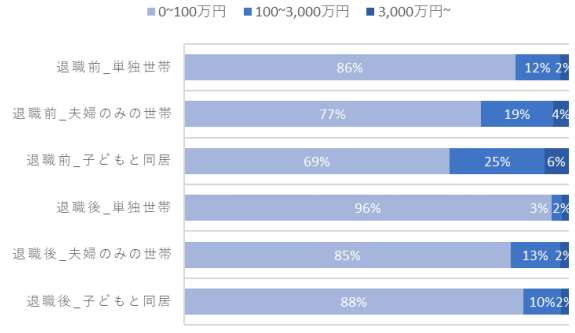


図 3c 世帯構成毎の承継予定の金融資産残高階層

＜投資スタンス・老後見通し＞

投資に対する意識について比較したものが図 4a である．いずれの区分も，現在投資している層は 5 割前後を占めている．退職の前後では，退職前の世

帯において，比較的投資にポジティブな回答をした割合が多い．

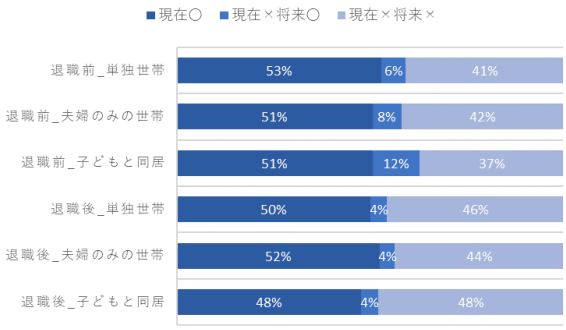


図 4a 世帯構成別の投資に対する意識（現在・将来）

次に，老後の見通しについて比較したものが図 4b である．退職の前後では，退職後の世帯において，「賄える」と回答をした割合が多い．また，退職前後の双方ともに，「夫婦のみの世帯」が相対的に「賄える」と回答した割合が多い．

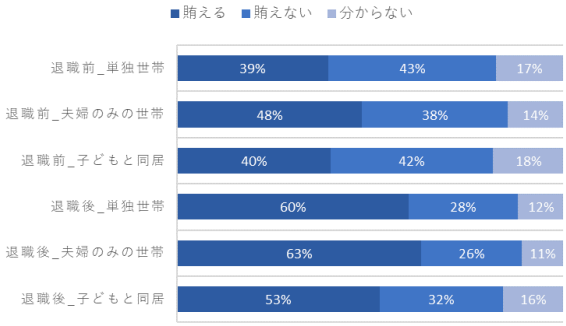


図 4b 世帯構成別の老後の見通し(収入で支出を賄えるか)

4 考察

ここでは，居住地域や世帯構成の違いにより，各種資産状況や投資に対するスタンス・老後の見通しがどのような傾向となるか考察を行う．

居住地域については，東京・名古屋・大阪など大都市を含む関東・中部・近畿地方において，現在の金融資産残高が多い傾向が見られた．一方で，承継予定の金融資産については，現在の金融資産残高とは違い，相対的に中国・四国地方が多い傾向が見られた．この点，地域の特性につき定性的な面からも検討を深めたい．投資に対する意識は，現在の金融資産残高と同様，大都市圏においてポジティブな層が多かった．居住者のフィナンシャルリテラシーとも関連する項目と考えられ，相互の連関性につき，今後分析を深化させたい．

世帯構成については，現在の金融資産残高や投資に対する意識など，カテゴリ毎の差があまり見られなかった．一方で，承継予定の金融資産額においては，特に退職前の「夫婦のみ世帯」「子どもと同居世帯」において、他カテゴリよりも資産承継を見込む層が多かった．この点，親類縁者を含む人間関係資本の厚さが寄与している可能性が示唆される．

5 まとめ

本稿では，退職前後世代の資産や家計収支の状況の特徴を分析するための第一歩として，個票のアンケートデータを用いた基礎的な分析を行った．当該アンケートデータの要約統計量を示した上で，アンケート回答者の居住地域や世帯構成毎にクロス集計を行った．今後の課題は，実証的なアプローチにより，さらに多面的な分析を加えることである．

参考文献

[1] 金融庁 金融審議会市場ワーキング・グループ報告書「高齢社会における資産形成・管理」，2019. https://www.fsa.go.jp/singi/singi_kinyu/tosin/20190603.html, last accessed 2019/12/25.

[2] 日経新聞「人生 100 年時代、2000 万円が不足 金融庁が報告書」，2019/5/15 付.

[3] 野尻哲史「高齢社会における金融サービスのあり方について」，金融庁金融審議会 市場 WG 資料，2018. https://www.fsa.go.jp/singi/singi_kinyu/market_wg/siryou/20181022/03.pdf, last accessed 2019/12/25.

[4] 横山重宏・小林庸平・大野泰資・古賀祥子「私的な資産形成に関する将来予測・政策シミュレーション分析」，『三菱 UFJ リサーチ&コンサルティング政策研究リポート』，2018.

[5] 森駿介「金融資産の保有状況で異なる老後資金問題」，『大和総研金融資本市場分析』，2019.

[6] 加藤康之「退職後の資産運用の枠組み」,『証券アナリストジャーナル』，Vol.56，No.8，2018，pp.19-28.

[7] 「退職前後世代の老後の生活に関する意識調査 ―老後生活収支に対する意識について―」, MUFG 資産形成研究所レポート, 2019. https://www.tr.mufg.jp/shisan-ken/pdf/kinnyuu_literacy_04.pdf

[8] 「退職前後世代の老後の生活に関する意識調査 ―介護・資産承継に対する意識について―」, MUFG 資産形成研究所レポート, 2019. https://www.tr.mufg.jp/shisan-ken/pdf/kinnyuu_literacy_05.pdf

[9] 総務省 「平成 26 年全国消費実態調査」，2014. https://www.stat.go.jp/data/zensho/2014/index.html, last accessed 2019/12/25.

Appendix

本稿で用いたアンケートの項目と選択肢については以下の通りである(表 A)：

表 A アンケートの項目と選択肢（抜粋）

項目分類 大	項目分類 中	項目名	尺度	カテゴリー
基本属性	－	性別	名義尺度	1) 男性, 2) 女性
		年齢	比例尺度	50~94歳
		地域	名義尺度	1)北海道, 2)東北地方, 3)関東地方, 4)中部地方, 5)近畿地方, 6)中国地方 ,7)四国地方, 8)九州地方
		職業	名義尺度	1)正社員, 2)非正規社員, 3)公務員, 4)自営・自由業, 5)専業主婦(夫), 6)無職, 7)その他
		世帯構成	名義尺度	1)単独世帯, 2)夫婦のみ, 3)子供と同居, 4)その他親族と同居, 5)非親族と同居, 6)その他
		住宅種別	名義尺度	1)持家戸建・自身, 2)持家マンション・自身, 3)賃貸・自身, 4)~6)同・親, 7)~9)同・子, 10)その他
		婚姻状況	名義尺度	1)未婚・離死別, 2)既婚
		子の人数	順序尺度	1)0人, 2)1人, 3)2人, 4)3人, 5)4人, 6)5人以上
資産状況	ストック	金融資産額(現状)	順序尺度	1)~100万, 2)100~200万, 3)200~300万,…,13) 2~3,000万,14) 3~5,000万,15) 5,000万~1億,16) 1億以上
		金融資産額(承継予測)	順序尺度	1)なし, 2)~100万, 3)100~300万,…,8) 2~3,000万,9) 3~5,000万,10) 5,000万~1億,11) 1億以上
		退職金額(予測/実績)	順序尺度	1)なし, 2)~1,000万, 3)1~1,500万, 4) 1.5~2,000万,5) 2~2,500万,6) 2.5~3,000万,7) 3,000万以上
	フロー	定期収入(現状)	順序尺度	1)~100万, 2)100~300万,…,7) 2~3,000万,8) 3~5,000万,9) 5,000万~1億,10) 1億以上
		定期収入(老後 予測/実績)	順序尺度	1)~20万, 2)20~30万, 3)30~40万, 4)40~50万, 5)50~60万, 6)60~70万, 7)70~80万, 8)80万~
		定期支出(老後 予測/実績)	順序尺度	1)~20万, 2)20~30万, 3)30~40万, 4)40~50万, 5)50~60万, 6)60~70万, 7)70~80万, 8)80万~
	その他	資産寿命	順序尺度	1)50~55歳, 2)55~60歳, 3)60~65歳, 4)65~70歳, 5)70~75歳, 6)75~80歳, 7)80~85歳,…, 11)100歳~
		資産取り崩しの見通し	名義尺度	1)定期的収入で賄え(てい)る, 2)定期的収入では足り(てい)ない, 3)あまり考えていない
健康状況	－	生命寿命	順序尺度	1)50~55歳, 2)55~60歳, 3)60~65歳, 4)65~70歳, 5)70~75歳, 6)75~80歳, 7)80~85歳,…, 11)100歳~
		健康寿命	順序尺度	1)50~55歳, 2)55~60歳, 3)60~65歳, 4)65~70歳, 5)70~75歳, 6)75~80歳, 7)80~85歳,…, 11)100歳~
投資状況	－	投資経験	名義尺度	1)現在有,2)現在無・過去有・将来有,3)現在無・過去無・将来有,…, 5)現在無・過去無・将来無
		リスク量	順序尺度	1) 0%, 2)~10%, 3)10~20%, 4) 20~30%, 5) 30%~40%, …, 11) 90%超

DEA によるシミュレーション・ログ集合の分類

A Classification of the Simulation Log Set by DEA

國上真章 [1]　菊地剛正 [2]　寺野隆雄 [3]

Masaaki Kunigami [1], Takamasa Kikuchi [2], and Takao Terano [3]

[1] 東京工業大学
[1] Tokyo Institute of Technology
[2] 慶應義塾大学
[2] Keio University
[3] 千葉商科大学
[3] Chiba University of Commerce

Abstract: This study proposes a method of using Data Envelopment Analysis (DEA) to classify a set of input/output logs in a social or organizational simulation and compares it with previous methods based on cluster analysis using examples. DEA is used mainly in policy analysis as a method to compare the efficiency based on multiple input and multiple output data. In DEA, the entire data is partitioned by the data that define efficiency corner points, called reference sets. In this study, we propose to use DEA as a classifier of simulation logs due to this property of reference sets. In this paper, we illustrate how DEA allows us to classify a set of simulation logs by their characteristics, using the SIR model of infectious disease spreading with additional variables multiple indicating measures. In addition, the results were compared with the results of the previous log classification using cluster analysis.

1 はじめに

本稿は，社会や組織のシミュレーションの入出力ログ集合の分類に包絡分析法 (Data Envelopment Analysis: DEA)の活用を提案するものである．DEA は多入力・多出力の経営データ間の効率性を相互に比較する手法として，政策分析等で利用されている[1]. DEA による分類では参照集合と呼ばれる効率性の端点を定める経営データによってデータ全体が分割される．本稿では，この参照集合の性質により DEA をシミュレーション・ログの分類器として用いる．

これまでに田中[2]は，組織シミュレーションのログをクラスター分析により分類するとともに，各クラスター毎にその性質を特徴づける意志決定を決定木により抽出する方法を提案している．また菊池[3]は，クラスター分析によりログを類型化し，形式化された仮想ケースとして記述している．

本稿ではシミュレーションの入出力ログをその振舞いに応じて分類するために DEA による方法を提案し，その特徴についてクラスター分析と比較しつつ考察する．

2 包絡分析法（DEA）

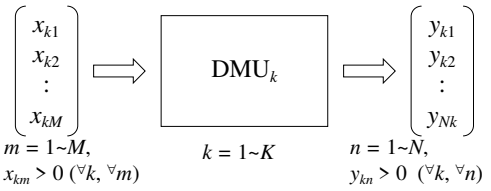
Charnes，Cooper，Rhodes が 1978 年に提案した包絡分析法 (Data Envelopment Analysis: DEA)[4]は，複数の入力と複数の出力をもつ意思決定主体（Decision Making Unit: DMU）の効率性を比較する手法である．DEA では各 DMU について，より最適化された効率性を持つ DMU からなる参照集合（Reference Set）による特徴づけを行うとともに，参照集合をつないで生成される包絡線（Envelopment）により，DMU の相対的な位置づけの比較を可能にする．

2.1 DEA の考え方

DEA の最も基本的な方法である CCR 法[4]の考え方は，次のとおりである．a) 複数の入出力をもつ意志決定主体（DMU）の重み付き効率性，b) 各 DMU の効率の重みの最適化を通した DMU 間の効率性の比較，c) 劣効率的な DMU に対しての最適な効率性を持つ DMU の集まりである参照集合（Reference Set）による特徴づけである．

まず，複数（K 個）の意志決定主体のそれぞれ DMU_k (k = 1~K) が，M 成分の入力 $\mathbf{x}_k = \{x_{km} > 0 \mid m =$

1~M} から N 成分の出力 $\mathbf{y}_k$ $\{y_{kn} > 0 \mid n = 1\sim N\}$ を生成するとする．この時，各 DMU_k の効率性 Θ_k は，入力 $\mathbf{x}_k$ と出力 $\mathbf{y}_k$ に重み $\boldsymbol{\xi}_k = \{\xi_{km} > 0\ (m = 1\sim M)\}$, $\boldsymbol{\eta}_k$ $(\eta_{kn} > 0\ (n = 1\sim N))$をつけて，$\Theta_k \equiv (\boldsymbol{\eta}_k \cdot \mathbf{y}_k) / (\boldsymbol{\xi}_k \cdot \mathbf{x}_k) = \sum_{n,m}\{(\eta_{kn} y_{kn}) / (\xi_{km} x_{km})\}$と定義できる．（図１）



$$\Theta_k \equiv \frac{\boldsymbol{\eta}\cdot\mathbf{y}}{\boldsymbol{\xi}\cdot\mathbf{x}} = \sum_{n=1}^{N}\eta_{kn}y_{kn} \Big/ \sum_{m=1}^{M}\xi_{km}x_{km}$$

図１：意志決定主体 DMU と効率性Θ.

このとき，重み $\boldsymbol{\xi}_k$, $\boldsymbol{\eta}_k$ はいくらでも動かせるので，効率性 Θ_k は変わってしまう．そこで相互に比較するので効率性の上限は１ $(\Theta_k \le 1)$ で共通化した上で，それぞれの DMU_k の効率性 Θ_k を最も有利になるように重み $\boldsymbol{\xi}_k$, $\boldsymbol{\eta}_k$ を最適化する．ただしこの最適化の際，その重み $\boldsymbol{\xi}_k$, $\boldsymbol{\eta}_k$ で他の全ての DMU_h 効率性も 1 を超えない $(\Theta_h \le 1)$ ように最適化するのがポイントである．（図２）

$$\max_{\eta,\xi}\ \Theta_k = \sum_{n=1}^{N}\eta_n y_{kn} \Big/ \sum_{m=1}^{M}\xi_m x_{km}$$

（DMU_kの効率性Θ_k）

$$\text{s.t.}\begin{cases} \displaystyle\sum_{n=1}^{N}\eta_n y_{hn} \Big/ \sum_{m=1}^{M}\xi_m x_{hm} \le 1 \quad (h = 1\sim K) \\ \eta_n \ge 0 \quad (n = 1\sim N) \\ \xi_m \ge 0 \quad (m = 1\sim M) \end{cases}$$

（そのη, ξで他のDMUの効率性Θ_hが1を超えないよう制限）

（重みは全て非負）

$$\Rightarrow \xi_{kn} = \xi_n^*,\ \eta_{kn} = \eta_n^*$$

図２：DMU の効率性Θを求める最適化問題.

もし最適化した重みでもでも効率性 $\Theta_k = 1$ にできず他の $\mathrm{DMU}_{k'}$ の $\Theta_{k'}$ に劣れば，DMU_k は効率的でないといえる．このとき，DMU_k を最適化した重みにおいてもその効率性 Θ_k を上回り最適な効率性が 1 となる $\mathrm{DMU}_{k'}$ を DMU_k の参照集合 E_k といい，DMU_k にとっての効率性改善の参考となる．（図３）

$$E_k = \left\{ k' \,\middle|\, \sum_{n=1}^{N}\eta_{kn}y_{k'n} \Big/ \sum_{m=1}^{M}\xi_{km}x_{k'm} = 1 \ \ (1 \le k' \le K) \right\}$$

（DMU_kを最適化してもΘ$_k$<1のときは，Θ$_k$=1を妨げている他の$\mathrm{DMU}_{k'}$（kのための最適化の重みでも効率1となる）が存在する）

図３：DMU_k の参照集合 E_k.

ここまでに見た DEA の最適化（図２）は，分数計画問題として定式化されているが，実用上は次の等価な線形計画問題に変換して解かれる．（図４）

$$\max_{\eta,\xi}\ \Theta_k = \sum_{n=1}^{N}\eta_n y_{kn}$$

（効率性の分母を1に固定する）

$$\text{s.t.}\begin{cases} \displaystyle\sum_{m=1}^{M}\xi_m x_{km} = 1 \\ \displaystyle\sum_{n=1}^{N}\eta_n y_{k'n} \le \sum_{m=1}^{M}\xi_m x_{k'm} \quad (k' = 1\sim K) \\ \eta_n \ge 0 \quad (n = 1\sim M) \\ \xi_m \ge 0 \quad (m = 1\sim M) \end{cases}$$

（制約式を等価な不等式に変更）

$$\Downarrow$$

$$E_k = \left\{ k' \,\middle|\, \sum_{n=1}^{N}\eta_n y_{k'n} = \sum_{m=1}^{M}\xi_m x_{k'm} \ \ (k' = 1\sim K) \right\}$$

図４：分数計画問題（図 2）は等価な線形計画問題によって解くことができる.

DEA を用いることの利点の一つが，このように多入力多出力系の効率性を比較する際の重みをどう決めるかという問題が，上記の最適化のプロセスで自動的に解決されることにある.

2.2 DEA による分類

DEA においては，ある DMU_k にとってその参照集合 E_k は，DMU_k にとって改善の参考となる対象であり，効率的な DMU の参照集合は自分自身である．また，入出力変数で張られる空間において DMU の散布図を描くと，効率的な DMU は全 DMU の包絡線となる．また効率的な DMU と原点で分割される領域には，その効率的な DMU を参照集合とする効率的でない DMU が含まれる．よってこれらの DMU 達は同じ参照集合を共有するという性質で分類されていることになる．（図 5）

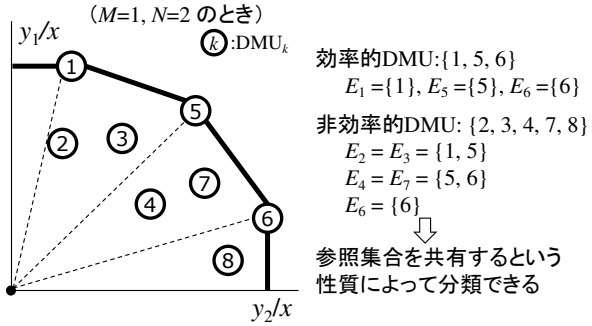


図５：効率的な DMU（参照集合）により包絡線を分割すると非効率な DMU を分類できる.

同じ参照集合を共有する DMU は，効率性の改善について同じ方向（参照集合の元）を共有している．このように，分類群の中において分類群を特徴付ける元（効率的な DMU～参照集合の元）と，その他の

元（非効率的な DMU）の関係が，自動的に与えられることが DEA を用いることの今一つの利点である．

また部分的にしか一致しない参照集合をとるグループ間においては，共通する部分とその多少によってグループ間の類似性を考察することができる．これも分類器としての DEA を見たときの利点である．

3 シミュレーションへの適用

簡単なシミュレーションを用いて，その入出力データを DEA により分類してみる．シミュレーションモデルとしては，感染症伝播に関する SIR モデル[5]を用いる．ただし SIR モデルは，独立な自由度が２しかないため，出力の自由度とともに政策的な入力変数を追加して DEA を適用する．

3.1 SIR モデルの拡張

本稿で例題として用いる拡張した SIR モデルは以下の状態遷移（図６）を持つ差分方程式系（式(1)～(3)）である．

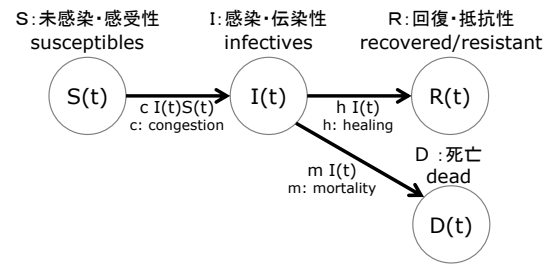


図６：拡張された SIR モデルの状態遷移と
差分方程式系および制御入力

$$\begin{cases} S(t+1) = S(t) - c(t)S(t)I(t), \\ I(t+1) = I(t) + c(t)S(t)I(t) - h(t)I(t) - m(t)I(t), \\ R(t+1) = R(t) + h(t)I(t), \\ D(t+1) = D(t) + m(t)I(t). \end{cases} \quad \cdots\cdots(1)$$

$$\begin{cases} P(t) = S(t) + I(t) + R(t), \quad \text{population alive}, \\ P(0) = P_0, \quad \text{initial population}, \\ S(0) = P_0 - \varepsilon, \quad I(0) = \varepsilon, , \ R(0) = 0, \ D(0) = 0. \end{cases} \quad \cdots\cdots(2)$$

$$\begin{cases} c(t) = \dfrac{c_0}{1 + \alpha_u \cdot u(t)}, \quad : \text{contagion}, \\ h(t) = h_0(1 + \alpha_v \cdot v(t)), : \text{heeling}, \quad \begin{cases} u(t), v(t), w(t) \ge 0 \\ u(t) + v(t) + w(t) \le 1 \end{cases} \\ m(t) = \dfrac{m_0}{1 + \alpha_w \cdot w(t)}, \quad : \text{mortality}. \end{cases} \cdots(3)$$

式(1)は状態変数の時間変化を，式(2)は従属変数と状態変数の初期値を，式(3)は制御変数による係数の時間変化を表す．

状態変数は，未感染の感受性人口を $S(t)$，感染者人口を $I(t)$，回復した非感受性人口を $R(t)$ とし，さらに死亡人口を $D(t)$として追加した，従属変数 P(t)は生存者人口である．また感染率 c，治癒率 h，死亡率 m を，制御変数 $u(t), v(t), w(t)$ によって変化させられる変数とした．ここで正定数 c_0, h_0, m_0 は制御入力がない時の $c(t), h(t), m(t)$ の値を示し，正定数 $\alpha_u, \alpha_v, \alpha_w$ はそれぞれ $u(t), v(t), w(t)$の感度である．

制御入力 $u(t), v(t), w(t)$について，$u(t)$は感染率 $c(t)$を抑制する施策を，$v(t)$は治癒率 $h(t)$を向上させる施策を，$w(t)$は死亡率 $m(t)$を抑制する施策をそれぞれ表す．これらの費用と効果はそれぞれ異なるが費用で規格化することで一般性を失わずにその違いを感度 $\alpha_u, \alpha_v, \alpha_w$ のみで表し，費用制約は $u(t) + v(t) + w(t) \le 1$ としている．

3.2 DEA による結果の分類

前節で導入した拡張された SIR モデルを例として用いて，シミュレーションの入出力データの集合を DEA によって分類する．共通する初期値，係数等は表１のとおりである．

Initial Conditions	$S(0)$: Susceptibles	$I(0) = \varepsilon$: Infectives	$R(0)$: Recovered	$D(0)$: Dead
	9990	10	0	0
	$P(0)$: Population Alive =10000			
Coefficients	c_0 Contagion	h_0 Healing	m_0 Mortality	
	0.00002	0.01	0.01	
Control Sensitivities	α_u	α_v	α_w	
	10	15	20	

表１：シミュレーションの初期値および係数

次に，シミュレーション実行(RUN)設定は，表２に示す 10 通りの制御入力の割付に従った．

		Stage-0			Stage-1			Stage-2			Stage-3		
		t = 0 ~ 29			t = 30 ~ 89			t = 90 ~ 179			t = 180 ~ 500		
		$u(t)$	$v(t)$	$w(t)$	$u(t)$	$v(t)$	$w(t)$	$u(t)$	$v(t)$	$w(t)$	$u(t)$	$v(t)$	$w(t)$
RUN # (DMU #)	0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.4	0.3	0.3
	1				1.0	0.0	0.0	1.0	0.0	0.0			
	2				1.0	0.0	0.0	0.0	1.0	0.0			
	3				1.0	0.0	0.0	0.0	0.0	1.0			
	4				0.0	1.0	0.0	1.0	0.0	0.0			
	5				0.0	1.0	0.0	0.0	1.0	0.0			
	6				0.0	1.0	0.0	0.0	0.0	1.0			
	7				0.0	0.0	1.0	1.0	0.0	0.0			
	8				0.0	0.0	1.0	0.0	1.0	0.0			
	9				0.0	0.0	1.0	0.0	0.0	1.0			

表２：シミュレーションの実行設定．

上記の設定から得られた実行結果の要約値は表３のとおりである．要約値を用いたのは，本例では実行数が少ないためと，DEA の性質から出力変数は正の望大特性であることが必要なためである．

		Input:		OutPut:		
		Cs	Cm	BtmActP	AvgActP	AlvPE
		Social Cost:	Medical Cost:	Bottom Active Population:	Average Active Population:	Alive Poplation at the end:
		接触・感染防止の社会的コスト: Avg[$P(t)\cdot\alpha_u\cdot u(t)$]	感染者の回復と救命コスト: Avg[$I(t)\{\alpha_v\cdot v(t)+\alpha_w\cdot w(t)\}$]	活動可能な人数が最も落ち込んだ時の値: Min[$P(t)$-$I(t)$]	活動可能な人数の平均値: Avg[$P(t)$-$I(t)$]	最終的に生き残った人の数: Min[$P(t)$]
Run # (DMU #)	0	13,574	231	1,915	5,112	5,296
	1	49,693	340	8,176	8,590	8,638
	2	34,737	617	8,371	9,011	9,072
	3	34,481	19,220	2,514	7,974	8,967
	4	41,879	912	8,650	9,560	9,603
	5	24,652	919	8,650	9,575	9,619
	6	24,328	6,776	6,505	9,147	9,488
	7	33,203	15,244	1,906	6,589	7,235
	8	24,088	15,932	1,906	8,663	9,399
	9	24,418	29,615	1,906	8,082	9,524

表３：シミュレーションの実行結果
（入出力要約値）

表４は，シミュレーションの入出力要約値のデータ（表３）から DEA（CCR 法）によって求めた，最適な効率性Θ（DEA Score），参照集合（Reference Set），入出力の要約値への最適化された重み（Input Weight, Output Weight）である．なお CCR 法の計算については，並木[6]による Python コードを用いた．

DEA_CCR		DEA Score Θ	Reference Set	Input Weights		Output Weights		
				Cs	Cm	BtmActP	AvgActP	AlvPE
Run # (DMU #)	0	1.0	{0, 1, 2}	0.4	22.4	1.8	0	1.2
	1	1.0	{1}	0	29.4	1.2	0	0
	2	1.0	{1, 2}	0.1	8.4	1.2	0	0
	3	0.6665	{8, 5}	0.3	0	0	0	0.7
	4	0.7903	{0, 2, 5}	0.1	5.3	0.5	0.4	0
	5	1.0	{5}	0.4	0	1.2	0	0
	6	0.9995	{8, 5}	0.4	0	0	0	1.1
	7	0.5584	{8, 5}	0.3	0	0	0	0.8
	8	1.0	{8}	0.4	0	0	0	1.1
	9	0.9997	{8}	0.4	0	0	0	1.0

表４ :DEA による表３の入出力要約値の分析結果

DEA の結果（表４）によると，効率的な実行結果として，{Run 0, 1, 2, 5, 8}が得られた．一方，非効率な実行結果としては，{Run 3, 6, 7}が{Run 5, 8}を参照集合として共有する分類群となっており，入出力要約値の重みづけも互いに近い傾向を示している．

また効率的な実行結果である，{Run 0, 1, 2, 5, 8}については，Run 8 と 5 にはこれらを参照集合として共有するグループがあり，Run 0, 1, 2 は互いを参照集合としているなどそれぞれの間で違いが少ないことを示している．反面，Run 8 と Run 0, 1 の間にはこれらを共有する非効率集合が無いなど，これらの間で違いが大きいことを示している．

これらの結果を模式的にまとめたのが図 7 である．効率的な実行結果（参照集合）の共有によって，実行結果がグループ化されている，これによりグループ内での効率性の改善の参考となるデータが明らかになっている．また最適な入出力要約値の重みについて，グループの特徴づけがされている．さらに，参照集合の共有の有無によってグループ間の近さも読み取ることができる．

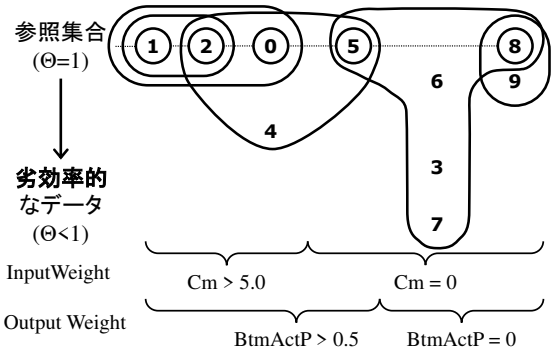


図７：DEA による入出力要約値のグループ化

3.3 クラスター分析による分類

次に，同じ入出力要約値（2 入力 3 出力）について，クラスター分析による分類を行う．分類は階層的クラスタリングの Ward 法により，データ間の距離は Euclid 距離を用いた．

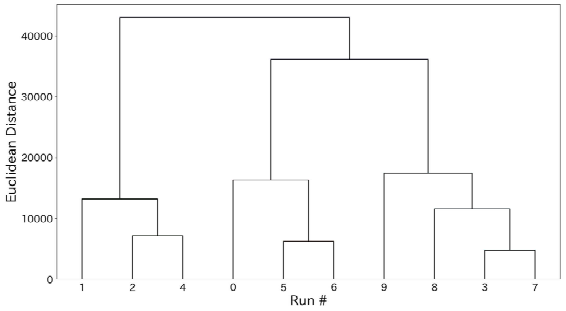


図８：入出力要約値のクラスタリング

クラスタリングによる，入出力要約値の類似性（ベクトルとしての距離の近さ）による樹形図は図８のとおりである．

4 考察

これまでに見たように DEA による分類とクラスター分析では，同じデータに対しても異なった分類を示す. ここでは DEA とクラスター分析のアプローチについて分類器としての特徴という視点でから追加的に考察する．

まず基本的なアプローチについて，DEA ではグループの分類と特徴づけは，参照集合によって行われる．参照集合はグループの端点であり，グループ内でも特異な元である．このためクラスター分析のようなグループ内の平均あるいは重心付近の元という意味は持たないが，グループ内の元の効率性の方向性でのトップという意味は明確である．

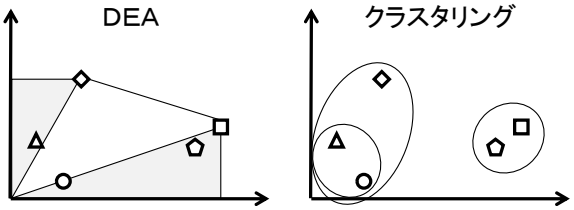


図８：クラスタリングが互いに近いものをまとめるのに対し DEA は端点でグループ化する．

DEA による分類においては，入出力の重みは自動的に最適化される．このためクラスター分析でのデータの重みをどう決めるかという問題は，DEA 任せにできる．一方 DEA には，分類に特化した方法ではないための使い勝手の悪さも存在する．例えば，クラスター分析では，生成するクラスターの階層数を選ぶことができる．これ対し DEA では，グループの数は DEA 任せであり，一つのグループしか生成されないことや，すべてのグループがひとつの参照集合しか含まないということもあり得る．また DEA では，入力は正の望小特性，出力は正の望大特性である必要があり，負の値や望大特性と望小特性が入り交ざっていると事前にデータの変換処理が必要になる．

次に，データの追加に対する安定性について比較してみる．まず，グループの内側に非効率なデータが追加される場合，DEA によるグループ分けは影響を受けない．一方クラスター分析ではクラスター間に中間的なデータが追加されるとクラスタリングが大きく影響される可能性がある．（図９上）

他方，新たな包絡線の外側に位置するようなデータが追加された場合，DEA によるグループ分けは大きな影響を受ける．一方，クラスター分析では大きな外れ値は独立したクラスターとなり，その他のクラスタリングには影響しない．（図９下）

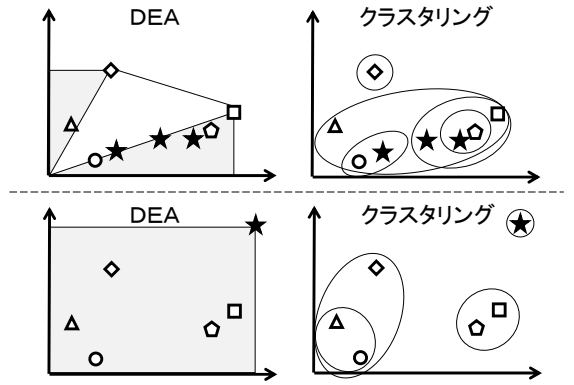


図９：新しいデータの追加に対する安定性：図８のグループの内部に追加（★）された場合（上）とグループの外側に追加（★）された場合（下）

このように DEA による分類は，クラスター分析に対して相反する性質を有する．このことから DEA による分類は，クラスター分析を超えるものというよりは，互いに補う様に使うことでより分析の幅を広げられる可能性を持つと考えられる．

5 まとめ

本稿は，シミュレーション入出力ログ集合の分類法としての DEA の活用を提案し，シミュレーションによる例題と，クラスター分析との比較を考察した．DEA による分類は，クラスタリングと相互補完的な分類器となると考えられる．ログの本格的な分類器として活用するためには，Window 分析等の時系列の DEA も必要であり，今後検討を深めていきたい．

謝辞

本稿の着想の多くは，竹内俊彦先生の解説[7]に拠っている，この場で深く感謝を表したい．また，本研究の一部について，公益財団法人科学技術融合振興財団の調査研究助成を受けている．

参考文献：

[1] Cooper, WW., Seiford, LM., Tone, K.: Data Envelopment Analysis, A Comprehensive Text A Comprehensive Text with Models, Applications, References and DEA-Solver Software (Second Edition), Springer, (2007).

[2] 田中祐史, 國上真章, 寺野隆雄: エージェントシミュレーションにおけるログクラスターの系統的分析からわかること，シミュレーション＆ゲーミング，Vol.27, No.1, pp. 31-41, (2017)
DOI https://doi.org/10.32165/jasag.27.1_31

[3] 菊地剛正, 國上真章, 高橋大志, 鳥山正博, 寺野隆雄 ビジネスケースの形式的記述のためのシミュレーション結果の類型化手法, 経営情報学会論文誌 研究ノート Vol.29, No.3, (to be published 2020)

[4] Charnes, A., Cooper, WW., Rhodes, E.: Measuring Efficiency of Decision Making Units, European Journal of Operational Research, Vol.2 , pp.429-444, (1978).

[5] Kermack, WO., McKendrick, AG.: A Contribution to the Mathematical Theory of Epidemics, Proceedings of the Royal Society, vol.115A, pp.700‑721 (1927),
DOI https://doi.org/10.1098/rspa.1927.0118,
(reprinted Bulletin of Mathematical Biology vol.53, pp.33-55 (1991))

[6] 並木誠: Python による数理最適化入門，朝倉書店，pp.70-71, (2018)

[7] 竹内俊彦：DEA 入門・泉恵女学園物語，http://blog.livedoor.jp/lionfan/archives/52682140.html, (2020.8 閲覧)

暗号資産がポートフォリオのリスクリターン特性に与える影響に関する分析

Analysis of the impact of crypto assets on portfolio risk return performance

刘梦垚 [1]　上瀧 弘晃 [2]　高橋大志 [3]

Mengyao Liu[1], Hiroaki Jotaki[2], Hiroshi Takahashi [3]

[1,2,3] 慶應義塾大学大学院経営管理研究科

[1,2,3]Graduate School of Business Administration, Keio University

Abstract: In this research, we analyze the impact of crypto assets on portfolio construction. Through this research, we attempt to clarify the risk return characteristics of portfolios that include crypto assets as investment targets. The analysis period is from 2014 to 2020, and in addition to the Cryptocurrency Index (CRIX), S&P500, Gold, Crude oil, PE, REITs, Goldman Sachs Commodity Index (GSCI), etc. Consider effectiveness and challenges.

はじめに

近年、暗号資産に関する議論が関心を集めている。ビットコインなどをはじめとする暗号資産の多くは、従来の資産とは異なる特徴を有しており、利点および課題を含め、数多くの議論が行われている[１]。

暗号資産に関する議論は、いくつかの視点から行われているが、投資対象としての視点も一つの主要な議論に挙げられる[２]。暗号資産は、ビットコインだけでなく、多くのデジタルマネーが開発されており、Cryptocurrency Index （CRIX）などの暗号資産の推移を示す指数も報告されている[２][４]。本研究では、暗号資産を投資対象に含めた際のポートフォリオ特性に与える影響について分析を行う。

次章において、先行研究について触れた後、分析手法、データ、分析結果について説明する。最後に、まとめおよび今後の課題を記す。

先行研究

暗号資産に関する研究は、ビットコインを投資対象とした分析が数多く報告されている。例えば、ビットコインと債券インデックスを投資対象とし、ビットコインのリスクヘッジに関する研究などが報告されている[５]。

また、近年、多様な暗号資産への投資が行われるようになっており、ビットコイン以外の暗号資産を投資対象とした分析も報告されている[６]。例えば、外貨、商品、株式、ETF および暗号資産のビットコイン、リップル、ライトコインを投資対象として、投資ポートフォリオ作成する分析を行う研究もある。暗号通貨のポートフォリオが実際にポートフォリオの有効性を向上させることを示している[７]。そして、最近では暗号資産のインデックスを投資対象とした分析を行う論文も増えている。例えば、暗号資産をポートフォリオ管理の研究対象とし、暗号資産は従来の資産と比較して流動性が低いため、ポートフォリオに追加する際には、流動性限定リスクリターン最適化（LIBRO）アプローチを提案する研究がある[８]。また、Chuen/Guo/Wang (2018)は、2014 年 8 月 11 日から 2017 年 3 月 27 日までの暗号資産のインデックス（CRIX）を対象とした分析を行っており、DCC などのモデルを利用して、CRIX および暗号通貨がポートフォリオのリスク分散に優れた投資資産となる可能性があることを示している[９]。

これらの研究を背景とし、本研究においては、ポートフォリオの特性に与える影響に焦点をあて分析を行う。とりわけ、本研究では、近年拡大している市場価格変動を考慮した分析についても試みる。

目的

本研究では、暗号資産を含むポートフォリオのリスク・リターン特性を明らかにすることを目的とする。分析においては、株式や債券などの伝統的は資産に加え、暗号資産が投資対象に含まれた場合における、投資ポートフォリオについて分析を行う。こ

れら分析を通じ、投資クラスとしての暗号資産の可能性および課題について検討する。また、先行研究には暗号資産の発展がここ最近であることから、分析対象期間が 3 年程度の研究が数多いが、本研究では分析対象期間を直近まで追加することで約 6 年とし、分析結果を先行研究と比較する。さらに、市場が急変動した場合における暗号資産の効果についても検証を行う。

データ

本研究では、資産運用における代表的な資産を対象として分析を行う。一番数多く代表的な資産を分析した Chuen/Guo/Wang (2018)の先行研究を参考にして、7 種類の資産の各代表的な指数を利用する。本分析に用いた資産をテーブル 1 に示す。本分析では、暗号資産の動きを示す指標として Cryptocurrency index(CRIX) を用いた。同指標は、暗号資産の代表的な指数の一つに挙げられる[１０]。[1]本研究の分析対象期間は、2014 年 7 月 31 日から 2020 年 4 月 22 日とし、日次データを用いた。

Table 1: 本分析に用いた指数

1	S&P500
2	Crude Oil
3	Gold
4	S&P Listed Private Equity
5	MSCI U.S. REIT Index
6	Goldman Sachs Commodity Index (GSCI)
7	Cryptocurrency index(CRIX)

分析手法

マーコヴィッツ現代のポートフォリオ理論アプローチを使用して、平均分散と有効フロンティアアプローチで、伝統的な資産と暗号資産から構成されるポートフォリオを構築する。

具体的な分析方法としては、モンテカルロシミュレーションを行って、暗号資産が含まれてないと含まれた場合の最小分散ポートフォリオと最大シャープレシオを持つポートフォリオの各資産の投資比率を比較する。暗号資産を含むと、ポートフォリオのリスク・リターンパフォーマンスが向上かどうかを検証する。

[1] Kim/Trimborn/Härdle [2019]らは，暗号資産の指数としてのCRIX の正当性について議論を行っている。

分析結果

Fig. 1 は､暗号資産を含むポートフォリオのリスクリターン特性を示したものである。表の横軸は、ボラティリティ、縦軸はリターンを示している。

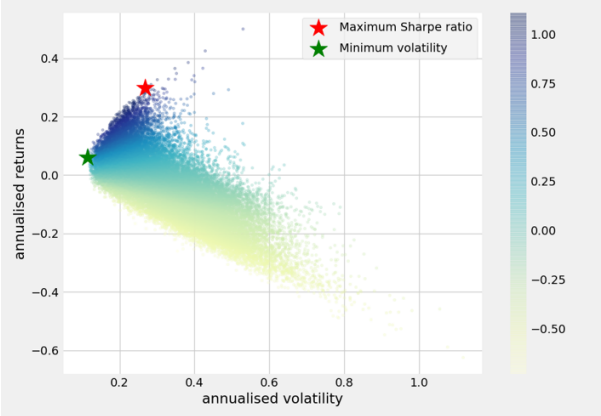


Fig. 1　ポートフォリオのリスク・リターン特性

図中の赤い点は最も高いシャープレシオを持つポートフォリオを示すものである。同ポートフォリオ内において、暗号資産のウェイトが 34.78%である。これらの結果は、暗号資産がポートフォリオのリスク・リターン特性の改善に貢献している可能性を示すものである[2]。

まとめ

本研究では、暗号資産がポートフォリオのリスク･リターン特性に与える影響に関し検討を行った。本稿では、基礎的な分析として、資産運用における伝統的な資産と暗号資産から構成されるポートフォリオを構築し、その特性について議論を行った。より詳細な分析は今後の課題である。

参考文献

[１] Lam, P.N., and D.K.C. LEE. “Introduction to Bitcoin.” Handbook of Digital Currency, edited by D.K.C. Lee, pp. 5-30. San Diego: Elsevier, 2015.

[２] Burniske, C., and A. White. “Bitcoin: Ringing the Bell for a New Asset Class.” Research white paper, 2017.

[３] Simon Trimborn, Wolfgang Karl Härdle. “CRIX an Index for cryptocurrencies” Journal of Empirical Finance 49 (2018) 107-122

[2] 緑色の点は、最小分散ポートフォリオであるが、暗号資産のウエイトは 4.36%であった。

[４] Chen, S., C.Y.H. Chen, W.K. Härdle, T.M. Lee, and B. Ong. “A First Econometric Analysis of the CRIX Family.” Working paper, (2016)

[５] Md Akhtaruzzaman, Ahmet Sensoyc, Shaen Corbetd. “The influence of Bitcoin on portfolio diversification and design”. Finance Research Letters.2019.

[６] Halaburda, H. “Digital Currencies: Beyond Bitcoin.” DigiWorld Economic Journal, 103 (2016), pp. 77-92.

[７] Yanuar Andrianto, Yoda Diputra. “The Effect of Cryptocurrency on Investment Portfolio Effectiveness”. Journal of Finance and Accounting.2017; 5(6): 229-238

[８] Trimborn, S., M. Li, and W.K. Härdle. “Investing with Cryptocurrencies—A Liquidity Constrained Investment Approach.”2017.

[９] David LEE Kuo Chuen, Li Guo and Yu Wang. “Cryptocurrency: A New Investment Opportunity?” The Journal of Alternative Investments Winter 2018, 20 (3) 16-40

[１０] Alisa Kim, Simon Trimborn, Wolfgang K. Härdle, “VCRIX - A Volatility Index for Crypto-Currencies” (2019)