

Kaolid: 飲料の味変化のための口中香を実現する 蓋型嗅覚デバイスの設計と基礎評価

真弓大輝^{†1}

奈良先端科学技術大学院大学

中村優吾^{†2}

九州大学

松田裕貴^{†3}

奈良先端科学技術大学院大学

三崎慎也^{†4}

奈良先端科学技術大学院大学

安本慶一^{†5}

奈良先端科学技術大学院大学

1. はじめに

糖分入り飲料 (SSB) は、私たちの食生活において主要な糖分摂取源となっている。特に低所得国や中所得国においては、消費パターンの増加は都市化や経済発展に伴い、これら飲料の入手可能性が増加したため、多くの低所得国や中所得国で SSB の摂取量が増加している [1, 2, 3, 4]。また習慣的な SSB の摂取は、肥満 [5, 6, 7, 8] や糖尿病 [5, 9]、心臓病などの心疾患のリスクが高まるため [10, 11]、SSB の摂取量を低減させる取り組みが必要である。これまで SSB の摂取量問題を解決するための施策がいくつか提案されている [12, 13, 14, 15]。これら手法により、一時的に SSB の摂取を低減することは可能であるが、砂糖が少ないまたは無糖の飲料水は、味が物足りなく感じられることが多い。結果的に、糖分含有量が少ない飲料を選択するという習慣が長続きせず、甘い飲み物を常飲する元の生活に戻ってしまうケースが多い。

これまで、我々のグループでは、「味」として認識する大部分を香りが占めている [16, 17, 18] ことに着目し、飲み物を飲む際に甘い香りを提示することで、知覚する甘さを増幅するマグカップ型の嗅覚デバイスの研究開発に取り組んできた [19, 20, 21]。本デバイスにより、香りによって知覚するコーヒーの味の甘さ増幅を実現しているが、味の満足度向上には至っていない。

そこで我々は、嗅覚へのアプローチ方法を模索する。人には鼻先から香る「鼻先香 (Orthonasal smell)」と口の中から鼻を抜ける香り「口中香 (Retronasal smell)」の 2 つの鼻腔経路が存在し、鼻先香と口中香の嗅覚情報は、脳で異なる方法として処理されることが示されている [22, 23, 24]。



図 1 Kaolid の概要 ((a) ストロー型デバイス, (b) コップ型デバイス)

特に、食べ物や飲み物の風味に寄与する口中香は、味覚との関連が強いことが知られている [25]。

しかし、既存アプローチは、鼻先香の提示に留まっており、口中香を提示できていないという点で、味覚の変化を促すには十分な設計ではない可能性がある。そこで本研究では、口中香を実現し、飲料の味覚を変化させるスマートな蓋型嗅覚デバイス「Kaolid」を提案する。図 1 に示すように、Kaolid は小型で軽量な嗅覚デバイスを搭載し、口中香を実現するストロー型デバイスとコップ型デバイスである。ストロー型デバイスは、ストロー中間部分から香りをチューブ内に注入し、飲料と一緒に香りをユーザの口内に届ける。コップ型デバイスは、蓋からコップ内に香りを噴射し、ユーザが飲む際に香りを口内に届け、口中香を実現する。

本稿では、Kaolid のプロトタイプシステムの設計・実装について述べ、口中香を提示した際の飲料の温度の違いが味覚に与える影響の調査結果について報告する。

2. 関連研究

2.1. SSB の代替飲料

SSB の摂取量を低減させるために、代替飲料として人工甘味飲料が注目されている。人工甘味料の普及に伴い、日常生活で人工甘味料が添付された製品に触れる機会が多くなっているが、人工甘味料で体重を減らすことは難しい。Fowler らの研究 [26] では、1980 年代に 7~8 年間に渡り、3682 人の成人を対象に調査を行なっている。結果として、人工甘味料入り飲料を摂取した人は、摂取量に依存し、高い

Manuscript Format for Kickoff Symposium of IPSJ Behavior Transformation by IoT (BTI)

^{†1} DAIKI MAYUMI, Nara Institute of Science and Technology

^{†2} YUGO NAKAMURA, Kyushu University

^{†3} YUKI MATSUDA, Nara Institute of Science and Technology

^{†4} MISAKI SHINYA, Nara Institute of Science and Technology

^{†5} KEIICHI YASUMOTO, Nara Institute of Science and Technology

BMI を示した。また、103 名の青少年を対象に行った研究でも、砂糖入り飲料をダイエット飲料に置き換えたが、25 週間後の BMI は減少しなかった [27]。結果の要因として、砂糖または人工甘味料によってもたらされる甘味が人間の食欲を高めることが関連している [28]。

100% フルーツジュースが SSB の代替品として受け入れられるかどうかについても、関心が寄せられている。ほとんどのフルーツジュースはビタミンや栄養素を含んでいるため、健康に良いと見なされる。しかし一部のフルーツジュースには、果物に含まれる天然糖が SSB と同程度のカロリーと糖分が含まれている。フルーツジュースの摂取が体重増加 [3] および 2 型糖尿病のリスク [29] と関連していることが報告されている。

一方、水には砂糖やカロリーが含まれておらず、水分補給に最適な飲料と考えられている。水の需要が高まるにつれて、様々な種類のスパークリングやフレーバーのオプションが利用できるようになり、習慣的に SSB を消費する人が水に切り替えるのに役立つ可能性がある。

2.2. 嗅覚デバイスに関する研究

近年、様々な種類の嗅覚デバイスとそれを用いたインタラクションが提案されている。例えば、ウェアラブルデバイスで香りを提示するヘッドマウント・ディスプレイ [30, 31] や、首からかけるウェアラブルデバイス [32, 33] などが開発されている。Wang らの研究 [34] では、ピアス型、ネックレス型、顔に装着するオーフェイス型の嗅覚デバイスを提案している。提案システムは日常生活で利用しやすいデザインと小型化を実現している。日常生活に溶け込んだデザインでユーザが日常的に持ち運びが可能な設計である。提案システムの社会的受容性、快適性および装着者と観察者の双方が感じる匂いの強さなどを評価している。Amores らの研究 [32] では、スマートフォンから遠隔操作可能で、生体情報や文脈情報に基づいて香りの強さや頻度を変えることができるネックレス型の嗅覚デバイスを提案している。Dobbelstein らの研究 [33] では、日常生活で利用可能なウェアラブル嗅覚デバイスを開発し、個人の香りの通知を受けることが可能である。Choi ら [35] は、フレームの端に香りを放出するための加熱モジュールを埋め込んだ 3D プリントメガネを開発している。

2.3. 嗅覚が味覚に与える影響

人の味覚のうち 75~95 % は香りによる影響であると言わされており、一般的に我々が「味」として認識している大部分は匂いが占めている [16, 17, 18]。つまり、香りを制御することで人の味覚を変化させることができると可能性がある。また風邪を引いて鼻が詰まったり、鼻を摘んで食事を行なっ

た際、味を感じなかったり、感じ取りにくい経験はないだろうか。これは味覚が嗅覚に依存していることがわかる身近な体験の 1 つである。また、嗅覚が味覚に与える影響の例としてかき氷のシロップが挙げられる。かき氷で使用される代表的なシロップとして、いちご味・レモン味・ブルーハワイ味などがある。これら 3 種類のシロップは甘味、旨味、塩味、苦味、酸味の数値は全てほとんど同じ数値であるが味わいがそれぞれ異なる。赤いシロップといちごの香りでいちご味を表現するなど、視覚および嗅覚情報の提示方法の違いによって異なる味わいを表現している。

Ranasinghe らの研究 [36, 37] では、AR や VR 空間での飲料体験を拡張することを目的として、電気刺激や香り、視覚効果を実現するデバイスを制作し、各刺激が味覚に及ぼす影響を調査している。結果として、異なる感覚相互作用によって味覚の拡張を実現しているが、甘味增幅に関しては実現できていない。また香りの提示方法は、鼻にチューブを直接挿す仕組みとなっており、ユーザが日常的に利用しやすいデザインではない。

また我々の研究グループは、香りによって甘さを增幅させるマグカップ型の嗅覚インターフェースを提案してきた [19, 20, 21]。被験者 33 名に対して実証実験を行い、香りによって知覚する味の甘さが増幅することを確認したが、味の満足度は向上しなかった。

つまり、SSB の代替飲料として人工甘味料を用いることは不十分であり、糖分やカロリーを含まない水の飲料促進を行うことが望ましい。我々は飲料の甘さを増幅させる手法の一つとして、香りによるアプローチを行ってきたが、現在の設計では味の甘さ増幅は実現するものの味の満足度向上は難しい。そこで本研究では、人の鼻腔経路の違いに着目し、異なる嗅覚のアプローチを考える。人には鼻先から香る「鼻先香 (Orthonasal smell)」と口の中から鼻を抜ける香り「口中香 (Retronasal smell)」の 2 つの鼻腔経路が存在し、鼻先香と口中香の嗅覚情報は、脳で異なる方法として処理されている [22, 23, 24]。特に、食べ物や飲み物の風味に寄与する口中香は、味覚との関連が強いことが知られている [25]。本論文では、口中香を提示可能とするデバイス「Kaolid」のプロトタイプシステムの設計・実装を提案し、口中香を提示した際の飲料が味覚に与える影響の調査結果について報告する。

3. 蓋型嗅覚デバイス Kaolid の設計と実装

3.1. アプローチ

糖分摂取量を低減するために、砂糖が入っていない飲料の摂取が重要である。しかし、これまで砂糖入り飲料を多く摂取していた人にとって、糖分のない水に代替することは

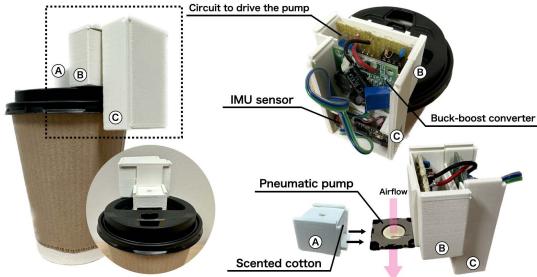


図 2 Kaolid のシステム設計

難しい。そこで味として認識される大部分が匂いが占めていること、食べ物や飲み物の風味に寄与する口中香は味覚との関連が強いといふこれら知見に基づき、飲料を飲む際に口中香を提示することで味覚を変化させ、知覚する味の甘さ増幅と味の満足度向上を図る。我々は飲み物を飲む際に日常的に使用する道具であるストローおよびコップに着目し、飲むという行為の動作センシング機能と、味覚を変化させる香り情報の提示機能を備えた嗅覚デバイス「Kaolid」を提案する。

3.2. システムの設計

本システムは口中香を実現する機構として、空気圧ポンプと空気圧ポンプを駆動するための回路で構成されている。

香りを噴射する空気圧ポンプは一方から空気を吸引し、一方から空気を吐出するデバイスとなっている。本研究で用いた空気圧ポンプは圧電式（ピエゾ）ダイヤフラムポンプの原理を利用しておらず、超音波振動を応用することで小型かつ薄型でありながら、高い圧力・流量を実現している。

図 2 に示すように、口中香を実現するための機構として、蓋上部からコップ内に香りを噴射する機構を設計した。香りを噴射する図 2・A は、空気圧ポンプと香りが染み込んだコットンを備えるため、3D プリンタでコンテナを作成し、蓋上部に取り付けた。

また図 2・BC に示すように、空気圧ポンプを駆動させるための回路はコンパクトに備えるために蓋上部に取り付ける設計にした。空気圧ポンプは入力電圧によって出力する圧力が変化する特性を持っているため、本デバイスでは空気圧ポンプへの入力電圧を 15V に設定し、空気圧は約 1.6kPa で実験を行なった。これはデバイスから 30cm ほど離れた箇所でほんのり香る程度の空気圧である。また香りにはアロマオイルを使用し、香りを染み込ませたコットンをコンテナ内に収めている。これによりコンテナに溜まった香りを空気圧ポンプによって噴射する機構となっている。

3.3. プロトタイプ

本研究では、口中香を実現するデバイスの設計を目的として、香りをストロー内に噴射するストロー型デバイスとコッ



図 3 実験の様子

プ内に香りを噴射するコップ型デバイスを提案する。冷たい飲み物と温かい飲み物に対応するため、飲み物を飲む際に日常的に使用する道具であるストローとコップに着目した。

図 1 (a) に示すストロー型デバイスは、ストローの中心部から別口が生えており、別口から香りをストロー内に噴射する機構で口の中に香りを届ける。また図 1 (b) に示すコップ型デバイスは、コップ上部に香りを噴射するデバイスを設置し、コップ内に香りを噴射する機構することで飲料時に口の中に香りを届ける。

4. 調査実験

本研究では、Kaolid のプロトタイプを用いて各種フレーバーが味覚に及ぼす影響を調査するため、感覚的な官能評価実験を実施した。また Kaolid と各種フレーバーで実現する味の評価を行うためにアンケートを実施した。

4.1. 参加者

本実験では、所属研究室の学生 10 名が参加し、年齢は 20 代（平均年齢 23.9 歳）であった。被験者の健康状態に問題はなく（風邪や発熱などの症状はない）、味覚や嗅覚についても正常な状態で実験は行われた。

4.2. 実験概要

本実験では、Kaolid のプロトタイプ別に異なる飲料を用意した。冷たい飲み物で利用するストロー型デバイスには、香りのない炭酸水を用いた。また温かい飲み物で利用するコップ型デバイスには、白湯を用いた。香りは人が感じることのできる 10 種類の匂いのグループ [38] から炭酸水、白湯と相性が良いとされる柑橘系のグループからオレンジ、パイナップルを選択した。利用するデバイスの順番によるバイアスを防ぐため、順番が異なるように 2 種類の実験パターンを用意し実施した。また香りによるバイアスを防ぐため、各刺激の間に 5 分間休憩し水を飲んでもらった。各試行後、試飲に対するアンケートに回答してもらった。アンケートは、9 段階のヘドニックテストを採用した。被験者には、飲

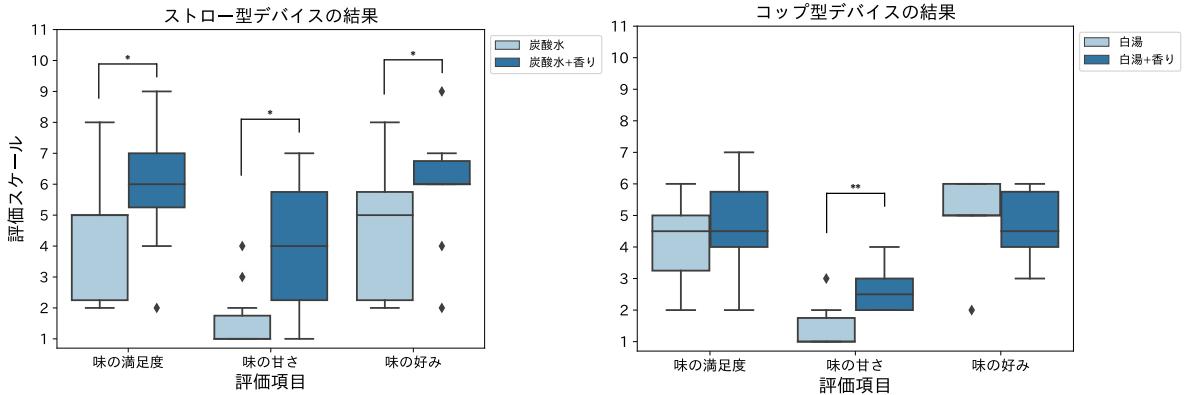


図 4 全体の結果（左・ストロー型、右・カップ型）

なんだ際に知覚した味の満足度、味の甘さ、味の好み、具体的な甘さの評価を行ってもらった。具体的な甘さの評価では、図 3 に示すように、被験者の前に糖分量の異なる砂糖水を 10 カップ用意し、飲んだ際に感じた甘さに近いものを選択してもらった。異なる砂糖水 10 カップは、「1」は砂糖が含まれない水であり、「9」は一般的な炭酸飲料の砂糖含有量 11.3g/100ml の砂糖が入った砂糖水を用意した。その他は等分になるように砂糖水を用意した。「10」は一般的な炭酸飲料よりも甘く感じた際に選択してもらうために用意した。これを行うことで Kaolid を利用した際に感じた味の甘さと実際の糖分量で比較することができる。

5. 実験結果

はじめに、本実験で用いた香り（オレンジ、パイナップル）の結果を統合し、評価を行った。次に香りによる味の評価を比較するため、本実験で用いた香り（オレンジ、パイナップル）別に評価を行った。最後に具体的な味の甘さの評価結果を示す。

5.1. 全体結果

図 4 にそれぞれストロー型デバイス、カップ型デバイスの結果を示す。ストロー型デバイスを用いた味の評価では、香りを付与していない場合に比べ、全ての評価項目（味の満足度、味の甘さ、味の好み）で評価が高くなる傾向が見られた。特に味の甘さに関しては、味のない炭酸水に口中香を提示することで大幅に知覚する甘さが増幅することが確認された。つまり、ストロー型の Kaolid によって口中香を実現し、ユーザの飲料体験を拡張することが可能であることが示された。

カップ型デバイスを用いた味の評価では、香りを付与することで味の満足度、味の甘さに関しては、香りを付与していない場合に比べ評価が高くなる傾向が見られた。しかし、味の好みに関しては香りがない方が評価が高くなつた。

カップ型デバイスでは、味の満足度や味の甘さは上昇する傾向が見られたが、ユーザの味の好みには合わない結果となつた。つまり、カップ型の Kaolid によってユーザの飲料体験を拡張するかに関しては今後議論していく必要がある。

また味の変化に関しては、ストロー型デバイスの方がカップ型デバイスよりも変化すると感じたユーザが多いことがわかった。被験者のコメントでは、「(ストロー型デバイスに用いた飲料が) 炭酸だったので、より味の変化を感じられた。」というコメントがいくつか見られた。他には、「ストローの方が匂いが無駄なく口に入る分、味の変化が大きい結果になると感じた。」と回答する人もいた。またカップ型に関しては、「味の変化はストロー型の方が感じられたが、口当たりの好みはカップ型であった。」と回答する人がいた。

つまり、ストロー型デバイスの方が香りが直接口に運ばれるため、味の変化を感じやすいことが被験者のコメントより確認された。しかし、ストロー型デバイスは香りがストローを伝って口に入るため、普段とは飲み心地が異なつた結果、カップ型の方を好む人も見られた。

続いて、2組の標本に対応があるときに有意差検定として用いられる Wilcoxon の符号順位検定を使用し、香りの有無による各評価項目を定量的に比較した。各評価項目の検定結果を表 1 にまとめる。検定の結果、ストロー型デバイスでは、すべての評価項目で有意差が認められた。またカップ型デバイスでは、味の甘さの評価で有意差が認められた。

5.2. 香り別の分析結果

次に香りの違いによる味覚の調査を行うため、オレンジとパイナップルの香り別に評価を行った。

5.2.1 オレンジの香りの結果

オレンジの香りを用いた際の実験結果を図 5 に示す。ストロー型デバイスに関しては、全体の分析結果同様、全ての評価項目（味の満足度、味の甘さ、味の好み）で香りを付与していない場合に比べ評価が高くなる傾向が見られた。ま

表 1 香りの有無による各評価項目の検定結果

比較する 2 組		p 値	有意差 ($p \leq 0.05$)
ストロー型デバイス	味の満足度 (炭酸水) & 味の満足度 (炭酸水 + 香り)	1.032×10^{-2}	✓
	味の甘さ (炭酸水) & 味の甘さ (炭酸水 + 香り)	1.755×10^{-2}	✓
	味の好み (炭酸水) & 味の好み (炭酸水 + 香り)	2.778×10^{-2}	✓
コップ型デバイス	味の満足度 (白湯) & 味の満足度 (白湯 + 香り)	2.720×10^{-1}	-
	味の甘さ (白湯) & 味の甘さ (白湯 + 香り)	9.765×10^{-3}	✓
	味の好み (白湯) & 味の好み (白湯 + 香り)	3.506×10^{-1}	-

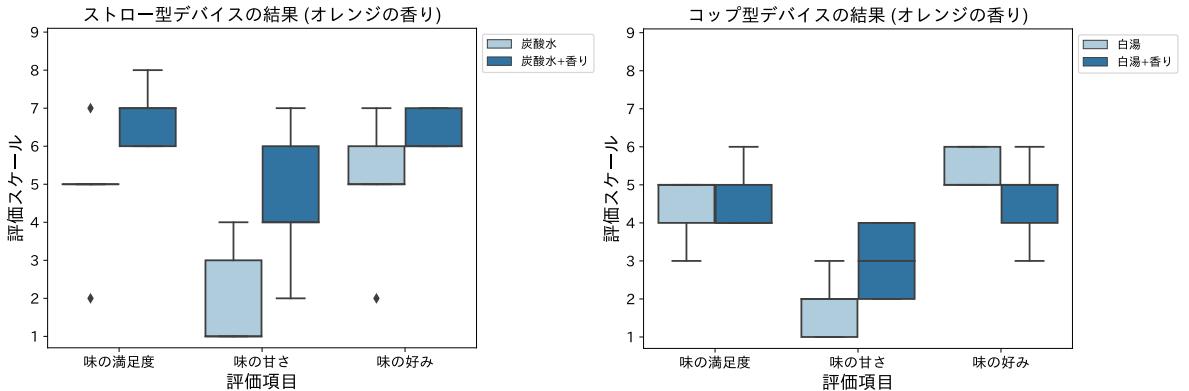


図 5 オレンジの香りの結果 (左・ストロー型, 右・コップ型)

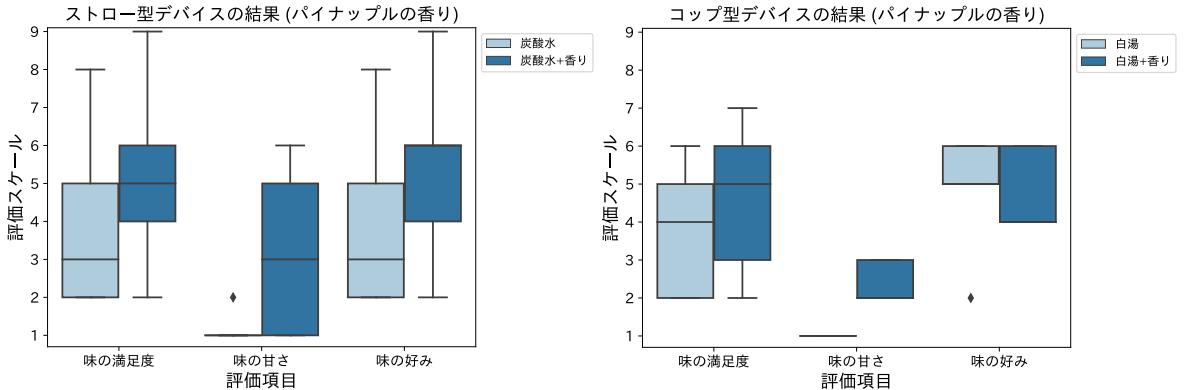


図 6 パイナップルの香りの結果 (左・ストロー型, 右・コップ型)

た味の満足度と味の好みに関しては、標準偏差の結果より、散らばりが少なく評価が高い傾向が確認された。味の甘さに関しては、全体分析結果に比べさらに味の甘さが増幅する傾向が確認された。

一方、コップ型デバイスでは、味の甘さに関しては増幅する傾向が確認されたが、味の満足度は香りがない場合に比べてほとんど変わらない結果となった。味の好みに関しては、全体分析結果同様、香りを付与することで評価が下がる結果となった。

味の変化に関しては、ストロー型デバイスの方が大幅に変化する傾向が見られた。また標準偏差の結果より、データの散らばりも少ないことが確認された。

5.2.2 パイナップルの香りの結果

パイナップルの香りを用いた際の実験結果を図 6 に示す。ストロー型デバイスでは、オレンジの香りの結果に比べ、全ての評価項目で低い傾向が確認された。また、各評価項目でデータのばらつきが大きく見られ、個人の好みの差が大きく見られた。コップ型デバイスでは、オレンジの香りと同様の傾向が確認された。

香り別に分析を行った結果、ストロー型デバイスではパイナップルの香りよりオレンジの香りの方が好まれる傾向が見られた。これは本実験で使用した飲料が影響していると考えられる。ストロー型デバイスには、飲料として炭酸水を

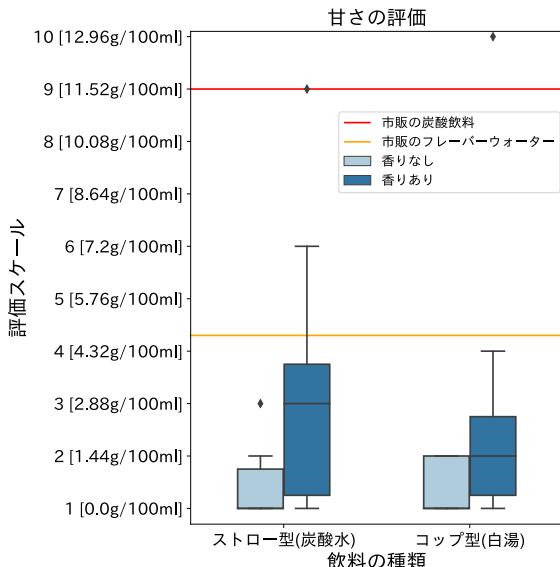


図 7 甘さの評価結果

用いたが、炭酸水と相性が良い香りとしてオレンジであることが確認された。本実験で用いた香りの選定には、人が感じることのできる 10 個の匂いのグループ [38] から 2 つのグループから選定を行った。つまり、オレンジのグループの香りが炭酸水と相性が良い可能性が示された。

またカップ型デバイスでは、味の甘さは増幅する傾向が確認されたが、香りの種類に関係なく満足度や味の好み、味の変化は有意に上昇する傾向は確認されなかった。つまり、現状のカップ型デバイスでは、口中香の実現は難しい可能性が示された。

5.3. 味の甘さの評価

具体的にどれくらいの味の甘さを感じるかを評価した結果を図 7 に示す。市販の炭酸飲料の糖分量を赤線、市販のフレーバーウォーターの糖分量をオレンジ線で示す。結果より、炭酸水を用いたストロー型デバイスの場合では香りを提示することで平均でスケール約 2 分の味の甘さ増幅を記録した。評価スケール 1 つで 1.44g/100ml の糖分量であったため、中央値で比較した場合、ストロー型デバイスでは約 2.88g 分の甘さ増幅を実現したことがわかる。また被験者によっては、市販の炭酸飲料の甘さと同等の評価をしていることがわかった。また白湯を用いたカップ型デバイスの場合では、ストロー型デバイスには劣るが、同様に甘さ増幅の傾向が見られた。カップ型では、一般的な炭酸飲料よりも甘く感じたと評価した被験者も見られた。つまり、Kaolid を用いることで味の甘さ増幅を実現し、糖分摂取量低減を補助する可能性が見られた。

6. おわりに

本実験により、香りがない場合に比べ、Kaolid を利用することで味の満足度、味の甘さが大幅に上昇する傾向が確認された。また検定の結果、被験者数は 10 名であったが、ストロー型デバイスではすべての評価項目で有意差が確認され、カップ型デバイスでも味の甘さで有意差が確認された。今後はさらに大規模に実験を行うことで、Kaolid の有効性を示していく。

また香り別に分析を行うことで、飲料によって相性の良い香りが変わる可能性が示唆された。本実験では、香りにオレンジ、パイナップルを用いたが、人が感じることのできる 10 個の匂いのグループ [38] に基づき、飲料と相性の良い香りを探索する必要がある。またデバイス別では、ストロー型デバイスは口中香を実現した可能性が見られたが、カップ型デバイスに関しては、口中香の実現はできていない可能性が見られた。今後は、白湯などの温かい飲み物でも口中香を実現するデバイスを模索する必要がある。

参考文献

- [1] Malik, V., Willett, W. and Hu, F.: Global obesity: trends, risk factors and policy implications, *Nature Reviews Endocrinology*, Vol. 9, No. 1, pp. 13–27 (2013).
- [2] World Health Organization: Diabetes, Available online: <https://www.who.int/news-room/fact-sheets/detail/diabetes>. Accessed: 2022-02-25.
- [3] Schulze, M. B., Manson, J. E., Ludwig, D. S., Colditz, G. A., Stampfer, M. J., Willett, W. C. and Hu, F. B.: Sugar-sweetened beverages, weight gain, and incidence of type 2 diabetes in young and middle-aged women, *Jama*, Vol. 292, No. 8, pp. 927–934 (2004).
- [4] Singh, G. M., Micha, R., Khatibzadeh, S., Shi, P., Lim, S., Andrews, K. G., Engell, R. E., Ezzati, M., Mozaffarian, D., of Diseases Nutrition, G. B. and (NutriCoDE), C. D. E. G.: Global, regional, and national consumption of sugar-sweetened beverages, fruit juices, and milk: a systematic assessment of beverage intake in 187 countries, *PloS one*, Vol. 10, No. 8, p. e0124845 (2015).
- [5] Gross, L. S., Li, L., Ford, E. S. and Liu, S.: Increased consumption of refined carbohydrates and the epidemic of type 2 diabetes in the United States: an ecologic assessment, *The American journal of clinical nutrition*, Vol. 79, No. 5, pp. 774–779 (2004).
- [6] Vartanian, L. R., Schwartz, M. B. and Brownell, K. D.: Effects of soft drink consumption on nutrition and health: a systematic review and meta-analysis, *American journal of public health*, Vol. 97, No. 4, pp. 667–675 (2007).
- [7] Te Morenga, L., Mallard, S. and Mann, J.: Dietary sugars and body weight: systematic review and meta-analyses of randomised controlled trials and cohort studies, *Bmj*, Vol. 346 (2013).

- [8] Malik, V. S., Pan, A., Willett, W. C. and Hu, F. B.: Sugar-sweetened beverages and weight gain in children and adults: a systematic review and meta-analysis, *The American journal of clinical nutrition*, Vol. 98, No. 4, pp. 1084–1102 (2013).
- [9] Imamura, F., LauraO' Connor, Ye, Z., Mursu, J., Hayashino, Y., Bhupathiraju, S. N., Forouhi, N. G. Consumption of sugar sweetened beverages, artificially sweetened beverages, and fruit juice and incidence of type 2 diabetes: systematic review, meta-analysis, and estimation of population attributable fraction, *Bmj*, Vol. 351 (2015).
- [10] Yin, J., Zhu, Y., Malik, V., Li, X., Peng, X., Zhang, F. F., Shan, Z. and Liu, L.: Intake of sugar-sweetened and low-calorie sweetened beverages and risk of cardiovascular disease: a meta-analysis and systematic review, *Advances in Nutrition*, Vol. 12, No. 1, pp. 89–101 (2021).
- [11] Malik, V. S. and Hu, F. B.: The role of sugar-sweetened beverages in the global epidemics of obesity and chronic diseases, *Nature Reviews Endocrinology*, Vol. 18, No. 4, pp. 205–218 (2022).
- [12] Popkin, B. M. and Hawkes, C.: Sweetening of the global diet, particularly beverages: patterns, trends, and policy responses, *The lancet Diabetes & endocrinology*, Vol. 4, No. 2, pp. 174–186 (2016).
- [13] Muth, N. D., Dietz, W. H., Magge, S. N., Johnson, R. K., Bolling, C. F., Armstrong, S. C., Haemer, M. A., Rausch, J. C., Rogers, V. W., Abrams, S. A. et al.: Public policies to reduce sugary drink consumption in children and adolescents, *Pediatrics*, Vol. 143, No. 4 (2019).
- [14] Hoenink, J., Stuber, J., Lakerveld, J., Waterlander, W., Beulens, J. and Mackenbach, J.: The effect of on-shelf sugar labeling on beverage sales in the supermarket: a comparative interrupted time series analysis of a natural experiment, *ISBNPA*, Vol. 18, No. 1, pp. 1–11 (2021).
- [15] Villinger, K., Wahl, D., Engel, K. and Renner, B.: Nudging sugar portions: a real-world experiment, *BMC nutrition*, Vol. 7, No. 1, pp. 1–5 (2021).
- [16] Lyman, B.: *A psychology of food: More than a matter of taste*, Springer Science & Business Media (2012).
- [17] Rosenblum, L. D.: *See what I'm saying: The extraordinary powers of our five senses*, WW Norton & Company (2011).
- [18] Stuckey, B.: *Taste what you're missing: the passionate eater's guide to why good food tastes good*, Simon and Schuster (2012).
- [19] 真弓大輝, 中村優吾, 松田裕貴, 安本慶一 Aromug : 糖分摂取量低減を補助するスマートマグカップの検討, 情報処理学会 IoT 行動変容学研究グループ キックオフシンポジウム, pp. 1–4 (2022).
- [20] 真弓大輝, 中村優吾, 三崎慎也, 松田裕貴, 安本慶一 Aromug: 糖分摂取量低減を補助するスマートマグカップの設計と基礎評価, マルチメディア, 分散, 協調とモバイル (DICOMO 2022) シンポジウム論文集, pp. 981–988 (2022).
- [21] Mayumi, D., Nakamura, Y., Misaki, S., Matsuda, Y. and Yasumoto, K.: Aromug: Mug-type Olfactory Interface to Assist in Reducing Sugar Intake, *The 5th International Workshop on Computing for Well-Being (WellComp '22)* (2022).
- [22] Heilmann, S. and Hummel, T.: A new method for comparing orthonasal and retronasal olfaction., *Behavioral neuroscience*, Vol. 118, No. 2, p. 412 (2004).
- [23] Small, D. M., Gerber, J. C., Mak, Y. E. and Hummel, T.: Differential neural responses evoked by orthonasal versus retronasal odorant perception in humans, *Neuron*, Vol. 47, No. 4, pp. 593–605 (2005).
- [24] Bojanowski, V. and Hummel, T.: Retronasal perception of odors, *Physiology & behavior*, Vol. 107, No. 4, pp. 484–487 (2012).
- [25] Rozin, P.: " Taste-smell confusions" and the duality of the olfactory sense., *Perception & psychophysics* (1982).
- [26] Fowler, S. P., Williams, K., Resendez, R. G., Hunt, K. J., Hazuda, H. P. and Stern, M. P.: Fueling the obesity epidemic? Artificially sweetened beverage use and long-term weight gain, *Obesity*, Vol. 16, No. 8, pp. 1894–1900 (2008).
- [27] Ebbeling, C. B., Feldman, H. A., Osganian, S. K., Chomitz, V. R., Ellenbogen, S. J. and Ludwig, D. S.: Effects of decreasing sugar-sweetened beverage consumption on body weight in adolescents: a randomized, controlled pilot study, *Pediatrics*, Vol. 117, No. 3, pp. 673–680 (2006).
- [28] Black, R. M., Leiter, L. A. and Anderson, G. H.: Consuming aspartame with and without taste: differential effects on appetite and food intake of young adult males, *Physiology & behavior*, Vol. 53, No. 3, pp. 459–466 (1993).
- [29] Bazzano, L. A., Li, T. Y., Joshipura, K. J. and Hu, F. B.: Intake of fruit, vegetables, and fruit juices and risk of diabetes in women, *Diabetes care*, Vol. 31, No. 7, pp. 1311–1317 (2008).
- [30] Yamada, T., Yokoyama, S., Tanikawa, T., Hirota, K. and Hirose, M.: Wearable olfactory display: Using odor in outdoor environment, *IEEE Virtual Reality Conference (VR 2006)*, IEEE, pp. 199–206 (2006).
- [31] Narumi, T., Nishizaka, S., Kajinami, T., Tanikawa, T. and Hirose, M.: Augmented reality flavors: gustatory display based on edible marker and cross-modal interaction, *Proceedings of the SIGCHI conference on human factors in computing systems*, pp. 93–102 (2011).
- [32] Amores, J. and Maes, P.: Essence: Olfactory interfaces for unconscious influence of mood and cognitive performance, *Proceedings of the 2017 CHI conference on human factors in computing systems*, pp. 28–34 (2017).
- [33] Dobbelstein, D., Herrdum, S. and Rukzio, E.: inScent: A wearable olfactory display as an amplification for mobile notifications, *Proceedings of the 2017 ACM International Symposium on Wearable Computers*, pp. 130–137 (2017).
- [34] Wang, Y., Amores, J. and Maes, P.: On-face olfactory

- interfaces, *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pp. 1–9 (2020).
- [35] Choi, Y., Parsani, R., Roman, X., Pandey, A. V. and Cheok, A. D.: Sound perfume: building positive impression during face-to-face communication, *SIGGRAPH Asia 2012 Emerging Technologies*, pp. 1–3 (2012).
- [36] Ranasinghe, N., Suthokumar, G., Lee, K.-Y. and Do, E. Y.-L.: Digital flavor: towards digitally simulating virtual flavors, *Proceedings of the 2015 ACM on international conference on multimodal interaction*, pp. 139–146 (2015).
- [37] Ranasinghe, N., James, M. N., Gecawicz, M., Bland, J. and Smith, D.: Influence of Electric Taste, Smell, Color, and Thermal Sensory Modalities on the Liking and Mediated Emotions of Virtual Flavor Perception, *Proceedings of the 2020 International Conference on Multimodal Interaction*, pp. 296–304 (2020).
- [38] Castro, J. B., Ramanathan, A. and Chennubhotla, C. S.: Categorical dimensions of human odor descriptor space revealed by non-negative matrix factorization, *PloS one*, Vol. 8, No. 9, p. e73289 (2013).

Vacuum Colorner: 複合現実を用いた掃除における行動変容を促すアプリケーションの提案

浜田惇矢 ^{†1}

慶應義塾大学総合政策学部

浜中智己, 佐々木航 ^{†2}

慶應義塾大学院政策・メディア研究科

大越匡 ^{†3}

慶應義塾大学環境情報学部

中澤仁 ^{†4}

慶應義塾大学環境情報学部

概要

家事は生活する上で日常的に多くの人が取り組んでいる一方で、煩わしさや面倒臭さ、ストレスを感じる場合がある。さらに家事に対するストレスは家庭内ストレスの一種である。家庭内ストレスを日常的に感じ続けることにより、人々の主観的健康にも影響を与えることがわかっている。本研究では、MR（複合現実）を用いて家事における掃除機がけをストレスフリーかつゲーム感覚で行えるゲーミフィケーションを応用したアプリケーション Vacuum Colorner を提案する。Vacuum Colorner を用いることで、掃除機の位置情報をもとに掃除をした箇所を青、赤、黄色、半透明の 4 色から好きな色で塗ることができる。これによりユーザーは掃除をした位置を MR 経由で視覚的に認識し、掃除に対する達成感を感じられる。本システムの有効性を検証した結果、定量的な比較項目である面積とストレス値において、視覚的な介入を行わなかった場合と比較して統計的に有意となった。よって、MR や色塗りといった視覚情報を介入方法として、家事における掃除を対象に行動変容を促せる可能性を見出した。

1. はじめに

家庭内での家事を苦痛に感じている人は多数存在する。リンナイ社のアンケート [1] によると家庭内で感じるストレスに関する意識調査において 25 % の人が家事にストレスを感じると回答した。また、家庭内ストレスは自覚的健康度に悪影響 [2] を与えているという事もわかっているため家事

のストレスはより健康的に幸福に生活することにおいて妨げになっている。また、家事がストレスとなる要因は評価されない事と、義務感 [3] である。本研究では家事をよりストレスフリーに娯楽感覚で行うことで義務感を排除することを目的としている。そのためにゲーミフィケーション [4] という手法を利用し、より家事を娯楽に近づける。ゲーミフィケーションとはタスクにゲームの要素を加えることで意欲の向上や、ストレスの軽減を図る手法である。本研究ではゲーミフィケーションの即時フィードバックと達成感の項目を利用し、掃除機をかけたところが色で塗られ、かけたところがすぐわかる即時フィードバック、かけ終わったタイミングで綺麗に、多く色が塗られていることを確認し達成感を味わうことのできるアプリケーションを開発した。視覚情報の追加には複合現実を用いた。複合現実は現実世界との相互的なインタラクションが可能であるが故に現実世界とシステムによる視覚情報の差が少ないため、我々の日常生活に溶け込みやすくかつ、新たな価値の創造が容易である。そこで、本研究を通して普段の掃除機掛けに違和感のない視覚情報を追加し、色が塗られていくという新たな価値を付随できるアプリケーションを提案し、複合現実を用いた掃除機行為への介入による行動変容について調査した。調査をおこなった掃除機行為に対する行動変容の項目は面積、距離、時間、ストレス値について評価を行い、面積とストレス値において無介入時とシステムによる介入時において統計的な有意差が確認された。

本論文は以下のように構成される。3 章では本研究の関連研究についてゲーミフィケーションを利用した関連研究と複合現実についての関連研究に分けて提示する。3 章では複合現実を用いた掃除活性化システム「Vacuum Colorner」についての詳細を述べる。4 章では評価実験の手法について述べる。5 章では評価に基づく考察を行い、6 章で本研究のまとめと、今後の展望を結論として述べる。

Vacuum Colorner: Proposal for an Application to Promote Behavior Change in Cleaning Using Mixed Reality

^{†1} ATSUYA HAMADA, Faculty of Policy Management, Keio University

^{†2} SATOKI HAMANAKA, WATARU SASAKI, Graduate School of Media and Governance, Keio University

^{†3} TADASHI OKOSHI, Faculty of Environment and Information Studies, Keio University

^{†4} JIN NAKAZAWA, Faculty of Environment and Information Studies, Keio University

2. 関連研究

本研究は介入方法にゲーミフィケーションとその手法として複合現実を使用している。掃除機行為に対するゲーミフィケーションの介入についての関連研究と複合現実による掃除機行為に対する介入についての関連研究の2種類に分けて考える。

2.1. ゲーミフィケーションを利用した家事促進

小笠原ら [5] 市村ら [6] 矢田ら [7] は、掃除機行為にゲーミフィケーションを応用した。また、小坂ら [8] は「掃除機をかける」という行為をゲームコンテンツに設定することで子供たちの行動を誘導、誘発し、楽しみながら掃除機かけを行うことを目指すシリアルゲーム「妖怪クリーナーズ」の提案を行った。ホコリセンサを利用して、ごみの中の妖怪を吸い込むという対戦型の掃除機がけをすることによるゲームの提案を行った。また、その表示は画面だけでなく、実際に掃除機にもエフェクトがあり、現実空間とゲームをうまく結びつけることで、掃除機によりゲームの感覚に落とし込んだ。また、他のゲーミフィケーションを盛り込んだ掃除機行為とは違い、現実世界へのアプローチがある。それは視覚的に掃除機で吸い込んだ時にノズルが光るという点である。やはりゲーム画面を通してではなく現実世界に影響することは直感的であり、人々(特に子供達)の楽しみを掻き立てるものであることが分かった。これら関連研究のようなゲーミフィケーションの掃除機行為への適応は多く行われてきた。その結果として、システム使用時の掃除機行為に対する意欲の向上や意識の変化が見られた。

また、これらの研究では掃除に対して加速度センサやプロジェクトを設置する過程が必要となっていることや対象が子どもであったり、普段の掃除機を利用した状態でシステムを利用する事が難しかった。本研究では、より掃除機行為に対するゲーミフィケーションを馴染みあるものにするため、掃除機への設置はQRコードのみとし、気軽に家庭でも再現可能な形をとることとした。また、複合現実を利用することでより現実に近い視覚情報を追加することなど、全年齢に向けて行動変容を図るようにした。

2.2. 家事行為への MR(複合現実) 利用例

本研究では掃除機行為を別のディスプレイを参照することなく直感的にストレスフリーに楽しくすることを目的としているため、複合現実を利用した。ロペズら [9] はMRを用いて掃除の面倒さや手間を解消し、掃除のモチベーションを向上して楽しく掃除を行うための掃除活性化システム「Dustoon」を提案した。掃除機状況の可視化することで、重複掃除の防止や、掃除不足箇所の特定を可能にした。塵濃度センサから取得した塵濃度を5秒おきにデバイスに送

信し、塵濃度をもとに色を変化させたオブジェクトを3次元マーカをカメラの画像認識により追跡し位置を合わせて描画するシステムであった。

本研究ではオブジェクトの位置合わせに誰でも簡単に設置や入手の可能であるQRコードを利用した。それにより自宅でMRデバイスに加えて、プリンタ等でQRコードを用意、配置すれば普段の掃除機を用いたとしてもシステムの利用を可能とした。また、QRコードは動画撮影用の撮影用カメラではなく空間認識カメラにより位置情報を取得することができる。そのため、撮影用カメラを作動させる必要がなく動作が軽く、かつ認識対照が掃除機のノズルの分遠距離になったとしてもスムーズな動作を可能とした。

3. 複合現実を用いた掃除活性化システム

本研究では掃除機をストレスフリーに行うために掃除機をかけたところが色が塗られるシステムを実装した。色を塗ること自体にストレス軽減の効果 [10] があるだけでなく、掃除機の重複や掃除残しを確認することで視覚的な楽しさ、達成感を味わうことのできる掃除機体験を提供する。システム構成図は以下の図1の通りである。

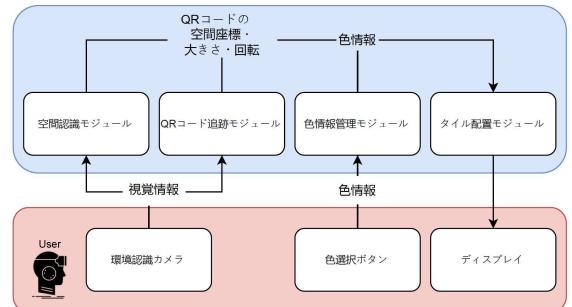


図1 システム構成図

本研究ではMRデバイスにMicrosoft社のhololens2を利用した。実装にはUnity(2020.3.38f1)を用いた。まず、環境認識カメラによって空間の座標を取得する。加えて環境認識カメラより、QRコードを追跡する。掃除機にQRコードを設置することで、実質的に掃除機の追跡を可能にした。これら二つの情報から、QRコードの空間内での座標、大きさ、回転を取得し、それをもとにタイルを設置する。タイルはQRコードと同じサイズのタイルを掃除機側に少しづらした位置に配置する。こうすることで、ユーザーまるで掃除機をかけたところが色が塗られているかのような体験を与える。また、ユーザーにとって色塗りが楽しい体験になるように好きな色で塗る機能を付した。色選択はアプリケーション開始時の目線の位置に常駐させる形で表示させ、色は4色(青、赤、黄色、半透明)を用意し、ボタンを押下することで、各色に変更することができる。以上のよ

うに 4 色の好きな色を選択し掃除機をかける事でその色で掃除機をかけた軌跡が見れるアプリケーションを実装した。使用時の外からの様子を以下の図 2 に示す。



図 2 システム使用の様子

3.1. 空間認識

本研究では hololens2 をデバイスとして使用しており、hololens2 は 2 種類のカメラを搭載している。1 種類目はデバイス中央の環境認識カメラ、2 種類目はデバイスの左右に 2 つ付いている 4 つの環境認識カメラである。空間認識は hololens2 を起動したタイミングから常に空間の状態を把握し続ける。これにより、オブジェクトを配置する際により自然に適切な位置に配置することができる。例えば机の裏にオブジェクトが回り込んだ場合には表示を止めることや、実世界の物体とオブジェクトが重ならないように実装が可能である。また、本研究には実装していないが、オブジェクトに重力をつけた際には床におち、重心の安定した状態になるなど、オブジェクトをより現実世界の物体に近いものにすることが可能である。本研究ではデバイスの空間認識による座標に対して、適切な位置に座標を合わせることにより、ユーザに直感的な視覚情報を与えている。

3.2. QR コードの座標取得

本研究では掃除機の検出に QR コードを用いた。掃除機を掃除機であると認識することは技術的に可能であるが、掃除機を検出する際に使用する撮影用のカメラを起動しなくてはならないために、動作が重くなる。MR は空間把握を正確に行うため、他の動的処理を加えると動作がスムーズにならない欠点がある。本研究ではなるべくデバイスへの負荷を削減するため、撮影用カメラを用いることな

く、掃除機の追跡を行うこととした。QR コードの追跡は Microsoft.MixedReality.QR.QRCodeWatcher という API プラグインファイル [11] を利用した。本研究では QR コードにマイクロ QR コードを使用すること、光の反射防止のため画用紙にペントアブラックを使用して QR コードを描画することで長い距離においてもサイズの小さい QR コードでの検出を可能としている。180cm の人が使用した際、掃除機のヘッドに取り付けた QR コードまでの距離は約 230cm である。この時過不足なく反応するマイクロ QR コードのサイズは約 15cm 四方であり、本研究では約 14cm 四方の QR コードを利用した。

3.3. 色選択

より色塗りを楽しい物にするために好きな色を選択できるようにした。色は全部で 4 色あり青、赤、黄色、半透明である。色選択はボタンができるようしている。使用感を hololens2 の使用時と同じようにするために MRTK のボタンオブジェクトを利用している。ボタンの利用感はハンドトラッキングにより認識している手の人差し指の先にポインタが表示されるようになっている。ポインタがボタンに乗っている時、ボタンの前方に透明のボックスが表示される。これにより、ボタン一つ一つが隣り合っていてどのボタンを参照しているのか分かりやすくなっている。さらに、ボタンを押して選択というテキストを表示し、UI を分かりやすくしている。ボタンを押した際のトリガーはボタンをタッチしている時、押し始めた時、押し切った時等あるが、本システムではボタンを押し切った時をトリガーとして、タイルのマテリアルの変更を行っている。ボタンは様々なジェスチャーにより出したりしまったりすることもできるが、掃除機行為の最中に意識せずにボタンが表示することを防ぐため、今回は常にボタンが出てる状態でボタンの位置は起動時の目線のポジションにする。そうすることで起動時に色選択ができるようにし、かつ掃除機行為の邪魔にならないようにした。ユーザ側の操作は単純にするためにボタン操作のみとしている。

3.4. タイル設置

空間認識により、周りの空間情報を取得し、QR コード追跡により掃除機の位置情報を空間上の座標として取得する。タイルは厚みのないタイルオブジェクトを使用し、QR コードよりも少し掃除機側にずらして設置する事で掃除機をかけた所が色が塗られているように見せる。また、タイルの大きさは設置した QR コードの大きさと同じにしてある。掃除機の大きさに応じて QR コードの大きさを変えることでどの掃除機に対してもシステムを利用出来るようにするためである。また、より自然に見せるために、重力や物体の衝

表 1 実験概要

	A 群(5名)	B 群(5名)	C 群(5名)	D 群(5名)	E 群(5名)	F 群(5名)
30 秒			唾液アミラーゼ検査(1回目)			
1 分半以上	普段通り	普段通り	デバイス装着	デバイス装着	システム使用	システム使用
30 秒			唾液アミラーゼ検査(2回目)			
1 分半以上	デバイス装着	システム使用	普段通り	システム使用	普段通り	デバイス装着
30 秒			唾液アミラーゼ検査(3回目)			
1 分半以上	システム使用	デバイス装着	システム使用	普段通り	デバイス装着	普段通り
30 秒			唾液アミラーゼ検査(4回目)			

突は行わないものとした。物体の物理演算を行わないことによりシステムをスムーズに動作させる目的もある。また、配置には Unity のプレハブにタイルを登録し、プレハブからオブジェクトを読み込むことで、タイルを配置している。色選択により、プレハブの色変更を行っているため、指定した位置に指定した色のタイルを配置している。使用時の様子を下記の図 3 に示す。

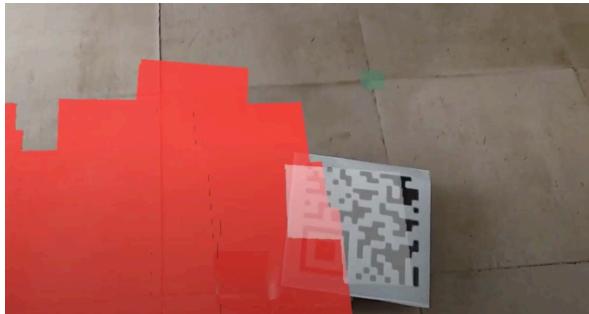


図 3 システム使用の様子

4. 評価実験

4.1. 実験設計

本研究では大学生 30 名の被験者を集め、実験を行った。実験は各工程に対し、1 教室 (4.05m, 8.55m) を自由に掃除してもらい、最低 1 分半、上限は満足いくまで掃除機をかけてもらうという実験を行う。

実験では唾液アミラーゼによるストレス評価と、掃除機をかけた面積、距離、時間を測る。各工程の前後でニプロ社の唾液アミラーゼモニターを利用して、唾液アミラーゼの値を取得する。中野ら [12] は唾液アミラーゼが急性のストレス評価に適していることを示し、快・不快の判別が可能であることを示した。唾液は非侵襲で、随時性、簡易性に優れ、血液等のように採取がストレス要因にならないことから、短期的かつ快・不快のストレス値を複数回取得する本研究と非常に相性がよいと考えたため、唾液アミラーゼをストレス評価指標として採用した。唾液アミラーゼ活性の生理的特徴として、朝低く午後に上昇し [13]、就寝中は再び

低い値 [14] となるという日内変動がある。その差には統計的な有意差が示されていない [15] が、個人差がある可能性を考慮し、唾液アミラーゼの絶対値ではなく、 SAA_0 をストレッサーを加える前の唾液アミラーゼ活性 (ベースライン kU/l)、 SAA をストレッサーを加えた後の唾液アミラーゼ活性 (kU/l) と定義したとき、

$$\Delta \frac{SAA - SAA_0}{SAA_0} \times 100 \quad (1)$$

にてアミラーゼ活性の絶対値から変化率 (%) を求め、その数値をストレス値として判定する。変化率が大きい方が、よりストレスを感じたことになる。

実験は 3 工程あり、3 工程を連続で行うため、唾液検査、1 工程目、唾液検査、2 工程目、唾液検査、3 工程目、唾液検査という順で唾液アミラーゼを取得する。そのため、1 回の実験で 4 回唾液アミラーゼを取得する、唾液アミラーゼからストレス値を算出する際絶対値と前の値からの差分を記録する。

また掃除機をかけた面積、距離は、掃除機にマウスを取り付けることによって座標を取得し、距離と面積を算出した。使用マウスは ROCCAT 社の Kone Pro Air[16] を使用した。dpi は専用の dpi 変更ソフトウェアの swarm を利用し、50dpi に設定をした。また、距離に関しては 1 フレームごとに x 座標、y 座標を前のフレームで取得した値と比較して距離を算出し、それを距離変数に加算していくことにより、掃除機を始めてから終えるまでのストロークの総距離を算出している。面積はポインタに対して、格子状に区切った部分にポインタが入った場合色を塗る仕様にし、色が塗られている総合計を取得できるようにした。重複を加味しており、実際に目線で見ている色の塗られた範囲に近似した面積を取得した。システム上のエラー等を確認するために、デバイスからのミラーキャストを用いて画面共有を行うことによって被験者の視界を確認しながら実験を進めた。以下の表 1 に実験概要をまとめた。

4.2. 実験後アンケート

実験終了後に評価に必要なアンケートを実施する。アンケート内容は年齢、性別、5 段階評価である

表 2 実験後に行った 5 段階評価のアンケート内容

1	普段の掃除機掛けは楽しいか
2	日常的に掃除機をかけるか
3	普段の家事にストレスを感じるか
4	普段掃除機を掛けることにストレスを感じるか
5	何もつけていない時の掃除機掛けは楽しかったか
6	デバイスをつけているのにストレスを感じたか
7	システムを使って達成感を感じたか
8	システムを使った掃除機は楽しかったか
9	色を塗ることで掃除が楽しくなったか
10	システムを使って日常的に掃除をしたいか

定性評価では定量評価では得ることのできない楽しさや、モチベーションについて調査することが主な目的である。よって、システムを使った際と普段の楽しさやモチベーションを比較できるように 5 段階の数値に落とし込んだアンケートとした。以下の表 2 に 5 段階アンケート内容をまとめめる。

また、5 段階評価では確認できない部分については記述式で解答を求めた。記述式アンケート項目は、

- システムの良かったところ
- システムの改善点
- 実験中にストレスを感じたことがあったか

から構成される。

5. 評価

5.1. 定性評価

実験参加者 30 名に対して 5 段階のアンケートを行った。本アンケートでは 1 がネガティブ、5 がポジティブな評価になるように設定している。以下の表 3 に得られた評価に対する平均値をまとめる。

表 3 5 段階評価のアンケート内容に対する評価

質問項目	平均値
普段の掃除機掛けは楽しいか	1.8
日常的に掃除機をかけるか	2.1
普段の家事にストレスを感じるか	3.8
普段掃除機を掛けることにストレスを感じる	3.3
何もつけていない時の掃除機掛けは楽しかったか	2.4
デバイスをつけているのにストレスを感じたか	2.8
システムを使って達成感を感じたか	3.2
システムを使った掃除機は楽しかったか	3.9
色を塗ることで掃除が楽しくなったか	3.7
システムを使って日常的に掃除をしたいか	2.8

楽しさのアンケートについて、システムによる介入時では中央値として 4 の評価が得られた。また、無介入時では

中央値として 1 の評価が得られた。介入時と比較すると 3 の差があり、通常の掃除機掛けよりもシステムを使った方が掃除機行為が楽しく感じられたことが分かる。普段の掃除機掛けは中央値として 1 の評価が得られ、また、面倒だと感じている(1 を選択した)被験者は 50 %存在し、掃除機がけを面倒だと考えている被験は非常に多いことが分かった。これによりシステムを使った掃除機掛けは普段の掃除機がけに比べて幾つか面白いものであるという評価が得られた。

また、モチベーションについてのアンケートでは普段は中央値として 2 の評価が得られた。対象者が学生であったためか掃除機を日常的にかけたいという被験者は少なかった。また、システムを使用した場合では中央値として 2.5 の評価が得られた。日常的に掃除機をかけるかの評価に対して 0.5 高い評価であり、モチベーションの向上についても効果がある可能性が示唆される。

システムを使って達成感を感じたかというアンケートでは中央値として 4 の評価が得られた。色を塗ることにより、掃除機を行った形跡が見え達成感が得られるのではないかと考えられる。また、色を塗ったことによる楽しさも中央値として 4 の評価が得られた。これにより、システムの色塗り要素に楽しさがあることが推測される。

記述式のユーザアンケートではシステムの良かった点、改善点、実験中にストレスに感じた点について解答を求めた。システムの良かった点については、

- ゲーム感覚で快感を感じながら掃除をすることができ、長時間の掃除機がけも苦ではなくなくなった
- 近未来的
- 掃除したところが可視化できたことで見落としていたところを直ぐに発見できる
- 自分の動きが可視化されて面白かった
- 掃除した軌道が分かりやすく見えたので、掃除し忘れた箇所がないか確認できた

などの回答が得られた。

システムの良かった点について、掃除漏れがなくなることにより、達成感や楽しさを感じるという意見が最も多かった。このアンケートより掃除機をかける行為はやはり綺麗に掃除できていること、目に見える達成感があることにより楽しく行うことができるのではないかと推測できる。また、視覚的な面白さを追加することで掃除機行為よりも色塗り行為を優先し、掃除という行為を家事のようなタスクではなく、色塗り行為を行った際のおまけのように感じられる人も存在した。色塗り行為をゲームのように感じるという回答も多数あり、本システムを使うことで、たまたま掃除機がかけられたという可能性もあると考えられる。

改善点については、

表 4 面積評価統計値

	面積			距離			時間			SRE		
	A	B	C	A	B	C	A	B	C	A	B	C
中央値	13.4	13.4	18.4	106.0	98.2	108.3	1.7	1.7	2.3	35.0	106.0	0.0
平均値	14.5	14.2	18.6	99.6	97.0	115.5	1.9	1.9	2.9	101.3	125.3	18.3
標準偏差	5.1	5.1	6.2	17.0	14.6	30.9	0.5	0.5	1.6	169.2	171.0	115.7

表 5 システムによる介入時と他群との u 検定結果と p 値

	面積	距離	時間	ストレス値
無介入	○	×	×	○
	p = 0.006	p = 0.102	p = 0.182	p = 0.042
装着時	○	○	×	○
	p = 0.002	p = 0.027	p = 0.066	p = 0.008

- 実際にゴミが落ちている時に、色とゴミが混ざってしまうのではないかと思った
- ちょっと重いので、軽かったらいいかなと思った
- 装着しているものが少し重かった
- 部分的ではなく、軌道に色をつけるといいと思った

等の意見が得られた。

もっと面白くできるという回答もあり、具体的には音楽をつけることが挙げられた。また、システムで色を塗る関係上その部分の視界が遮られることで、ゴミが見つけづらくなってしまうという意見がみられた。これについては、逆に掃除機をかけていない部分に色を塗り、掃除機をかけたところの色を消すことが考えられるが、実際に色を塗っていくアプローチを楽しい、面白いという意見があるため、この部分は個人ごとに一長一短である。しかしながら、デバイスのカメラから、ゴミの検知を行いアラートを出すなど、視界が悪くなることを補える要素をシステムに追加することができれば改善できると考えられる。デバイスの動作の関係上綺麗に色を塗ることは難しい、そこについて改善をしている回答が多くみられた。もしスマーズな動作を必要とするのであれば、違うアプローチが必要となってくるだろう。

実験中ストレスを感じた点についてアンケートをとったが、ほとんどがデバイスに対する不満であった。デバイスに対する不満ではやはり重量感や、装着時の不快感が挙げられた。やはりデバイスの装着自体にストレスを感じてた人は多かった。その中には不慣れであるからであるという理由も含まれていたが、初めて使うデバイスとしてはやはりストレスに感じると推察できる。

5.2. 定量評価

定量評価では主に何もつけていない時とシステム使用時を比べどのような変化があったかを検定を通して評価した。特に唾液アミラーゼなど、個人差が生じるデータを評価す

るため、統計的な有意差を示す。本研究では独立した 2 群比較かつ、データに必ずしも正規性があるとが限らない。それゆえ、本研究では正規性のない/分からぬ独立した 2 群比較を行うマン・ホイットニーの U 検定を用いて有意差を検証した。

データの評価値は以下の表 4 のとおりである。

- A: 無介入時
- B: デバイス装着時
- C: システム使用時
- SRE: ストレス値

数値で見るとデバイスをついている際と、何もつけていない時の差はあまり無いと考えられるが、システムを使った際には面積が大きくなっていることがわかる。また、統計値だけでは個人差につき偶然の結果が得られた可能性があるため、統計的有意差を示すために行ったマンホイットニーの u 検定の結果を下記の表 5 に示す。u 検定の帰無仮説は 2 群に差がないことであり、有意確率は 5 %とした。

システムを利用した際に掃除する面積が増えたことについては何もつけていない時、デバイスをついている時どちらと比べても統計的有意差が示され、システムの面積における優位性が示された。また、ストレス値においても何もつけていない時とシステム使用時、デバイス装着時とシステム使用時に統計的有意差があることが示された。これによりシステムを使用することでストレスが軽減されていることが統計的に示され、システムのストレスに対する有効性を明らかにしている。

6. 考察

6.1. 定性評価に対する考察

定性評価では大学生 30 名に対するアンケートにより、現状普段の家事や、掃除機行為に対してストレスを感じている人がほとんどであった。システムを利用した人は他の条

件と比較し楽しんだ傾向にあった。その要因として達成感や色塗りの要素を追加することで楽しさを感じたことが推測される。また、普段の掃除意欲とシステムを使った際の掃除意欲を比べると、システムを使った方が意欲が相対的に高い結果が得られた。これはシステムを使った掃除機行為が楽しいものであるという認識から意欲向上が見られたと考えられる。

システムの良い点として、軌道が見えることで掃除が丁寧になったり、達成感を感じたり、意識が変化することが挙げられた。それに対し、システムの改善点では動作が重いことに起因する色が綺麗に塗られないことや、デバイスとの相性の問題である、目がチカチカする、デバイスが重い、視界が狭いなどの回答があり、これらのほとんどはデバイスの改善により解決すると考えられる。

現状の段階ではスムーズな動きやデバイスの改善点を同時に解決できる技術は見受けられず、デバイスの軽量化、処理能力の向上を待つ他ない。しかしながら、本システムにおいて改善点がありながらも、通常の掃除機行為と比べると楽しさや意欲の点でポジティブな効果が見られており、MR や掃除行為に対しての色塗りのアプローチが効果的であると推測される。

6.2. 定量評価に対する考察

面積とストレス値においてシステムの優位性が示された。面積や距離においては色塗りの要素があるが故に逆に丁寧にかけている様子が見受けられたが、それ以上に掃除時間が長かったことから、面積や距離の値が大きくなったものであると考えられる。距離に関しては無介入時のストロークの速度が非常に速い被験者が数名いたため、統計的な有意差が出せなかったと考えられる。時間においては 1.5 以上の制約をかけてしまったが故にほとんどの被験者が 1.5 分で掃除機掛けをやめてしまったことが、統計的な有意差が出せなかった原因であると考えられる。しかしながら、距離、時間に關しても中央値や平均値だけでみると優位性を見てとれ、データの数を増やすなどすれば統計的な有意も示せるのではないかと考えられる。また、デバイスをかけた際のストレスは定性評価ほど大きいものは得られていないかったが、明らかにデバイスをつけた際の評価が低いことが多くこれは視界が狭まっていることや、デバイスへの不慣れさからくる物理的なものであるとも考えられる。

7. 今後の展望

本システムは少し早く掃除機掛けをすると、QR コードの検出が遅れてしまい、タイルとタイルに隙間が生まれてしまう。隙間が生まれるときれいに色が塗られているようには見えなくなってしまう。改善方法として空間メッシュを

取得し、掃除機をかけた位置座標により近い、メッシュに色付けを行う方法が考えられる。その場合、その地点が床である判定や掃除機の位置情報との位置合わせを行う必要がある。メッシュに色付けを行う際各メッシュの隙間を埋めるようにするとより、きれいに色付けをすることが可能となるはずだ。また本研究では実験の対象を大学生とした。しかしながら、掃除機掛けをした部分が塗られていくゲーム性は大学生には少し単純すぎたと考えられる。一方で小学 3 年生以下がシステムを使用した際、とても熱中して掃除機をかけている姿が見受けられた。本研究では主な評価指標を唾液アミラーゼによるストレス値としたが、意欲向上や、掃除を楽しいんでいるかという指標で、子供を対象とした実験であれば、MR や本システムの通常に比べた有効性がより大きく示せると示唆される。また、実験で対象とした教室が元々綺麗に保たれていたこともあり、実際に掃除機掛けによる達成感がより失われてしまっていたようを感じる。ゴミが少ない環境での掃除機掛けだったので実験自体を作業のような感覚で取り組み、その結果短時間で掃除機掛けを終わらす被験者が多かったのではないかと考えられるので、実験環境にゴミやゴミに類するものを用意することにより正確なデータが取れたと推測される。また、ゴミを用意することでゴミを何 % きれいにできたかなど違う評価指標で評価することも考えられる。

8. 結論

本論文は健康な生活を促進するために、家庭内ストレスの削減と家事の意欲向上を目的とし研究を行った。その手法としてゲーミフィケーションの要素である即時フィードバックや達成感の要素を加えることで目的を達成できるような行動変容を促した。ゲーミフィケーションの要素をあらゆる家事に応用できると考えたのが、色塗りにより掃除をした形跡をリアルタイムに確認できるという機能である。本研究ではその中で掃除機をピックアップし、掃除機をかけたところが色で塗られていくシステム (Vacuum Colorer) を提案した。システムは MR 技術を用いたもので hololens2 にて実装を行った。システムでは色選択を行ったのち、掃除機を掛けると掃除機に取り付けた QR コードを追跡し、QR コードから掃除機をかけた形跡を算出し、選択した色のタイルを配置していくというものである。空間認識に優れた MR デバイスで実装を行うことでより現実に近い視覚情報を与えた。システムの有効性を示すため 30 名の大学生を被験者とし実験を行った。実験では掃除意欲、掃除の楽しさ、面積、距離、ストレス値、時間について比較を行うため、定性的なアンケートと定量的な数値を用いてシステムの有効性を示した。意欲、楽しさ、ストレス要因については 5 段階のアンケートと記述式のアンケートにて評価を行った。面

積、距離は掃除機につけたマウスの座標データから値を算出し、ストレス値は唾液アミラーゼの変化率から値を求めた。以上の実験から、意欲や楽しさ、面積、ストレス値において何もつけていない普段の掃除機かけよりもシステムを使用した掃除機掛けに優位性が示された。これにより、MR や色塗り効果を交えたシステムは掃除機掛けに効果があり、これは他の家事にも応用可能であると考えられる。

謝辞 本研究成果は、国立研究開発法人情報通信研究機構の委託研究（05401）、JSPS 科研費 JP21K11853、JST RISTEX (JPMJRX21J1) の支援を受けたものである。

参考文献

- [1] リンナイ株式会社【熱と暮らし通信】(2018). <https://www.rinnai.co.jp/releases/2018/0515/>.
- [2] 卵津羅祥子, ウヅラサチコ職場, および家庭におけるストレス要因が自覚的健康度、心理学的健康度に及ぼす影響 (2002).
- [3] ゲンナイ製薬株式会社【プレスリリース・ニュースリリース】. <https://prtentimes.jp/main/html/rd/p/000000008.000038158.html>.
- [4] Seaborn, K. and Fels, D. I.: Gamification in theory and action: A survey, *International Journal of human-computer studies*, Vol. 74, pp. 14–31 (2015).
- [5] 小笠原遼子, 山木妙子, 塚田浩二, 渡邊恵太, 椎尾一郎インタラクティブな掃除機, エンタテインメントコンピューティング 2007 講演論文集 (2007).
- [6] 市村哲, 矢澤崇史, 戸丸慎也, 渡邊宏優ほか家事をゲーム化ケーション化する試み～掃除への適用～, マルチメディア, 分散協調とモバイルシンポジウム 2014 論文集, Vol. 2014, pp. 1285–1290 (2014).
- [7] 菅野恭平 川畠裕也畠田一貴 VACUUU(・▽・)UUUM, 日本バーチャルリアリティ学会大会論文集, Vol. 19.
- [8] 小坂崇之ほか妖怪クリーナーズ: 現実世界の掃除をゲームコンテンツにしたシリアルゲームの提案, エンタテインメントコンピューティングシンポジウム 2016 論文集, Vol. 2016, pp. 148–151 (2016).
- [9] Kaeriyama, Y., Yokokubo, A. and Lopez, G.: Dustoon: Support System for Improvement of Cleaning using Dust Sensor and Mixed Reality, *IEICE Technical Report; IEICE Tech. Rep.*, Vol. 120, No. 319, pp. 42–44 (2021).
- [10] 竹島由起, 山田晃子秋野暢子と一緒に楽しもう! 5 分でできるシンプル健康法 (8) ストレスが消える塗り絵, 日経ヘルス, Vol. 4, No. 13, pp. 52–54 (2001).
- [11] Microsoft: QR コード追跡の概要. <https://learn.microsoft.com/ja-jp/windows/mixed-reality/develop/advanced-concepts/qr-code-tracking-overview>.
- [12] 中野敦行, 山口昌樹唾液アミラーゼによるストレスの評価, バイオフィードバック研究, Vol. 38, No. 1, pp. 3–9 (2011).
- [13] Jenzano, J. W., Brown, C. and Mauriello, S. M.: Temporal variations of glandular kallikrein, protein and amylase in mixed human saliva, *Archives of Oral Biology*, Vol. 32, No. 10, pp. 757–759 (1987).
- [14] Kivelä, J., Parkkila, S., Metteri, J., Parkkila, A.-K., Toivanen, A. and Rajaniemi, H.: Salivary carbonic anhydrase VI concentration and its relation to basic characteristics of saliva in young men, *Acta physiologica scandinavica*, Vol. 161, No. 2, pp. 221–225 (1997).
- [15] Salvolini, E., Mazzanti, L., Martarelli, D., Di Giorgio, R., Fratto, G. and Curatola, G.: Changes in the composition of human unstimulated whole saliva with age, *Aging Clinical and Experimental Research*, Vol. 11, pp. 119–122 (1999).
- [16] roccat: kone pro air. <https://jp.roccat.com/products/kone-pro-air>.

水分補給支援システムと行動変容検証

元川錦 ^{†1}

青山学院大学

横窪安奈 ^{†2}

青山学院大学

ロペズギヨーム ^{†3}

青山学院大学

2023年2月17日

1. はじめに

水は体内において多方面に渡って重要な役割を担うため、人は一定の水分を補給しなければ生きていくことができない [1]。厚生労働省が推奨する1日の水分補給量は1.2Lである一方、人間は一度に大量の水を吸収することができず、1時間あたり250mlが限界である。水分補給の重要性は老若男女問わず広く知られているが、知らず知らずの内に水分不足になることが多く、水分補給の重要性は認知していくも、適切な水分摂取量を認知していないことが多い。2019年にサントリーが水分摂取に関する調査として、全国の20歳以上のオフィスワーカー1600人を対象に水分摂取量のアンケート調査を実施した。本調査によると、約9割のオフィスワーカーが自分自身が1日の中で摂取する飲料、すなわち水分量は500mlペットボトル2本分である1L以下であると回答した。これはオフィスワーカーの約9割が一日に必要な水分量を補給できていないことを示している。加えて、1日に必要だと思う水分量として、最も適当な水分摂取量を調査したところ、6割のオフィスワーカーが1日に摂取するべき水分量の不足を認識できていないことが明らかになった。総務省の発表 [2] によると、水分不足が大きく関与しているとされる熱中症による救急搬送人員は2020年6月から9月の全国において累計6万人超で、そのうち約半数が自宅にいたにもかかわらず発症しており、いつ、どこにいても水分不足になると考えられる。以上のことから、ユーザーの水分補給状態に加え、発汗などによる健康状態や、ユーザーがいる環境の温度・湿度などの環境情報を総合的に考慮したタイミングでのサポートが必要かつ効果的である。

2. 関連研究

水分補給や飲料容器に着目した研究は以下のようなものがある。Beddoeら [3] は Nudge Technology と呼ばれる、ユーザーに適切な選択や危険回避を人に促す仕掛けや手法 [4]

に注目しシステムを開発した。システムは適度なタイミングで水分補給できていないと、内容物が溢れてしまいユーザーに水分補給が不十分であることを想起させ、それ以上こぼさないように促す。本システムではアンケート結果から飲水量の向上が示唆された。長田ら [5] は、コップにセンサを埋め込んだインテリジェントコップを作成し、服薬の際に十分な量の飲水を促すシステムを開発した。その結果、飲水量の測定が可能になり、多くの被験者が指定された飲水量を飲み切った。これにより、患者への適切な服薬指導が可能であることが確認された。Lesselら [6] は重量から水分量を測定するコースター型デバイスを開発し、ユーザーの水分補給状態を反映したスマートフォンゲームを提案した。本システムでは、飲水量が向上し、水分補給の促進に効果があることが明らかになった。一方で、スマートフォンのゲーム機能を用いたフィードバックによる有意差はみられなかった。

以上のように、現在様々な手法を用いた水分補給支援システムが研究開発されており、ユーザーに水分補給不足を認識させ、促進させることが可能である。一方で、[3]のシステムではボトル内の水分があふれてしまうことによるリスクが大きく、このようなシステムを実際のオフィス環境で使用したり、持ち歩いたりすると、周囲の電子機器の故障につながる可能性があり、実用的では無い。[5] や [6] のような持ち運びに適さないシステムでは、日常生活を通しての支援が難しく、利用シーンが限定されてしまう。このように、これらの既存システムは、制約のない現実的な環境において、自然な形で水分補給を促すことを可能にするトラッキングシステムとしては多くの制限があると考える。そこで、本研究では、これらの制約や実用性の課題を踏まえて「日常生活どこでもいつでも使用できるシステム」を実現することでユーザーの水分補給を支援できると考える。加えて、上記の問題を解決するために、物理駆動型情報提示手法を採用する。物理駆動型情報提示は、ウェイトレスがコーヒーを運んでくるように自然でわかりやすく、有用であると考える。また、これまでに開発された水分補給新システムではユーザーの水分補給に基づいた支援はされているものの、水分補給支援で考慮されるべきユーザーの健康状態や環境情報は全く考慮されておらず、ユーザーの状態を踏まえた

Hydration Support System and Behavior Change Verification

^{†1} NISHIKI MOTOKAWA, nmotokawa@wil-aoyama.jp

^{†2} ANNA YOKOKUBO, yokokubo@it.aoyama.ac.jp

^{†3} GUILLAUME LOPEZ, guillaume@it.aoyama.ac.jp

支援がされていないのが現状である。そこで、本研究では、上記の課題を解決するために、ユーザの生体情報に基づくパーソナライズされた情報提示を用いることで、ユーザの生体情報とユーザのいる環境に応じてパーソナライズされた情報提示の実装を行う、パーソナライズされた情報提示は、運動などによる発汗で知らぬ間に消耗した水分を補給することができるため、有用だと考える。

3. HydReminder

3.1. システム概要

本研究では、水分補給不足を通知し、適切なタイミングで水分補給不足を解消するための水分補給支援システム HydReminder を開発し、その有用性を検証した。HydReminder は、ユーザの生体信号をもとにパーソナライズすることで、理想的なタイミングで水分補給を促すことができるウェアラブル・デスクトップ両対応型水分補給支援システムである。HydReminder は、水分補給の「Hydration」と思い出させるものの「Reminder」を組み合わせた名称である。コースター型デバイス・キャップ型デバイスの内側に搭載したセンサでユーザの水分補給状況を、同じくキャップ内側に搭載された環境センサでユーザのいる環境の温度と湿度を、そしてユーザの手首に装着したスマートウォッチでユーザの心拍数をそれぞれ測定する。ユーザの水分必要度に対し、水分補給が不足すると、ウェアラブル・デスクトップ両モードに対応したフィードバックを通して、ユーザへ水分不足を伝達する。HydReminder のシステム構成を図 1、デバイスを図 2、図 3、利用イメージを図 4 図 5 にそれぞれ示す。

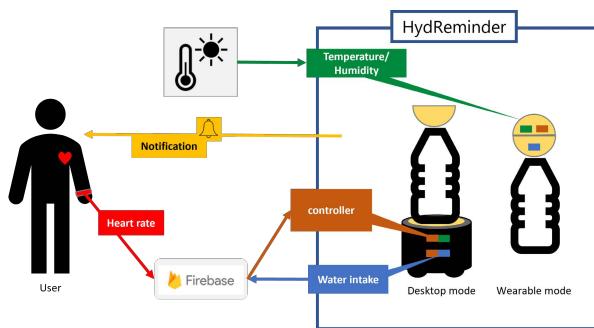


図 1 HydReminder のシステム構成

3.2. 水分補給必要度判定アルゴリズム

システム遷移図を図 6 に示す。制御部は、クラウドホスティングの NoSQL データベースである Firebase Realtime Database を介して、スマートウォッチからユーザの心拍数を取得する。取得した心拍数データは、ユーザの健康状態や



図 2 HydReminder デスクトップモード

運動状態として利用される。また、制御部と有線で接続された環境センサからユーザ環境の温度と湿度を取得し、熱中症の危険度や環境の極端な乾燥度合いを判断するために利用する。そして、これらの健康状態や運動状態、危険度を踏まえた水分補給ができているかを確認するために、デスクトップモードでは、ロードセルから水分量を、ウェアラブルモードでは、赤外線距離センサから容器内の水面までの距離を用いて、ユーザが十分に水分を補給しているかどうかを判断する。ユーザが十分な水分補給をしていない場合、キャップ型デバイスはアラート音を発し、スマートウォッチの画面には水分補給を促すメッセージが表示される。スマートウォッチ上の支援メッセージの様子を図 7 に示す。

3.3. HydReminder デバイス

HydReminder はコースター部とキャップ部の二種類のデバイスから構成される。コースター部分は車輪駆動型デバイス (toio core cube^{*1}) を固定するための部品 C. ロードセル、アナログ・デジタルコンバータ (HX711)、制御部 (M5 ATOM Lite) を固定するための部品 B. これらを格納する外壁の部品 A の 3 部品で構成されている (図 8)。コースター部分はいずれも 3D プリンタを用いて出力した。ロードセルは上下を 3D プリンタで出力した板で固定し、コースター型デバイスに置かれた飲料容器と中身の水分の重量を測定する。測定されたセンサデータはアナログ・デジタルコンバータ (HX711) を介して接続された制御部へ送信され重量へ変換される。制御部はロードセルと有線で、車

^{*1} toio core cube, Sony, <https://toio.io/>

Necessity level 1	Necessity level 2	Necessity level 3
Temperature ~25 °C Humidity 40 ~ 60%	Temperature 25 ~ 30°C Humidity ~ 40%	Temperature 30°C ~ Humidity ~ 30%
BPM Available : true <u>Let's drink!</u> The temperature and humidity in your location is appropriate. Stop Reading	BPM Available : true <u>It's about time.</u> It's going to be dry... Stop Reading	BPM Available : true <u>Drink and rest NOW.</u> Danger! Whenever possible, move to a cooler and Move to a place that is not dry. Stop Reading

図 7 スマートウォッチ上のフィードバックの様子



図 3 HydReminder ウェアラブルモード

輪駆動型デバイスと BlueTooth でそれぞれ接続しており、ロードセルから取得したセンサデータを重量に変換し、水分補給しているか否かを認識する。

ボトルキャップ部分は ToF センサと制御部を固定するための部品環境センサと制御部を固定するための部品ペットボトルキャップとして固定するための部品の 4 部品で構成されている(図 9)。コースター部分はいずれも 3D プリンタを用いて出力した。使用した 3D プリンタは XYZ Printing 社製の da vinci super であり、樹脂フィラメントを熱して出し造形することが可能である。パート C には車輪駆動

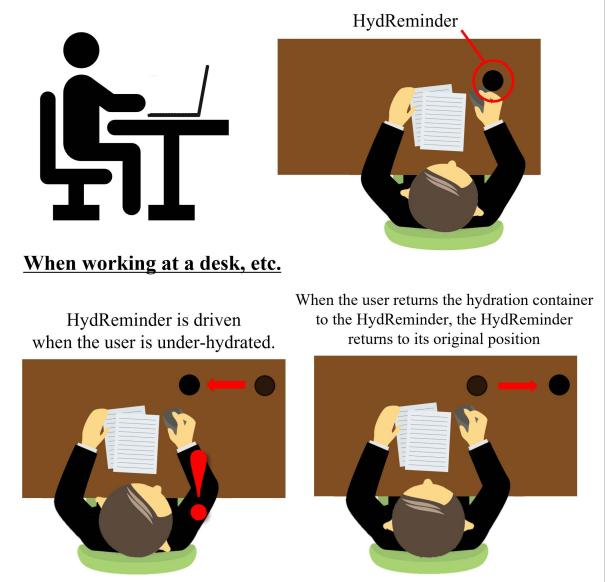


図 4 HydReminder の利用イメージ（デスクトップモード）

型デバイスを配置し、パート B にはロードセルとアナログ・デジタルコンバータ、制御部を設置した。ロードセルは上下を 3D プリンターで出力した板で固定し、コースター型デバイスに置かれた飲料容器と中身の水分の重量を測定する。測定されたセンサデータはアナログ・デジタルコンバータ (HX711) を介して接続された制御部へ送信され重量へ変換される。制御部はロードセルと有線で、車輪駆動型デバイスと BlueTooth でそれぞれ接続しており、ロードセルから



図 5 HydReminder の利用イメージ（ウェアラブルモード）

取得したセンサデータを重量に変換し、水分補給しているか否かを認識する。これらの部品を統合するために、「卓上でも邪魔にならない大きさ」「部品をすべて配置できる」を満たす最小サイズで外壁シェルを作成し、パーツを格納している。

3.4. HydReminder のモード

HydReminder はデスクトップモードとウェアラブルモードの 2 モードから構成される。デスクトップモードは主にデスクワーク中などで使用し、物理駆動型情報提示による支援を行う。ウェアラブルモードは移動時や運動時などで使用し、スマートウォッチ画面のメッセージとアラート音による支援を行う。デスクトップモード時のデバイスの様子を図 2 に示す。デスクトップモード時はコースター型デバイスにキャップ型デバイス球体上部をドッキングした状態で使用する。コースター内部に設置されたロードセルから圧力データを取得し、アナログ・デジタルコンバータ (HX711) を介して接続された制御部へ送信され重量へ変換される。制御部はロードセルが圧力を検知不能になった際、コースターデバイスからボトルが持ち上げられ飲水しているとする。また、ドッキングしている球体上部に設置した環境センサからユーザの周辺環境温湿度を計測し、クラウドホスティングの NoSQL データベースである Firebase Realtime Database を介して、スマートウォッチからユー

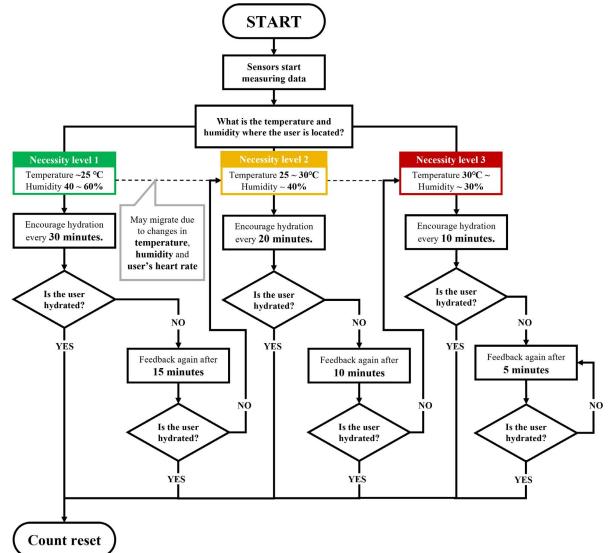


図 6 Power switch for the control unit attached to a wheel-driven device.

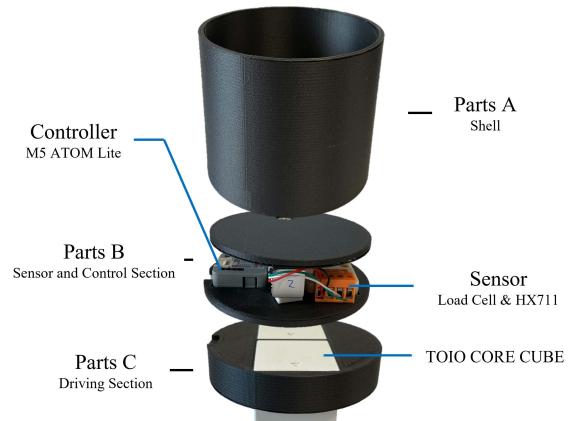


図 8 コースター部のデバイス構成

ザの心拍数を取得する。ユーザが十分な水分補給をしていない場合、bluetooth 接続している車輪駆動型デバイスを駆動しユーザへ水分不足を知らせる。その際、ペットボトルから水分補給をする際にペットボトルのふたを開け口元までもっていき、卓上へ置くまでの一連の動作にかかる時間を計測した結果、最短で 5 秒以上かかることから、不正防止を目的として 5 秒以上圧力が計測されなかった場合のみシステムはリセットされ、toio の駆動時には toio は停止する仕組みになっている。利用者が水分補給を終え、水分補給容器をコースターに戻すと、コースターは来た道を戻り、元の位置で待機する。

ウェアラブルモードのデバイスの様子を図 3 に示す。ウェアラブルモード時はコースター型デバイスを使用せず、

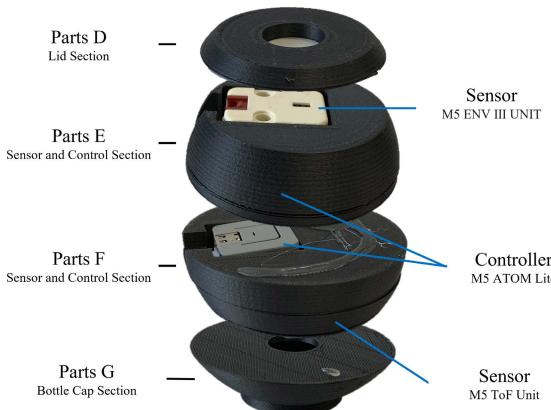


図 9 ボトルキャップ部分のデバイス構成

キャップは球体の状態で使用する。球体下部に配置された赤外線戦距離センサから容器内の水面まで距離を用いて、ユーザが十分に水分を補給しているかどうかを判断する。球体下部に配置された制御部は、デスクトップモード時同様に、Firebase Realtime Database を介して、スマートウォッチからユーザの心拍数を取得する。同様に球体下部に配置され、制御部と有線接続している赤外線戦距離センサから容器内の水面まで距離を計測し、ユーザが十分に水分を補給しているかどうかを判断する。また、球体上部に設置した環境センサからユーザの周辺環境温湿度を計測し、熱中症の危険度や環境の極端な乾燥度合いを判断するために利用する。ユーザが十分な水分補給をしていない場合、制御部からアラート音を発し、スマートウォッチ上に支援メッセージを表示する

4. ユーザビリティ検証実験

4.1. ユーザビリティ検証実験の概要

HydReminder のユーザビリティ検証実験では、10代から50代までの男女10名を被験者とした。被験者には2日間に分け HydReminder を使用した場合と使用しなかった場合で、1時間もしくは9:00～12:00・13:00～18:00の計8時間の日常生活中に普段通りの生活をしてもらった。実験は、被験者が普段生活している空間で行い、極端な乾燥や湿度、気温のある環境にいる場合もあったことから、周辺環境が水分補給に影響を及ぼす可能性があると考える。実験終了後にアンケートに回答してもらい、水分摂取量を実験担当者がデジタルスケールにて計測した。また、提案システムの印象を評価するために、HydReminder のユーザビリティの受け止められ方にについて測定するためにSUSを用いたアンケートおよび自由記述のアンケートを実施した。

4.2. HydReminder の有無による水分摂取量・回数の比較

HydReminder を1時間使用した場合と使用していない場合において水分摂取量および水分摂取回数の比較を行った。水分摂取量を比較した結果を図10、水分摂取回数を比較した結果を図11に記す。結果を見ると水分摂取量では10人中9人の水分摂取量が増加した。水分摂取回数では10人中7人の水分摂取回数が増加し回数が減った被験者はいなかった。

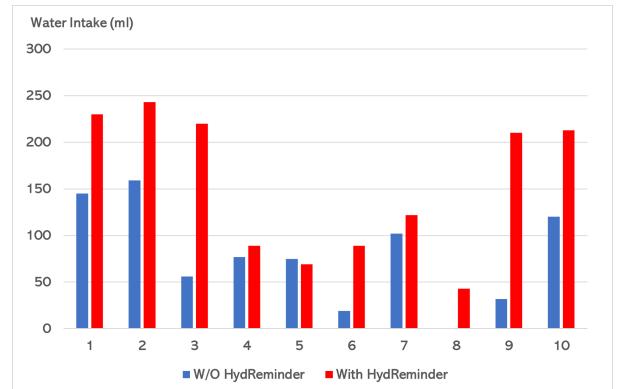


図 10 Comparison of the amount of water consumed by each subject (ml)

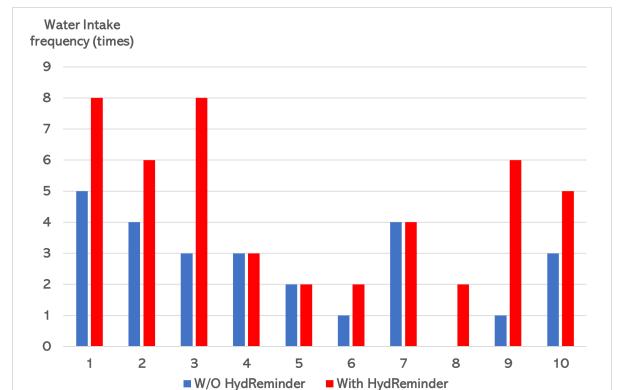


図 11 Comparison of water intake frequency per subject(times)

また、HydReminder を9:00～12:00・13:00～18:00の計8時間使用した場合と使用していない場合において水分摂取量および水分摂取回数の比較を行った。水分摂取量を比較した結果を図12、水分摂取回数を比較した結果を図13に記す。結果を見ると水分摂取量では10人中10人の水分摂取量が増加した。水分摂取回数では10人中9人の水分摂取回数が増加し回数が減った被験者はいなかった。

表 1 SUS score per subject used for 1 hour(points)

Subject	1	2	3	4	5	6	7	8	9	10	Ave
Score	87.5	80	90	77.5	75	97.5	75	92.5	87.5	90	85.25

表 2 SUS score per subject used for 8 hours(points)

Subject	1	2	3	4	5	6	7	8	9	10	Ave
Score	82.5	82.5	82.5	80	77.5	85	80	82.5	95	85	83.25

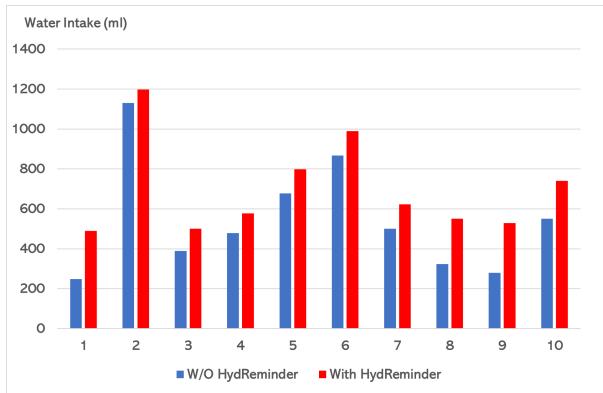


図 12 Comparison of the amount of water consumed by each subject (ml)

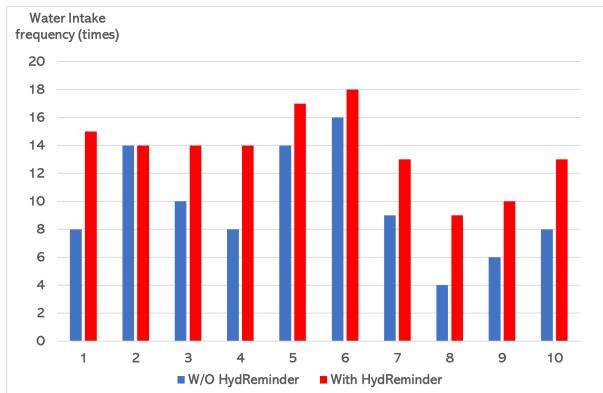


図 13 Comparison of water intake frequency per subject(times)

4.3. SUS による評価

HydReminder のユーザビリティ検証実験では SUS を用いたアンケートを実施した。それぞれの項目から 100 点を基準としたスコアを計算し、ユーザビリティを数値化した。その結果を表 1、表 2 にそれぞれ記す。結果を見ると、HydReminder を 1 時間使用した 10 人全員で SUS の平均スコアである 68 点を超えており、10 人の平均スコアは 85.25 点であった。また、HydReminder を 8 時間使用した 10 人全員で SUS の平均スコアである 68 点を超えており、

10 人の平均スコアは 83.25 点であった。以上の結果から、

HydReminder は非常に優れたユーザビリティであることが示された。

4.4. 考察

4.2 節で述べたように、HydReminder を使用した場合と使用していない場合において、HydReminder の有無による実験結果を比較したところ、HydReminder を 1 時間使用した被験者 10 名中 9 人の水分摂取量が増加し、10 人中 7 人の水分摂取回数が増加し回数が減った被験者はいなかった。また、HydReminder を 8 時間使用した被験者 10 名中 10 人の水分摂取量が増加し、10 人中 9 人の水分摂取回数が増加し回数が減った被験者はいなかった。この結果から HydReminder を比較的長時間使用することで、システム使用の効果がより現れる可能性を考えられる。加えて、システム未使用時に 8 時間で 300ml に満たない量しか水分補給をしていない被験者 2 人において、システムを使用することで 2 倍近く水分摂取量を増加させることができた。このことから、普段の水分摂取量が少ないユーザが HydReminder を使用することで、より効果が表れる可能性がある。また、表 1、表 2 からわかるように、HydReminder を 1 時間使用した被験者の SUS 平均スコアが 85.25 点、HydReminder を 8 時間使用した被験者の SUS 平均スコアは 83.25 点と SUS スコアの平均である 68 点を上回っていた。そのため、HydReminder は非常に優れたユーザビリティであることが示された。アンケートの自由記述欄からは「こんなに飲む必要があるとは知らなかつた」「スタイルッシュなのでオフィスでも使いやすい」「動くと飲まなきゃって気持ちになる」との意見があった。一方で、「ちょっと大きい」「わざわざドッキングさせなくてもいいかも」との意見があった。そのため、デバイスのデザインや仕様をより改善していく必要があると考える。以上の結果から HydReminder を使用することにより、ユーザへ適切なタイミングで水分補給の必要性を知らすことができ、水分補給を促進することができるようになると考える。

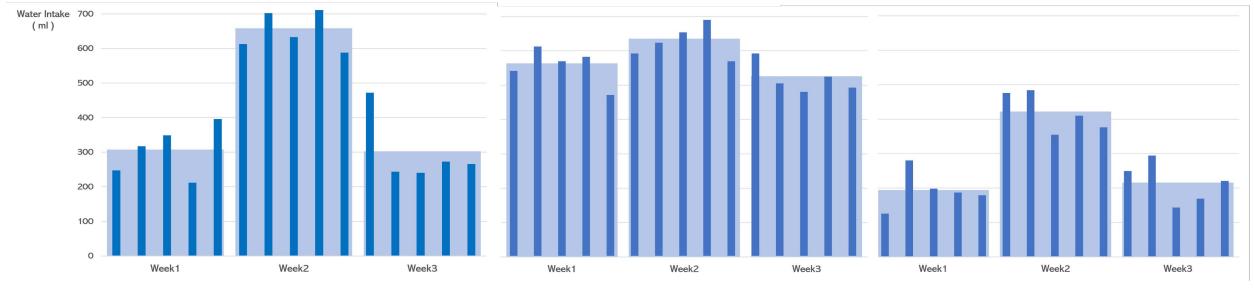


図 14 被験者 A の水分摂取量の推移

5. 行動変容検証実験

本研究では、HydReminder を習慣的に使用した際にユーザーの行動や習慣に変化が起こるかを検証するために水分摂取量の計測およびアンケート調査を行った。

5.1. 実験方法及び評価方法

HydReminder を習慣的に使用した際にユーザーの行動や習慣に変化が起こるかを検証するために水分摂取量の計測およびアンケート調査を行った。行動変容検証実験では、10代から30代までの男女3名を被験者とし、被験者には平日5日間を3週間計15日間にわたりHydReminderを使用した場合と使用しなかった場合で、9:00～12:00・13:00～18:00の計8時間日常生活に普段通りの生活をしてもらった。なお、実験1週目と3週目はHydReminderデバイスのみ使用し、特にフィードバックを行わず、実験2週目のみHydReminderのデバイスとシステムを使用してもらった。実験は、被験者が普段生活している空間で行い、極端な乾燥や湿度、気温のある環境にいる場合もあった。そのため周辺環境が水分補給に影響を及ぼす可能性があると考える。

5.2. 行動変容検証実験の結果

水分摂取量および水分摂取回数の比較を行った。結果を図14にそれぞれ示す。全被験者においてHydReminderを使用しなかった1週目に比べて使用した2週目の平均水分摂取量・摂取回数は増加するという結果になった。また、再びHydReminderを使用しなかった3週目はHydReminderを使用した2週目に比べて平均水分摂取量・摂取回数は減少し、同じくHydReminderを使用しなかった1週目と比べて大きな変化は見られなかった。一方で、HydReminderを使用しなかった1週目に比べてHydReminder使用後の3週目に水分摂取量が増加した被験者もいた。加えて、事後アンケートにおいて、「システムを使うことで水分補給への意識は変わりましたか?」という問い合わせに対し、被験者全員が「変わった」と回答した。アンケートの自由記述欄からは「特定のデバイスを使わなくてはならないという制限が少し煩わしい」「システムを使うことで『そろそろ飲んだほうがいいかも』と思うようになった」「頑張って飲んだ報酬とかがあれば続けられそう」などの意見があった。以上の結果から、ユーザーの普段の水分補給習慣によってはその後の水分補給習慣に変化を与える可能性が示された。

煩わしい」「システムを使うことで『そろそろ飲んだほうがいいかも』と思うようになった」「頑張って飲んだ報酬とかがあれば続けられそう」などの意見があった。以上の結果から、ユーザーの普段の水分補給習慣によってはその後の水分補給習慣に変化を与える可能性が示された。

5.3. 考察

5.2節で述べたように、HydReminderを一定期間使用した結果、水分補給行動に大きな変化は見られなかったものの、事後アンケートにおいて、「システムを使うことで水分補給への意識は変わりましたか?」という問い合わせに対し、被験者全員が「変わった」と回答した。ユーザーの普段の水分補給習慣によってはその後の水分補給習慣に変化を与える可能性が示された。アンケートの自由記述欄からは「特定のデバイスを使わなくてはならないという制限が少し煩わしい」「システムを使うことで『そろそろ飲んだほうがいいかも』と思うようになった」「頑張って飲んだ報酬とかがあれば続けられそう」などの意見があった。以上の結果から、大きな変化が見られなかった要因として、長期間使用することによってシステム使用の煩わしさを感じたり、より多く水分補給することのメリットを感じにくい点などが考えられる。このことから、システムをより使いやすくすることや使用期間に応じてユーザーへ報酬を与えるといった改良を行うことでHydReminderの使用による更なる行動変容を促せる可能性があると考える。加えて、水分摂取量増加によるメリットに合わせた目標設定などでユーザーの水分補給へのモチベーションを向上・保持することも重要であると考える。

また、今回の実験では実験環境やユーザーの行動にばらつきがあり、制限なく実施されたことから、周辺環境状態など外的要因による変化の可能性も考えられる。そのため、外的要因による影響を可能な限り排除した実験環境下での実験を行い、さらに多くのデータを取得する必要があると考える。

実験中のユーザーのいた環境の温湿度の割合を表3に示す。

結果を見ると、気温 25°C 以下、湿度 40% 以上のデータが多く、大半を占めている。この原因として、本実験は冬季に集中して実施したことや、ユーザがオフィス等の比較的空調整備された空間にいたためと考えられる。そのため、今後の実験を通して高温多湿時に HydReminder を使用してもらい有用性の検証をする必要があると考える。

表3 Percentage of temperature and humidity during the experiment (%).

Temp(°C)/Humi(%)	-30%	30%-40%	40%-
-25°C	22.3	17	49.2
25-30°C	2.1	1.9	6.6
30°C-	0.8	0.1	0

6.まとめ

本研究では、水分補給不足を通知し、適切なタイミングで水分補給不足を解消するための水分補給支援システム HydReminder を開発し、その有用性を検証した。HydReminder は、ユーザの生体信号をもとに支援間隔をパーソナライズすることで、ユーザに適したタイミングで水分補給を促すことができるウェアラブル・デスクトップ両対応型水分補給支援システムである。HydReminder を使用した際に水分補給が促進されているかの確認及び HydReminder の印象を評価するために、水分摂取量の計測および SUS を用いた印象評価を行った。結果を見ると、HydReminder を 1 日 8 時間使用した場合、10 人中 10 人の水分摂取量が増加し、10 人中 9 人の水分摂取回数が増加した。また、HydReminder を 8 時間使用した被験者 10 人 SUS の平均スコアは 83.25 点であった。以上の結果から、HydReminder は非常に優れたユーザビリティを持ち、水分補給支援システムとして有用であることが示された。また、HydReminder を習慣的に使用した際にユーザの行動や習慣に変化が起こるかを検証するために水分摂取量の計測およびアンケート調査を行った。結果を見ると、システムを一定期間使用することによる水分補給習慣に大きな変化は見られなかった。一方で、HydReminder 使用前と比較し、HydReminder 使用後に水分摂取量が増加した被験者もいた。加えて、事後アンケートにおいて「システム使用による水分補給への意識変化はあるか」という問い合わせに対し、被験者全員が「変わった」と回答した。以上の結果から、HydReminder の使用によりその後の水分補給習慣に変化を与える可能性が示された。

今後の展望として、現状では人間情報としてユーザの心拍数のみを活用していたが、皮膚電位や IMU といった複数の人間情報を組み合わせて用いることで、よりユーザの

状態を考慮したシステムへと改良できると考える。また、HydReminder を用いた行動変容検証実験のアンケートの自由記述欄からは「頑張って飲んだ報酬とかがあれば続けられそう」という意見があった。現状では、システムを終了してしまうとこれまでの水分補給結果が失われてしまうため、水分補給履歴をデータベースに保存し、月別、週別などで可視化することで長期的なモチベーションの向上につなげ、ユーザの水分補給習慣をより良くしていきたいと考える。

謝辞 本研究は、セコム科学技術振興財団の研究助成を受けたものである。

参考文献

- [1] サントリーサントリースタジオ「水育人間」と水の関わり、サントリー（オンライン）, <https://mizuiku.suntory.jp/kids/study/n004.html> 2021-7-12.
- [2] 総務省令和2年（6月から9月）の熱中症による救急搬送状況、消防庁（オンライン）, https://www.fdma.go.jp/disaster/heatstroke/items/heatstroke_geppou_2020.pdf 2021-7-12.
- [3] Beddoe, A., Burgess, R., Carp, L., Foster, J., Fox, A., Moran, L., Bennett, P. and Bennett, D.: Disruptabottle: Encouraging Hydration with an Overflowing Bottle, *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI EA '20, New York, NY, USA, Association for Computing Machinery, p. 1–7 (online), 10.1145/3334480.3382959 (2020).
- [4] 井原敏宏「IoT と AI 普及の鍵は Nudge」NTT ファシリティーズが実証、日経 XTECH（オンライン）, <https://xtech.nikkei.com/it/atcl/news/17/011802980/> 2021-7-12.
- [5] 長田拓也, 鈴木拓央, 中内 靖インテリジェント・カップによる服薬指導支援システムの提案、日本機械学会論文集, Vol. 83, No. 853, pp. 1–7 (2017).
- [6] Lessel, P., Altmeyer, M., Kerber, F., Barz, M., Leidinger, C. and Krüger, A.: WaterCoaster: A Device to Encourage People in a Playful Fashion to Reach Their Daily Water Intake Level, *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, CHI EA '16, New York, NY, USA, Association for Computing Machinery, p. 1813–1820 (online), 10.1145/2851581.2892498 (2016).

高齢者の介護予防に向けた行動変容を促す介入方法の推定

早川昭二^{†1} 渡辺一郎^{†1} 上村拓也^{†1} 前田一穂^{†1} 烏谷彰^{†1}
富士通株式会社^{†1}

1. 背景

厚生労働省が提唱する健康寿命延伸プランでは、2040 年までに健康寿命を男女とも 3 年以上延伸する目標と、その実現のために高齢者が要介護状態になることをできる限り防ぐ「介護予防」の重要性が示されている[1]。これに対し、先進的な自治体では、介護予防の施策として、行政と高齢者を繋ぐ「仲介者」となる行政保健師が直接自宅へ訪問して高齢者の状況を把握し、適切な介護予防行動を促す「アウトリーチ支援」が実施され始めている[2]。超高齢化社会の進展により、支援が必要な高齢者数は今後も増え続けることが予想されている一方で、「アウトリーチ支援」の担い手となる行政保健師数の伸びは少ない[3]。仲介者が増えない状況において、介護予防行動をする高齢者を増やすためには、少人数の仲介者により効果的に介入していく必要があり、仲介者の介入スキルの向上や業務効率化が必須となる。「アウトリーチ支援」で介護予防行動を促す訪問勧奨業務を支援する技術の実現により、地域住民の健康寿命延伸に寄与することを目指している。

2. 本研究の課題と解決アプローチ

従来の訪問勧奨業務では、保健師が訪問現場で得た介入の成功ノウハウは、保健師の職場での事例共有会などにおいて「成功事例」として共有されていたが、データとしてチーム内に共有する仕組みがなかった。そこで、このような成功ノウハウを共有し、他の保健師が効果的にそのノウハウを活用可能にするシステムを提供することで、訪問勧奨業務を支援することを考えた。

2.1 行動変容を促すための介入プロセスに関する分析

介入の成功ノウハウを共有する仕組みの実現のため、訪問現場において、仲介者である行政保健師や社会福祉士がどのように介入手段を選択し行動変容を促しているのかをヒアリング調査した。その結果、訪問先の部屋に置かれている物や生活状況の聞き取りといった情報から、その人の特性や傾向、行動の障壁となっている問題(阻害要因)などを探り、適切な介入方法を選択・実施することで、高齢者の意識や状況を変えて行動変容を促している実態を把握した。ここで得た実態を踏まえ、図 1 に示すような 3 段階からなる介入選択プロセスの仮説を立てた。

このプロセスは①訪問時に着眼した高齢者個々人の特性

を表す要因群を見極める過程（特に行動変容の成否への影響の大きい要因）、②要因群に基づき有効と判断した介入を選択する過程、③介入による意識変化や行動変容の成否結果、からなる。ここで、②から③の間において、介入により直接的に行動を変容させるのではなく、介入によって高齢者の阻害要因を取り除き「これなら自分にもできる」と自信を持たせるなど、外的・内的要因を変化させて行動を変容させることを考える。この介入と行動変容との間を工学的に表現することが、介入の成功ノウハウをチーム内に共有・活用する仕組む実現する上での技術課題となる。

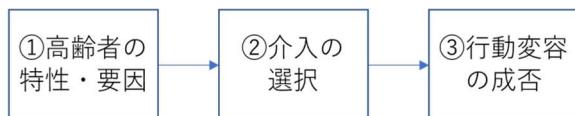


図 1：仲介者の介入方法選択のプロセス

2.2 介入により行動変容するメカニズムについての仮説

介入と行動変容の間を繋ぐメカニズムを工学的に表現するため、行動要因という説明変数を導入した。行動要因とは、高齢者それぞれの自己効力感やモチベーション、知的・身体的能力、社会的支援の状況などのように、目的の行動を起こす際に、十分であれば行動を起こしやすくなり、不足すれば行動を阻害するような要因のことである。これらの行動要因は、図 1 中の①、②の過程で、仲介者が実際に介入方法を選択する際にも考慮している重要な要因であり、BJ Fogg の MAP モデル[4]や、社会的認知理論[5]、COM-B[6]などの行動変容理論でも用いられているものである。これらの解釈可能な要因群について、介入による改善量を推定・提示し、行動変容の成否を説明可能とすることで、支援システムを利用する仲介者の納得性も高められると考えた。

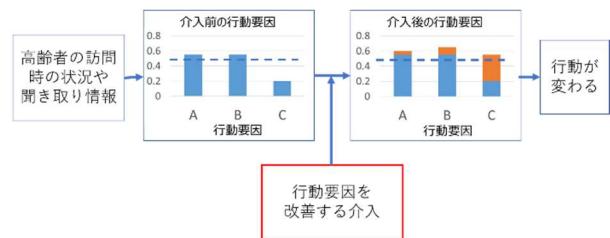


図 2：介入による行動変容メカニズムの仮説

導入した行動要因が介入により改善し、行動変容するメカニズムを工学的に表現した仮説を図 2 に示す。図では、訪問時の高齢者の状況や聞き取りといった情報から、介入前の行動要因 A,B,C を推定したとする。ここで、要因 C(例

Estimation of Intervention Methods to Promote Behavior Changes for Care Prevention in the Elderly

†1 SHOJI HAYAKAWA, Fujitsu Ltd.

†1 ICHIRO WATANABE, Fujitsu Ltd.

†1 TAKUYA KAMIMURA, Fujitsu Ltd.

†1 KAZUHO MAEDA, Fujitsu Ltd.

†1 AKIRA KARASUDANI, Fujitsu Ltd.

えば自己効力感とする)が行動に移すのに必要なしきい値(ここでは便宜的に 0.5 であるとする)を下回り、これが阻害要因となって行動できていなかったことを想定している。

それに対して、「行動要因を改善する介入」を行うことで、要因 C の値が改善してしきい値を超える(自己効力感が改善される)ことで、行動が変わるメカニズムを考えた。

すなわち、介入前の行動要因に、介入によって決まる改善量を加算した結果、行動要因がしきい値を超えて行動が変わるメカニズムである。

我々の研究[7]では本メカニズムを検証する予備実験として、訪問勧奨時に取得したアンケート等の業務データから行動変容に与える影響が大きいと考えられる行動要因を抽出・定量化し、算出した行動要因から、新規の高齢者であっても行動状態を予測できる可能性を示した。本稿では、訪問勧奨時に、対象高齢者に対して行った介入方法と、保健師が判断した介入前後の行動変容ステージ情報を用いて、介入により行動要因の値が改善し、その結果として行動状態が変容するメカニズムを統計的にモデル化する手法を提案し、実データを用いてその検証を行ったので報告する。

3. 介入による行動変容のモデル化

本章では、前章で示した行動変容メカニズムのモデル化方法について述べる。

3.1 行動状態の識別モデルと行動要因の改善量推定モデル

2.2 節では二つの仮定を設定した。一つ目は、「対象者の行動要因によって行動状態が決まる」であり、二つ目は、「介入によって行動要因が改善し、その結果として行動変容する(行動状態が介入前と変わる)」である。それぞれに対応した、以下に示す二つのモデルを考えた:

モデル 1: 行動要因から行動状態を識別する機能

モデル 2: 介入による行動要因改善量を推定する機能

モデル 1 は行動要因を入力し、その行動状態を出力する識別モデルである。このモデルは過去の発表[7]において報告済である。モデル 2 は、高齢者の行動要因と介入方法の情報を入力し、各行動要因の改善量を推定する回帰モデルである。

3.2 モデル全体の構造

図 3 にモデル 1 とモデル 2 を連結した、介入による行動変容メカニズムのモデル全体を示す。図 2 で示した介入前の行動要因の値に、介入による改善量が加算されるメカニズムを図 3 中の加算器で実現している。すなわち、介入が行われると、図中のスイッチが入り、介入前の行動要因に、モデル 2 によって推定された改善量が加算器で加わり、その結果、ノードが行動要因に対応した入力層に介入後の行動要因が入る仕組みである。

以降、モデル 1 とモデル 2 それぞれの学習方法について

述べる。まずモデル 1 のみを単独で学習し、次にモデル 1 のパラメータを固定した状態で、モデル 2 の学習を行う。なお、以下の図や説明では、行動状態の識別について、行動変容ステージ[8]の「無関心期」なのか、「関心期以上」なのかを識別することを例題としている。

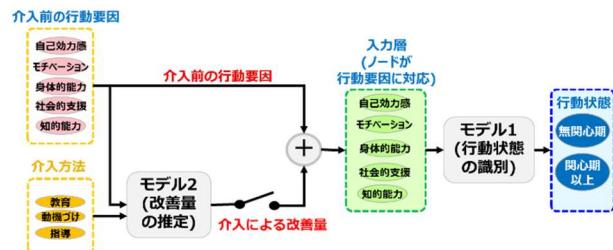


図 3 : 行動要因の改善量を推定するモデル

3.3 行動状態の識別モデル(モデル 1)の学習

モデル 1 は、高齢者それぞれの介入前の行動要因を入力ベクトルとし行動状態の識別結果を出力する、いわば静的な関係を表す関数である。そこで行動要因を説明変数として、無関心期の高齢者の集団と、すでに関心期以上である高齢者の集団を識別するモデルを学習する。図 4 に多クラス分類ロジスティック回帰であるモデル 1 を示す。モデル 1 は重み W_1 、バイアス b_1 、活性化関数としてソフトマックス関数および”無関心期”と”関心期以上”的二つのノードを持つ出力層で構成される。パラメータ W_1, b_1 は誤差逆伝播法により学習する。

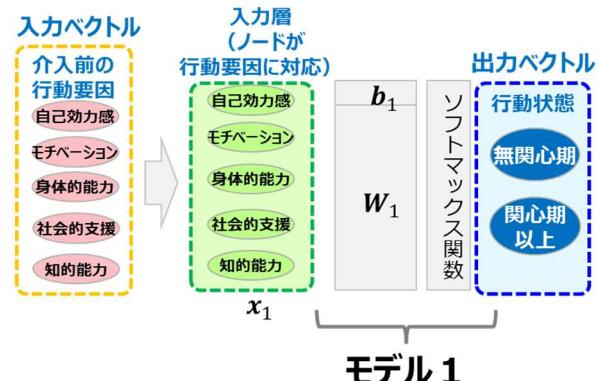


図 4 : モデル 1 の学習

3.4 行動要因の改善量推定モデル(モデル 2)の学習

モデル 2 は、学習済のモデル 1 を活用し、図 3 のスイッチをオンにして、ニューラルネットワークとして学習を行う。図 5 に学習時の構成を示す。モデル 1 のパラメータを固定した状態で、入力ベクトルを、介入前の行動要因の値と実施された介入方法(1種類の場合は one-hot ベクトル、2種類の場合は”two-hot”ベクトル)とし、出力ベクトルは実際の介入後の行動状態を教師データとして学習を行う。

次にモデル 2 の内部構造を説明する。図 5 の①部分は介入前の行動要因と介入方法を一緒に入力して、交互作用の効果を算出する部分である。次に②部分は①の算出結果に

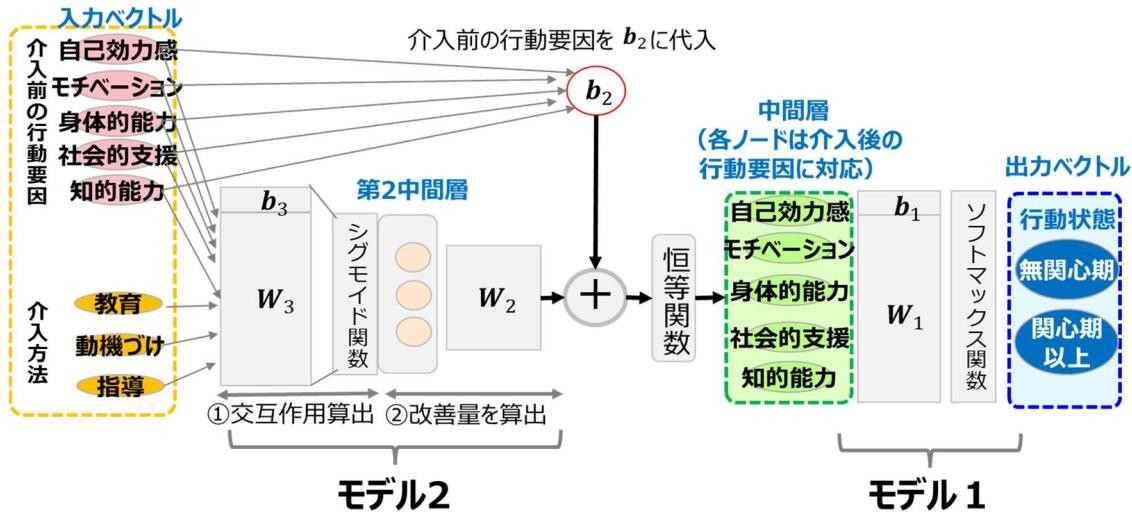


図 5：モデル 2 学習時のニューラルネットワークの構成

基づいて、各行動要因に対する介入による改善量を求める部分である。学習時において、3.3 節で学習済のモデル 1 は誤差をモデル 2 に伝播するために用い、モデル 2 のパラメータ W_2, W_3, b_3 を誤差逆伝播法により更新した。

モデル 1,2 が学習できれば、新規の高齢者に対して、行動変容する確率の高い介入方法を推定できる。次節では、提案したモデル 2 を実データで検証する。

4. 実験

本章では、実際の保健勧奨の現場において高齢者から聞き取ったアンケートデータを用いて、前章で提案したモデルにより効果的な介入方法を推定できるか検証した結果について述べる。

4.1 実験データ

実験データには、ある自治体で行われた訪問勧奨業務の対象となった高齢者 70 人分のデータを用いた。データには、高齢者への介入方法（複数種の介入も含む）と、仲介者が判断した介入前後の行動変容ステージの情報を含んでいる。その内訳を表 1 に示す。また実験に用いた行動要因の値は、研究発表[3]に記載した、健康状態や日常生活に関するアンケートデータを SEM で分析し因子得点化する方法で算出した。

表 1：実験に用いた高齢者の内訳

モデル 1 の学習に用いた高齢者数	<ul style="list-style-type: none"> 介入前が無関心期: 37 名 介入前が関心期以上: 33 名
モデル 2 の学習・評価に用いた高齢者数	<ul style="list-style-type: none"> 介入前が無関心期で介入後も無関心期: 29 名 介入前が無関心期で介入後に関心期以上へ移動: 8 名

4.2 学習条件

モデルの学習・評価は少数データであったため Leave-one-out CV(LOOCV)により実施した。モデル 1 は無関心期と関

心期以上を識別する必要があるため、表 1 に示す計 70 名分のデータを用いて学習した。またモデル 2 は、介入前の行動状態が無関心期の高齢者が介入によって関心期以上に遷移するかを学習するため、表 1 に示す加入前が無関心期の高齢者計 37 名のデータを用いて LOOCV 評価した。各モデルの学習条件は以下の通りである：

モデル 1 のエポック数:500

モデル 1 の学習率:0.01

モデル 2 のエポック数:1000

モデル 2 の学習率:0.1

4.3 介入後の行動状態の推定結果

モデル 2 による介入後の行動状態の推定精度、すなわち、介入前の行動要因と実施された介入方法を入力することで、正解データ通りの行動状態を出力できるかを、評価した。結果を表 2 に示す。正解率は 68%(25/37)、再現率は 75%(6/8)、適合率は 38%(6/16) という結果となった。

表 2：介入後の行動状態の推定結果

		推定結果	
		無関心期	関心期以上
正解	無関心期のまま	19人	10人
	関心期以上に遷移	2人	6人

4.4 考察

前節で学習したモデルが、図 2 で示したような介入により行動要因が改善されるメカニズムを実現できているか、図 5 の中間層に入力される介入前の行動要因と改善量の値を取り出して確認した。図 6 に介入により関心期以上に遷移した高齢者の例と、介入しても無関心期のままの高齢者の例をそれぞれ 2 名について、介入前の行動要因と加えた改善量を 5 種類の行動要因ごとに示す。介入により関心期以上へ遷移した高齢者 2 例については、ほとんどの行動要因において改善量が加算されていた（図 6 上）。また、例 1 で

は「動機づけ」の介入によりモチベーションが改善しているのに対し、例 2 では「訓練」の介入により身体的能力が改善しており、介入によって改善する行動要因が異なる傾向が見られた。一方、介入しても無関心期の高齢者 2 名に対しては、改善量が小さいか、マイナス方向に加算されていた（図 6 下）。実際の介入後の行動要因と比較を行う必要はあるものの、図 2 で提案した、介入による行動要因の改善量を介入前の行動要因に加算するモデルにより、行動変容のメカニズムを表現できる見込みを得た。また、介入前の行動要因と介入方法を入力することで、意図通りに改善量の推定ができるということから、事前に対象者の行動要因が分かれば、その対象者を行動変容されるうえで有効な介入方法を予測できる可能性を示すことができた。

5. まとめ

本稿では、介入により行動変容するメカニズムを、説明変数として行動要因を導入したモデルで表現する方式を提案し、実際の訪問勧奨業務において収集されたデータを用いて検証した。実験の結果、介入により無関心期から関心期以上にステージが遷移している高齢者は、説明変数である行動要因が改善されており、行動変容のメカニズムの確からしさを裏付けることができた。また、新規の高齢者に対して、ステージ遷移に有効な介入方法を確率的に予測できることを示した。これにより、他の保健師の介入方法のノウハウをモデルで表現でき、新規対象者に対しても他の保健師が成功率の高い介入方法を知ることができるようになり、ノウハウの共有に有用であると考える。今後は、高齢者数を増やして実験を行い、精度改善を図る。また、推定した行動要因と似た高齢者の過去事例の検索可能なアプ

リケーションの検討など[9]、訪問勧奨業務に従事する保健師を支援可能なシステム検討を進める。

参考文献

- [1] 厚生労働省 HP, “現在の健康づくりの取り組みと今後の施策について”, 厚生労働省健康局健康課, 2021 年 3 月 19 日
- [2] 厚生労働省 HP, “高齢者の保健事業と介護予防の一体的な実施について”, 2018 年 9 月 6 日
- [3] 厚生労働省 HP, “令和 2 年衛生行政報告例(就業医療関係者)の概況”, 2022 年 1 月 27 日
- [4] BJ Fogg 著、須川訳, “習慣超大全”, ダイヤモンド社, p. 50, 161 (2021).
- [5] 中澤ら, “社会的学習理論から社会的認知理論へ-Bandura 理論の新展開をめぐる最近の動向-”, 心理学評論, vol. 31, no. 2, pp. 229-251 (1988).
- [6] Michie, S. et.al., “The behaviour change wheel: A new method for characterizing and designing behaviour change interventions,” *Implementation Science*, vol. 6, no. 42, (2011).
- [7] 早川ら, “高齢者の介護予防行動を決定する要因に関する研究”, マルチメディア, 分散, 協調とモバイルシンポジウム 2022 論文集, pp. 590-594 (2022).
- [8] Prochaska JO, et al., “The transtheoretical model of health behavior change,” *American Journal of health promotion AJHP*, vol. 12, no. 1, pp. 38-48 (1997).
- [9] 田口ら, “業務経験の分析に基づく適切な参照面談事例の提示による保健師の育成支援”, 情報処理学会第 85 回全国大会講演論文集 (in printing) (2023).

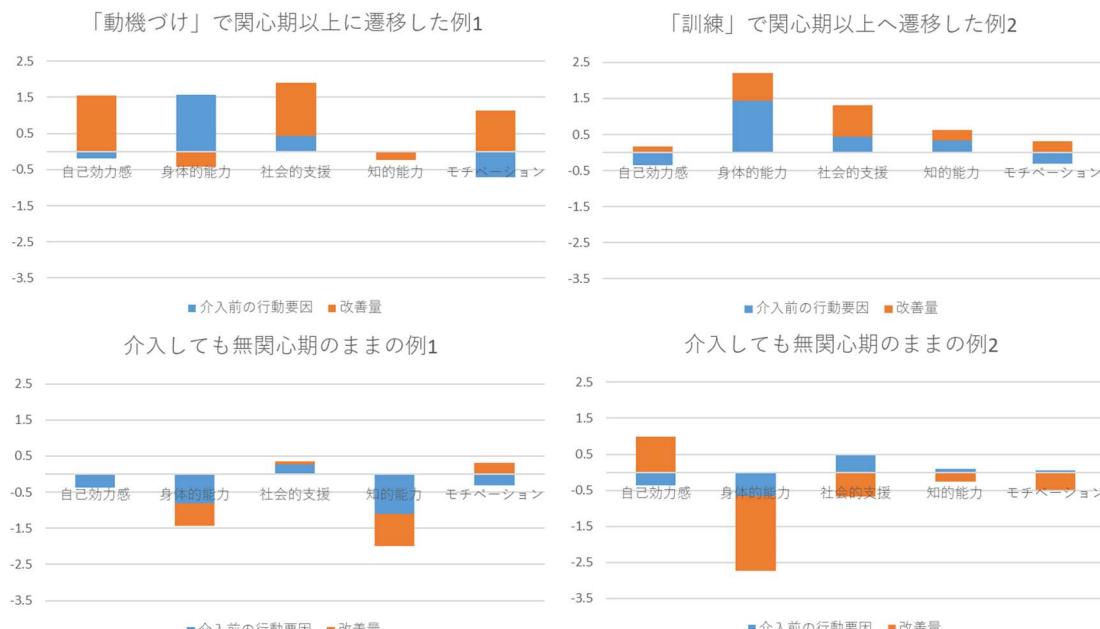


図 6: 介入前の行動要因と改善量の例
(上: 介入により関心期以上へ遷移した高齢者 2 例、下: 介入しても無関心期のままの高齢者 2 例)

目標宣言共有型プラットフォームを用いた ソーシャルナッジの量的質的効果の評価

濱谷尚志^{†1}
NTT ドコモ^{†1}

落合桂一^{†2}
NTT ドコモ^{†2}

山田渉^{†3}
NTT ドコモ^{†3}

檜山聰^{†4}
NTT ドコモ^{†4}

白井拓也^{†5}
NTT ドコモ^{†5}

荒川豊^{†6}
九州大学^{†6}

概要

健康寿命延伸の実現のため、健康的な行動を習慣化することが重要である。習慣化を支援するためスマートフォンのアプリケーションを用いた取り組みが数多くなされている。本研究では、ソーシャルネットワーキングサービスでの繋がりを通じて、仲間との関わりを通じた習慣化支援技術について検討を行う。具体的には、健康的な行動の目標を立てて、仲間から応援を受けることにより目標達成が支援されること、さらに、応援を受ける仲間との親密さに応じて、同じ応援1回でも目標達成後押しに対する効果が異なるという仮説を検証する。実際に大学生504人の約3ヶ月半にわたる歩数データ、歩数目標宣言データを取得し評価を行った結果、目標宣言を行うこと自体に歩数を増加させる効果があること、目標宣言に対して応援を受けることで歩数が増加すること、および応援を受ける相手との過去のコミュニケーション履歴に応じて、歩数の増加効果の大きさに差異があることが示唆された。

1. 研究概要

高齢化の進展および疾病構造の変化により、日本では生活習慣病の予防、社会生活を営むために必要な機能の維持および向上等による、健康寿命延伸の実現が求められている[1]。生活習慣病の予防、身体機能の維持のためには、健康的な生活習慣の定着が欠かせない。習慣化のために、身近なデバイスであるスマートフォンやスマートウォッチ等のモバイル端末を活用した支援が重要となってきている。

スマートフォンを活用した健康的な習慣づくり支援の取り組みとして、毎日の歩数に応じてポイントを付与するサー

ビスが提供されている^{*1}。また、健康に限らず、一人ひとりが望む習慣化を支援するため、同じ目標を掲げるユーザ同士でグループを形成し、グループ内のユーザ間で日々の目標達成状況を共有することで、習慣化を支援するサービスが存在する^{*2}。このサービスでは、ユーザに対してポイントなど金銭的価値を持つ報酬を付与することなく、仲間との関わりを通じて習慣づくりを支援している点が特徴的である。

我々は、仲間との関わりを通じた習慣化支援に関してより追求できる観点があると考える。具体的には、ユーザ同士の親密さに応じて、相手からの応援行為による行動変容への寄与度が変動しうること、すなわち、応援を行うかどうかの量的な観点だけでなく、どういった相手から応援を受けるかどうかの質的な観点が重要であるという仮説を立てた。この仮説について、著者らの先行研究[2]により、ソーシャル・ネットワーキング・サービス(SNS)の公開データを用いた検証により、運動継続に関して応援(いいね!)の量だけでなく誰からの応援なのかも影響しうることを確認しているが、相関分析にとどまっており、因果が明らかになっていない課題が残っていた。

本研究ではこの課題解決を目的として、応援による運動促進効果を因果も含めて分析にするため、学生を対象として、目標を宣言し他者と共有することができるSNSサービスを実装し、約3ヶ月にわたる目標宣言、ならびに歩数のデータを収集した。データ収集期間において、目標宣言をサービス内の機能として実装する前後のデータを用いることで、目標宣言を行うこと自体の効果、ならびに目標に対して応援されることの量および質的な効果を明らかにすることを目的とし、取得したデータについて目標宣言機能開始前後の差分の差分析、および回帰分析の一手法の固定効果モデルを用い、目標宣言の実施や応援を得るかどうかの変数の回帰係数値として定量的な効果を推定した。

その結果、歩数目標宣言の実装後に歩数目標宣言を実施したユーザ群において、以降の期間において対照群に対し歩数の平均値が有意に高いことを確認し、歩数を増加させ

Evaluation of Quantitative and Qualitative Effect of Social Nudge Platform

^{†1} TAKASHI HAMATANI, NTT DOCOMO, INC.

^{†2} KEIICHI OCHIAI, NTT DOCOMO, INC.

^{†3} WATARU YAMADA, NTT DOCOMO, INC.

^{†4} SATOSHI HIYAMA, NTT DOCOMO, INC.

^{†5} TAKUYA SHIRAI, NTT DOCOMO, INC.

^{†6} YUTAKA ARAKAWA, Kyushu University

^{*1} dヘルスケア, <https://health.docomo.ne.jp/>

^{*2} みんチャレ, <https://minchalle.com/>

る効果を確認するとともに、歩数の増加効果が9週間にわたって継続することを確認した。さらに、固定効果モデルによる検証により、宣言した目標に対する他者からの応援に関して、応援1回ごとに歩数の平均値が約289から571歩増加すること、ならびに、過去のコミュニケーション回数が多いユーザからの応援は、過去にコミュニケーション履歴がない相手と比較して歩数を約2倍増加させることを確認し、応援の量的・質的な効果を確認した。

2. 関連研究

運動や読書などの習慣化のために、エーテンラボ株式会社より「みんチャレ」というサービスが社会実装され、習慣化に挑戦したいユーザを中心に活用されている。このサービスでは、読書の習慣化や体重の減少など同じ目標を目指すユーザが匿名で5人集まり、チームで報告を行ったり、挑戦が途切れる日が続くとチームを脱退させられるなどの仕掛けにより習慣化を高める支援を行っている。「みんチャレ」では同じ目標を持った仲間からの働きかけのより社会的な繋がりを用い行動変容の後押し（本研究ではソーシャルナッジと呼ぶ）を実現している。

みんチャレではチームの仲間の頑張り度合いが自身の行動変容にも影響すると考えられる。例えば、継続率が高いチームにおいては、皆が頑張っているという同調圧力（ピアプレッシャー）が強くなることに対し、継続率が低いチームにおいてはピアプレッシャーが弱まると考えられるため、どういう仲間とチームを組むかが重要な要素となりうる。先行研究においても、チームを組むだけでなく競い合う設計にすることが重要であることや[3]、行動変容を目的としたチームの設計において感情的なサポートや情報的なサポートが求められていることなどが明らかになっている[4]。

著者らの先行研究[2]では、チームという概念を拡大し、SNSの一つであるTwitterの公開情報を用いた検証により、ダイエットの継続において応援（いいね！）の量だけでなく誰からの応援なのかという質的な観点も影響しうることを確認した。この研究では継続と応援の相関性($|R| \approx 0.357$)を認めたものの、継続したことに対する称賛なのか、称賛を受けたことによる継続成功なのかという因果が明らかにならない課題があった。

3. 問題設定

本研究では、計測が容易であること、および一日単位で行動変容を起こすことが容易であると考えられることなどから、運動の指標として歩数に着目する。さらに、歩数を単に計測するだけでなく歩数の目標宣言や応援を行う機能を提供する研究目的のアプリケーションを配布することで、目



図1 研究用アプリケーションの概要

標宣言や応援の実績データを収集する。このとき、目標宣言や応援を行う機能を時間差で提供し、歩数を宣言しない期間と、歩数を宣言することができる期間を意図的に作り出し、これらの期間の差分を比較することで、目標宣言や応援が歩数に与える因果的効果を分析することを目指す。

アプリケーションでは、行動変容のための仕掛け（ナッジ）を取り入れている。例えば、ナッジのフレームワーク EAST [5]に基づき、他の人の目標宣言を見られるようにすること、他の人の目標宣言に応援やコメントを付けられるようにしておくことで、Socialの観点を取り入れる。また目標を宣言すること自体も MINDSPACE フレームワーク [6]におけるコミットメントとして知られている。さらに、歩数などの目標宣言を自動的に作成することだけでなく、他の人の目標に相乗りすることでより容易に宣言する機能を取り入れている（Easyの観点）。さらには、プッシュ通知によって応援されたことを知らせる機能により、Timelyにユーザに働きかけることを目指す。

さらに、すべての機能を実装したアプリケーションを配布、ユーザが利用開始するのではなく、当初歩数の計測だけを行う期間を設け、その後、目標宣言などの社会的な行動変容のための機能を提供することで、ベースラインの歩数を取得し評価に組み込むことで、目標宣言などの機能の追加に伴う因果的効果の分析を行う。

4. システムを利用した実データ収集

歩数と応援の関係分析に必要なデータを取得するため、本研究では九州大学の学生を対象に研究用アプリケーションを提供する。アプリケーションの概要を図1に示す。アプリケーションでは、歩数を計測する機能と、歩数の目標をリストから選んで宣言する機能（歩数宣言）、他の人の目標宣言に対して相乗り（同じ目標を宣言する機能）や、応援やコ

メントを送る機能を有する。また、歩数に限らず、勉強などの利用者自身が達成したい目標を宣言し、手動で達成を報告する機能（自由目標宣言機能）を有する。歩数を計測する機能では、各 OS で提供されている計測用の API を利用し歩数を取得する。歩数の目標を宣言する機能では、歩数を 4,000 歩、6,000 歩、8,000 歩、10,000 歩の中から選択し、自由記述のコメントと合わせて宣言することができる。宣言された目標は、他のユーザがアプリケーションを利用したときに最初に表示される画面（タイムライン）において表示され、目標を見たユーザは応援（ハートマーク）やコメント（自由記述）を送ることができる。コメントや応援を受領したユーザには、OS のプッシュ通知機能により通知が送信される。

上記のアプリケーションを、研究参加およびデータ提供の同意を得た被験者を対象に配布し、2022 年 7 月 1 日よりデータの収集を開始した。後段で因果的効果を検証するため、アプリケーション提供開始当初は歩数宣言は実装せず、歩数はバックグラウンドでデータ取得するのみとした。したがって、当初は自由宣言機能と応援・コメント機能のみのアプリケーションをユーザは利用し、応援やコメントなどのコミュニケーションを行いつつ、自己の目標管理のためにアプリケーションを利用した。なお、自由宣言は一日に何回も実施可能、歩数宣言は一日一回のみ宣言可能であり、研究参加者には研究参加に対する謝礼として、アプリケーションの利用日数に応じたポイントを配布した。ポイントの配布は自由宣言や歩数宣言、歩数といったアプリケーション内での行動と独立して行うことで、歩数の宣言や歩数といった行動実績とインセンティブは独立に設計した。その後、8 月 16 日より歩数宣言機能を提供開始し、歩数の目標宣言ならびに応援によるデータの取得を開始した。データの収集は 2022 年 10 月 17 日までを行い、研究期間中の参加者も含め研究参加者は最終的に 504 人となり、のべ 26,456 日の歩数データを収集した。8 月 1 日から 8 月 15 日は、アプリケーションで想定外のバグが生じており、歩数データを連携できない事象が発生していたため、次章の評価においては、8 月 1 日から 15 日のデータは含めないものとする期間中の歩数宣言の総数は 112 人による合計 374 件であり、自由宣言の総数は 313 人による合計 3,516 件であった。歩数宣言への応援は 158 人による 5,961 件、コメントは 29 人による 178 件、自由宣言への応援は 294 人による 18,092 件、コメントは 92 人による 1,483 件であった。

以上の特性から、研究参加者のうち約 22% が歩数宣言を利用するが、他のユーザは歩数を宣言せず、自由宣言や他のユーザの宣言にコメントをするなどの利用にとどまっていることがわかる。また、応援機能とコメント機能を比較すると、応援機能の利用数がコメント機能を大きく上回っており、簡単な操作（応援アイコンのタップ）のみで実施でき

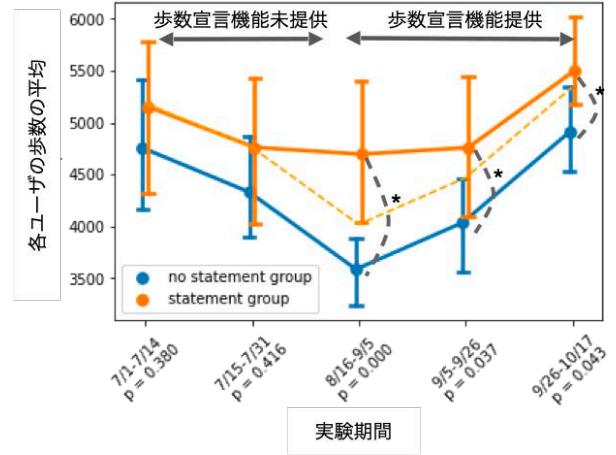


図 2 歩数宣言の有無による影響

る応援行動が支持されていることがわかる。

5. 評価結果

前章で述べたデータセットを用いて、歩数の宣言や応援という行動と歩数の関係性を分析し、ソーシャルナッジの効果を明らかにするため、本章では以下 2 つの方針で分析を行う。

1. 歩数目標の機能の有無が分かれる期間を比較することにより、歩数宣言自体の効果を測る。
2. 歩数宣言がある条件のもとで、宣言の有無や応援の有無による効果を測る。

5.1. 歩数宣言自体の効果の評価

歩数宣言自体の大規模的な効果を評価するため、歩数宣言機能をリリースする前後における歩数を比較する。今回のデータ取得期間は 7 月から 10 月と夏季から秋季にわたり、この時、歩数には季節的な影響が含まれるものと考えられるため、単に歩数宣言機能のリリース前後を比較すると季節的な影響が混在する評価となってしまう。そのため、本研究では差分の差分法 [7] を用いて、歩数宣言機能の実装後に歩数宣言機能を利用した群 (N=112) と利用しなかった群 (N=392) の差を時系列的に比較することで季節的影響の除去を試みる。評価においては、歩数宣言機能リリース前の二群のトレンドが並行しているかどうかを確認するため 2 週間ごとに 2 つの期間に分け、さらに、歩数宣言機能リリース後の持続効果を検証するため、リリース後の期間を 3 週間ごとに 3 つに分けて評価を行った。

図 2 に 5 つの期間における歩数宣言実施群と不実施群の各ユーザの歩数の中央値について、中央値と信頼区間を示す。歩数は天気や気温、その人の用事などの外的要因の影響

を受けやすい値であるため、本研究ではユーザの内的なモチベーションによる歩数の傾向がより現れやすいと考えられる中央値を各ユーザの代表値として用いた。さらに、各群における中央値の分布に関して、マンホイットニーの U 検定を用い、統計的有意な差が生じているかどうかを確認している。

図 2 の結果より、歩数宣言のリリース前の 2 期間においては、2 群における中央値の分布には有意差が生じていないことが確認できた。一方で、信頼区間としては重なりがあるものの、傾向としては宣言ありの群のほうが歩数の中央値が大きいユーザが多いとは解釈できる。これは、歩数宣言を行うユーザの中には、もともと歩数が多いから宣言を行なう、すなわち宣言を行うから歩数が上がることと逆の因果を伴っている可能性があることが示唆されている。

さらに、歩数宣言リリース前の 2 期間を比較すると、2 群の平均値が並行に推移していることが確認できる。このことにより、歩数宣言をリリースした以降の期間も並行するトレンドを仮定することができるため、破線で仮定した並行トレンドを示す。破線で示した並行トレンド（宣言実施の群において、宣言を実施しなかった場合の仮定）と、宣言実施の群を比較すると、差が生じていることがわかる。この差について、季節的な影響を除去した上で歩数宣言によつてもたらされた効果と捉えることができる。また、破線と宣言実施群の差を比較すると、歩数宣言実装直後の 3 週間ににおいて宣言実施群の歩数が大きく上回っており、宣言実施群と未実施群でも歩数の差が大きく開いていることが分かる。このことより、歩数宣言を提供開始したことにより、歩数が増加したことが示唆される結果となった。加えて、マンホイットニーの U 検定による検定結果によると、歩数宣言実装後の 3 期間（9 週間）にわたって統計的に有意な差が継続していることが分かる。

以上の結果より、歩数宣言の実施前後において、歩数宣言の実施前は 2 群間の歩数分布に有意差が生じていなかったことに対して、歩数宣言実施後は 2 群間の歩数分布に有意差が生じ、かつ 9 週間に渡って有意差が継続していることから、アプリケーション内で歩数を宣言し他のユーザから見える状態にする行為によって歩数が増加する効果があつたものと考えられる。このことから、ナッジのフレームワークとして知られる MINDSPACE[6] における一つの手法であるコミットメント（目標を公言すること）が、歩行という行動変容達成においても効果をもたらしたといえる。

5.2. 歩数宣言における宣言有無や応援の効果の評価

前節の評価により、「歩数宣言を行うか否か」という観点において、歩数宣言を行うことの効果は差分の差分析によって明らかになった。本節では、歩数宣言によって歩数を増加

させる効果が生じるという前提の元で、歩数宣言に付随する応援やコメント、相乗りといった他のユーザからの働きかけの効果を定量的に測ることを目標とする。そこで、本研究ではユーザのソーシャルナッジに関する行動を変数化したうえで、パネルデータ分析でよく用いられる固定効果モデルを用いたフィッティングにより行動の効果を回帰式の係数として定量化する。

ソーシャルナッジに関する行動は、歩数宣言の実施だけでなく、自身の目標宣言に相乗りされること、応援されること、およびコメントされることが考えられる。上記の行動のうち、相乗りとコメントに関してはそれぞれ 28 件、178 件と応援の 5,961 件と比較して数が少ないとから、分析の簡単のため相乗りとコメントは本研究における分析対象からは除外することとした。さらに前節の評価により、歩数宣言を行うユーザと行わないユーザの数に偏りがあることが確認できたため、サンプル数の調整を行うことで偏りの影響を低減する。具体的には、歩数宣言を行うかどうかをアウトカム、歩数の平均値と標準偏差、研究参加日数を共変量として傾向スコアによるマッチングを行い、歩数宣言を行うユーザとそうでないユーザの数を均一にした上で、固定効果モデルへのフィッティングを行う。

各行動の重みを詳しく分析するため、歩数宣言については初回の宣言と、二回目以降の宣言かどうかを別々のダミー変数として定義する。また、応援に関しては、先行研究において誰から応援されるかが重要であると知見を得ていることから、応援された相手がどういう相手かを考慮するため、過去にも応援されたことがある相手かどうかを考慮する。具体的には、各ユーザの過去の歩数宣言だけでなく自由宣言への応援も含め、「過去に応援されたことがない相手かどうか」「過去に 1 度応援されたことがある相手かどうか」「過去に 2 度以上応援されたことがある相手かどうか」の 3 つに場合分けを行った。ここで、閾値としては 1 回と 2 回を設けているが、これは今回収集したデータに基づき、それぞれの場合分けされたサンプル数が可能な限り均等になるようを選定した。

以上の条件のもと、具体的には以下の回帰式を用いる。

$$\begin{aligned} Steps_{it} = & \alpha_0 + \alpha_1 \cdot Statement_1 + \alpha_2 \cdot Statement_2 \\ & + \alpha_3 \cdot Like_0 + \alpha_4 \cdot Like_1 + \alpha_5 \cdot Like_2 \\ & + \sum \beta_i \cdot Individual_i + \sum \gamma_t \cdot Day_t + \epsilon_{it} \end{aligned}$$

ここで、 $Statement_1$ と $Statement_2$ は初回の歩数宣言であること、2 回目以降の歩数宣言であることであることを表すダミー変数であり、 $Like_0$, $Like_1$, $Like_2$ はそれぞれ過去に応援されたことがない相手からの応援の数、過去に 1 回応援されたことがある相手からの応援の数、過去に 2 回以上応援されたことがある相手からの応援の数をそれぞれ表す。また、 $Individual_i$ は各ユーザを表すダミー変数、 Day_t

表 1 歩数宣言実施と歩数宣言への応援が歩数に与える影響

	Parameter	Std. Err.	T-stat	P-value	Lower CI	Upper CI
α_0	6381.0	119.71	53.306	0.0000	6146.4	6615.7
α_1	-211.92	235.49	-0.8999	0.3682	-673.53	249.69
α_2	-528.76	278.84	-1.8963	0.0579	-1075.3	17.810
α_3	289.61	125.08	2.3154	0.0206	44.428	534.78
α_4	370.06	157.52	2.3493	0.0188	61.294	678.83
α_5	571.85	72.457	7.8923	0.0000	429.82	713.87

は各日付を表すダミー変数である。以上の式に実際のデータを当てはめることで回帰係数 α_0 から α_5 を推定する。各ユーザを表すダミー変数を導入することで主体固定効果、日付を表すダミー変数を導入することで時間固定効果を考慮する。

表 1 に歩数宣言の実施、ならびに歩数宣言への応援が歩数に与えた影響を示す。表では、Parameter 列に推定されたパラメータ（歩数）を表す。すなわち、 α_0 の行では、切片として 6381 歩、すなわちユーザの個人差や日付による差を除いた上で、すべてのユーザはそれぞれの日で、平均的に 6381 歩歩いていることを示している。その前提のもとで、歩数宣言を実施するかどうかの係数 α_1 と α_2 は負の値を示しており、歩数宣言を行うことは歩数を下げる方向に寄与するよう見える。その一方で、応援の係数 α_3 , α_4 , α_5 が正の値であること、ならびに応援は歩数宣言を行っていないと実施できないことから、これらは打ち消し合って、応援の数に応じて歩数が増加していくことを示していると考えられる。また、宣言したにも関わらず、応援がゼロであると、最終的な歩数については負の影響を与えるということも示している。ただし、p 値は有意水準 5% の条件のもとでは有意な効果であることは認めない。

応援の係数 α_3 , α_4 , α_5 に着目すると、p 値は有意水準 5% を下回っており、応援が歩数に正の効果をもたらすことは有意であるといえる。さらに、パラメータの値の大小関係が $\alpha_3 < \alpha_4 < \alpha_5$ と過去の応援の量に応じて大きくなっていることから、応援される相手との過去のコミュニケーションの量に応じて、応援が歩数にもたらす質的な効果が増幅されることが示唆されている。以上の結果を踏まえると、歩数目標を宣言すること自体の効果は有意ではないが、宣言した目標に対して他者より受け取る応援については、歩数に対し正の影響を与えることを定量的に確認できた。

6. まとめ

本研究では運動の習慣化を目的とし、普及している SNS サービスを利用したソーシャルナッジの手法を検討するた

め、目標宣言型プラットフォームを用いて歩数の目標宣言、応援のデータを収集、分析を行った。実際に 504 人のべ 26,456 日のデータを取得し、歩数宣言を行うかどうかによる歩数への効果、ならびに歩数宣言に対する他ユーザからの応援が歩数に与える効果を検証した。その結果、歩数目標宣言の実装後に歩数目標宣言を実施したユーザ群において、対照群に対し歩数が増加する効果があり、歩数の増加効果が継続することを確認した。さらに、固定効果モデルによる検証により、宣言した目標に対する他者からの応援に関して、応援 1 回ごとに歩数の平均値が増加すること、および過去のコミュニケーション回数が多いユーザからの応援が歩数の平均値をより多く増加させることを確認し、応援の量的・質的な効果を確認した。

謝辞 本研究は九州大学倫理審査委員会による承認のもと行われたものである（研究課題名：人々を望ましい行動へと導く行動変容技術に関する研究、承認番号：シス情認 2022-02）。本研究で利用したアプリケーションの開発、運用に協力した株式会社イマーゴ iQLab プロジェクト関係者に感謝の意を表す。

参考文献

- [1] 厚生労働省: 国民の健康の増進の総合的な推進を図るための基本的な方針, https://www.mhlw.go.jp/stf/seisakunitsuite/bunya/kenkou_iryou/kenkou/kenkounippon21.html.
- [2] 濱谷尚志, 三村知洋, 山田渉, 落合桂一: Twitter を活用したソーシャルナッジのためのいいね！と運動習慣化成功との関係の分析、情報処理学会 IoT 行動変容学研究グループキックオフシンポジウム (2022).
- [3] Nishiyama, Y., Okoshi, T., Yonezawa, T., Nakazawa, J., Takashio, K. and Tokuda, H.: Toward health exercise behavior change for teams using lifelog sharing models, *IEEE Journal of Biomedical and Health Informatics*, Vol. 20, No. 3, pp. 775–786 (2015).
- [4] Luhanga, E. T., Hippocrate, A. A. E., Suwa, H.,

- Arakawa, Y. and Yasumoto, K.: Identifying and evaluating user requirements for smartphone group fitness applications, *IEEE Access*, Vol. 6, pp. 3256–3269 (2018).
- [5] 厚生労働省: 明日から使えるナッジ理論, <https://www.mhlw.go.jp/content/10900000/000506624.pdf>.
- [6] Dolan, P., Hallsworth, M., Halpern, D., King, D. and Vlaev, I.: MINDSPACE: influencing behaviour for public policy (2010).
- [7] Conley, T. G. and Taber, C. R.: Inference with “difference in differences” with a small number of policy changes, *The Review of Economics and Statistics*, Vol. 93, No. 1, pp. 113–125 (2011).

短時間パターン認識結果の投票的統合によるドア利用者識別

豊浦正広 ^{†1}
山梨大学 奥田雅也 ^{†1}
山梨大学

1. はじめに

カメラ画像からの人物識別は、ほぼ 100% の精度を実現する手法が多く登場している [1, 2, 3, 4]. 一方で、プライバシ保護や機密情報保護の観点から、家庭内や医療・介護施設内、企業内の IoT システムではカメラ画像の利用が避けられることも多い。本研究では、ドアの利用者識別に焦点を当てて、カメラ画像を用いた人物識別を目指す。

カメラ画像を用いないことで容姿や表情などの不要な情報を獲得せずに済み、心理的障壁を下げることができる。初めから情報を獲得しないことによって、もしデータが漏洩したときにも被害を少なくすることができる。カメラを通して顔画像などの取得を制限する動きは世界中で広まりつつある。たとえば、EU は法執行当局による顔認証を含む生体認証の使用を禁止する規制案を発表しており [5]、この規制は EU に居住する人には適用されるため、EU 域外においても規制が適用される。

SNS を通しての企業の情報流出も問題になっており、SNS にアップロードされた画像に移りこんだ個人情報の取り扱いについて、社会的な関心も高い [6, 7]. カメラは人間にとっても直感的に状況を認識しやすい一方で、認識のしやすさ故に、情報が流出したときには問題となることが多い。本研究では、カメラなしで利用者の特定ができるようにする立場に立ち、その中でも屋内での人の行動追跡に有用であるドアの利用者を認識する手法を提案する。

本研究では 9 軸センサによる利用者の認証を目指す。9 軸センサは軽量で低電力であることから、利用者に意識されにくく、電力確保もそれほど問題にならない。この点においても、9 軸センサは IoT デバイスとしてカメラよりも優れているといえる。個人の認証には、ID カードや生体認証、PINなどを求める方法も考えられるが、これらはいずれもドアを開けるための手間を増やすことになり、紛失や盗難の問題も起こる。本研究で行うドアの利用者認証は、これらの認証に比べるとセキュリティ的には強固とはいえないものの、家庭内や企業内で利用者が限られた空間で用いて人の流れを追跡するために向いている。認証の手間を増

やさないので、人の流れを妨げることもない。

ドアに取り付けた加速度からの利用者の識別を機械学習を用いて行う研究には先例 [8, 9, 10, 11] があるものの、精度は 90% 程度までにとどまっていた。精度が高くならなかったひとつの原因是、SVM や古典的な DNN が採用されていたためである。我々は Transformer[12] を含む新しいネットワーク構造を導入することで精度を高めた。もうひとつ的原因は、学習のためのサンプル数が十分に集められなかっただためである。深層学習のためにはアノテーションの付けられた大量のデータが必要となるのが一般的である。従来手法では、ドア 1 回の開閉で 1 つのサンプルとしていたために、現実的に十分なデータが得られなかった。この結果、十分にネットワークの訓練ができなかつたものと考えられる。

我々はこの問題を解決するために、ドアの開閉による時系列 9 軸データを微小な運動の連続として捉えて、ネットワークの訓練に十分なサンプル数を確保する手法を提案する。1 回のドアの開閉の中で投票を行うことで、最終的な利用者の識別を行う。

本論文の貢献は以下の通りであり、概要を図 1 に示す。

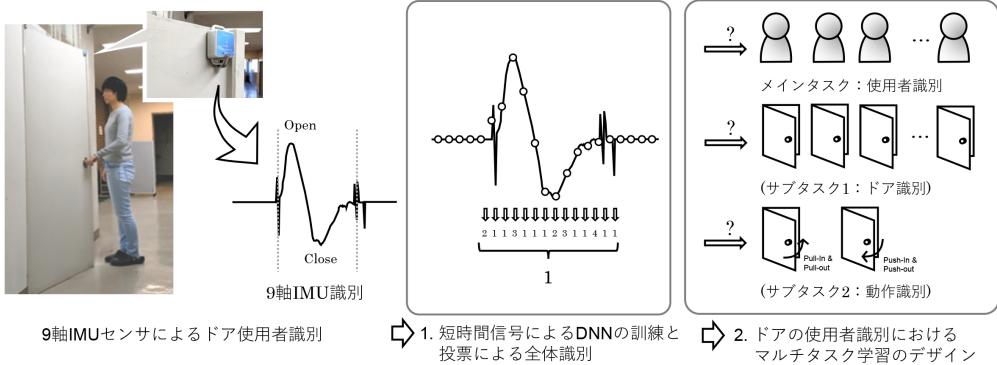
1. ドアの 1 回の開閉で 1 つのサンプルしか得られなかつた従来の訓練方法に対して、ドアの開閉を微小な運動の連続と捉えることで、DNN の訓練に十分なデータを確保できるようにした。 (4.1)
2. 複数のドアに取り付けて、複数の利用者を認識するための、マルチタスクネットワークをデザインした。ネットワークには Transformer を含み、時系列信号に対する効率的な訓練を実現した。 (4.2)

本論文の構成は以下の通りである。2. では、ドアのユーザーを識別する手法に関する関連研究を紹介する。3. では、実験を行った環境や条件を紹介し、データセットをいかに作成したかについて説明する。4. では、Transformer を含むマルチタスク学習によるドアの利用者の手法を紹介する。5. では実験結果を示し、6. で結論を述べ、今後の研究の方向性を述べて論文を締めくくる。

Door User Identification with Voting Short-term Pattern Categorization

^{†1} Masahiro TOYOURA, University of Yamanashi

^{†1} Masaya OKUDA, University of Yamanashi



2. 関連研究

ドアの利用者識別のために、カメラ [13], ID カード, ビーコン認証, 生体認証, PIN, スマートフォン認証 [14]などを利用する方法がある。カメラは利用者に特別なツールの持ち運びを要求しない利点があるが、映像が流出した場合に、格好や持ち物、表情などといった不要な情報が含まれ、会社内や家庭内では使いにくいことがある。カメラ以外のツールでは、一般的にドアを開ける以外の動作を必要とするために、さりげなく行動を管理したい IoT アプリケーションには向かない。スマートフォンによる認証 [14] では、ドアに取り付けたビーコンによって、ドアに近づくだけで認証する試みを行っているが、6 名の利用者で 69%, 10 名の利用者で 62% の精度であり、高い精度であるとはいえない。ただし、この研究では、ドアに近づくときだけにドアを開けるという機能が実現できており、むしろこちらに注力したように見える。認証ツールを紛失したり忘れたりしたときに、ドアを開けられなくなるリスクは残る。

ドアの利用者識別のために、ドアに取り付けた加速度センサを用いる先行研究がある。Piltaver ら [8] は、ドアに取り付けた加速度センサのデータから、登録されているデータとの照合を行うことによって、12 人の被験者を 89.9% の accuracy で識別した。Dynamic Time Warping(DTW)によって時間方向の変化を吸収し、得られる特徴量を k-NN によって照合することで、識別精度を向上した。Fukao ら [9] は、ドアノブに取り付けたジャイロセンサから利用者の識別を行い、4 名の識別に入室時には 0.73、退室時には 0.90 の F-measure を得た。彼らの手法では、ドアノブを回す動作だけで利用者を認識するので、認識結果によってはドアを開けないシステムを開発することもできる。同じ研究グループ内の Ishida ら [10] は、冷蔵庫のドアの開閉についても似た技術で 7 名を Bayesian Network によって 91.9% の

accuracy で識別した。Vegas ら [11] も、ドアハンドルの端に加速度センサとジャイロセンサを取り付けてユーザの識別を行った。DTW で加工した特徴量を SVM や ANN を用いて学習し、47 人の利用者に対して 0.912 の AUC 値を得ることができている。

いずれの研究においても、機械学習でそれほど高い精度を得ることができていなかった。いずれも単一のドアに対する識別であり、ドアを変更したときの識別率については言及されていない。深層学習に耐えるだけのドアの開閉データのサンプルを多数集めることは難しく、比較的簡単な学習器によって識別を試みているところに問題が残ると考えられる。また、網羅的に列挙したハンドクラフト特徴量では、特徴量の次元を低く抑えることはできるものの、データの特徴を十分に表現できなかつたのではないかと、我々は推察する。本研究では、少数の試行からデータを大量に集めて、深層学習の自動特徴抽出によって、認識精度を向上させることを目指す。

加速度センサからの人間行動推定では、運動の種類を識別するがあり、本研究で利用する Transformer を含めて、深層学習も積極的に取り入れられている [15, 16, 17, 18, 19]。ミーティング中の人間行動に特化して、その行動の検出を目指した研究もある [20]。Fitbit や GARMIN, Apple watch といったような民生用のウェアラブルウォッチでも、睡眠や歩行、自転車等の生活行動の認識が提供されており、生活の中に浸透している IoT デバイスとなっている。これらの手法は加速度センサを人が装着することを想定しており、モノに取り付けて受動的に人間の行動を推定する本研究とは立ち位置が異なるものの、タグ付けや認識手法の一部は大いに参考になる。Hassan ら [19] は、ウェアラブルセンサの信号から人間行動認識を目指し、2.56 秒分の信号波形から、平均値や絶対値、周波数成分の歪度・尖度、エントロ

ピーなどの 17 のハンドクラフト特徴量を設計した。識別器についても、1 層のニューラルネットワーク、SVM、3 層のニューラルネットワークについて試行を行っている。最終的には、この 17 の特徴量を 3 層のニューラルネットワークで学習することによって、立つ・座る・歩く・横たわるなどの 12 種類の行動を 89.61% で認識した。カテゴリによっては十分なサンプル数がないとも論文数に書かれている。サンプル数が大量に得られないときには、自動で抽出される特徴量を得ることも、より深いニューラルネットワークを用いることも難しくなる。5. に示す実験では、手動抽出の特徴量の精度の違いについても、検討する。

3. ドア信号の取得

ドアに簡易に取り付けた 9 軸センサによって、ドアの利用者の識別を目指す。

3.1. ドア動作による 9 軸信号の特性

ドアには図 2 に示すように、ドアフックに取り付ける形で 9 軸 IMU センサ WitMotion WT901SDCL を取り付けた。ドアの上部に取り付けることで、利用者は特に意識することなく、ドアの開閉をすることができる。同製品は SD カードを挿入すれば 9 軸データの収録を開始し、SD カードを取り出せばデータの収録を停止する。一度収録を開始すれば、数時間はバッテリーが持続するので、複数人の実験にも利用可能である。将来的にはソーラーバッテリーなどへの接続によって充電を不要にし、Bluetooth 接続などでセンサに触れることなくデータを取り出すことを目指す。同社製製品には Bluetooth 接続のものがあるが、獲得できる信号周波数が低いことなどから、今回の製品を採用した。

9 軸センサによって、200Hzまでの加速度・角速度・地磁気のデータを取得することができる。Bluetooth で通信する装置では、バッテリーの減りが早くなり、データの欠損も起こりやすくなる。データ欠損を避けるためには、周波数は 200Hz 未満に落とすことも可能であるが、従来研究 [19] に従い、デバイスの上限速度である 200Hz とした。通信によるデータ損失を避けるために、装置に内蔵された SD カードに信号を保存し、あとからデータを手動で取り出すこととして、実験を進めた。図 2 はドアの外から入ってドアを開め、中から再度ドアを開けて外に出てドアを閉めたときに、取得できる信号の例を示す。

ドアの利用者を認証して制限するような機能はないため、ドアを利用した後に利用者を識別するにとどまる。ドアノブの動きを抽出することによって、ドアの開閉前に利用者の認証をすることもできる [9] が、本論文では高度なセキュリティまでは要求しないような、建物内の人の動きの記録を想定して、ドアノブではなく、ドア上部の動きの推定を

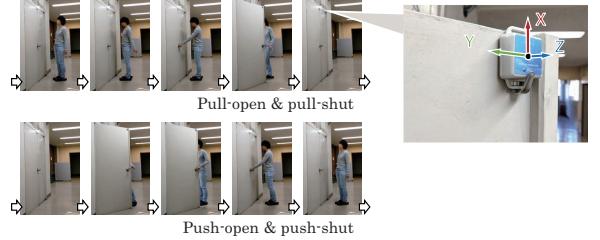


図 2 各軸の向き。X 軸はドアの上方向、Y 軸はドアの水平方向、Z はドアの手前方向とした。実験の様子。ドアの上部の最も移動量が大きくなる位置に 9 軸センサを取り付けている。軽量であり、利用者はセンサを意識することなくドアを開閉することができる。

行った。

3.2. ドア信号に対するアノテーション付与

収録したドアセンサのデータを教師あり学習のための訓練データにするためには、どの時刻のデータが、どの利用者・どのドア・どの動作に対応するのかのラベルを付与する必要がある。どの利用者であるか、どのドアであるかは比較的明確にラベルを付けることができる。実験時には、ある実験参加者だけが、あるドアを開閉するためである。動作には、引いて開けて引いて閉める、押して開けて押して閉めるの 2 通りがあることになる。こちらもこの 2 つの動作が繰り返し出現するので、どちらから始めたかだけを記録しておけば、容易にラベルをつけることができる。

ドアが開き始めてから閉まり終わるまでを特定することは、意外に難しい。センサ信号のパワーを見ることで、ドアに人が触れているかを推定することはできるものの、ドアから手が離れてしばらくの間には、ドアが振動してパワーが十分に小さくならないためである。また、センサにはジッターノイズが載るために、実際にはドアが揺れていなくても、センサ信号のパワーは大きくなることがある。ドアから手が離れたものに対して、利用者が誰であるとか、動作がどうであるかというラベルを振ってしまうと、利用者に依らずに同じようなパターンを描くことになるために、ネットワークの訓練がうまく行かないことが予想される。ドアに手が触れている時間帯をできるだけ正確に判定したい。

センサ信号を 1 動作ごとに切り分けるためには、図 3 に示すように、センサ信号をまず平滑化して、ジッターノイズを除去する。平滑化のためには、従来手法 [19] でも用いられているバターワースフィルタを利用した。平滑化した上で、パワーに閾値を設けて、ある閾値を超えたときに動作が開始し、ある閾値を下回ったときに動作が終わったもののみを抽出する。ドアの振動が十分に収まらないうちに、次の開

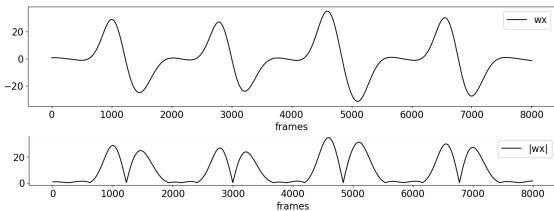


図 3 x 軸回りの角速度の解析によるドア信号の切り分け

閉動作が始まってしまうときには、閾値処理によってうまく信号が切り分けられないことがあるので、ドアの開閉の回数が記録したものよりも少なくなるときには、時系列信号を目視で確認して、パワーの谷となるところで、動作が完了し、すぐに始まったものとみなして、時間帯の設定を修正することとした。

4. Transformer を含む DNN による信号解析

4.1. 短時間分割サンプルと投票

従来の IoT センサパターン認識では、1 パターンの開始から終了までの特徴を捉えることで認識を行ってきた。この方法の問題はいくつかある。まず、(1) 時間長の異なるパターンを扱う必要があることである。次に、(2) 開始と終了の時刻が正しく得られないと、パターンが異なって見えることである。さらには、(3)1 回の動作で 1 サンプルしか得られないために、DNN の訓練に必要なサンプル数を準備するのに労力がかかることがある。

(1) に対しては、パターンに対してフェーズを手動で定義して、それぞれのフェーズの継続時間や波形を解析する手法が提案されている [11]。パターンをモデル化して認識するので、モデル化からずれたパターンでは認識がうまくいかない。ハンドクラフト特徴量を設定せずに、データのみから特徴を抽出する深層学習の利益も得られにくい。

(2) に対しては、3.2 で述べる通り、単純なドアの開閉パターンであっても、開始と終了の時刻は判断がしにくい問題がある。データセットを作成するときには手動でタイミングは設定できるが、実行時には自動でタイミングが得られる必要があり、これに失敗すると、その後の認識はさらに難しくなる。

(3) に対しては、実験参加者により多くの試行を重ねてもらう必要がある。ドアの種類が増えれば、またそれだけ試行回数を増やす必要がある。ファインチューニングなどによって新しいドアに対するネットワークを獲得しようとするときには、新しい利用者に多くの試行を再度重ねてもらう必要があり、サービスの実現が難しくなる。

本研究では、時系列センサ情報を細かく分割して、それぞ

れの分割データに対して認識を行うようにすることで、(1)から(3)の問題を解決する。最終的には、開始から終了までの認識結果で投票を行うことによって、1 動作に対する認識結果を得るようにする。この方法だと、パターンの時間長に影響されにくく、開始・終了の時刻のずれに対する影響も少なくて済む。細かく分割したデータの数は大量になるので、実験参加者には少数回の試行を行ってもらうだけで、訓練に耐えうるサイズのデータセットを準備することができるようになる。

実行時には、一度に複数回の開閉が連続することは考えにくいので、このような時間帯の分割の失敗が起こることは、あまり想定する必要はないものと考える。もし分割に失敗しても、後に述べる短時間ごとの認識とその投票のアルゴリズムはそのまま使えるので、大きな問題にはならないものと考える。それでも失敗するような一連のセンサデータに対しては、認識をあきらめて、次の動作の認識を試みるのがよいと考えることとした。

4.2. マルチタスク学習と Transformer による認識

観測される時系列信号には、利用者の特徴の他に、ドアの特徴や動作の特徴が反映される。たとえば、同じ利用者であっても、ドアが異なれば観測される信号が変わる。言い換えると、利用者を認識するためには、ドアの影響を考慮する必要がある。従来手法では、どんなドアであっても人が正しく認識できるように、様々なドアで実験を繰り返して、ドアの影響を排除することを目指してきた。この方法では、多くの試行を行う必要があり、実際に利用するドアが少数であっても、多数のドアでの実験が必要となる。

この問題を解決するために、マルチタスク学習を導入した [21]。メインタスクに利用者の認識を設定し、サブタスクにドアと動作の認識を設定した。マルチタスク学習では、サブタスクを考慮した上でメインタスクを行うことによって、メインタスクで考慮すべきことが深層ネットワーク上の重みとして実現される。動作は、ドアを押して開けて押して閉める (push-in-push-out) と引いて開け引いて閉める (pull-in-pull-out) の 2 種類とした。ドアに対してどちらの動作をするかによって、観測される時系列信号が異なることが予想されるためである。また、この動作が認識できれば、ドアのどちらから入ってどちらから出たかも把握でき、実用上も意味がある。

具体的なネットワークとして、図 4 に示すネットワークを構築した。入力される時系列信号から、人・モノ・動作の 3 つの認識結果が得られるように構成されている。ネットワーク内には、時系列信号をうまく扱える Transformer[12] を採用している。これによって、1 時刻だけの信号だけではなく、その時刻に至るまでの一連の時系列信号を考慮して、

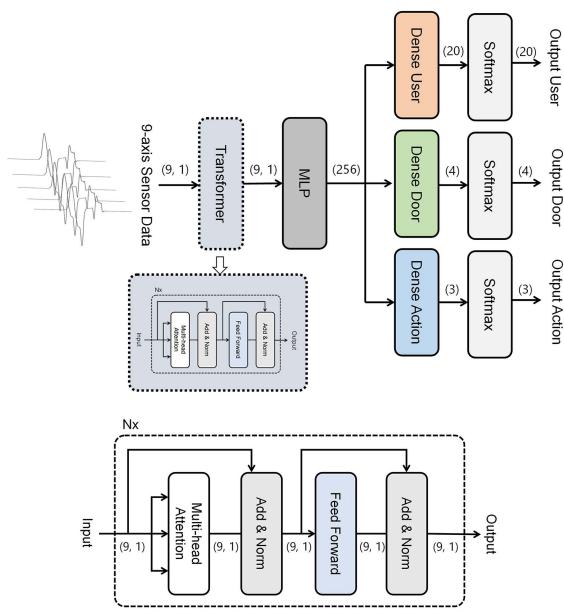


図4 マルチタスク学習に用いたネットワーク。人・ドア・動作のそれぞれの認識を行う。

最終的な認識結果を得ることができるようになる。3. で述べたデータセットの信号が入力されると、対応する正解タグが教師ラベルとなって、深層学習ネットワークの訓練が行われる。

5. 実験

人・ドア・動作がそれぞれ認識できるかを確かめるために、実験協力者を集めて実験を行った。

5.1. データセットの収集

実験では、5つのドアの上部の端から3cmの位置に9軸IMUセンサを取り付けて、20名の実験協力者(11~42歳、男性15名、女性7名、3名は利き手が左手で残り17名は右手)がドアを開閉するときのセンサ信号を収録した。各被験者にはそれぞれのドアを30回ずつ開閉してもらった。ドアの開閉回数は、横で見ている実験実施者がカウンタを使って数え、規定回数になって時点で実験参加者に伝えるようにした。毎回の開閉後にはドアノブから手を離して、一呼吸おいてもらうように指示した。

ドアから手を離してしばらくすると、9軸IMUセンサで得られる信号のパワーが小さくなることが期待できる。少しの間はドアの振動が残っているために、センサ信号のパ

ワーは残っている。信号のパワーに閾値を設けることで、1回ずつのドアの開閉の信号を切り分けた。

ドアが開閉する時系列信号に対して、利用者・ドア・動作のタグが1つずつ与えられる。実験では1つのドアに対する30回の開閉を1~14回目の14回を学習用、15~20回目の6回をテスト用、21~30回目の10回を検証用に分割した。競争相手となるニューラルネットワーク以外の学習器に対しては、1~20回目の20回を学習用、21~30回目の10回を検証用とした。

5.2. 特徴量および学習器の違いによる精度比較

提案する深層ネットワークでどれほど正しく利用者を識別できるかを検証した。比較のために、先行研究で実績があった3層のニューラルネットワーク(FCN)、ランダムフォレスト、最近傍法によっても、利用者の識別を試みた。実装はそれぞれPythonとTensorflowを用い、Google Colaboratoryで実行した。従来手法と比べて、論文に記載された値に対して上下するものがあるが、処理対象のデータセットが異なるためである。特に、従来手法では異なるドアによって実験を行っているものではなく、異なるドアの動きから同じ利用者であることを識別するのは、より難しいタスクとなっている。また、短時間識別結果を多数決して得たことも結果に影響している。

まずは、入力にHassanら[19]が提案した17次元のハンドクラフト特徴量を採用したときの結果を示す。入力信号は短時間信号に分割したうえで、17次元のベクトルに変換されて、学習器に入力される。最終的には、一動作に対する複数の識別結果を多数決して結果を得る。従来手法[19]で良好な結果を得ていたFCNに対しても、0.040だけ優れた結果が得られている。さらに、Transformerを導入することによって、時系列の変化をよく捉えることができ、結果がさらによくなったと考えられる。

表1 学習器の違いによる精度比較。Hassanら[19]が導入したハンドクラフト特徴量を学習器の入力とした。

学習器	正解率	適合率	再現率	F値
FCN	0.898	0.899	0.898	0.898
Random Forest	0.897	0.899	0.898	0.897
Nearest Neighbor	0.718	0.734	0.716	0.717
Proposed	0.939	0.939	0.939	0.938

次に、入力に9軸信号を直接採用するときの結果を示す。提案手法のみで、ハンドクラフト特徴量を入力とするよりも、よりよい結果が得られている。その他の手法では、時系列信号の特徴を扱う部分がなくなるので、結果は悪くなっている。提案手法では、深層学習で得られる時系列特

徵がネットワークで表現されており、よい結果を得ることができている。ドア信号を表現するためには、17 次元の手作り特徴量よりも直接に 9 軸信号を直接入力したときの方が、優れた学習器を作ることができた。

表 2 学習器の違いによる精度比較。時系列の原信号を入力とした。

学習器	正解率	適合率	再現率	F 値
FCN	0.800	0.800	0.800	0.799
Random Forest	0.875	0.874	0.875	0.874
Nearest Neighbor	0.649	0.654	0.649	0.648
Proposed	0.973	0.973	0.973	0.973

従来研究のうちでドアに対して利用者の識別を行ったものを選び、得られる正解率 (accuracy) を表 3 にまとめた。複数のドアに対して利用者識別を試みたものはなかった。Vegas らの手法 [11] の検証のためには、この論文の実験よりも多くの利用者を識別しており、単純に優劣を比較はできない。その他の手法に対しては、多くの利用者と多くのドアに対して、より高い精度を示しており、提案手法の方がよい結果を出している。

6. Discussions and future work

本研究では、カメラ画像に依らずにドア利用者の識別する手法を提案した。9 軸センサの信号からの識別によって、カメラ画像を用いるときには問題となるプライバシ保護や機密情報保護に関する問題を解決した。

ドアの開閉を微小な運動の運動の連続として捉えることで、DNN の学習に十分なデータを確保できるようにし、Transformer を含むマルチタスク学習ネットワークを採用することによって、実験参加者 20 人を 0.973 の正解率で識

表 3 各手法による認識結果

手法	対象	条件	正解率
Piltaver et al. [8]	Door	12 users, 1 door	0.899
Vegas et al. [11]	Handle	47 users, 1 door	0.912
Fukao et al. [9]	Knob	4 user, 1 door (F 値)	0.900
Ishida et al. [10]	Fridge	7 users, 1 door	0.919
Proposed	Door	27 users, 5 doors	0.973

表 4 ドアと動作の認識結果

	クラス数	正解率
ドア	5	0.965
動作 (pull-open and push-open)	2	0.687

別することができた。

ドアや動作の種類の推定精度についても議論することができる。提案手法では、5 つのドアの認識精度と、pull-open と push-open の認識精度は、表 4 に示す通りであった。ドアの種類はよく推定できるが、どの方向から開けたかはそれほどわからないことが、表から読み取れる。種類の数から考えると、それぞれ 0.2 と 0.5 がベースラインとなる。センサを取り付けるときにどのドアであるかはわかるので、推定することに意味があまりない状況も多い。利用者が 1 つのセンサを別のドアに付け替えたり、ドア以外のモノに取り付けるときには有用になる場面があるかもしれない。

実験では、通信によるデータ損失を避けるために、SD カードに保存した 9 軸信号をあとから取り出す方法を採用了。通信によって、バッテリーの消費が早くなるという問題も生じる。実験では従来手法に倣って 200Hz での信号解析を行ったが、これほど高い周波数でなくとも解析ができる可能性があり、より小さな周波数を設定することで、データの損失やバッテリーの消費を軽減できる。今後、人を識別するのに必要な下限の周波数について検証を進めたい。また、Bluetooth などによる通信によって、実時間で利用者を識別するドアの開発を行いたい。

謝辞

本研究は JST さきがけ JPMJPR213 の支援を受けた。

参考文献

- [1] F. Schroff, D. Kalenichenko, J. Philbin, “Facenet: A unified embedding for face recognition and clustering,” Proc. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 815-823, 2015.
- [2] Fu, M. Fu, S. Sun, H. Gao, D. Wang, X. Tong, Q. Liu, Q. Liang, “Improving Person Re-identification using A Self-focusing Network in Internet of Things,” IEEE Internet of Things Journal, 2021.
- [3] Z. Cheng, X. Zhu, S. Gong, “Surveillance Face Recognition Challenge,” 2018.
- [4] S. Afra and R. Alhajj, “Early warning system: From face recognition by surveillance cameras to social media analysis to detecting suspicious people,” Physica A: Statistical Mechanics and its Applications, vol. 540, pp. 123151, 2020.
- [5] EU FRA, “Facial Recognition Technology: Fundamental Rights Considerations in the Context of Law Enforcement,” ed: Publications Office of the European Union Luxembourg, 2019.
- [6] K. Sarikakis and L. Winter, “Social Media Users’ Legal Consciousness About Privacy,” Social Media + Society,

- vol. 3, no. 1, , pp. 1-14, 2017.
- [7] I. Mutambik, J. Lee, A. Almuqrin, W. Halboob, T. Omar, A. Floos, "User concerns regarding information sharing on social networking sites: The user's perspective in the context of national culture," PLOS ONE, vol. 17, no. 1, pp. e0263157, 2022.
- [8] R. Piltaver, H. Gjoreski, and M. Gams, "Identifying a person with door-mounted accelerometer," Journal of Ambient Intelligence and Smart Environments, vol. 10, pp. 361-375, 2018.
- [9] A. Fukao, K. Futami, and K. Murao, "An Individual Recognition Method based on Door Opening and Closing Motion using a Gyroscope," Applied human informatics (AHI), vol. 3, no. 1, pp. 1-11, 2021.
- [10] A. Ishida, K. Murao, T. Terada, and M. Tsukamoto, "A user identification method based on features of opening/closing a refrigerator door," in 2017 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops), 2017, pp. 533-538: IEEE.
- [11] J. Vegas, C. Llamas, M. A. González, and C. Hernández, "Identifying users from the interaction with a door handle," Pervasive and Mobile Computing, vol. 70, p. 101293, 2020.
- [12] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A.N. Gomez, L. Kaiser, I. Polosukhin, "Attention is All You Need," Advances in neural information processing systems, vol. 30, 2017.
- [13] M. Ye, J. Shen, G. Lin, T. Xiang, L. Shao, and S. C. Hoi, "Deep learning for person re-identification: A survey and outlook," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 44, no. 6, pp. 2872-2893, 2021.
- [14] W. Shao, T. Nguyen, K. Qin, M. Youssef, and F. D. Salim, "Bledoorguard: a device-free person identification framework using bluetooth signals for door access," IEEE Internet of Things Journal, vol. 5, no. 6, pp. 5227-5239, 2018.
- [15] A. Bayat, M. Pomplun, and D. A. Tran, "A study on human activity recognition using accelerometer data from smartphones," Procedia Computer Science, vol. 34, pp. 450-457, 2014.
- [16] A. Ignatov, "Real-time human activity recognition from accelerometer data using Convolutional Neural Networks," Applied Soft Computing, vol. 62, pp. 915-922, 2018.
- [17] S.-M. Lee, S. M. Yoon, and H. Cho, "Human activity recognition from accelerometer data using Convolutional Neural Network," Proc. IEEE International Conference on Big Data and Smart Computing (bigcomp), pp. 131-134, 2017.
- [18] Y. Shavit and I. Klein, "Boosting Inertial-Based Human Activity Recognition With Transformers," IEEE Access, vol. 9, pp. 53540-53547, 2021.
- [19] M. M. Hassan, M. Z. Uddin, A. Mohamed, and A. Almo-
gren, "A robust human activity recognition system using smartphone sensors and deep learning," Future Generation Computer Systems, vol. 81, pp. 307-313, 2018.
- [20] A. Ohnishi, K. Murao, T. Terada, and M. Tsukamoto, "A method for structuring meeting logs using wearable sensors," Internet of Things, vol. 5, pp. 140-152, 2019.
- [21] Y. Zhang and Q. Yang, "A Survey on Multi-Task Learning," IEEE Transactions on Knowledge and Data Engineering, vol. 34, no. 12, pp. 5586-5609, 2022.

Stand-up Activity Prediction from Textile Pressure Sensors

Tahia Tazin^{†1}

Kyushu Institute of Technology

John Noel Victorino^{†2}

Kyushu Institute of Technology

Yu Enokibori^{†4}

Nagoya University

Sozo Inoue^{†5}

Kyushu Institute of Technology

1. Abstract

This paper presents a stand-up activity prediction system using textile pressure sensors to reduce unexpected falls. To achieve this, we equipped a wheelchair with textile pressure sensors to collect data on wheelchair users' activities, such as sitting, exercising, and trying to stand up. In a lab environment, we collected activity data from three users to forecast sudden attempts to stand up, which may lead to falls.

In this study, the effectiveness of the personalized model and the leave-one-subject-out (LOSO) method were compared. Also, we compared the performance with different feature sets. With four general statistical features, we achieved an 84% accuracy and an 81% F1-score to predict trying to stand-up activity in the one-person leave-out method. Then, the accuracy and F1-score increased to 86% and 83%, respectively with six additional features. Finally, personalized models produced the highest accuracy and F1-score of 91% and 94%, respectively for the stand-up activity.

Overall, the results demonstrate the potential of predicting wheelchair activities using textile pressure sensors, contributing to a lower risk of falling from the wheelchair.

2. Introduction

Along with longevity, labor shortages are becoming a severe problem in today's society, especially in nursing homes. Caregivers are not growing at the same rate as the aging population, making accurate monitoring difficult. It is not possible for a caregiver to monitor every move-

ment of the elderly all the time, especially when they are sitting in a wheelchair. Accidents are more likely to occur when they try to stand up from the wheelchair. Also, many times these accidents occur in a location where it is difficult for a nurse to monitor these issues at all times.

On the other hand, in addition to the increase in the elderly population in Japan, dementia patients are also increasing rapidly. Based on data analysis results from research by Kyushu University, it is estimated that by 2025, 1 in 5 adults over 65 will be affected by dementia [1]. Movement and balance-related brain regions can be impacted by dementia. Many people with Alzheimer's disease or another kind of dementia eventually lose the capacity to move and carry out daily duties. Also, patients with dementia often forget their current state. As a result, they can fall when they suddenly try to stand up from their sitting position. In nursing homes, people with dementia experience fall an average of 4.05 times a year compared to 2.33 times annually for other patients [2]. To prevent these accidents, it is vital to establish the proper provisions after completing assessments on the relevant patients. Using a textile pressure sensor-based prediction system that alerts caregivers when patients attempt to stand up from their wheelchairs is one solution to this issue.

Electronic textile structures have attracted much attention over the past 20 years in research and development due to their innate softness, breathability, and flexibility. These features produce a satisfying platform for sensing various inputs like pressure, temperature, and strain [3, 4, 5]. However, to support the daily lives of the elderly, professional life-logging technologies have also started to arise, and context-aware solutions with sensor devices are being introduced. Attaching several sensors to the body makes it possible to obtain much information about human movement and activity. Wearable technology should be incorporated into daily life use products,

Stand-up Activity Prediction from Textile Pressure Sensors

^{†1} Tahia Tazin, Kyushu Institute of Technology

^{†2} John Noel Victorino, Kyushu Institute of Technology

^{†4} Yu Enokibori, Nagoya University

^{†5} Sozo Inoue, Kyushu Institute of Technology

as attaching several devices may involve labor or additional expense. At the same time, having many sensors attached to the body can feel unsteady for elderly or dementia patients. As a result, we concentrated on wheelchair cushions since they are commonly used for elderly and dementia patients and can be used to monitor body pressure and postural position, allowing us to get many data on human behavior.

If activity were predicted from the pressure sensors built into cushions, it would be possible to naturally sense human behavior during movement. This capability would be helpful in many situations, not only in daily life but also in nursing homes or hospitals. Also, such cushion-type technology would become comfortable for wheelchair users who are not used to charging or discharging devices, so they might be employed in technologies used for dementia patients or the elderly.

Currently, there are many methods of using wearable sensors for fall detection; however, the sudden effort of the elderly to stand up has never been predicted. Also, the ideal sensor configuration for predicting states must be thoroughly researched. The rate of unexpected falls will drastically decrease if we can investigate the posture movement before standing up and alert the nurses via the alarm system.

This paper proposes a method for predicting body postural movements before standing up from a wheelchair using e-textile pressure sensors embedded in wheelchair cushions. We designed a supervised machine learning model to address the trying-to-stand-up activity prediction system utilizing textile pressure sensor data. In a lab setting, we started by collecting a series of raw pressure data. The instances of raw data were then aggregated, with each instance labeled with the activity that took place while the data was being collected. Then, using machine learning techniques, we created predictive models for stand-up activity prediction systems.

The rest of the paper is organized as follows: Section 3 discusses the related research, then Section 4 describes the details of the proposed method, Section 5 analyzes the results, and finally, Section 6 summarizes the entire paper.

3. Related Research

Nowadays, detecting the posture and daily activities of the elderly is a hot research topic worldwide. These

activities, or posture detection, are done with different sensors, such as accelerometers, gyroscopes, force, and pressure sensors. Because pressure sensor-based smart cushions are unobstructed, they have been used in various related studies to track the comfort and well-being, posture, fatigue, and physical activity of regular chair and wheelchair users [6, 7]. On the other hand, researchers have proposed many methods by which it is possible to detect the fall of the elderly. In this case, also, fall detection is done by various wearable sensors and cameras [8, 9, 10].

Xu et al. [11] proposed the smart textile-based sensor in the cushion to recognize human sitting postures. The binary representation of a gray-scale image was evaluated using data on the binary pressure distribution that had been collected. By employing dynamic temporal warping to analyze the pressure distributions, they could pinpoint specific postures of the sitting position. Also, Kamiya et al. [12] used an 8 x 8 pressure sensor matrix placed in a chair cushion to recognize the sitting postures. They used Support Vector Machine (SVM) with a radial basis function for classification.

On the other hand, Khan et al. proposed a camera-based technique for an abnormal human activity detection system [13]. They proved that this method could be used for elderly care using Kernel discriminant analysis and the Hidden Markov model. Similarly, Pavan et al. analyzed the video signals by applying manifold learning to identify human activities [14].

Sensor systems with early bed-leaving behavior predictions have been researched to prevent fall accidents. Asano et al. [15] proposed a depth camera-based detection system to avoid fall-related incidents. After optimizing parameters and motion variables with distinct participants' body size, location, and orientation, they used SVM to recognize bed-leaving behavior patterns. After 68 iterations, their experimentally produced result had a 92.65% recall rate. On the other side, the authors in [16] provided the bed-leaving behavior identification system. They employed bed pad sensors on a bed and a rail sensor placed inside a rail to detect bed leaving behavior. Additionally, they generated a benchmark dataset of ten subjects' continuous and discontinuous behavioral patterns. According to the experimentally collected results, Random forest achieved 91.1% accuracy in their benchmark dataset.

However, different bed-leaving sensors are now widely

available from manufacturers. For instance, clip sensors, infrared (IR), and mat sensors are often used in hospitals and nursing homes. The most practical method currently available, clip sensors, is simple to use, but because they are fastened directly to the patient's nightwear, care recipients are confined by sensor wires. Furthermore, a care recipient's neck could get wrapped around a sensor wire. Consequently, the usage of clip sensors has recently been discouraged. Due to the lack of constraints, mat sensors which are more affordable than clip sensors, are frequently utilized in clinical settings. One drawback of mat sensors is that they take a long time to detect and react when a care recipient places their feet on the mat while seated at the bed terminal.

As clip and mat sensors have some drawbacks, and wearable sensors are always difficult to use for older adults, we proposed a textile pressure sensor-based stand-up activity prediction system that can prevent any accidental event in a nursing home. Our proposed method has been discussed in the following section.

4. Proposed Method

The smart wheelchair contains an e-textile pressure sensor in the seat cushion and the backrest. These sensor data can be used to recognize postural movement activities such as sitting, exercising, and trying to stand up. The pressure sensors must be connected to a Raspberry Pi via Bluetooth to get the raw activity data. When the device is connected, it can even detect very slight postural movement. Figure 1 shows the overview of our smart wheelchair. In this section, we discuss our proposed system architecture.

4.1. System Architecture

As it is a textile pressure sensor-based smart wheelchair system, the pressure sensor can detect when someone is sitting in a wheelchair. After that, the pressure sensor starts collecting each postural movement data. So, from that data, it is easily possible to recognize the users' regular activities. This system cannot be able to detect when someone is standing up from the wheelchair and fall immediately. Moreover, many established methods are available for fall detection-related alarm systems in nursing homes or hospitals. Nevertheless, when patients already stand up from their wheelchairs may already be too late for caregivers to take necessary actions. Keeping



図 1 e-textile 壓力センサを搭載した車いす

Figure 1 Wheelchair with e-textile pressure sensors

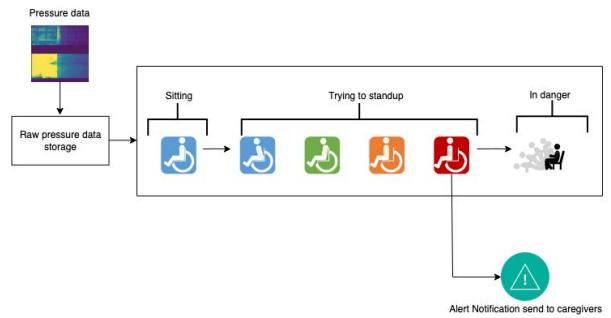


図 2 織物圧力センサを用いた転倒防止システムのアーキテクチャ提案

Figure 2 Proposed architecture of the textile pressure sensor-based fall prevention system

this in mind, we have proposed a system that can give an alarm when someone is trying to stand up from their sitting position.

Figure 2 shows our proposed system architecture. Seat and backrest pressure sensors gather pressure data when a person is seated in a wheelchair. An older person or a patient's posture frequently alters as they attempt to stand up from a wheelchair. Their backrest pressure sensor and seat pressure sensor both experience a gradual reduction in pressure as they lean forward. They consequently slowly adopt a dangerous position. Our system alerts the caregivers to take action since the patient could unexpectedly stand up from the wheelchair or fall while seated in a risky position.

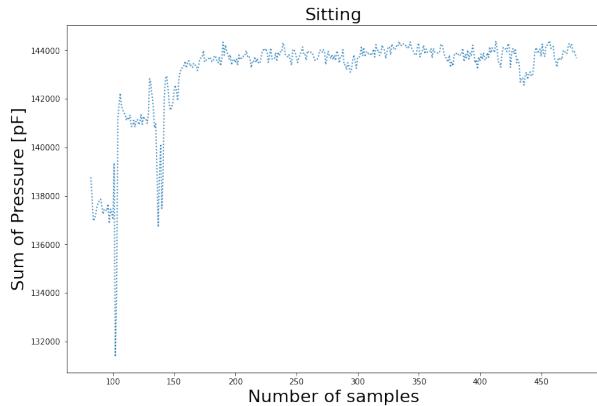


図3 座位動作の圧力センサデータ（単位：約ピコファラド、pF）

Figure3 Pressure sensor data for sitting activity (approximately in picofarad, pF)

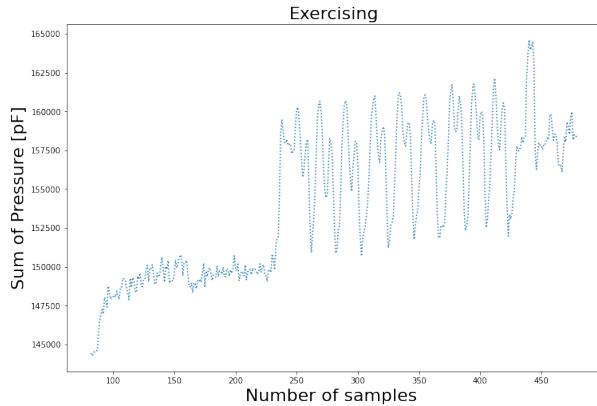


図4 運動時の圧力センサデータ（単位：約ピコファラド、pF）

Figure4 Pressure sensor data for exercising activity (approximately in picofarad, pF)

4.2. Experimental Setup

We collected a sequence of three activity data such as sitting, exercising, and trying to stand up using wheelchair pressure sensors. We analyze those data using a machine learning algorithm to predict the activity in different methodologies. The time duration for conducting each activity is 60 seconds. Data were collected at a constant rate of 8Hz from e-textile pressure sensors. For the first experiment, in the laboratory environment, we collected three subjects' data, and we considered 2-person data for training and different 1-person data for testing. Figures 3-5 show the sum of each activity's pressure data (approximately in picofarad, pF).

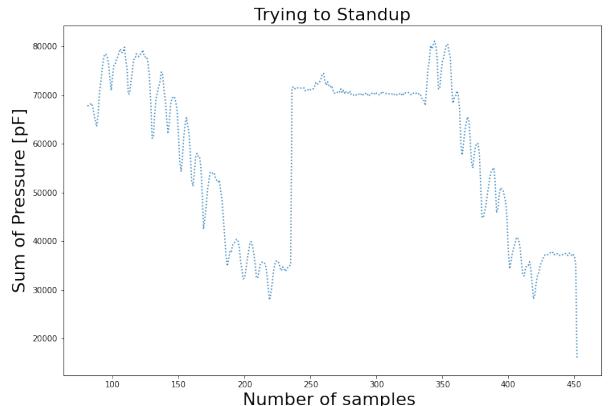


図5 立ち上がり動作に挑戦する圧力センサデータ（単位：約ピコファラド、pF）

Figure5 Pressure sensor data for trying to stand-up activity (approximately in picofarad, pF)

We obtained 16,000 columns of textile pressure sensor data for each activity in our experiments. Since the pressure sensor data is complicated, we did three experiments for better accuracy. We extracted mean, variance, minimum, and maximum features from the raw pressure sensor data for the first experiment. Using a 1-second window size, we set the features for 50% overlapping data. We used one traditional machine learning model K-Nearest Neighbor (KNN), for predicting the activity. In the second experiment, we extracted ten statistical features, i.e., mean, median, standard deviation (STD), minimum, maximum, median absolute deviation (MAD), interquartile range (IQR), Kurtosis, skewness, and variance. Then for the third experiment, we again collected data with one person on a different date. Moreover, we used the same subject data for training and testing, but they were collected on different dates.

5. Result Analysis

In the first experiment with four statistical features and one person leave out the method, we achieved 84% accuracy. For trying to stand up activity, we got 81% F1-score. Then we extracted six more statistical features, and with them, our model achieved 86% accuracy and 83% F1-score for trying to stand up activity prediction. In the personalized experiment, we achieved 91% accuracy and 94% F1-score for trying to stand-up activity prediction. The result of our experiment with trying to standup prediction demonstrates that adding more sta-

表1 マルチクラス分類の結果

Table 1 Multi-Class Classification Results for three experiments, showing Macro average and weighted average for precision (Prec.), recall (Rec.), and F1-Score (F1)

Activities	Experiment 1: Leave-one-out 4 features			Experiment 2: Leave-one-out 10 features			Experiment 3: Personalized models 10 features		
	Prec.	Rec.	F1	Prec.	Rec.	F1	Prec.	Rec.	F1
Sitting	0.95	0.93	0.94	0.96	0.93	0.95	0.86	0.93	0.89
Exercise	0.87	0.68	0.76	0.86	0.75	0.80	0.90	0.93	0.91
Trying to Stand-up	0.73	0.90	0.81	0.78	0.90	0.83	1.00	0.88	0.94
Accuracy	0.84			0.86			0.91		
Macro Average	0.85	0.84	0.84	0.87	0.86	0.86	0.92	0.91	0.91
Weighted Average	0.85	0.84	0.84	0.87	0.86	0.86	0.92	0.91	0.91

Prec.: Precision, Rec.: Recall, F1: F1-Score

tical features has increased the accuracy of our model. In particular, we observed that the accuracy of our model improved from 84% to 86% when utilizing a one-person leave-out method by increasing the number of statistical features from four to ten. Table one shows the detail prediction result. From experiment 1 and 2, we observed that with one person leave out method, exercise data are mostly misclassified with trying to stand up activity data. However, we achieved the highest accuracy with the personalized model because of less misclassification between exercise and trying to stand up data. Figure 6-8 shows confusion matrix for each experiment, respectively.

6. Conclusion

The smart technology described in this paper uses e-textile pressure sensors and machine learning algorithms to detect attempts to stand up by wheelchair users. This system's deployment attempts to reduce the possibility of unintentional falls happening in healthcare facilities, including hospitals and nursing homes. The personalized model showed the highest level of accuracy at 91% by combining a pressure matrix array configuration with the KNN algorithm. According to these findings, a wheelchair system with an e-textile pressure sensor is a helpful tool for preventing unexpected falls. The suggested technology is intended to anticipate attempts to stand up from a wheelchair, which can also carry a significant risk of injury.

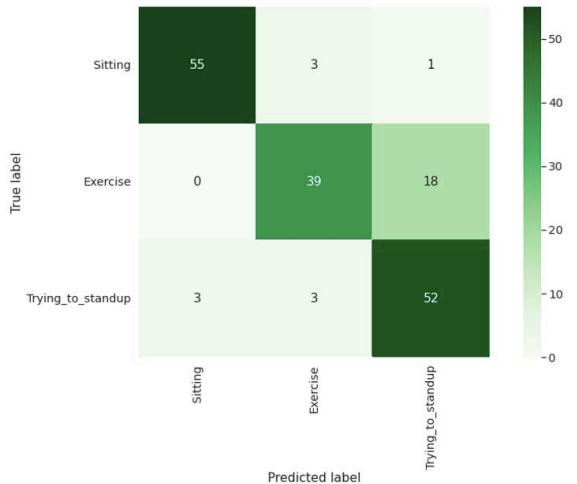


図6 4つの特微量の場合の混同行列

Figure 6 Confusion matrix with 4 statistical features

In summary, the safety and well-being of older adults and dementia patients at risk of falling can be significantly improved using e-textile pressure sensors in fall prevention systems. These sensors can recognize movement patterns and pressure variations, enabling early identification of falls and other abnormalities. These devices can assist in preventing falls and lowering the risk of injuries by providing real-time feedback and alarms to caregivers and healthcare professionals. E-textile pressure sensors have the potential to be an essential factor

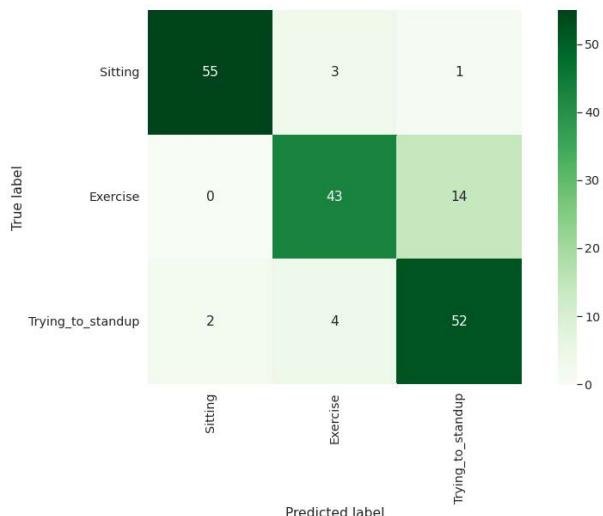


図 7 10 個の特徴量の場合の混同行列

Figure 7 Confusion matrix with 10 statistical features

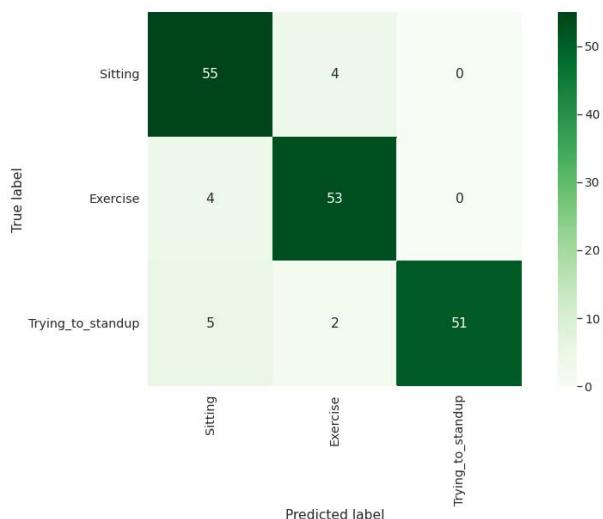


図 8 10 個の特徴量の場合の個人化モデルの混同行列

Figure 8 Confusion matrix for personalized model with 10 statistical features

in the ongoing endeavor to enhance fall prevention techniques and support healthy aging. In the future, we will collect real-field data from nursing homes to make this system more robust and valuable.

References

- [1] Ninomiya, T.: A study on future estimates of the elderly population with dementia in Japan, *Research report for 2014. Grant-in-aid for scientific research on health, labor and welfare. Special Research Project on Health, Labor and Welfare Science* (2015).
- [2] van Doorn, C., Gruber-Baldini, A., Zimmerman, S., Hebel, J., Port, C., Baumgarten, M., Quinn, C., Taler, G., May, C. and Magaziner, J.: Dementia as a Risk Factor for Falls and Fall Injuries Among Nursing Home Residents, *Journal of the American Geriatrics Society*, Vol. 51, pp. 1213–8 (online), 10.1046/j.1532-5415.2003.51404.x (2003).
- [3] Islam, G. M. N., Ali, M. and Collie, S.: Textile sensors for wearable applications: a comprehensive review, *Cellulose*, Vol. 27, pp. 1–29 (online), 10.1007/s10570-020-03215-5 (2020).
- [4] Liu, S., Ma, K., Yang, B., Li, H. and Tao, X.: Textile Electronics for VR/AR Applications, *Advanced Functional Materials*, Vol. 31 (online), 10.1002/adfm.202007254 (2021).
- [5] Atalay, O., Kalaoglu, F. and Bahadir, S.: Development of textile-based transmission lines using conductive yarns and ultrasonic welding technology for e-textile applications, *Journal of Engineered Fibers and Fabrics*, Vol. 14, p. 155892501985660 (online), 10.1177/1558925019856603 (2019).
- [6] Arnrich, B., Setz, C., Marca, R., Tröster, G. and Ehlert, U.: What Does Your Chair Know About Your Stress Level?, *IEEE Transactions on Information Technology in Biomedicine*, Vol. 14, pp. 207–214 (online), 10.1109/TITB.2009.2035498 (2010).
- [7] Congcong Ma, Wenfeng Li, J. C. S. W. L. W.: A Fatigue Detect System Based on Activity Recognition, pp. 303–311 (2014).
- [8] Chen, T., Ding, Z. and Li, B.: Elderly Fall Detection Based on Improved YOLOv5s Network, *IEEE Access*, Vol. PP, pp. 1–1 (online), 10.1109/ACCESS.2022.3202293 (2022).
- [9] Vadivelu, S., Ganesan, S., Murthy, O. and Dhall, A.: Thermal Imaging Based Elderly Fall Detection, pp. 541–553 (online), 10.1007/978-3-319-54526-4_40 (2017).
- [10] Wang, X., Gao, C. and Guo, Y.: Elderly fall detection using SIFT hybrid features, p. 96751W (online), 10.1117/12.2199683 (2015).
- [11] Xu, W., Huang, M.-C., Amini, N., He, L. and Sarrafzadeh, M.: eCushion: A Textile Pressure Sensor Array Design and Calibration for Sitting Posture Analysis, *Sensors Journal, IEEE*, Vol. 13, pp. 3926–3934 (online), 10.1109/JSEN.2013.2259589 (2013).

- [12] Kamiya, K., Kudo, M., Nonaka, H. and Toyama, J.: Sitting Posture Analysis by Pressure Sensors, pp. 1–4 (online), 10.1109/ICPR.2008.4761863 (2008).
- [13] Khan, Z. and Sohn, W.: Abnormal Human Activity Recognition System Based on R-Transform and Kernel Discriminant Technique for Elderly Home Care, *IEEE Transactions on Consumer Electronics - IEEE TRANS CONSUM ELECTRON*, Vol. 57, pp. 1843–1850 (online), 10.1109/TCE.2011.6131162 (2011).
- [14] Turaga, P., Veeraraghavan, A., Srivastava, A. and Chellappa, R.: Statistical Computations on Grassmann and Stiefel Manifolds for Image and Video-Based Recognition, *IEEE transactions on pattern analysis and machine intelligence*, Vol. 33, pp. 2273–86 (online), 10.1109/TPAMI.2011.52 (2011).
- [15] Asano, H., Suzuki, T., Okamoto, J., Muragaki, Y. and Iseki, H.: Bed Exit Detection Using Depth Image Sensor, *J. TWMU*, Vol. 84, pp. 45–53 (2014).
- [16] Madokoro, H., Nakasho, K., Shimoji, N., Woo, H. and Sato, K.: Development of Invisible Sensors and a Machine-Learning-Based Recognition System Used for Early Prediction of Discontinuous Bed-Leaving Behavior Patterns, *Sensors*, Vol. 20, p. 1415 (online), 10.3390/s20051415 (2020).

イスの揺れからの人物と活動の識別

望月康平^{†1} 細川健人^{†2} 豊浦正広^{†3}
山梨大学 山梨大学 山梨大学

1. はじめに

人間の行動をセンシングすることでデータ化するときに、得られたデータを人間が理解できる形に解釈することによって、人間の活動を支援する材料に昇華することができるケースがある。IoT センサの性能の進化とコモディティ化によって、多種多様なセンサを利用できるようになってきたものの、人間の行動への解釈の側面においては、まだ不十分な点も多い。

我々は、人間の活動の手がかりになる振動を観測できるイスに着目する。人間の活動の推定はカメラ映像の解析によって活発に行われているものの、カメラ映像にはプライバシーに関わる余計な情報まで取得してしまう問題がある。この点において、人間が触れたときにだけ人間の活動を観測できるイスは有望である。EU で 2024 年の施行が計画されている AI 規制法案では、監視カメラの利用を制約する条項を含み、公権力によるむやみなカメラ映像の取得が制限される。日本でも、個人情報保護法の規制によって、監視カメラの映像を無加工で保存したり転送したりすることは、問題になることがある。これらは、そもそも人間が撮影されることに忌避感を覚えることが発端であり、可能であるならばこの状況は避けるべきである。この点において、振動センサの信号に対しては、現在のところ規制する法律はなく、また、人間が感じる忌避感も映像に比べると小さい。さらに、イスに取り付けた振動センサであれば、人間が触れていないときにはデータが発生せず、余計な情報を取得しなくて済む利点がある。

本研究では、イスに取り付けた 9 軸センサから得られる信号により、活動に加えて人物も識別することを目指す。人物の識別では、特定の個人であることを識別するものではなく、別のデータと同一の人物から発生したものであることを識別するものとし、データ間や活動間の紐づけを行うために利用できる。

本研究の貢献は以下の通りである。

1. Transformer を含む深層学習ネットワークによって、イ

スに取り付けた 9 軸センサの信号から、人物と活動の両方を推定する手法を提案する。

2. イスに座る人物を撮影し、同時にセンサの信号を収録することによって、センサ信号と人物および活動が紐づけられたデータセットを構築する。
3. 着席中の信号によって活動を認識するのとは別に、離着席時の信号によって人物の認識を高い精度で行えることを示す。

以下、2. では関連研究を紹介する。3. では着席中に観測される 9 軸センサの信号からからイスに座った人の活動認識する手法を提案する。4. ではさらに、離着時の大きな信号に着目することで、人物を高い精度で認識する手法を提案し、実験結果と考察を加え、5. で本研究をまとめる。

2. 関連研究

Mizumoto ら [1] は、8 つの加速度センサを内蔵したイスの設計を行うことで、業務を妨げることなく人物の姿勢を認識する手法を提案した。この手法により、座っている人の姿勢の検出と共に、在席時間や場所から着席時の姿勢の改善に活用できる。しかし、コストの面から必要なセンサの数はより少ないことが望ましい。比較的安価なセンサを使用した研究として、Yoo ら [2] らによる、近接センサを使用した研究がある。近接センサは感応式のセンサであり、スマートフォンやパソコンといったタッチパネルに利用される。Yoo らは、背もたれに取り付けられた近接センサから座姿勢を推定し、ヘッドレストに取り付けられたスピーカーから警告を出すイスを提案した。提案手法は、通常のイスと比較して、視覚ディスプレイによる作業中の人物の姿勢を改善する効果があることが示された。また、Sztyler ら [3] は、加速度計、ジャイロ、および地磁気のセンサーデータを使用し、ウェアラブルデバイスから位置情報を含む人の行動を認識する新しいモデルを提案した。センサによる行動認識は高い精度で可能であることが示され、特に腰部に取り付けた場合に高い効果が得られるという結果が示された。

深層学習を人物の行動認識に利用した研究に、Chen ら [4] や Wang ら [5] の研究がある。Chen ら [4] は、単一の加速度センサから行動を表す特徴を抽出し、ニューラルネットワー

User and Action Identification from Chair Viblation

^{†1} Kouhei MOCHIZUKI, University of Yamanashi
^{†2} Taketo HOSOKAWA, University of Yamanashi
^{†3} Masahiro TOYOURA, University of Yamanashi

クを訓練することで行動の分類を行った。実験から、従来手法に比べて高い精度で人物の行動の認識が可能となったことが示された。Wang ら [5] は、ウェアラブルセンサから得られたデータを使用して、深層学習モデルである LSTM を訓練させることで行動識別を行った。HAR データセットを使用した実験では、高い精度での行動の識別が可能であることが示された。Chen らや Wang らの研究成果により、スマートフォンやウェアラブルデバイス等のモバイルデバイスへ拡張の幅が広がった。画像処理の分野で Hu ら [6] は、航空写真のマルチスペクトル画像から空間的な特徴を抽出するために、Convolutional Neural Network(ConvNet) と LSTM を結合したモデルである ConvLSTM を提案し、より高度な土地利用状況の認識を可能とした。

Shinmoto ら [7] は、対象者の腹部に大きな RFID タグを貼り付けることで、バッテリーを装着することなく、室内の対象者の行動を認識する装置を提案した。壁に配置した複数の RFID リーダによって、各時刻での RFID の位置と姿勢を推定することができるようになった。室内にいさえすれば位置や姿勢がおおまかにわかることは望ましい性質ではあるが、腹部に取り付けた RFID の物理的な大きさは無視することができず、利用できない場合も考えられる。イスに取り付ける装置では、対象者にバッテリーの管理をさせる必要はなくなり、明るい室内であればソーラーバッテリなどでも駆動することが可能である。

Ishac ら [8] は、イスに取り付けた 9 つの圧力センサによって、座っている人の前傾や後傾などの姿勢を推定し、健康に悪い姿勢であった場合に警告を出すシステムを提案した。人によっては背もたれから離れて座ることもあり、その場合には正しく姿勢が推定できなかった。本研究では IMU センサを用いるために、イスのどこかに体が接していないければいけないといったような制約を置く必要はなくなる。

Enokibori ら [9] は、胴体の姿勢を推定するために導電性繊維を織り込んだベストを作り、電気容量の変化によって姿勢を推定した。姿勢の変化によっては活動の推定をすることもできるものの、ベストを常に身に着けるのは活動の記録に対する費用対効果が小さい。

スマートフォンに内蔵されたセンサを用いて活動を認識する手法は多く提案されている [10]。スマートフォンを持つ位置に頑健に活動を推定する方法は提案されているものの、スマートフォンを常に身に着ける必要があり、介護対象者や就学児などには適用が難しい。イスに取り付けたセンサであれば、イスに座ったときの信号を取り逃すことはない。

また、人物や行動を識別する手法として、映像を使用する手法も活発に提案されている [11][12]。Bremond ら [11] は、カメラ映像を使用した人物の行動認識として、映像の各フレームを静止画像として特徴抽出を行い、その後分類

されたフレームをつなぎ合わせることで、一連の行動として認識を行った。人物の行動を自動的に認識する Bremond らの手法は、自動運転の安全性の向上などに利用されている。しかし、カメラ映像を使う場合には、プライバシ保護の問題が起こる。

以上から、本研究では、単一のセンサをイスに取り付けて使用することで、簡易で安価な、プライバシに配慮した人物と活動の識別を目指す。

3. 着席中の信号からの人物と活動の識別

3.1. 識別ネットワーク

Transformer は、Vaswani ら [13] によって発表された自然言語処理や時系列に有効とされるモデルである。時系列信号であるセンサデータを効率的に扱うため、本研究では Transformer を採用した。Transformer は意味抽出を行う Encoder 部と、抽出した意味から新たな出力へと変換する Decoder 部から構成される。特に Encoder 部の中で、複数の異なる Attention 機構を用いて入力データの前後関係の情報を取り出す Multi-Head Attention 機構が、意味抽出において大きな役割を果たしている。そこで本研究では、時系列の前後関係を抽出するために Multi-Head Attention 機構を含む Transformer の Encoder 部を使用した。Encoder 部を含むネットワーク全体は、図 1 のように構成される。入力は 3 軸加速度・3 軸角速度・3 軸地磁気の 9 次元の時系列である。図 1 の右図は Transformer の Encoder 部の詳細である。Multi-Head Attention と Feed-Forward 層に、それぞれ残差接続と正規化処理が行われる。Encoder 部は、多層パーセプトロン、Dense 層へと接続される。本モデルは、人物と活動を同時に学習するマルチタスク学習を行った。そのため、出力は、人物と活動それぞれの Softmax 関数から得られる識別クラスの尤度である。

3.2. 着席中の行動のセンシング

着席中のイスのセンシングによる人物と活動の識別を行った。まず、イスを利用したデータ収集のために、着席して複数の活動を行ってもらう実験を行った。実験の様子を図 2 に示す。イスの座面裏に 3 軸加速度・3 軸角速度・3 軸地磁気を計測する 9 軸 IMU センサを設置した。イスの座面裏に取り付けることで、活動の妨げにならず、センサの存在を意識せずに使用することができる。センサの方向関係は、X 軸が正面方向、Y 軸が左右方向、Z 軸が上下方向とした。本実験では、9 軸 IMU センサに ATR-Promotions 製の TSND121 を使用した。同製品は、サンプリング周波数が 50Hz であり、専用のソフトウェアである SensorController を介して信号を取得する。また、9 軸センサと合わせてウェ

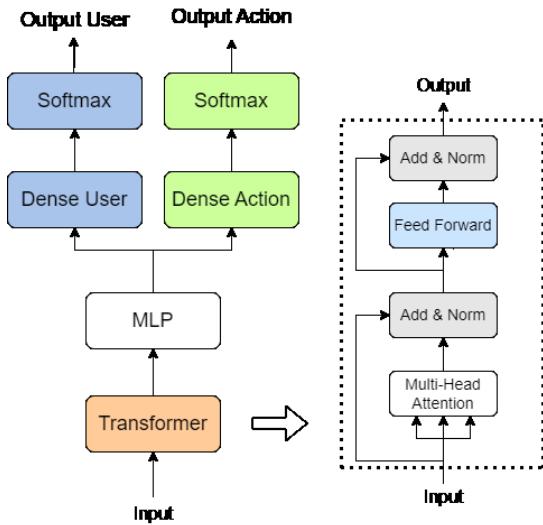


図 1 Transformer を含む識別ネットワーク

Figure 1 Proposed identification network including Transformer



図 2 実験 1: 活動計測実験の様子

Figure 2 Experiment 1: Activity Measurement

ブカメラによる撮影を行った。録画映像は、DNN の訓練時の正解データの作成に使用した。

3.3. データセット作成

男性 14 人、女性 8 人を対象とした実験を行った。実験協力者には、3.2 の実験環境下で、表 1 のスケジュールで決まった手順で活動を行ってもらった。活動は、座る・立つといった着席前後の 2 種類の活動、および、PC 作業・ボールペンによる書き作業・静止といった着席時の 3 種類の活動の全 5 種類である。計測は、各人ごと 4 セットを行い、各セットを学習・検証・テスト用に分割して使用した。また、録画映像から人物と活動のタグ付けを行った。図 3 は、9 軸センサから得られた 1 セット分のイスの回転方向の角速度の例である。座る・立つといった離着席動作と比較して、PC 作業・静止・書き作業はパワーが小さいことが見て取れる。本研究では、着席中の活動に対して識別を行うため、PC 作業・書き作業・静止の 3 種類の活動の信号のみを使用した。映像によるタグ付け後、3.1 の Transformer を含むネットワークを訓練させ、人物と活動の識別を行った。

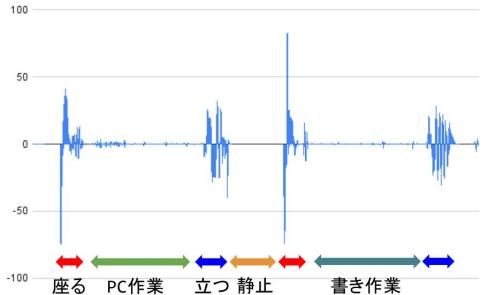


図 3 実験 1: 1 セット分の信号例

Figure 3 An example of Signal in Experiment 1

表 1 実験 1: 実験手順

Table 1 Overview of Experiment 1

経過時間	活動
0:00	計測開始
0:10	座る
0:20	PC 作業開始
1:20	PC 作業終了
1:30	立つ
1:40	座る
1:50	書き作業開始
2:50	書き作業終了
3:00	立つ
3:20	座る
3:30	計測終了

3.4. 実験結果

訓練したネットワークによって、22 人の人物を 0.1565、3 種類の活動を 0.8142 の正解率で識別することができた。時系列に有効とされる Transformer を深層学習ネットワークによる識別は、着席中の信号による数種類の活動に対して有効であると示された。一方で、識別するクラス数の多い人物の識別は困難であるという結果となった。これは、着席中のパワーの小さい信号では、人物のような多クラスの識別の場合、クラスごとの特徴がデータの差として現れづらいことが原因であると推察される。

4. 離着席動作による人物識別

着席中の信号を用いた人物識別の正解率の低いことを受けて、手法の改善を行った。3.3 で取得したデータから、着席中のパワーの小さな信号より、立つ・座るといった、離着席時の大きな信号の方が人物の特徴が表れているのではないかと考えた。そこで、着席中の活動の小さな信号ではなく、イスを離着席する動作の大きな信号に注目して識別を

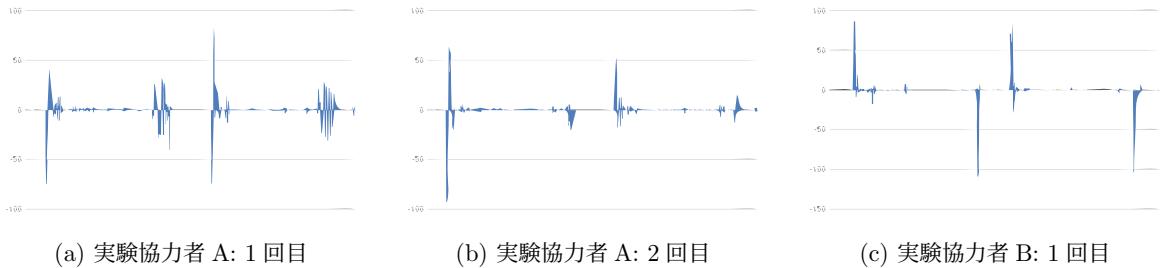


図 4 実験協力者の離着席時の計測データの観察 (イスの回転方向の角速度)

Figure 4 Observed sensor data of participants from chairs (Angular velocity in the direction of chair rotation).

行った。

4.1. データ分析

3.3 で使用したデータを分析することで、離着席動作が人物識別に有効であるか検証した。図 4(a), (b) から、実験協力者 A では 2 回の計測で着席時には大きな信号が見られ、離席時は比較的小さな信号が共通して見て取れる。また、図 4(c) と比較して、実験協力者 B では離着席時に正負の片方に対して大きな信号が見られる。また、信号の大きさについても、実験協力者 A は離着席動作で差が見られるが、B は離着席共に平均 $\pm 100[\text{rad}/\text{s}]$ 程度であるという違いも見られる。この違いは、離着席時にイスをどの程度回転するかという、使用者ごとの特徴に起因すると考えられる。着席中の信号と比較して、離着席動作には、このような人物ごとの特徴が現れやすいことから、離着席動作による人物識別は可能であると判断した。

4.2. 離着席動作のセンシング

離着席動作におけるイスのセンシングによる人物とイスの識別を行った。まず、データ収集のために、複数種類のイスで離着席動作を行ってもらう実験を行った。実験の様子を図 5 に示す。本実験では、9 軸 IMU センサ WitMotion WT901SDCL を使用した。WT901SDCL は、200Hz までの 3 軸加速度・3 軸角速度・3 軸地磁気を計測することができる。本実験では、信号の周波数はデバイスの上限である 200Hz とした。同製品は、SD カードを挿入することで計測を開始し、SD カードを取り出すことで計測を終了する。計測した信号は、SD カードにバイナリデータとして記録され、WitMotion の設定を行うソフトウェア (MiniIMU) 等により、各時刻ごとの 9 軸の信号を確認することができる。センサは、3.2 と同様に、イスの座面裏に取り付けた。

4.3. 前処理

図 6 は、1 回の離着席動作を行ったイスの上下方向の加速度の信号例である。最初の着席時は負の方向に、その後の



図 5 実験 2: 離着席動作

Figure 5 Experiment2: Seating and leaving actions

離席時は正の方向に大きな信号が得られていることがわかる。この様な信号に対して、前処理として、まず欠損値の除去を行った。また、センサから得られる信号には、少なからずノイズが含まれる。そこで、ノイズによる影響を低減するために、バターワースフィルタによる平滑化を行った。また、加速度・角速度・地磁気はそれぞれ取り得るデータの範囲が異なるので、データの標準化を行った。その後、離着席動作を行う部分に対して人物とイスのタグ付けを行った。タグ付けは、9 軸の信号の内、イスの上下方向である加速度 Z 軸と、回転方向である角速度 Z 軸に特徴が含まれていると考え、その 2 つの二乗和が閾値を超えたときにイスを使用しているとして判定した。

4.4. 計測実験

男性 22 人、女性 2 人を対象とした実験を行った。実験協力者には、静止したイスに着席し、数秒静止した後、イスから離席するという動作を 20 回行ってもらった。実験協力者には、着席時にイスの正面の机で作業を行う想定で座ることと、離席時にイスの向きを正面に戻すことの 2 点を指示した。実験協力者には、規定回数の離着席動作後に終了であることを伝えた。また、実験の疲れによるデータへの影響を無くすため、各イスが終了するごとに休憩時間を設けた。計測した 20 回分のデータは、4.3 で述べた前処理を行

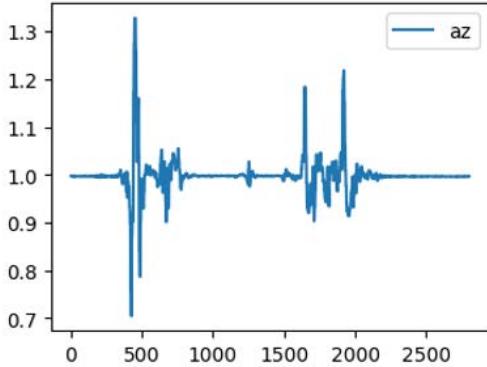


図 6 1回分の離着席動作の信号

Figure 6 Signal of a pair of seating and leaving actions



図 7 実験 2 で使用した 3 種類のイス

Figure 7 Three chairs used in Experiment 2

い、図 6 の様に 1 回分ごとに分割して使用した。

図 7 は実験に使用した 3 種類のイスである。イス 1 と 2 は異なる商品だが、イス 1 と 3 は同じ商品である。1 と 3 で同じ商品を使った理由は、同じ商品でも、スプリングやキャスターの動かしやすさ等の違いから識別できると考えたからである。また、システムを導入することを想定した時、教室や会社では同じイスを大量に使用することが考えられるため、今回イス 1 と 3 は同じ商品を使用した。

4.5. 区間内投票による識別

本実験では、時系列を扱うことから、3.1 で述べた Transformer を含むネットワークを使用した。出力は、人物とイスのマルチタスク学習により、各クラスの Softmax 関数の出力が得られる。また、新たに、各クラスの出力から、指定した区間内における投票を行った。出力では、各時刻ごとのクラスの識別結果の尤度が得られるが、本研究では、時系列データでは前後関係が重要であることから、指定した区間内で各クラスの尤度の平均を取り、区間内で最も尤度が高いクラスを最終的な出力として採用した。例えば、20 人の人物識別の場合、9 次元のデータの入力に対し、出力はデータごとの 20 人の Softmax 関数による尤度が得られる。

区間を 512 とすると、前後 512 個の 20 人分の尤度から、各人物ごとの尤度の平均を求め、区間内における尤度の平均が最も高いクラスを、その区間内全てのデータにおける識別結果とした。

4.6. 実験結果

取得した 24 人、3 種類のイス、20 回分のデータを、3.1 で述べた Transformer を含むネットワークで訓練させた。また、マルチタスク学習により、人物とイスの同時識別を行い、4.5 の処理により最終的な出力を得た。訓練したネットワークによって、24 人を 0.9768、3 種類のイスを 0.9702 の正解率で識別することができた。結果から、3.4 の考察の通り、着席中のパワーの小さな信号と比較して、離着席動作のパワーの大きな信号には、人物ごとの特徴が多く現れ、大人数の識別に対して有効であることが示された。また、イス 1・3 から、同じ製品のイスの識別も可能となった。

5. 結論

イスを利用する人物と活動を同時に識別することを目指した。まず、着席中の活動の信号により、人物と活動の識別を行った。識別には、時系列に有効なモデルである Transformer を採用した。実験では、22 人の 3 種類の活動に対し、22 人の人物を 0.1585、3 種類の活動を 0.8142 の正解率で識別できた。さらに、人物の正解率が低いことを受けて、手法を改善した。着席中の小さな信号からは人物の識別は難しいと考え、離着席動作の大きな信号から人物の識別を行った。実験では、24 人の 3 種類のイスを 20 回離着席したデータを、Transformer を含むネットワークで訓練させた。結果、24 人の人物を 0.9768、3 種類の活動を 0.9796 の正解率で識別できた。

以上から、離着席動作と着席中の信号で 2 段階の識別を行うことで、イスを利用する人物と活動を同時に識別することが可能となった。

今後の課題として、活動やイスの種類を増やした場合の正解率への影響を調べることを挙げる。イスの種類については、3 種類でしか検証を行っておらず、提案手法がどのような種類のイスに対しても有効であるかは確かめられていない。特に、ガスチェア以外の回転しないイスに対して、どれほど提案手法が有効に働くかについては、興味がある。また、活動の種類についても、3 種類のみを識別するにとどまっている。これまでのデータの観測からは、静止時と会話時にはほとんど信号に差がないことが確かめられており、活動によっては識別が難しいものがあると考えている。実際の利用で必要とされる活動の種類を踏まえながら、識別ネットワークの精度を高め、手法の限界を検証したい。

謝辞

本研究は JST さきがけ JPMJPR213 の支援を受けた。

参考文献

- [1] Mizumoto, T., Otoda, Y., Nakajima, C., Kohana, M., Uenishi, M., Yasumoto, K. and Arakawa, Y.: Design and implementation of sensor-embedded chair for continuous sitting posture recognition, *IEICE Transactions on Information and Systems*, Vol. 103, No. 5, pp. 1067–1077 (2020).
- [2] Yoo, W.-g., Yi, C.-h. and Kim, M.-h.: Effects of a proximity-sensing feedback chair on head, shoulder, and trunk postures when working at a visual display terminal, *Journal of Occupational Rehabilitation*, Vol. 16, pp. 631–637 (2006).
- [3] Szytler, T., Stuckenschmidt, H. and Petrich, W.: Position-aware activity recognition with wearable devices, *Pervasive and mobile computing*, Vol. 38, pp. 281–295 (2017).
- [4] Chen, Y. and Xue, Y.: A Deep Learning Approach to Human Activity Recognition Based on Single Accelerometer, *IEEE International Conference on Systems, Man, and Cybernetics*, pp. 1488–1492 (2015).
- [5] Wang, L. and Liu, R.: Human activity recognition based on wearable sensor using hierarchical deep LSTM networks, *Circuits, Systems, and Signal Processing*, Vol. 39, pp. 837–856 (2020).
- [6] Hu, W.-S., Li, H.-C., Pan, L., Li, W., Tao, R. and Du, Q.: Spatial-spectral feature extraction via deep ConvLSTM neural networks for hyperspectral image classification, *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 58, No. 6, pp. 4237–4250 (2020).
- [7] Shinmoto Torres, R. L., Visvanathan, R., Abbott, D., Hill, K. D. and Ranasinghe, D. C.: A battery-less and wireless wearable sensor system for identifying bed and chair exits in a pilot trial in hospitalized older people, *PLOS ONE*, Vol. 12, No. 10, p. e0185670 (online), 10.1371/journal.pone.0185670 (2017).
- [8] Ishac, K. and Suzuki, K.: Lifechair: A conductive fabric sensor-based smart cushion for actively shaping sitting posture, *Sensors*, Vol. 18, No. 7, p. 2261 (2018).
- [9] Enokibori, Y., Ito, Y., Ikeda, K., Suzuki, A., Shimakami, Y., Kawabe, T. and Mase, K.: Respiratory Volume Estimation by a Stretchable Textile Sensor, *Advances in Science and Technology*, Vol. 80, Trans Tech Publ, pp. 136–141 (2013).
- [10] Wang, A., Chen, G., Yang, J., Zhao, S. and Chang, C.-Y.: A comparative study on human activity recognition using inertial sensors in a smartphone, *IEEE Sensors Journal*, Vol. 16, No. 11, pp. 4566–4578 (2016).
- [11] Brémond, F., Thonnat, M. and Zúñiga, M.: Video-understanding framework for automatic behavior recognition, *Behavior Research Methods*, Vol. 38, No. 3, pp. 416–426 (2006).
- [12] Ye, M., Shen, J., Lin, G., Xiang, T., Shao, L. and Hoi, S. C.: Deep learning for person re-identification: A survey and outlook, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 44, No. 6, pp. 2872–2893 (2021).
- [13] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L. and Polosukhin, I.: Attention is all you need, *Advances in neural information processing systems*, Vol. 30 (2017).

顔表情解析による自然な笑いと作り笑いの識別

中嶋桃香 ^{†1}

慶應義塾大学環境情報学部

浜中智己, 佐々木航 ^{†2}

慶應義塾大学大学院政策・メディア研究科

大越匡 ^{†3}

慶應義塾大学環境情報学部

中澤仁 ^{†4}

慶應義塾大学環境情報学部

概要

表情と感情の結びつきに着目した先行研究によると、基本 6 感情を表す普遍的な表情があると明らかになっている。しかし、現実世界では表情と感情は必ずしも一対一の関係ではない。例えば喜びのようにポジティブな感情を表す笑顔と、ネガティブな感情を表す作り笑いがある。特に作り笑いはうつ病の症状として見られることがある。そのため、心療内科では笑顔状態で元気な印象であっても注意深く診療を行っている。そこで、笑顔表出時の感情を推定することを目的として、顔表情データから機械学習モデルを構築することによって自然な笑いと作り笑いの識別をする手法を提案した。その結果、最大で 97.0% の識別精度を達成した。

1. はじめに

人間はコミュニケーションを取る際に、表情や声色といった情報から相手の感情や内面状態を推定している。様々な種類の表情の中で、特に笑顔は対人関係を円滑に進め、日常生活で頻繁に見られる幸福の表象とされている [1]。感情と表情の結びつきに関する研究の第一人者として知られる Ekman らの研究によると、図 1 のように基本的な 6 つの感情(怒り、嫌悪、恐怖、喜び、悲しみ、驚き)を表す普遍的な表情があることが示されている [2]。

李・渋谷 [4] らの研究では、社会的笑いに関する心理学研究を笑いの分類、社会的笑いの機能、自然な笑いと作り笑いの差異、笑いの発達、笑いの識別、笑いの 6 つの観点から概観した。この研究では、快感情を伴わない笑いを社会的な笑いとして述べている。例えば、眞の感情を隠蔽、擬装



図 1 Ekman らが提唱した基本 6 感情に結びつく普遍的な表情 [3]

するための笑い、面白く感じていないのに空笑いする擬態、悲しみや恐怖の感情を他人に悟られないようにするための笑いである。社会的笑いについて、コミュニケーションを円滑に進めるために笑いがどのような感情の下で生じたのかを理解することが重要だと述べている。

そして、心理臨床場面において快感情を伴わない笑い、内面と表出のずれがうつ病患者の特徴として研究されている。うつ病の中でも特に若い女性に多いのが「ほほえみうつ病」「仮面うつ病」だと言われている [4]。この病気の特徴は周囲の人々に笑顔や明るい表情を出して周りから悟られないように振る舞うことである。また、他人の視線を意識しやすい傾向や自己の姿勢や行動への意識は心理学の研究では公的自己意識と呼ばれているが、この公的自己意識が高いほど作り笑いをしやすいという研究もされている [5]。

このように、作り笑いが注目される中で、自然な笑いと作り笑いを識別することを目的とした研究が行われている。しかし、既存研究ではセンサーが付属した眼鏡型のデバイスを用いて表情の動きをトラッキングしているため実用性を伴わず、日常生活に応用するのは難しい [6]。そこで、本研究は PC やスマートフォンのカメラで撮影される顔表情から Action Unit ごとに表情の動きの強度を特徴量として、機械学習によって自然な笑いと作り笑いを識別する手法の提案

Facial Expression Analysis to Discriminate Between Natural and Fake Smiles

^{†1} MOMOKA NAKAJIMA, Faculty of Environment and Information Studies, Keio University

^{†2} SATOKI HAMANAKA, WATARU SASAKI, Graduate School of Media and Governance, Keio University

^{†3} TADASHI OKOSHI, Faculty of Environment and Information Studies, Keio University

^{†4} JIN NAKAZAWA, Faculty of Environment and Information Studies, Keio University

をする。本研究で行った顔表情のデータ収集実験において顔表情データを収集し、Openface を用いて顔表情を Action Unit に分割した。その Action Unit の動きの強度を特徴量として笑いを識別するモデルの学習を行った。AUC を用いたモデルの性能評価において 95.7% を達成した。

本論文は以下のように構成される。2 章では自然な笑いと作り笑いの差異、識別に関する関連研究を提示する。3 章では自然な笑いと作り笑いを識別するための詳細なアプローチについて述べる。4 章では機械学習モデルを構築するためのデータセットを収集する実験について述べる。5 章ではデータ収集実験の結果や実験内で実施したアンケート結果から性格と作り笑いの関連性について述べる。6 章では自然な笑いと作り笑いを識別するモデルを学習させるためのデータセット構築や精度評価について述べる。8 章では結果を元にした結論と考察を述べる。そして 7 章で考察とともに今後の展望を述べる。

2. 関連研究

2.1. 自然な笑いと作り笑いの差異

フランスの神経学者デュシェンヌ (Duchenne) は、ポジティブな情動から表出される笑顔を「真の笑顔」、ポジティブな情動とは関係なく表出される笑顔を「偽の笑顔」として、2 つの笑顔を解剖学的特徴から違いを明らかにした [7]。2 つの笑顔の一例を図 2 に表す。



図 2 左が真の笑顔、右が偽の笑顔を表す [8]

真の笑顔はデュシェンヌスマイルとしても知られており、大頬骨筋による頬の上昇と眼窩部眼輪筋による口角の上昇が特徴として挙げられる。菅原の研究 [8] では、笑顔は目元と口元が魅力の要因であり、知覚において他の顔要素よりも好印象を与える際に有意に働くということを述べている。そして、目元と口元のどちらか片方の動きだけでは不十分であり、目と口の両方にデュシェンヌスマイルの特徴が見

られてこそ好印象を与えるのである。また、人種や民族の違いによって顔のパーツは個人によって大きく異なるが、喜怒哀楽などの感情を表情から認識する際には感情の共通認識がある程度得られることも述べている。

2.2. 自然な笑いと作り笑いの識別

齊藤、正井、杉本らの研究 [6] は反射型光センサアレイを装着した眼鏡型のデバイスから得られるセンサーデータで自然な笑いと作り笑いを識別する研究を行っている。表情筋の動きによって皮膚変形が生じると、図 3 のような眼鏡型デバイスに取り付けられた 16 個の反射型光センサーと皮膚表面の間の距離に変化が生じる。その距離によって光の反射強度が変わるために、センサーデータから表情を識別する設計である。

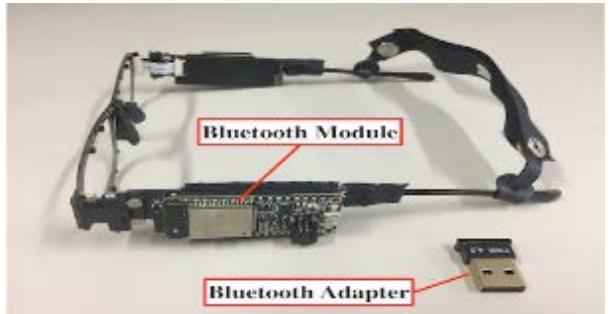


図 3 反射型光センサアレイを装着した眼鏡型のデバイス [6]

この研究では、コメディ動画によって誘発された快感情を伴う自然な笑いとコンピュータからの指示によって意図的に作らせた快感情を伴わない作り笑いとを区別できるかを検証した。センサーから得られた反射強度の分布である幾何学的特徴と時間的特徴の両方から学習された機械学習モデル (SVM) で 94.6 % と非常に高い平均識別精度を記録した。しかし、この研究は眼鏡型のデバイスから得られるセンサーデータに依存するものである。そのためこの提案された手法に基づいて笑顔の識別を日常生活に応用することは難しい。

3. アプローチ

自然な笑いと作り笑いとを識別するためのアプローチを以下の図 4 に表す。

本研究は、特殊なデバイスを用いることなく笑いを識別することを目的とした。日常生活にこの識別技術を応用することを見込んで、機械学習モデルの構築という手法を採用した。そして、モデルの構築においては笑顔が表出しているときの表情データセットを作成するところから行った。

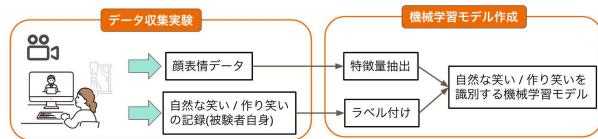


図 4 本研究のアプローチの概観図

笑顔状態の表情データは基本的にポジティブな感情から表出するものを取り上げている。本研究が着目しているような内面状態と表出事項の不一致、作り笑いのようにネガティブな感情を含む笑顔のデータセットは存在しない。そのため、データセット構築から取り組む必要がありデータセット構築は本研究の大きな新規性の一つである。

4. データセット作成のための実験

本実験は、会話上から生まれる自然な笑いと作り笑いの顔表情データを取得することを目的としたものである。男子大学生 5 名、女子大学生 3 名を対象に、30 分ほどの実験を行った。実験は、被験者 2 名と筆者自身の合計 3 名で構成され被験者 2 名の組み合わせはランダムに行った。実験中は筆者自身がファシリテーターとなり、オンライン上で質問に回答したり雑談を行った。そして、実験中に会話する様子を録画し、実験後被験者に録画を配布して各被験者に自然な笑い/作り笑いをした時間について動画を見ながら記録してもらった。また、先行研究でパーソナリティと作り笑いの相関について述べられていたため、個人の性格と作り笑いの表出について検証すべく、被験者ごとに Big-Five に基づく性格尺度を測るために TIPI-J に回答してもらった。

被験者自身に作り笑いの記録をしてもらうにあたって、以下のような作り笑いの定義を共有した。これは先行研究からまとめたものである [4]。

1. 擬態：初対面同士の会話、意見が対立した際に対立を避けるために、また緊張を軽減するための笑い
2. 雰囲気操作：場が凍っている時、白けてしまった時に場の雰囲気を和ませるための笑い
3. 感情制御：立場や雰囲気が原因で本音を言えない時、社交辞令によりネガティブな感情を隠蔽するための笑い

実験手順の詳細は以下の通りである。

1. 実験の概要と作り笑いの定義について説明を受ける。
2. アイスブレイクのために設定された自己紹介を行う。
3. 実験実施者からのコミュニケーションに関する質問に回答する。
4. 実験実施者からの与えられたお題について被験者同士で話す。
5. 実験終了後、実験実施者が提供する録画データを見な

がら、「自然な笑い/作り笑い」をした場面について記録を行う。

6. 「自然な笑い/作り笑い」の場面を記録したものと、TIPI-Jへの回答結果を Google form で送信する。

4.1. 被験者に行った質問項目

被験者に会話のきっかけとして行った質問項目を以下に示す。

1. 対面コミュニケーションにおいて自分が気まずいな、困ったなと思う瞬間を教えてください。
2. オンラインコミュニケーションにおいて自分が気まずい、困ったなと思う瞬間を教えてください。
3. 自分の感情とは反して笑顔を無理矢理作る、作り笑いをしたことがありますか？ またどのような場面でしましたか？
4. SFC にコース制度を導入するしたら賛成ですか？ 反対ですか？ 理由も教えてください。

これらは筆者が答えやすい質問と答えにくいと想定したものである。答えやすい質問として (1)～(3)、答えにくいと想定される質問として (4) を設定した。これらの質問は筆者が被験者に 1 人ずつ聞いていき、1 対 1 の対話形式とした。(1)～(3) には作り笑いの定義を実験の冒頭で行った際に、コミュニケーションにおける意思疎通の齟齬や作り笑いについての例を述べたため、実験の流れに沿った質問であり、答えやすいと考えた。(4) は唐突に実験の主題とは異なった質問で、理由も述べなければならないことから深く考える必要があり答えにくいと考えた。

4.2. アンケートによる笑顔のラベル付け

笑顔のラベル付けにおいては、実験終了後に Zoom の録画を配布して、自然な笑い（ポジティブな感情を含む）をした時間と作り笑い（ポジティブな感情を含まない）をした時間について記録するよう依頼した。記録を依頼する前に作り笑いの定義に関して被験者に再度補足の説明を行い、実験動画を視聴しながら記録してもらった。このアンケートデータから筆者が自然な笑いと作り笑いの場面だけ動画をトリミングし、機械学習モデル作成のために、データ整形を行った。

5. 実験結果

本節では第 4.1 章で述べたデータ収集実験において被験者から得られた回答について考察を述べる。また性格と作り笑いの表出頻度についての検証結果も本章で記す。

実験から自然な笑いの動画データが全体で 120 秒、作り

笑いが 251 秒収集できた。特徴化する過程については第 6 章で詳細を述べるが、データセットのカラム数としては自然な笑いのデータが 2986 行、作り笑いは 6291 行となった。

実験中に行った(1)~(4)の質問の回答結果を以下に記す。

5.1. 対面コミュニケーションにおいて気まずさを感じる点

対面コミュニケーションにおいて気まずさを感じる点について、「自分が原因で仲違いしてしまった相手とばったり会ったときのコミュニケーション」「初対面同士でグループワークを行い、意見が対立したとき」、「言葉が急に出てこなくなり、沈黙状態を生み出してしまうとき」、「自分が話しているのに相手から会話を早く切り上げたい雰囲気を感じ取ったとき」、「顔見知り程度の人とばったり会って挨拶以外に会話をしようとも話題選定が難しいとき」、といった回答が得られた。5名の被験者から初対面や顔見知りといったようにそこまで関係性を築けていない人とのコミュニケーションにおける話題が挙げられた。また2名の被験者から、先輩や上司のように立場が上の人と話す際は関係性が築いていても気まずさを感じることが多いという回答が得られた。

5.2. オンラインコミュニケーションにおいて気まずさを感じる点

オンラインコミュニケーションにおいて気まずさを感じる点について、「複数人で話すとき、お互いに話始めるタイミングを伺っているような雰囲気があるとき」、「グループワークでカメラオフの状態で話す人がいると表情が見えないから話し始めるタイミングが伺えない」、「深く考えずにテキストチャットを送信してしまい、意図が正確に伝わっていないかわからないとき。口頭で簡単な補足をすることができないから」、「大人数でのグループワークで沈黙状態から話し始めるハードルが高い」、「話し出しが被ったとき、会話を始める順番の譲り合いのようなことが始まり気まずい」、「グループワークでのアイスブレイクが気まずい、当たり障りのない授業に関する話題を持ちかけて頑張ることが多い」といった回答が得られた。

4名の被験者からオンライン上で会話する際に話し始めるタイミングについて回答が得られた。また2名の被験者から、相手がカメラオフだと表情が見えず話しづらい、感情が読み取りづらい、といった回答が得られた。

5.3. 作り笑いをする経験や場面について

作り笑いをした経験について8名中7名が作り笑いをすると回答した。7名の被験者からは、「目上の立場の人と会話をするときは対面でもオンラインでも作り笑いを頻繁に

してしまう」、「自分が興味のない話題でも相手が楽しく話せるように、作り笑いをして場を和ませる」、「相手の話が全く分からなくても、ひたすら相槌と作り笑いをして話を聞き続けたことがある」といった回答が得られた。作り笑いの経験があると回答した7名のうち全ての被験者が、会話する相手との関係性について述べていた。関係性について、目上の人に対しては頻繁に作り笑いをすると回答していたが、友人や後輩のような親しい関係性では作り笑いをすることなく、ありのままに感情を表している。といった回答が得られた。作り笑いをあまりしないと回答した1名の被験者は「感情が表情や言動にストレートに出てしまうので、あまり作り笑いはしない」と回答していた。

5.4. 性格と作り笑いの表出頻度の関連性

本研究において、被験者ごとの性格を表す指標としては Big-Five を採用した。Big-five は Goldberg [9] が提唱した、人の性格が5つの因子の組み合わせで決まるという学説である。そして、実験において信頼性が担保される形で被験者に負担をかけずに性格尺度を測定するために日本語版 Ten Item Personality Inventory (TIPI-J) を採用した [10]。これは Big-five の5因子に対して、各2項目の設問から測定するものである。5つの因子を以下に記す。

- Openness (開放性)
- Conscientiousness (誠実性)
- Extraversion (外向性)
- Agreeableness (調和性)
- Neuroticism (情緒安定性)

TIPI-J を作成する際に信頼性・妥当性の検討について、予備調査と本調査を大学生を対象に行っており、青年期以降に適用することが可能と述べていることから本研究の実験でも Big-five の5特性を測る際に妥当な指標だと考える。TIPI-J の回答から Big-five の5因子に対して尺度を測る式を以下に記す [10]。

- 開放性 = (項目 5 + (8 - 項目 10)) /2
- 誠実性 = 項目 3 + (8 - 項目 8)) /2
- 外向性 = (項目 1 + (8 - 項目 6)) /2
- 調和性 = ((8 - 項目 2) + 項目 7) /2
- 情緒安定性 = (項目 4 + (8 - 項目 9)) /2

以上の計算式から被験者ごとに、開放性、誠実性、外向性、調和性、情緒安定性を算出した結果を表1に示す。

そして、TIPI-J から算出された各5項目の素点と実験内における自然な笑いの発生回数と作り笑いの発生回数について相関係数を算出した結果を表2に表す。

1番強い相関が見られた項目が外向性との相関関係で相

被験者番号	自然な笑い	作り笑い	開放性	誠実性	外向性	調和性	情緒安定性
1	1	7	6.0	2.5	5.5	5.5	5.0
2	3	5	5.0	6.0	6.5	5.0	4.0
3	3	1	6.0	5.0	3.0	3.0	5.5
4	1	3	5.0	5.0	4.5	4.5	4.0
5	6	5	5.0	5.5	6.0	5.5	4.0
6	2	3	4.5	6.0	5.0	4.5	5.0
7	2	1	6.0	7.0	6.0	6.0	1.5
8	12	5	6.5	8.5	5.0	6.5	4.0

表 1 TIPI-J に基づいて被験者ごとに算出した性格の項目

	自然な笑い	作り笑い
開放性	0.4375	1.480297e-16
誠実性	0.6899	-0.3161
外向性	0.0307	0.4821
調和性	0.5041	0.4738
情緒安定性	-0.0397	0.2904

表 2 Big-Five と自然な笑いと作り笑いの表出に関する相関

関係係数は 0.482154 となった。次いで強い相関が見られた項目は調和性で相関係数が 0.473895 となった。また、性格の尺度と自然な笑いの表出回数にも相関が見られなかった。

自然な笑いの表出回数と誠実性の項目で相関係数が 0.689952 となった。データ数が 8 であることから、十分な相関関係にあるとはいえないと考えられる。次いで自然な笑いと強い相関関係が見られた項目は調和性で相関関係が 0.504178 となった。

このような結果になった 1 つの原因としてデータ不足が挙げられる。今回の実験では被験者 8 名を対象にしている。そもそも、被験者をさらに増やしてデータ数を増やすべきという反省点が考えられる。

6. 機械学習モデルの評価

6.1. 特微量抽出

実験で収集した動画を特微量化するにあたって、Tadas Baltrušaitis とケンブリッジ大学 Multicomp group が開発した Openface2.0 [11][12] を用いた。このツールキットは顔のランドマーク検出、頭の姿勢推定、表情変化の認識、視線推定をリアルタイムで行うものである。顔表情を解析するツールにおいて、顔のランドマーク検出から姿勢推定まで全てをカバーするソフトは少ない。また、このソフトは Windows, OSX, Ubuntu といったクロスプラットフォームに対応しており GPU を必要としないリアルタイム性能を持つ。この Openface を用いて、顔の筋肉の動きを客観的に分割する Facial Action Unit(AU) を検出した。Openface

の処理によって笑いの動画は 1 フレーム 0.04 秒単位で各 AU の動きが 1-5 の 5 段階評価で表される強度を特微量化した。その結果、データとして自然な笑いは 2986 フレーム、作り笑いは 6291 フレーム集まった。特微量化されたデータは以下のような構成になった。

カラム名
face_id
timestamp
confidence
success
AU01_r - AU45_r
AU01_c - AU45_c

表 3 Openface による Action Unit 抽出

face_id は、1 つの動画内に複数人が映っている場合に人を区別するためのデータである。confidence は、顔のランドマーク検出におけるトラッキングの信頼度を表すものである。実験内で実際にあったものとして、口元を手で隠すような動作がある際に顔のランドマークが検出できずにトラッキングの信頼度が 0 に近い値を取っていた。そのため今回実験で撮影された動画からデータセットを作成する際には、特微量抽出を行った際に confidence の値が 0.90 を下回ったカラムを削除する処理を行った。success はランドマーク検出における顔部位のトラッキングが成功しているかどうかを表すものである。主にフレーム内に顔が写っているかどうかを表している。AU_r は、AU の強度について表すものである。例えば、AU01_r のとき内眉の上がり具合について 0 (存在しない), 1 (最小強度で存在), 5 (最大強度で存在) の範囲で連続した値が outputされる。AU_c は、AU が存在するかどうかを表すものである。例えば、AU01_r のとき内眉の上がり具合が 0 (存在しない), 1 (存在する) が outputされる。

そして、実際に機械学習モデルを構築するにあたって Action Unit の動きの強度である [AU01_r] から [AU45_r] のみを説明変数として使用した。

今回作成したデータセットでは、実験内で被験者に回答

手法	Precision	Recall	F1 Score	AUC
Logistic Regression	0.895	0.910	0.902	0.840
Random Forest Classifier	0.972	0.973	0.973	0.957
Support Vector Classifier	0.893	0.916	0.904	0.840

表 4 教師あり学習の各手法とその性能評価

してもらった TIPI-J の性格尺度は使用していない。被験者はそれぞれ実験内で複数回作り笑いを行っているため 1 人の被験者につき約数百行の顔表情データがある。しかし顔表情のデータに性格尺度のラベルを付けるとしても数百行のデータに一意のラベルをつけることになる。そのため今回作成するデータセットでは Action Unit の強度のみを採用した。

6.2. 機械学習モデルの構築と精度評価

本節では機械学習モデルの手法について述べる。

まず実験で作成したデータをモデル構築用データ（学習データ）と予測精度検証用データ（検証データ）に分割する。今回は 8:2 の割合で学習用データと検証用データを分割した。

そして、実際に自然な笑いと作り笑いとを分類する機械学習の手法に Logistic Regression・Random Forest Classifier・Support Vector Classifier を用いる。今回はデータ数も各クラスにつき数千行とそこまで大きくなないことから、学習速度が早いこと、また 2 値分類に優れているモデルを採用した。

今回識別したモデルの自作データセットにおいて、自然な笑いのデータが全体の 32%、作り笑いのデータが全体の 68% とやや不均衡なデータセットとなっている。そのため今回は特に不均衡なクラスの問題において有用だとされる AUC (Area Under the ROC Curve) を用いて、識別モデルの全体の性能評価を行うことにした。AUC は、ROC 曲線下の面積を示し、0.5 以上が望ましい結果で 1 に近いほど識別能力が高いことを表す。ROC (Receiver Operating Characteristic) 曲線は、二値分類モデルの識別能力を表すグラフで、真陽性率 (True Positive Rate, recall) を縦軸に、偽陽性率 (False Positive Rate, precision) を横軸にとって、その結果をプロットしたものである。自然な笑いと作り笑いの識別モデルの性能について、表 4 に表す。

AUC で比較すると、ランダムフォレストによる手法において 95.7% を達成した。今回、学習手法として用いた Logistic regression, Random Forest Classifier, Support Vector Classifier において Precision, Recall, F1-Score, AUC の全ての指標においてランダムフォレストが 1 番高い性能を示した。ランダムフォレストはアンサンブル学習により複数の決定木が異なる特徴量を使用して学習し、その結果を組み合わせるため予測精度が高くなりやすいという

性質を持つ。そのためこのように高い精度で予測できたと考えられる。

また、機械学習モデルの汎化性能を検証するために k=8 の k 分割交差検証を用いて精度の評価も行ったところ、精度は 97.0% と非常に高い精度で識別できることが示された。

1 番高い精度で分類することができたランダムフォレストは決定木からなる学習手法であるため、各特徴量がどれだけモデルの精度向上に寄与しているのかを示す Feature Importance を算出し、以下の図 5 に示す。

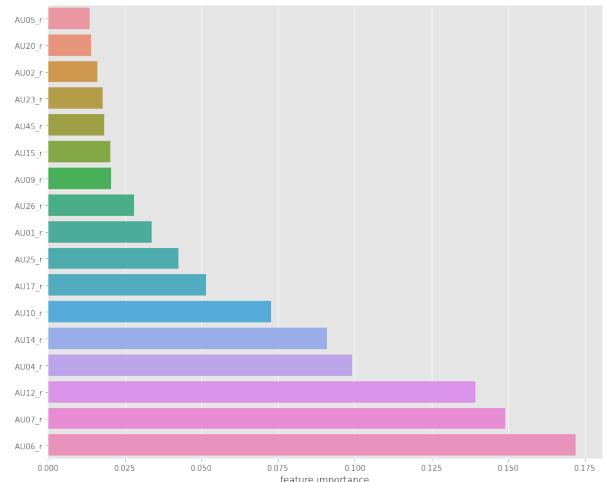


図 5 特徴量の重要度

特徴量の重要度において AU06 と AU07 と AU12 に重要度の高い傾向が見られた。この結果から作り笑いの表出時には、眼輪筋の動きによる頬の上がり方と瞼の動き、唇の両端の引き上がり具合が特徴的であり、これらの特徴量が識別予測に寄与していると考えられる。しかし、第 1 章で基本 6 感情にはそれぞれ普遍的な表情が存在するという主張を取り上げたが、佐藤らの研究 [2] では、基本 6 感情を表出す普遍的な表情の理論は日本人に当てはまらない可能性を示唆しており、実証に基づいて修正される必要性を述べている。そのため自然な笑いと作り笑いの識別に関して、Feature Importance から識別予測に寄与すると取り上げた特徴量が日本人特有の傾向という可能性もある。そのため世代や国籍を限定せず、自然な笑いと作り笑いの特徴について幅広く掘り下げていくためにはよりデータセットを充実させる必要があると考えられる。

7. 今後の展望

今後の展望として、作り笑いの深掘りと自然な笑いと作り笑いの識別技術の応用という観点から述べる。

7.1. 自然な笑いと作り笑いの識別技術のユースケース

自然な笑いと作り笑いの識別技術の応用について、この技術を実用段階に持っていくためには、手元にある任意の画像や動画からリアルタイムで自然な笑いと作り笑いを識別することが必要になると考えられる。リアルタイムで識別する技術をシステムとして実装することで、応用先の一例として挙げたカウンセリングや真の感情を隠すことが求められる場面への練習といった日常生活にも活用されると考えられる。

例えば営業や面接の場で、緊張する気持ちや気まずい気持ちを隠すために笑顔になることがあるだろう。相手にはネガティブな感情を悟られないように、表情の工夫を行うためにこの技術が応用できると考えられる。しかしコミュニケーションにおいて、自分の感情が隠されることなく相手に伝わったら不都合なこともある。例えば立場上の都合で本音と立前を使い分けて会話をすることが強いられる場面が挙げられる。そのため自然な笑いと作り笑いの識別技術は、コミュニケーションにおいて幅広く応用することを想定はしていない。あくまでカウンセリングのように内面状態(感情)と表情のずれを識別してメリットが生じる場合の実用を期待している。

7.2. 作り笑い時の感情分析

作り笑いについて、本研究のデータ収集実験では、被験者に作り笑いの定義を事前に共有した上で、作り笑いをした場面を申告してもらった。その実験内で被験者に行った質問から日常生活において作り笑いはさまざまな場面で見られることがわかった。例えば人と話していて「興味がない、つまらない」といったネガティブな感情を隠蔽するための笑いや初対面同士での会話で場を和ませるために笑いといったものである。このように、場の状況がさまざまあり感情も様々である。特に笑顔は通常嬉しい、楽しいといったポジティブな感情を結びつけられることが多いため、ネガティブな感情に基づいた笑顔(作り笑い)を識別することに意義があると考えられる。また、自然な笑いと作り笑いの識別だけでなく笑顔表出時の感情についても顔表情の動きから推定することができるのではないかと考える。作り笑い時のネガティブな感情の中でも辛い、苦しいといった感情を推定することが可能になれば、本研究の応用先として考えるうつ病患者のカウンセリングといった医療的な場面での実用に役立つことが考えられる。

また人間の目視では笑顔状態の人がポジティブな感情を

抱いているのかネガティブ、中立な感情を抱いているのか即座に判断することが難しいケースが想定される。商業施設や観光地において施設のサービスの満足度を測るために、現地でのインタビューや紙面でアンケートを依頼することが一般的な方法である。しかし対面でのインタビューは非常に時間もコストもかかる。そのため、1つの代替案として商業施設や観光地において各施設を訪問した人の帰るタイミングの表情をモニタリングし、笑顔であればそこからポジティブな感情を抱いているかどうかを測り、満足度の評価につなげられるのではないかと考える。

そして、コロナ禍においてオンラインでの対話やコミュニケーションが珍しくない。オンライン上ではカメラ越しの音声や動画といったように限られた情報から、相手の感情を汲み取る必要があることから表情から感情を推定する技術に、より注目が高まると考えられる。そのため、今後の展望として技術を幅広い分野で使用できるように、顔表情からではポジティブと捉えられるがちな笑顔状態における感情推定を行うことが挙げられる。

8. 結論

本研究ではデータセットの作成からモデルの評価までを行ったため、各ステップごとに結論と考察を述べていく。

8.1. データセット作成

まずデータセットに関する結論を述べる。本研究で作成したデータセットは自然な笑いと作り笑いのデータ比率が1:2となった。やや作り笑いに偏ったデータセットになってしまった原因として、実験環境の雰囲気が挙げられる。実験内容について第4.1章で述べたように、自然な会話を再現するために答えやすい質問と答えにくい質問を用意した。オンライン環境での実験ということで、そもそも硬い雰囲気を作り出してしまった、話すのが気まずかったというような原因が考えられる。また今回行った実験では質問内容は固定であるものの、会話内容は被験者によって異なり、雰囲気も様々であった。そのため場の雰囲気や会話のコンテキストを加味した上で表情データを収集することで、自然な笑いと作り笑いだけでなく、笑い表出時の感情を推定することも可能になると考えられる。またデータセット収集実験の被験者に関して、今回は大学生を対象にして行った。被験者が8名だったため、さらに被験者を増やすこと、また表情を収集する世代も幅広くすることでより充実したデータセットになることが考えられる。また、職業によって真の感情を隠して表情を作り出すことが得意になるケースも想定される。例えば役者や営業職のサラリーマンなどは自然な笑いと作り笑いの表出にあまり差分が見られないことも想定される。そのため世代や職業など幅広く被験者を募

集することで、自然な笑いと作り笑いの表出に関する差についてより深く知見が深まることが想定される。

8.2. 機械学習モデルの評価

最後に自然な笑いと作り笑いの識別モデルについて、ランダムフォレストによる学習において 95.7% という高い精度で識別することができた。特徴量の重要度を算出したところ、眼や頬の動き、口角といった部位に高い傾向が見られた。しかし、この特徴は大学生の会話時に限定した自然な笑いと作り笑いの識別に重要な特徴だと考えることもできる。今回は各被験者から収集した実験の録画データから顔表情を特徴量化して 1 つのデータセットを作成した。前の節でも述べたように被験者の対象を性年代や職業、国籍といった項目ごとにデータセットを分割することで表情の出方に特徴が見られる可能性もある。そのため本研究で作成したデータセットにおいては、非常に高い精度を達成することができたが、今後データセットを拡張した上でも同程度の精度を達成できるようにする必要があると考えられる。

謝辞 本研究成果は、国立研究開発法人情報通信研究機構の委託研究 (05401), JSPS 科研費 JP21K11853, JST RISTEX (JPMJRX21J1) の支援を受けたものである。

参考文献

- [1] 菅原 徹, 笠井直子, 佐渡山亜兵, 上條正義, 細谷 聰, 井口竹喜笑顔の多様性と印象の関係性分析, 感性工学研究論文集, Vol. 7, No. 2, pp. 401–407 (オンライン), 10.5057/jjske2001.7.401 (2007).
- [2] Ekman, P.: Universals and cultural differences in facial expressions of emotion, *Nebraska Symposium on Motivation. Nebraska Symposium on Motivation*, Vol. Vol. 19 (1971).
- [3] Ekman, P.: Paul Ekman Group — Universal Facial Expressions. <https://www.paulekman.com/resources/universal-facial-expressions/>.
- [4] 珊 李, 昌三渋谷社会的笑いに関する心理学研究の動向, 目白大学心理学研究 = Mejiro journal of psychology, Vol. (7), pp. 81–93 (オンライン), <https://cir.nii.ac.jp/crid/1050001338378901632> (2011).
- [5] 押見輝男公の自己意識が作り笑いに及ぼす効果, 心理学研究, Vol. 73, No. 3, pp. 251–257 (オンライン), 10.4992/jjpsy.73.251 (2002).
- [6] 千紗齊藤, 克俊正井, 麻樹杉本反射型光センサアレイを用いた眼鏡型装置による作り笑いと自然な笑いの識別, 情報処理学会論文誌, Vol. 62, No. 10, pp. 1681–1690 (2021).
- [7] Duchenne de Boulogne, G. B.: *The Mechanism of Human Facial Expression*, Studies in Emotion and Social Interaction, Cambridge University Press (1990).
- [8] 菅原 徹笑顔の形状と表情筋活動の分析, 可視化情報学会誌, Vol. 34, No. 133, pp. 14–19 (オンライン), 10.3154/jvs.34.14 (2014).
- [9] Goldberg, L. R.: An alternative "description of personality": the big-five factor structure., *Journal of personality and social psychology*, Vol. 59 6, pp. 1216–29 (1990).
- [10] 小塙真司, 阿部晋吾, Cutrone, P. 日本語版 Ten Item Personality Inventory (TIPI-J) 作成の試み, パーソナリティ研究, Vol. 21, No. 1, pp. 40–52 (オンライン), 10.2132/personality.21.40 (2012).
- [11] Baltrušaitis, T.: OpenFace 2.2.0: a facial behavior analysis toolkit. <https://github.com/TadasBaltrusaitis/OpenFace>.
- [12] Baltrusaitis, T., Zadeh, A., Lim, Y. C. and Morency, L.-P.: OpenFace 2.0: Facial Behavior Analysis Toolkit, *2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018)*, pp. 59–66 (online), 10.1109/FG.2018.00019 (2018).

Basic experiment related to overconfidence detection method from PC interaction data

Picard Sébastien、 上村拓也、 烏谷彰
富士通株式会社

1. Background

Overconfidence can lead to poor task performance or failure to achieve one's goals and can also prevent the implementation of optimal behavior change strategies. Overall, it can negatively impact a various range of industries from health promotion to fitness and exercise, to business and learning among others.

2. Problems with existing work

Typically, the way to approach the issue for the above-mentioned cases is as follows: analyze if the user has met their goal or not and compare that with a questionnaire about their expectation of success with regard to the goal [1], this allows to measure to what extent the user is overconfident. After that it is possible to take into account the overconfidence level for subsequent objectives and customize programs. However instead of waiting until the user has either achieved or failed at a first goal, it is desirable to detect overconfidence before adverse outcomes such as program abandonment or low achievement occur to be able to implement countermeasures.

3. Proposed method

We envisioned scenes where the user is interacting with a PC to carry out their work or to study. For such use cases, we used a definition of overconfidence similar to the over-estimation category from [2]. Specifically, we defined overconfidence as 1) the user will not be able to perform their task correctly or solve problems correctly 2) the user is confident of their future success and do not realize they will fail.

For this paper we focused on 1). Our hypothesis is that we can infer whether the user will achieve their goal or not by detecting behavior that indicates the user is experiencing difficulty.

We focused on the aspect that when a user is operating a PC, if they hesitate because they are experiencing difficulty, the typing patterns will be slower than usual, also the time needed to make decisions will be longer (for example they will pause longer before clicking validation buttons) than usual. We conducted a principle experiment to verify the feasibility of our method. To reflect the above-mentioned behavioral hypothesis, we implemented a

memory task and computed features such as follows: the duration of keystrokes, the duration between keystrokes, the duration hovering over a button before clicking it. We hypothesized that the more the user will experience difficulty recalling a combination from memory, the larger the feature values will be. Figure 1 illustrates this hypothesis.

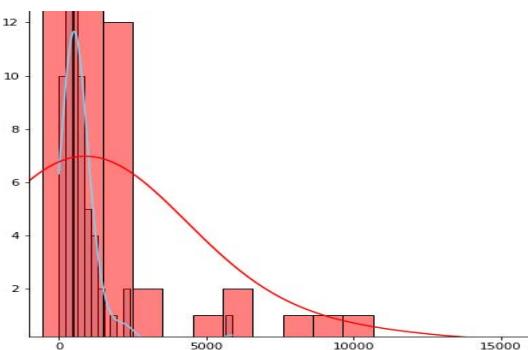


Figure 1): distribution of average duration in milliseconds between consecutive keystrokes. Blue line = 1st iteration (when the user is experiencing no difficulty), Red line = last iteration (when user cannot remember recall the combination)

In this experiment we started from simple combinations of digits, thus we could compute a baseline for each user (we assume the users experience no problem remembering the first easy combinations).

Based on the data of all the other participants in the experiment we built a model to assess whether the user will achieve their goal (until how many digits they can remember combinations) or not based on feature data relative to the baseline.

4. Experiment design

We devised a principle experiment to verify the feasibility of our method to detect whether a user operating a PC in an office setting with achieve their goal or not.

4.1 Task and experiment scenario:

- Display a random combination of digits for 5 seconds
- The user memorizes the combination and interacts with a custom UI requiring keyboard input and mouse movement before submitting their answer (this is called

an iteration for the rest of the paper)

- Starting from a combination with less than 5 digits, the combination grows in length with a 1-digit increment, either to the left or right to the previous combination
- This iteratively continues until the user either gives up or fails to remember up to 3 times (we call this series of iteration a sequence for the rest of the paper)
- Before the first combination is displayed the user is asked to input the expected length of the combination they can remember (We call this goal for the rest of the paper)
-

4.2 Example of features used as input for the model

Keyboard typing speed, average duration between keystrokes, time spent hovering on the submit button after typing the last digit.

5. Data collection

16 users responded to a pre-experiment survey and indicated their ascent to participate in the experiment. Explanations about how to install the program and run the user interface (UI) was provided to all participants. The task consisted of running the UI daily for 2 weeks (excluding weekends) and report their performance in a shared file so as to elicit overconfidence. 11 participants engaged with the program and we discarded the first sequence (used for practice). From the data obtained, we restricted the analysis to the iterations where the number of digits is lower than the goal, 66 sessions and 525 iterations.

6. Analysis

6.1 Features

We computed a battery of features meant to reflect our hypotheses stated in 3. Proposed method section. Those features included features computed from keyboard as well has mouse hovers over buttons to highlight the slowing of typing patterns expected to arise when hesitation due to inability to remember kicked in.

6.2 Normalization

Our hypothesis relies on comparing behavior with a normal level, we believe that changes in the above features are more meaningful than absolute values. To verify this, in the result section we compared performance for absolute values and data normalized as follows: a) the first 5 iterations of the second sequence (the first sequence being excluded from the data) is used to compute the minimum and maximum values independently for each participant. All features measuring durations (e.g. mean or standard deviation of the inter-keystroke duration) are normalized by subtracting the minimum and dividing by (maximum-minimum) independently for each user.

6.3 Model

To estimate whether a user is going to achieve or fail at their goal of remembering a certain number of digits, we used a statistical model which takes multiple features described in the previous section as input and outputs a binary label where 1 indicates the user will not meet their goal.

We used a random forest algorithm after early results (not reported in this paper) showing an improvement over other model tested (included KNN, logistic regression) and assumed that the in-built feature selection exhibited by random forest is responsible for the improvement of performance compared to other models. In this experiment we adopted a randomized search for the selection of hyper-parameters combined with a cross validation scheme. Further, to minimize the impact of overfitting due to personal differences we put emphasis in selecting folds such that data from participants cannot be both in the training and validation set. Additionally, we reported results where data from one user is exclusively used in the test set, this means that for user i in test, we trained the model based on the remaining 10 users, then we repeated for $i = 1$ to 11 and aggregated the results.

7. Results

From the computation of features related to the task where users remember a combination of growing length, we evaluated our method and reported results about the feasibility to predict whether the user will achieve their goal (meet their expectations) or not. At first, we focused on each iteration separately. This task is rather ambitious but constitute the fundamental building block of our method.

We report the following results in figure 2): samples where the user does not meet their goal are labeled as 1 and samples where the user meets their goal are labeled as 0. The rows of the confusion matrix indicate the true label and the columns the predicted labels. We report the accuracy as well as balanced accuracy which is the average accuracy per class, we consider it is a better indicator in case of imbalanced data sets such as we have here. The models are also trained to maximize the balanced accuracy. For the rest of the paper, we report performance statistics rounded to the fist decimal. Figure 2a) shows the results for the basic model with no normalization. Figure 2b) show results with normalization where the user goal, iteration number of current number of digits are also added as features.

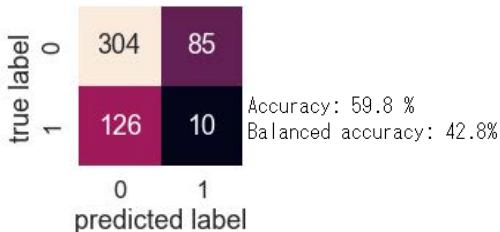


Figure 2a) confusion matrix without normalization

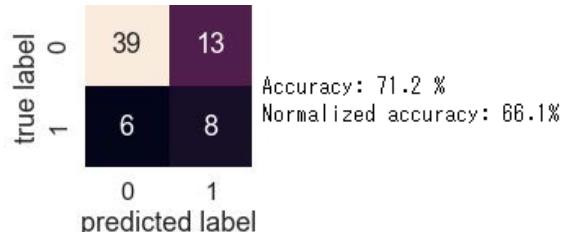


Figure 3): sequence level results

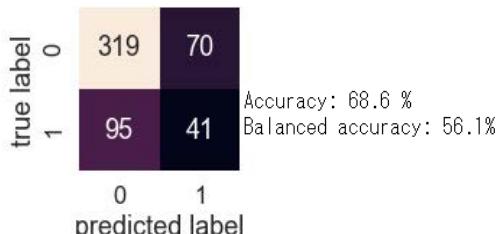


Figure 2b) confusion matrix with normalization

The results shown in figure 2a) confirm the feeling that the problem is extremely ambitious. Without normalization the balanced accuracy falls below 45% and reaches 56% with the normalization scheme. The improvement in balanced accuracy is substantial, especially in the latter case recall ($41/136 = 30\%$) for class 1 is much higher compared with the former case ($10/136 = 7.4\%$), however it is still much below 50%. For class 0, that is, when user will meet their goal, recall is $85 / 389 = 78.1\%$ in the latter case and $319 / 389 = 82\%$. The results obtained here confirm the hypothesis that absolute feature values are less meaningful than relative values and we keep the normalization scheme for the rest of the analysis.

To verify our claims, we then evaluated the performance of our method to predict sequence labels, focusing on making predictions at an early stage (when the number of digits is lower than the goal). Due to the low number of sequences we decided to harness the iteration labels instead of computing features at sequence levels. However, with binary labels the information loss would be too significant, so a more fine-grained approach is preferable to predict sequence labels from iteration labels. Finally, for ease of comparison with the results described in figure 2, we output the sequence classification results in binary format.

The implementation steps are described here:

Step 1: modify the iteration model to output 4-class labels (L0: outperformed goal by more than 10 digits, L1: by 5 to 10 digits, L2: met goal or exceeded by less than 4 digits, L3:user didn't meet their goal).

Step 2: obtain the 4-level labels for the first 5 iterations of each sequence and return the maximum value as intermediate result.

Step 3: return to a binary label by converting L0, L1, L2 to label 0, L3 to label 1.

Figure 3) shows more promising results with a balanced accuracy reaching 66%, also recall for class 1 exceeds 50% with $8 / 14 = 57.1\%$ but decreases for class 0 with $39 / 52 = 75.0\%$. The results obtained here show that it is possible to obtain better performance at sequence level by harnessing the results from the first few iterations.

We consider further tests with larger data sets are needed to confirm the reproducibility of the results and improvements in terms of prediction performance are required for various applications.

8. Conclusions and future work

Our experiment results indicate that it is possible to infer from PC interaction data the first part of our overconfidence definition, namely 1) the user will not be able to perform their task correctly or solve problems correctly.

We need to address the second part of our definition of overconfidence, namely 2) the user is confident of their future success and do not realize they will fail. We consider this may be done using a questionnaire, when the user is expected to be unable to reach their goal.

Overall, our vision is that overconfidence may be detected early enough to modify intervention strategies, benefit users and make significant impact for all sorts of applications. Future work includes various steps towards the realization of this vision, for example the implementation of the 2-step overconfidence detection in real-time and the evaluation of countermeasures for different applications.

9. References:

- [1] Ifcher, John, and Homa Zarghami. "Affect and Overconfidence: A Laboratory Investigation." *Journal of Neuroscience, Psychology and Economics* 7.3 (2014): 125-50. Print.
- [2] Moore, Don A., and Paul J. Healy. "The trouble with overconfidence." *Psychological review* 115.2 (2008): 502

行動変容を促すナッジ文言選定のための アプリログによる属性情報の補完手法

吉川 裕木子[†] 勝間田 優樹[†] 鈴木 喬[†] 山田 晓[†]
株式会社 NTT ドコモ[†]

1. はじめに

近年、スマートフォンの普及に伴い、個人が持つ端末に情報を送信することで、ユーザー毎に異なる内容を伝達する手法が広く実施されている。また、情報を配信する際に、音・振動・短い文章で知らせることでユーザーの興味を惹き、配信した情報の認知率を向上させるプッシュ通知が注目されている。よって、プッシュ通知によるレコメンデーションにおいて、端末を利用しているユーザーを分析し、配信内容やタイミングをユーザーの興味に沿うように最適化することで、ユーザーの情報認知率を最大化させるという営みが様々な企業で行われている。

また、ユーザーに情報を提供することが意思決定に与える影響について、行動変容に関する分野の研究が注目されている。ここで述べる行動変容とは、行動経済学でいわれる人間の行動理論に基づいて働きかけることで、その人の行動を変化させることができるものである。このような働きかけを活用して、環境問題や都市渋滞など様々な社会問題の解決に取り組んだ実例研究がおこなわれている[1][2]。

また、行動変容をさらに促す手法として、ナッジ[3]が注目されている。ナッジとは、人間の思考の癖を利用し、ユーザーにとって望ましい選択をするように促す行動変容に関する手法の一つであり、ナッジを活用した研究や実例が多く実施されている。例えば、プッシュ通知によって表示されるメッセージにナッジを適応することでプッシュ通知の開封率を向上させるという効果が挙げられている[4][5]。

ナッジには複数のフレームワークが存在する。例えば、人間の得をすることより損をすることの方を大きく評価するという損失回避バイアスという心理特性を活用した「損失回避フレーム」や、他人と同じようにふるまうことを好むという同調性バイアスという心理特性を活用した「同調フレーム」などが例に挙げられる[6]。

本研究では、ユーザーのもつ心理特性によって、異なるフレームワークによって設計されたナッジ種別の効果が異なることに着目し、ユーザーの心理特性の度合いを推定することで、ユーザーに最も効果のあるナッジの種別を推定することができると考える。さらに、その推定に基づいて、最適

なナッジを選定し、プッシュ通知に活用することで、プッシュ通知の開封率の向上をはかることができる。しかし、このナッジ種別の推定にはユーザーの属性情報が必要であり、その情報量はナッジ種別の推定に精度に大きくかかわる。こういったシステムを実現する際には、推定精度の担保のため多くの情報を収集する必要があり、その情報収集にかかるコストがシステム化する際の課題となることが予想される。

よって本提案手法では、属性情報を持たないユーザーのナッジ種別の推定に必要な情報の拡張手法を提案する。具体的には、属性情報が取得できている一部のユーザーを用いて、属性情報を持たないユーザーの属性情報を補完する。本提案手法により、すべてのユーザーに対して属性情報を獲得することが可能となり、これは最適なナッジ種別の推定に効果的であると言える。

2. 関連研究

ユーザーの意思決定の癖である心理バイアスとユーザーのパーソナリティの関係性は、過去の研究で示されている。例えば、時間割引率に関するアンケート分析[7]では、年齢や性別、婚姻の有無などが時間割引率に対して影響があることを示した。このような研究から、ユーザーが持つ心理バイアス特性を調査できるようにデザインされたアンケートを用いることで、ユーザー自身の属性情報との関係性を分析できることが分かる。

一方、心理バイアスと対応したフレームワークを用いたナッジとユーザーのパーソナリティの関係性についても、過去の研究で示されている。例えば、学習を促すメッセージ配信において、事前に測定したユーザーの性格特性と、メッセージ配信に用いた2種類のナッジに対する反応に相関関係があったという実例が挙げられている[8]。このような研究から、ユーザーのもつパーソナリティに応じたナッジを用いることで、行動変容率の向上が期待できることを示している。

これらの従来研究から、ユーザーのパーソナリティを用いて学習モデルを構築することで、そのユーザーの行動変容を促すのに最適なナッジを推定することが可能だといえる。しかし、こういった学習モデルの構築に必要な、ユーザーのパーソナリティは、年収や性別などの属性情報に関するア

Proposal and Evaluation for Nudge Optimization Recommendation System Using Extending Attribute Information with Application Logs
† YUKIKO YOSHIKAWA, YUKI KATUMATA, TAKASHI SUZUKI, AKIRA YAMADA, NTT DOCOMO Ltd.

ンケートや、くじの購入金額の質問[9]のようなユーザがもつ心理バイアス特性を特定するようなアンケートを実施することで収集する、といったことが従来の研究では行われてきた。しかし、こういったアンケートの質問設計は大変複雑であり、またユーザから有効な回答が得られる割合多くはない。よって、情報収集の点でナッジのユーザ最適化を実サービスとして実現するのが難しいといえる。

ユーザのパーソナリティを示すデータの収集方法に、ユーザのスマートフォンなどから収集できるログを活用して推定する研究がおこなわれている。例えば、携帯端末から収集できる位置情報などからわかる、日常的な行動パターンから職業属性が推定できる[10]。また、Webサイトにおける時系列的なアクセス数から、年齢などのユーザ属性に基づいたクラスタを推定することができる[11]。

これらのことから、パーソナリティデータを収集するために、スマートフォンから自動的に収集できるログを活用することで、アンケートなどから収集できる情報相当を推測できると考えられる。

よって、本研究では、一部のユーザに実施したアンケートなどから収集した情報と、全てのユーザから自動的に取得可能なスマートフォンからのログを用いて、ナッジのユーザ最適化に有効な属性情報を補完する手法を提案する。また、ナッジのユーザ最適化において本提案手法が有効であることを示す。

3. 提案手法

3.1 ユーザが持つ情報種別について

本研究では、図1のように2種類のユーザ群を対象ユーザとする。ユーザ群Aは属性情報とアプリログの2種類のデータをもち、ユーザ群Bはアプリログのみをもつ。また、属性情報とアプリログはベクトル表現に変換されており、ベクトル間距離を算出することで、ユーザ間の類似度を示すことができる。

ここで属性情報は、ユーザ群Aの過去のアンケート回答や他のサービス利用から推測される、性年代や趣味趣向などを指し、アプリログは本システムが導入されたアプリケーションから取得できる、各画面の閲覧時間やボタン押下回数などの操作ログを指す。

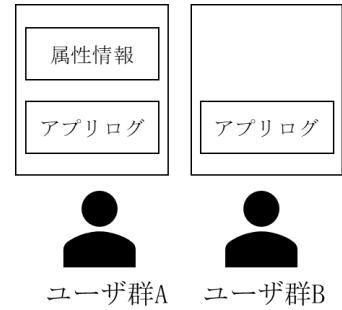


図1 本研究で用いたユーザ群と保持するデータ

3.2 アプリログのベクトル空間距離の算出

2で述べたスマートフォンからパーソナリティを推定できるという従来研究より、本研究における属性情報とアプリログには相関関係があり、アプリログが類似しているユーザ同士は属性情報も類似していると考えられる。本提案手法では属性情報とアプリログの類似度をベクトル間距離と定義する。

まず、属性情報とアプリログのベクトル間距離を定義する。ここで、ユーザ*i*とユーザ*j*間の属性情報 X_d の距離は以下の式で与えられる。

$$distance_{Xd} = \sqrt{(x_i^d - x_j^d)^2} \quad (1)$$

また、アプリログ X_a の距離を以下の式で与える。

$$distance = \sqrt{(x_i^a - x_j^a)^T W (x_i^a - x_j^a)} \quad (2)$$

$$W = \begin{bmatrix} W_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & W_n \end{bmatrix} \quad (3)$$

ここで、 W はアプリログの重み付けパラメータであり、アプリログの距離が近いと属性情報の距離が近いという関係が成り立っている。なお、パラメータ W の学習は以下の式を用いて、対象ユーザの全通りの組み合わせで行われる。

$$Loss Function = distance_{Xd} - distance \quad (4)$$

$$W' = W + \frac{\partial Loss}{\partial W} \quad (5)$$

次に、属性情報とアプリログの両方をもつユーザ群Aを用いて、パラメータ W を学習する。式(2)に学習済みの W を用いて、アプリログのユーザ間距離が近いユーザは同じクラスタになるようクラスタリングを行い、各クラスタに所属するユーザの平均ベクトルを獲得する。

3.3 属性情報の補完手法

アプリログのみをもつユーザ群BのユーザBの属性情報を補完する手法を示す。ユーザBのアプリログと3.2で獲得したクラスタの平均ベクトルとの距離を、式(2)を用いて

算出し、ユーザ B に最も近いクラスタを選択する。

次に、選択されたクラスタ内からユーザ群 A に所属するユーザ A をランダムで抽出し、そのユーザがもつ属性情報をユーザ B の属性情報とする。なお、ランダムで選ぶ理由として、ユーザ群 A とユーザ群 B にアプリログの傾向が異なる可能性が考えられるため、最も近いユーザを選んだ場合、常に同じユーザが選ばれてしまうことを避けるためである。

4. 評価手法

4.1 実証実験

本提案手法の評価は、2022年2月1日から2022年3月21日の期間で実施された、幕張新都心エリアにおける実証実験で収集されたデータを用いて行った^{*1}。本実証実験は、地域住民の回遊性向上を目的とした実証実験であり、実証参加者にはアプリケーション「まくはり MaaS」をインストールしてもらい、それを活用しながら幕張新都心エリアを周遊する。図2のように、アプリケーションからは、エリアを周遊するための交通手段の予約や、周遊エリアのスポット情報にアクセスすることができる。また、本評価手法で用いるユーザ数の内訳を表1に示す。

4.2 プッシュ通知について

実証実験で用いたプッシュ通知システムについて説明する。本システムでは、ユーザの状況をトリガとして、店舗と文言を該当のユーザへプッシュ通知を行う。ユーザの状況とは、アプリケーションをインストールした端末から取得できる位置情報や、特定の時刻、交通手段の予約履歴などのことを示す。ユーザ状況をトリガにしたプッシュ通知は、「幕張駅に着いた時に、周遊スポットをおすすめ」や、「お昼時に飲食店をおすすめ」といった内容のレコメンドをユーザに提供することが可能である。

また、メッセージ確認画面は図3のようになっている。プッシュ受信時には、通知タイトルに「あなたへのおすすめがあります」、通知文に「あなたにおすすめの店舗でお食事はいかがですか？」のプッシュ通知が通知エリアに表示される。なお、通知文にはプレーン・損失回避フレーム・同調フレームを活用して設計された3種類のナッジ文言を採用した。



図2 「まくはり MaaS」の画面イメージ

表1 評価で用いるユーザ数内訳

		1回以上開封有のユーザ数			
	総ユーザ数	文言1	文言2	文言3	合計
ユーザ群A	835	118	119	112	245
ユーザ群B	3125	345	378	366	773



図3 メッセージ確認画面

表2 アプリログの例

プッシュ開封時間	カーシェア利用回数
アプリ起動回数	シェアサイクリング利用回数
プッシュ開封率	タクシー配車サービス利用回数
クーポン利用回数	バス利用回数
メッセージ詳細閲覧回数	バス予約時間/乗車時間
スポット詳細閲覧回数	ルート検索条件(歩く速度・並び替え)
スポット検索回数	ルート検索使用回数

4.3 最適なナッジ文言の推定

本研究の評価で利用する実証実験の開封履歴について、実証実験中に介入があったことを説明する。具体的には、本実証実験において、ユーザ群 A に所属するユーザの半数には属性情報に応じて推定された文言のプッシュ通知を、もう一方にはランダムで配信を行った。なお、ユーザ群 B に対しては、すべてのユーザに対してランダムで配信した。

ユーザ毎の最適なナッジ文言の推定には、文脈付きバンディットアルゴリズム[12]を採用し、属性情報と過去の開封履歴から学習を行った。バンディットアルゴリズムは、

*1 https://www.ntt.com/jp/binary/pdf/info/news_release/topics_220125_00.pdf

複数の選択肢から利得を最大にするのを目的としており、新しい情報の探索と、既存情報に基づく選択肢の実行を同時に実施するものである。本実証実験では、選択肢を 3 種類の文言、情報の探索と実行をプッシュ通知、利得をその開封と定義して学習を行う。具体的には、過去の開封実績とユーザの属性情報に基づいて、文言のユーザ毎評価値を算出する 3 つのパラメータの更新を夜間に行い、日中はトリガが発火した際に、その評価値に基づいた確率でプッシュ通知を実施するものである。また、ユーザ群 A における属性情報に応じたナッジ文言の推定による介入効果を検証したところ、有意差 10%で優位に開封率が向上したことがわかった。

4.4 属性情報とアプリログ

実証実験の参加者から、事前同意を得て取得した属性情報とアプリログについて説明をする。属性情報は、NTT ドコモが過去に実施したアンケートなどから収集したデータであり、性年代や職業といったものが挙げられる。このような情報を、141 次元のベクトル表現にしたもの用いる。

一方、アプリログは、4.1 で紹介したアプリケーションから特徴量を作成した。アプリケーションは、各種モビリティサービスの利用ボタンや、経路検索機能、クーポン利用画面を所持しており、ユーザによって操作の差異が現れると考える。表 2 のような情報を、44 次元のベクトル表現にしたもの用いる。

4.5 プッシュ通知開封の推定による評価

本提案手法の評価は、実証実験の開封履歴を用いたオンライン検証によって行う。また、本研究の目的であるユーザ毎の最適なナッジ文言の推定を、ユーザ毎の各文言のプッシュ通知の開封を予測する問題として定義し、本提案手法の評価を行う。この予測問題に対して、Python のオープンソースライブラリである PyCaret[13]によって、ナッジ文言毎の機械学習モデルを構築する。また、機械学習モデルの説明変数として、アプリログと 3.3 で補完した属性情報を用い、目的変数としてプッシュ通知に対する開封を用いる。

本研究における評価値の指標として、PyCaret の機能によって選定された最も性能のよいモデルによる予測値と、正解値との相関値を評価値とする。

4.6 プッシュ通知の開封の算出について

目的変数に、ユーザ毎に本実証実験期間での合計開封回数を用いることで、予測値を算出することができる。なお、プッシュ通知回数が極端に少ないユーザや、開封自体が少ないといったユーザが存在するため、開封毎に何回目のプッシュ通知かを考慮した重み付けを行う。ここで、図 4 に本実証実験のプッシュ通知の結果を集計した結果を示す。

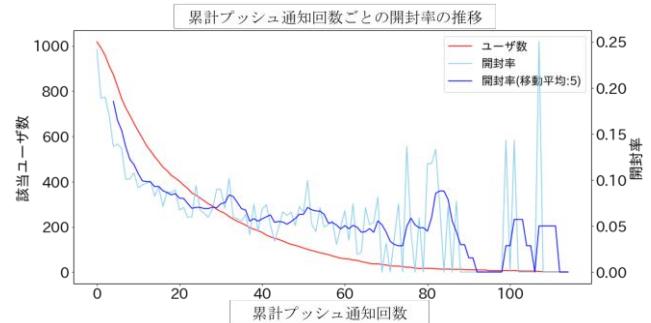


図 4 累計プッシュ通知回数ごとの開封率の推移

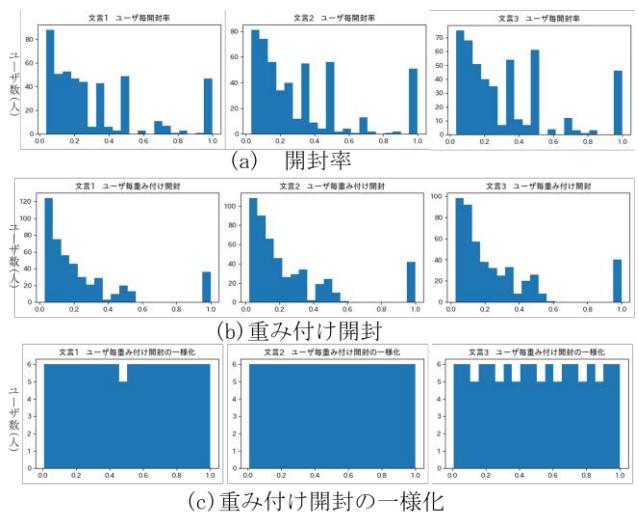


図 5 重み付けによる開封 Y の変化

図 4 で示すように、プッシュ通知回数が 20 回以下の場合、プッシュ通知の開封率が減少傾向にあるが、20 回を超えると定常化している。これは、プッシュ通知が少ない間は配信に興味を惹かれ開封するが、何度も配信されるとプッシュ通知に対する興味が低下し、反応しなくなるという事が言える。一方、配信回数に関わらず開封し続けるユーザが一定数いることも示している。

累計プッシュ通知回数が少ないとより、多いときに開封したもののが、1 回の開封の価値が大きくなるような重み付けを行った結果を図 5(b)に示す。また、重み付けは以下の式で行う。

$$Y^{j,a} = \frac{1}{n^{j,a}} \sum_{i:\pi_i^a=j} \frac{\log(1/p_i)}{\sum_{k:\pi_k^a=j} \log(1/p_k)} \quad (6)$$

ここで、Y はユーザ a の文言 j の重み付け開封である。n はユーザ a の文言 j の開封回数を、P は全ユーザの k 回目の平均開封率の移動平均を、 $\pi_i^a=j$ はユーザ a の i 回目のプッシュ受信の文言が文言 j であることを示す。なお、60 回以上のユーザ数が少ないとから、60 回以上のプッシュ通知に関しては無効とする。さらに、重み付け開封 Y が一様分布になるように、Y の大きさの順序を合計ユーザ数で割った Y' を予測値として用いる。一様化を行った結果を図 5(c)

に示す。

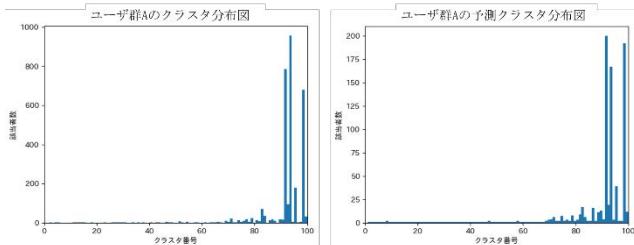


図 6 ユーザ群 A の予測クラスタ分布

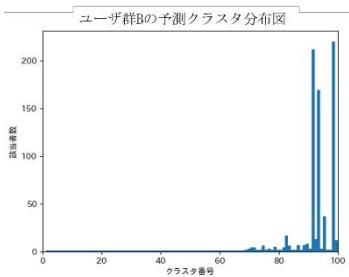


図 7 ユーザ群 B の予測クラスタ分布

	ユーザ群 A	ユーザ群 B
ユーザ数	835	3167
ユーザ所属クラスタ数	100	65
クラスタ所属平均ユーザ数	8.35	48.07
標準偏差	34.431	172.114
変動係数	4.123	3.580

表 3 ユーザ群 A とユーザ群 B の予測クラスタ結果

5. 評価結果

5.1 ユーザのクラスタリングに関する評価

ユーザ群 A に所属する 835 人のユーザの 141 次元の属性情報ベクトルと 44 次元のアプリログベクトルを用いて、式(2)のパラメータ W の学習を行った。学習した W を用いて 100 個のクラスタにユーザ群 A をクラスタリングした結果とユーザ群 A の所属クラスタを予測した結果を図 6 に示す。本提案手法による所属クラスタの予測正解率はミクロ正解率で 94.7% であった。しかし、図 6 のクラスタ分布図からわかるように、多くのユーザは 4 つのクラスタに所属しているため、正解率が高くなっているともいえる。

次に、ユーザ群 B に所属するユーザ 3125 人を用いて、所属クラスタの予測を行った結果を図 7 にしめす。また、表 3 に示すように、ユーザ群 A とユーザ群 B のクラスタリング結果の比較をしたところ、ユーザ群 A とユーザ群 B の傾向に大きく乖離がないところから、3.3 の提案手法による所属クラスタの予測が、ある程度可能であるといえる。

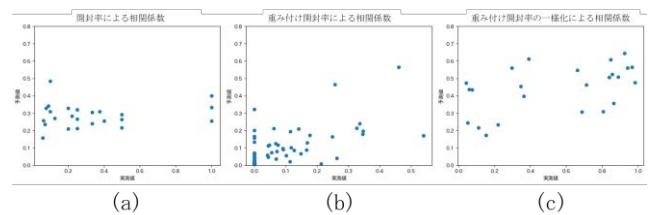


図 8 重み付けによる予測値 Y の分散変化

表 4 ユーザ群 A : 学習と評価に用いたユーザ数

	文言 1	文言 2	文言 3
ユーザ数(人)	118	119	112
学習データ	94	95	89
評価データ	24	24	23

表 5 ユーザ群 A : 属性情報による予測結果

	文言 1	文言 2	文言 3
Best model	K Neighbors Regressor	Random Forest Regressor	Ridge Regression
Correlation	0.289	0.123	0.211
MSE	0.120	0.090	0.129

表 6 ユーザ群 A : 属性情報とアプリログによる予測結果

	文言 1	文言 2	文言 3
Best model	K Neighbors Regressor	AdaBoost Regressor	Ridge Regression
Correlation	0.502	0.290	0.339
MSE	0.101	0.082	0.121

5.2 ユーザ群 A に関する評価

4.6 で示した開封に対する重み付けの処理による効果を、ユーザ群 A を用いた文言 1 における開封 Y の予測結果によって示す。予測に用いたユーザは、1 回以上開封を行ったユーザとする。これは、表 1 で示すように、プッシュ通知を一度も開封していないユーザの割合が多いため、全ユーザを用いて学習をしてしまうと、開封 Y の予測が難しくなるためである。説明変数として、属性情報とアプリログを用い、アプリログは開封率と強い相関のある特徴量を除いた 22 次元で推測を行った。

重み付けと一様化を目的変数である開封率に実施し、学習した場合の予測結果を図 8 に示す。予測値が低い値に集中することなく予測が行えていることが分かる。

次に、属性情報とアプリログが、ナッジ種別推定に有効であることを示すために、ユーザ群 A の属性情報のみを用いた場合と、属性情報とアプリログの両方を用いた場合の予測を行った。表 4 に示すように、モデルの学習と評価に用いたユーザは、各文言プッシュ通知を 1 度でも開封したことのあるユーザのみで行い、学習データと評価データは 8:2 で分割した。

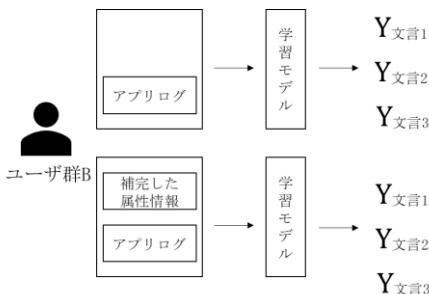


図 9 ユーザ B 群を用いた提案手法の評価

表 7 ユーザ群 B : 学習と評価に用いたユーザ数

	文言 1	文言 2	文言 3
ユーザ数(人)	345	378	366
学習データ	276	302	292
評価データ	69	76	74

表 8 ユーザ群 B : アプリログによる予測結果

	文言 1	文言 2	文言 3
Best model	Random Forest Regressor	Ridge Regression	Ridge Regression
Correlation	0.310	0.158	0.270
MSE	0.07	0.08	0.07

表 9 ユーザ群 B : 属性情報とアプリログによる予測結果

	文言 1	文言 2	文言 3
Best model	AdaBoost Regressor	Random Forest Regressor	Random Forest Regressor
Correlation	0.312	0.196	0.379
MSE	0.08	0.08	0.07

属性情報のみを用いた場合と、属性情報とアプリログを用いた場合の予測結果を表 5 と表 6 に示す。全ての文言において、属性情報とアプリログを用いた場合が、属性情報のみを用いた場合の相関値を上回っているため、ナッジの開封の予測には属性情報だけでなくアプリログも有効であることが分かる。

5.3 ユーザ群 B に関する評価

本提案手法の有効性の評価方法の概要を図 9 で示す。ユーザ群 B のアプリログのみを用いた場合と、3 で示した提案手法によって選定した属性情報とアプリログの両方を用いた場合の予測結果を比較することで有効性を示す。なお、予測に用いたユーザは、1 回以上開封を行ったユーザであり、モデルの学習と評価に用いたデータは表 7 に示す。表 8 と表 9 が示すように、2 つの場合を比較した結果、提案手法によって選定した属性情報を用いた場合、予測精度が改善したことがわかる。なお、文言によっては相関値の変化が小さい文言もあり、すべての文言に対して予測精度を

改善するためには、アプリログのベクトルの次元数をより増やすことで表現力を向上させる必要がある。

6. おわりに

本研究では、アプリログを用いて属性情報を補完する手法を提案と評価を行った。ナッジ毎の開封予測精度が向上したことから、本提案手法が有効であるという示唆を得られた。しかし、本研究で行った評価は、過去の実証実験のデータを用いたナッジ毎のプッシュ通知開封の予測によるものであり、本研究の目的であるナッジ種別の推定への有効性を 100% 示せるものではない。よって、各モデルが出した評価値の違いに基づいて、実際に文言を出し分けることで、その有効性を検証する必要がある。なお、評価値の違いに基づいてナッジの出し分けを行う点については、4.3 でユーザ群 A において効果があったことを述べている。今後は、本研究で未検証である内容について実証実験を実施することで、提案手法によるプッシュ通知の介入効果の向上の有効性を検証していきたい。

参考文献

- [1] Reddy, S. M. W., Montambault, J., Masuda, Y. J., Keenan, E., Bulter, W., Fisher, J. R. B., Asah, S. T., and Gneezy, A. “Advancing Conservation by Understanding and Influencing Human Behavior”, Conversation Letters, Vol.10, No.2, pp.248–256 (2017)
- [2] Xu, W., Kuriki, Y., Sato, T., Taya, M., and Ono, C. “Does Traffic Information Provided by Smartphones Increase Detour Behavior? An Examination of Emotional Persuasive Strategy by Longitudinal Online Surveys and Location Information”, Proc. 15th international conference on Persuasive Technology, pp.45–57 (2020)
- [3] Leonard, T. C.: Richard H. Thaler, Cass R. Sunstein, “Nudge: Improving decisions about health, wealth, and happiness” (2008)
- [4] Hyun Bae Kim, Toshiya Iwamatsu, Ken-ichiro Nishio, Hidenori Komatsu, Toshihiro Mukai, Yoko Odate, Masanobu Sasaki, “Field experiment of smartphone-based energy efficiency services for households: Impact of advice through push notifications”, Energy and Buildings, Vol. 223 (2020)
- [5] Carmina G. Valle, Brooke T. Nezami, Deborah F. Tate, “Designing in-app messages to nudge behavior change: Lessons learned from a weight management app for young adults”, Organizational Behavior and Human Decision Processes, Vol. 161, pp. 95-101, (2020)
- [6] 竹内幹, “社会保障のための行動経済学: 補正か誘導か?” 社会保障研究, Vol. 6, No. 3, pp. 233-244 (2021)
- [7] 池田新介・大竹文雄・筒井義郎, “時間割引率：経済実験とアンケートによる分析”, 淡路島行動経済学ワークショップ提出論文 (2005)
- [8] Yamauchi Taisei, Takami Kyosuke, Flanagan. Brendan, Ogata Hiroaki, “Nudge Messages for E-Learning Engagement and Student's Personality Traits: Effects and Implication for Personalization”, 30th International Conference on Computers in Education Conference, Proceedings (2022)
- [9] 池田新介, 筒井義朗, “アンケート調査と経済実験による危険回避度と時間割引率の解明”, 大阪大学社会経済研究所,
- [10] 佐藤一夫, “プライバシ保護を考慮したケータイ行動ログの利活用について”, 情報処理, vol. 50, No. 7 (2009)
- [11] 佐藤哲, “アクセスパターンのクラスタリングによる Web ログからのユーザ属性推定”, 情報科学技術フォーラム (2016)
- [12] Shipra Agrawal, Navin Goyal, “Thompson Sampling for Contextual Bandits with Linear Payoffs”, the 30th International Conference on Machine Learning, vol. 28, pp. 127-135 (2013)
- [13] <https://github.com/pycaret>

法人向け運動促進アプリケーションにおける インセンティブが身体活動に及ぼす影響の分析

斎藤虎太郎 ^{†1} 江口直輝 ^{†2} 荒川豊 ^{†3} 中村優吾 ^{†4}
九州大学 九州大学 九州大学 九州大学

1. はじめに

身体活動を習慣化することは、生活習慣病の予防、メンタルヘルス、生活の質 (QOL: Quality of Life) 及びウェルビーイングの改善において効果的である。厚生労働省が示した健康日本 21(身体活動・運動)[1] によると、日本人の 3 人に 1 人以上が運動不足であり、歩行を中心とした身体活動を増加させるように心掛けることで、日常生活において身体活動量を増やす必要があり、日本人の歩数を 1,200~1,500 歩/日増大させることが具体的な目標として示されている。また、令和元年国民栄養調査 [2] では、運動習慣改善の意思について、「関心はあるが改善するつもりはない」と回答した人の割合が最も高く、男性 23.9 %、女性 26.3 %と、約 4 人に 1 人が運動習慣について改善する意思がないことが示されており、日本人の歩数や身体活動の意欲を向上する方法について議論することは、現在も日本の健康問題上非常に重要なことであることが示唆されている。

こうした問題を解決するには、人々が何らかの動機づけによって、身体活動に関する行動変容を起こすことが必要である。厚生労働省の示す行動変容ステージモデル [3] によると、人の行動変容は、無関心期、関心期、準備期、実行期、維持期の 5 つのステージを通ると考えられている。人々が身体活動に関する行動変容を起こすように動機付けるには、その人が今どのステージにいるかを把握し、それぞれのステージに合わせた働きかけが必要とされている。

近年、身体活動に関する行動変容を促す手段として、スマートフォンのヘルスケアアプリなどにゲームの要素を取り入れるゲーミフィケーションが注目されている [4, 5]。スマートフォンアプリによる介入は身体活動量の増加に影響する可能性があることが示されている [6] が、研究対象のアプリの多くは研究用にデザインされたものであり、市場に出回るアプリの多くは効果が十分検証されていない [7]。つまり、ゲーミフィケーションを活用したアプリを、市場に出回

る程の規模感を持って社会実装した場合の介入の有効性については、十分明らかになっていない。

Agileware Inc. によって開発された KIWI GO^{*1}は、歩数データと心拍データを利用した多様なゲーミフィケーションによって、組織単位での身体活動促進を目的としたスマートフォン向けアプリケーションである (図 1)。KIWI GO は、組織単位で加入する福利厚生サービスの 1 つであり、選択されたプランに応じて、従業員に還元される総額インセンティブが異なっている。また、ここでの組織は、行動変容ステージモデルの 5 つのステージのいずれかに該当する人々で構成されていると考えられ、KIWI GO には、こうした様々な層の行動変容を動機づけるための多様な仕組みが実装されている。また、動機づけには、インセンティブの付与などによって得られる外発的動機づけと、その人自身の興味や関心、意欲によって得られる内発的動機づけの 2 種類が存在する。KIWI GO 提供するゲーミフィケーションの仕組みは、外発的動機づけに当たるものであるが、予算によってその上限は決められており、限界がある。提供できる外発的動機づけの中でやりくりして、内発的動機づけにつなげていく必要がある。

Agileware Inc. において、KIWI GO を導入したところ、実際に約 7 割の社員の運動量が増加したことが示されており、KIWI GO の利用が日々の身体活動の意欲向上に効果的であることが示唆されている。そこで本研究では、様々な仮説に基づいて設計された KIWI GO の多様なゲーミフィケーションやインセンティブが、実際にどのようにして利用者の身体活動の意欲向上に機能しているのかを明らかにする。

2. KIWI GO について

Agileware Inc. によって開発された KIWI GO は、歩数データと心拍データを利用した多様なゲーミフィケーションによって、組織単位での身体活動促進を目的としたスマートフォン向けアプリケーションである。KIWI GO の利用には組織単位での加入が必要となるが、KIWI GO には、利

Analysis of the Effects of Incentives on Physical Activity in a Corporate Exercise Promotion Application

^{†1} KOTARO SAITO, Kyushu University

^{†2} NAOKI EGUCHI, Kyushu University

^{†3} YUTAKA ARAKAWA, Kyushu University

^{†4} YUGO NAKAMURA, Kyushu University

*1 <https://kiwi-go.jp/>



図 1 KIWI GO のアプリ画面

用者の行動変容を動機づけるための多様なゲーミフィケーションの仕組みが実装されている。

ここでは、KIWI GO に実装されているインセンティブ機能について説明する。KIWI GO では、運動によってポイントが貯まるとともに、ポイントやアクティビティに応じた 5 つのインセンティブが用意されている。アプリ内の通貨は、100 コインで 1 円相当の価値である。

• 連続ログインボーナス

- 狹い：毎日アプリを起動させることで身体活動の意欲を高める
- 報酬：10 コイン ~ 100 コイン

• 目標達成ボーナス

- 狹い：目標達成の喜びによって運動強度の高い有酸素運動を促進する
- 1000 歩達成ごとに 400 コイン、運動強度の高い運動 5 分ごとに 800 コイン

• 1 日単位での目標達成ボーナス

- 狹い：目標達成の喜びによって身体活動の継続を促進する
- 5000 歩達成ごとに 1200 コイン

• ごほうびガチャ

- 狹い：ジュース一本程度のご褒美によって身体活動の意欲を高めたい
- 1 週間に 1 度ランダム抽選 (5000 コイン or 16000 コイン)

• ドリームガチャ

- 狹い：豪華なご褒美によって身体活動の意欲を高めたい 口コミによる利用者の拡大
- 1 月に 1 度ランダム抽選 (Amazon ギフト 10000

円 or はずれ)

月々に得られるコインの上限は、会社が契約しているプランに依存しており、月の途中で上限に達する利用者も多い。また、歩数のデータはウェアラブルデバイスやモバイル端末から取得される。ポイントインセンティブの他にも、ギルド機能やギルド内でのイベント機能なども実装されており、同じ組織に属する利用者同士を結びつけることで、利用者の行動変容につなげることも、KIWI GO の狙いである。

3. 仮説と分析手法

本章では、本研究における仮説と分析手法について説明する。

3.1. 仮説

本研究は、様々な仮説に基づいて設計された KIWI GO の多様なインセンティブやゲーミフィケーションが、実際にどのようにして利用者の身体活動の意欲向上に機能しているのかを明らかにするために、以下の 2 つの仮説を立て、検証した。

- 仮説 1：アプリをインストールすることで身体活動の意欲が高まる
- 仮説 2：ドリームガチャに当選することで身体活動の意欲が高まる

3.2. 対象データと分析手法

本研究では、2021 年 9 月から 2022 年 9 月の 1 年間で KIWI GO を利用した 1,253 名のアプリ利用データを対象とし分析を行った。アプリ利用データは、利用者の歩数データ、ガチャの当選データ、ポイント交換の履歴などから構成されている。利用したデータや外れ値として除外したデータ、解析手法は検証する仮説ごとに異なるため、以降で説明する。

3.2.1 仮説 1

仮説 1 の検証では、日々の運動レポートのデータを利用した。日々の運動レポートには記録された日付、合計歩数、ワークアウトの合計時間などのデータが含まれている。利用者がアプリをインストールする際にモバイル端末に蓄積されている過去 30 日分の歩数データがアプリのデータベースに収集されるため、アプリインストール前の 1 日ごとの歩数データも運動レポートに含まれている。ウェアラブルデバイスやモバイル端末を所持していない日があることを考慮して、1 日ごとに集計された歩数データのうち上下 5 % を削除した。さらにアプリインストール前のデータに関して、アプリの不具合により数日しか歩数データを収集でき

ていない利用者も含まれていたため、2週間以上の歩数データがある 273 名の利用者のデータを分析対象とした。これらのデータに対し、アプリインストール前後で歩数分布と平均歩数の変化の比較を行った。また、アプリインストール後の経過に関する分析として、0日から 30 日後、31 日から 60 日後、61 日から 90 日後の 4 つの時期の平均歩数について比較する分析を行った。さらに、対象の 273 人に対し、アプリインストール後の歩数の推移が、アプリインストール前の歩数によって変化しているかについての分析を行った。具体的にはアプリインストール前の利用者ごとの歩数データを平均歩数によって 4 分割し、それらの利用者群がアプリインストール後の 0 日から 30 日後、31 日から 60 日後、61 日から 90 日後の 4 つの時期において平均歩数がどのように推移していくのかについて分析を行った。

3.2.2 仮説 2

仮説 2 の検証では、KIWI GO 利用者の日々の運動レポートとドリームガチャ当選者のデータを利用した。日々の運動レポートには記録された日付、合計歩数、ワークアウトの合計時間、ドリームガチャ当選者のデータには、当選した日付、当選した利用者の ID、景品の種類などのデータが含まれている。これらのデータに対してまず、ドリームガチャ当選者と非当選者の群の代表値の比較する分析を行った。具体的には、2021 年 9 月から 2022 年 9 月の 1 年間で、当選者と非当選者の群の 1 日あたりの平均歩数に有意な差が認められるかどうかについて検証した。ウェアラブルデバイスやモバイル端末を所持していない日があることを考慮して、1 日ごとに集計された歩数データのうち上下 5 % を削除し、残った 988 名の利用者のデータを分析対象とした。分析対象とした 988 名のうち、当選者は 142 名、非当選者は 846 名であった。次に、ドリームガチャ当選データに対して、ドリームガチャ当選後の歩数の推移について分析を行った。具体的には、ガチャ当選日前、ガチャ当選日の 0 日から 30 日後、31 日から 60 日後、61 日から 90 日後の 4 つの時期の平均歩数について比較する分析を行った。2021 年 9 月から 2022 年 9 月の 1 年間において、ドリームガチャの当選件数は合計 166 件であった。そのうち、当選日から 90 日以上の当選後の経過を分析するのに十分なデータが得られている 100 件の当選データを対象に分析を行った。さらに、ドリームガチャ当選データを用いて、ドリームガチャ当選後の歩数の推移が、当選前の歩数によって変化しているかについて、分析を行った。分析対象としたデータは前述のものと同様である。具体的には、当選者の群をガチャ当選日前の平均歩数が少ない方から多い方まで 4 分割し、4 つの群それぞれにおいて、ガチャ当選日前、ガチャ当選日の 0 日から 30 日後、31 日から 60 日後、61 日から 90 日後の 4 つの期間における、平均歩数について比較する分析を行った。

3.3. 統計処理

2 群間の値やベースラインの値、ベースラインからの変化量の比較には、データに対応があるパラメトリックのデータには対応のある t 検定を、データに対応があるノンパラメトリックのデータには Wilcoxon の符号順位検定を、データに対応がなく、2 群間の分散に差がないパラメトリックのデータには Welch の t 検定をそれぞれ用いた。また、結果は対応のないデータ群同士の比較では平均値、対応のあるデータ群同士の比較では平均値 ± 標準誤差で示した。

4. 結果と考察

本章では、各仮説の検証において得られた結果について説明する。

4.1. 仮説 1 の分析結果

インストール前後の歩数データに対し、インストール後 0 日から 30 日後、31 日から 60 日後、61 日から 90 日後の 4 つの時期の平均歩数について比較した結果を図 2 に示す。各期間ごとに一つ前の期間と比較したところ、アプリインストール前とインストール後 0 日から 30 日後の 1 日あたりの平均歩数に 894 ± 108 歩の有意な増大が認められた ($P < 0.01$)。さらに、31 日から 60 日後と 61 日から 90 日後の歩数について、1 日あたりの平均歩数に 112 ± 81 歩の有意な増大が認められた ($P < 0.05$)。一方で、0 日から 30 日後と 31 日から 60 日後においては有意な差は認められなかった。本結果より、アプリを導入したと直後は身体活動の意欲が高まることが明らかとなった。しかし、31 日から 60 日後と 61 日から 90 日後の歩数平均には有意差があることから、インストール直後の 30 日間をピークにして、次第に運動意欲が低下していく可能性があることが明らかとなった。

次に、アプリインストール前において、利用者ごとの歩数データを平均歩数によって以下のように 4 分割し、それらの利用者群がアプリインストール後の 0 日から 30 日後、31 日から 60 日後、61 日から 90 日後の 4 つの時期において平均歩数を推移を分析した。

- Q1 : 972 歩以上 3352 歩未満 69 人
- Q2 : 3352 歩以上 4913 歩未満 68 人
- Q3 : 4913 歩以上 6647 歩未満 68 人
- Q4 : 6647 歩以上 11512 歩未満 68 人

結果を図 3 に示す。歩数平均によって 4 分割し、さらにそれらの群において、各期間ごとにひとつ前の期間と比較した。その結果、Q1 には、インストール前とインストール後 30 日間において、1 日あたりの平均歩数に 1993 ± 128 歩

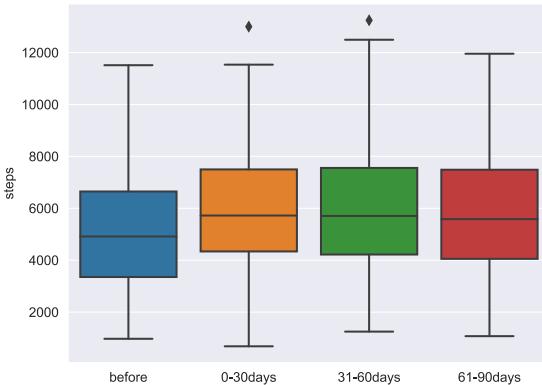


図2 アプリインストール前後の歩数推移

($P < 0.01$) 有意な増大が見られた。Q2には、インストール前とインストール後30日間において、1日あたりの平均歩数に 954 ± 89 歩 ($P < 0.01$) 有意な増大が見られた。Q3には、インストール後31日から60日後と61日から90日後の間ににおいて、1日あたりの平均歩数に 298 ± 80 歩 ($P < 0.01$) の有意な増大が見られた。Q4には、インストール後0日から30日後と31日から60日後の間ににおいて、1日あたりの平均歩数に 170 ± 88 歩 ($P < 0.05$) 有意な増大が見られた。本結果より、アプリ導入時において最も効果を発揮したのはアプリインストール前に平均歩数が最も少なかった群であるQ1であることが明らかになった。さらに、4群のうち2番目に平均歩数が少なかった群であるQ2においても、平均歩数が増加していた。これらの結果を踏まえ、KIWI GOをインストールすることは、元々それほど歩数が多くない利用者に対して、より効果的に身体活動を促進できるということが明らかとなった。一方で、アプリインストール前に最も平均歩数が多かった群であるQ4に対しては、アプリインストール前と0日から30日後で平均歩数が低下していることから、平均歩数が多い利用者には、アプリをインストールすることによる効果は少ないということがわかった。以上の結果を踏まえ、本アプリをインストールすることは、インストール前の身体活動量が少ない利用者に対しては、身体活動の促進に大きな効果を発揮するが、インストール前の身体活動量が多い利用者に対しては、身体活動の促進には効果を発揮しないことが明らかとなった。よって、本検証において、アプリをインストールすることで身体活動の意欲が高まるという仮説1は、インストール前の身体活動量が少ない利用者に対しては、正しいと言えると結論づけた。

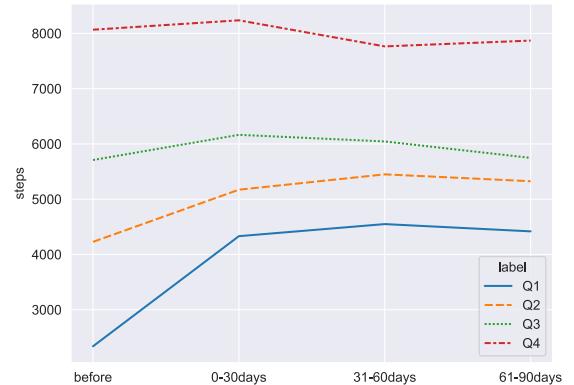


図3 Q1～Q4 の各群のアプリインストール前後における経過日数ごとの1日あたりの平均歩数

表1 当選者と非当選者の各群の代表値の比較

代表値	全体	当選者	非当選者
対象人数(人)	n=988	n=142	n=846
平均歩数/日(歩)	4694 ± 57	5029 ± 130	4638 ± 63

4.2. 仮説2の分析結果

当選者と非当選者の各群の代表値の比較を行った結果を表1に示す。

非当選者の群と比較し、当選者の群の1日あたりの平均歩数は391歩有意に大きかった($P < 0.01$)。

ドリームガチャ当選データに対して、当選日前、当選日の0日から30日後、31日から60日後、61日から90日後の4つの時期の平均歩数について比較した結果を図4に示す。

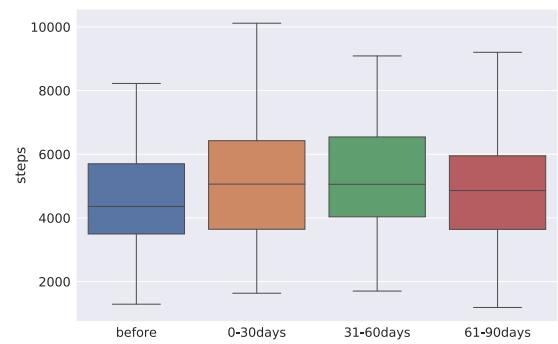


図4 ドリームガチャ当選後の経過日数ごとの1日あたりの平均歩数

各期間ごとにひとつ前の期間と比較したところ、当選日

前と 0 日から 30 日後の期間の 1 日あたりの平均歩数に 568 ± 142 歩の有意な増大が確認できた ($P < 0.01$)。そのほかの期間については有意な差は見られなかった。

当選者の群をガチャ当選日前の平均歩数によって以下のように 4 分割し、各群に関してガチャ当選日前、ガチャ当選日の 0 日から 30 日後、31 日から 60 日後、61 日から 90 日後の 4 つの時期の平均歩数について比較を行った。

- R1 : 1290 歩以上 3496 歩未満 25 人
- R2 : 3496 歩以上 4360 歩未満 25 人
- R3 : 4360 歩以上 5702 歩未満 25 人
- R4 : 5702 歩以上 11932 歩未満 25 人

結果を図 5 に示す。R1 群において、ガチャ当選前とガチャ当選後 30 日間の平均値には 1450 ± 247 歩 ($P < 0.01$) の有意な増大が認められた。R2 群において、ガチャ当選前とガチャ当選後 30 日間の平均値には 690 ± 196 歩 ($P < 0.01$) の有意な増大が認められた。それ以外のデータ群と期間に組み合わせには有意な差は認められなかった。

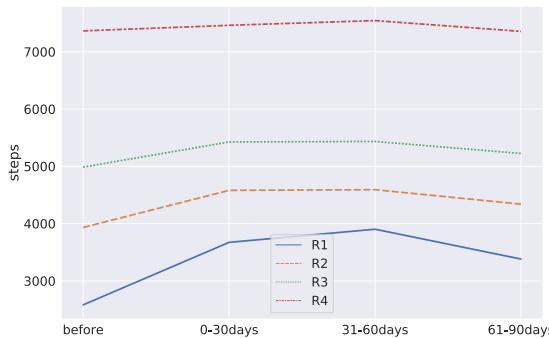


図 5 R1～R4 の各群のドリームガチャ当選前後における経過日数ごとの 1 日あたりの平均歩数

パ・リーグ 6 球団公式アプリ「パ・リーグウォーク」の利用が毎日の歩数に与える影響を調査した研究 [5] では、パ・リーグウォークのインストール前後において、アプリ利用者の 1 日あたりの平均歩数が 579 歩増加したことが報告されており、ドリームガチャ当選者の歩数増加と同程度であった。ポケモン GO が身体活動に与える影響について調査した研究 [8] では、アプリインストール直後は歩数増加が見られるが、時間経過とともに歩数増加が次第に弱まり、平均的には 6 週間で利用開始前の水準の歩数に戻るなど、効果が一部の利用者に限られている可能性も指摘されており [8]、本検証の結果と類似した傾向であった。また、本検証の結果は、ドリームガチャ当選後 30 日間での歩数の増加量が、当選前の歩数が少ない R1 群と R2 群において特に大きいことを示

しており、ドリームガチャの当選は、当選前にそれほど身体活動が活発でなかった利用者の身体活動意欲をより促進しているということが明らかになった。よって、本検証において、ドリームガチャに当選することで身体活動の意欲が高まるという仮説 2 は、人々身体活動を活発に行っていた利用者に対しては、本結果からは正しいとは言えず、人々身体活動が活発ではなかった利用者に対しては正しいが、一時的な効果にとどまるものであると結論付けた。

5. まとめと今後の展望

本章では、本研究におけるまとめと今後の展望に述べる。

5.1. まとめ

本研究では、Agileware Inc. によって開発された KIWI GO アプリの多様なゲーミフィケーションが、実際にどのようにして利用者の身体活動の意欲向上に機能しているのかについて、2 つの仮説を立て分析を行った。

分析の結果、アプリをインストールすることで身体活動の意欲が高まるという仮説 1 は、インストール前の身体活動量が少ない利用者に対しては正しく、ドリームガチャに当選することで身体活動の意欲が高まるという仮説 2 は、人々身体活動を活発に行っていた利用者に対しては、結果からは正しいとは言えず、人々身体活動が活発ではなかった利用者に対しては正しいが、一時的な効果にとどまるものであるという結論を得た。

日常歩数は身体活動量の指標の一つであり、死亡リスクや循環器疾患発症リスクとの関連が知られている [9]。歩数が 1,000 歩／日多いと 15 % 死亡リスクが低いという報告もあり [10]、今回 KIWI GO のアプリインストールやドリームガチャ機能の利用によって、歩数の増加が確認されたことは、健康寿命延伸の観点からも意義が大きい。また、KIWI GO が利用者に対してアプリをインストール、ドリームガチャに当選する、特に、人々身体活動が活発でない利用者において、身体活動を大きく促進していたことは、健康格差是正の観点からも特筆すべき点である。

5.2. 今後の展望

本研究では、KIWI GO のインストールと多様なゲーミフィケーションのうちの一つであるドリームガチャが身体活動に及ぼす影響について分析したが、KIWI GO にはドリームガチャの他にも、連続ログインボーナスや目標達成ボーナス、ギルド機能やイベント機能といったゲーミフィケーションの仕組みが実装されている。連続ログインボーナスや目標達成ボーナスは、利用者の身体活動の習慣化を促進する狙いのもと実装されている。本研究のドリームガチャ

に関する分析では、ドリームガチャは利用者の継続的な身体活動の促進には影響しているとは言えないことが明らかとなった。そのため、KIWI GO が利用者の継続的な身体活動の促進に影響しているかどうかを調べるには、連続ログインボーナスや目標達成ボーナスの機能に関して新たな仮説を立て、検証する必要がある。また、ギルド機能やイベント機能は、ギルド参加者の間での心理的つながりを利用し、行動変容を起こそうという狙いのもと実装されている。心理的つながりによる動機づけは、自己決定理論における外発的動機づけのうち、最も自己決定の程度が高いとされており、利用者の身体活動に与える影響が大きいことが推察される。そのため、ギルド機能やイベント機能に関しては、新たな仮説を立て、利用者の身体活動への影響度を検証したいと考えている。

KIWI GO には、本研究で着目した機能のほかにも、分析対象とすべき機能が多く残されている。今後は、本研究で扱えなかった機能に関しての分析にも取り組みたいと考えている。本研究での取り組み成果や今後の取り組み成果が、将来的に、より多様な人々の身体活動を促進するための新たなアイデアが生まれるきっかけとなれば幸いである。

謝辞 本研究の遂行にあたり、多数のデータや資料の提供、サポートをいただいた Agileware Inc. の皆様に深く感謝を申し上げます。

参考文献

- [1] 厚生労働省健康日本 21(身体活動・運動) (2000). https://www.mhlw.go.jp/stf/newpage_14156.html.
- [2] 厚生労働省令和元年度国民栄養調査 (2019). https://www.mhlw.go.jp/stf/newpage_14156.html.
- [3] 厚生労働省 e-ヘルスネット行動変容ステージモデル (2019). <https://www.e-healthnet.mhlw.go.jp/information/exercise/s-07-001.html>.
- [4] Mitesh, S., Patel, M., Small, D. and Harrison, JD, e. a.: Effectiveness of Behaviorally Designed Gamification Interventions With Social Incentives for Increasing Physical Activity Among Overweight and Obese Adults Across the United States The STEP UP Randomized Clinical Trial, *JAMA Internal Medicine* (2019).
- [5] Matsumoto, N., Kamada, M., Hayasi, H., Kawachi, I., Hirayama, T., Negishi, T. An official app for the Pacific Baseball Leagues “Pa-League Walk” : boost the team with your daily steps ! , J-STAGE, 日本健康教育学会誌 (2021).
- [6] Roxanne, G., Anne, M. M., Elon, J., van, O., Monique, S. and Evelyn, M.: The Effect of Physical Activity Interventions Comprising Wearables and Smartphone Applications on Physical Activity: a Systematic Review and Meta-analysis, *Sports Medicine-Open* (2018).
- [7] Amelia, R., Sarah, E., Ronald, P., Rachel, C., Jillian, R., Ilea, S., Alyson, C. and Carol, M.: Can Smartphone Apps Increase Physical Activity? Systematic Review and Meta-Analysis, *Journal of Medical Internet Research* (2019).
- [8] Katherine, B., H., Christian, S., Peter, U., Daniel, H., Ichiro, K. and Eric, B., R.: Gotta catch’ em all! Pokémon GO and physical activity among young adults: difference in differences study, *BMJ* (2016).
- [9] Pedro, F., Saint, M., Richard, P., Troiano, David, R. and Bassett, J.: Association of Daily Step Count and Step Intensity With Mortality Among US Adults, *JAMA* (2020).
- [10] I-Min, L., MBBS, Eric, J., Shiroma and Masamitsu, K.: Association of Step Volume and Intensity With All-Cause Mortality in Older Women, *JAMA Internal Medicine* (2019).

インセンティブとの組み合わせを前提としたナッジ効果測定のための実験設計方法

酒井 亮勢 * 落合 桂一 * 井手 崇博 * 大滝 亨 * 山田 曜 * 鈴木 喬 * 佐藤 弘之 *

川上 博 * 矢野 達也 * 大川 博生 * 和久井 圭祐 * 白井 拓也 * 荒川 豊 † 星野 崇宏 ‡

株式会社 NTT ドコモ * 九州大学 † 慶應義塾大学 ‡

【キーワード】ナッジ、インセンティブ、行動変容、行動経済学、行動インサイト

1. はじめに

2019年12月頃に発生した新型コロナウイルス感染症(COVID-19)は、2023年1月現在も収束の見通しが立たない[1]。国や各自治体、メディアからアルコール消毒やマスク着用等の感染症予防対策が推奨されている他、クラスターの発生防止の観点から厚生労働省より3密回避[2]が推奨されている。

各自治体や各企業がテレワークによる外出や対面の会議を避けることを推進しているものの、社会経済活動を継続するためには外出抑制は限度があり、感染症拡大の抑制と社会経済活動の両立が重要である。一定の外出量がある前提で3密が重ならないようにするために、密集場所の回避、言い換えると混雑回避が必要不可欠である。具体的な混雑回避の手段としては、移動時や商業施設利用時の混雑時間帯の回避や飲食店でのテイクアウト等が考えられる。つまり、これらの行動変容を促すことで社会全体の混雑を緩和させることができると可能である。

しかし、混雑回避を目的とした単なる呼びかけでは十分な行動変容を促すことは難しい。行動変容を促すための従来研究は様々あり、インセンティブ[3][4]によるものや人の心理に働きかける行動経済学の手法であるナッジ[5][6][7]、ゲームデザインやルールをゲーム以外に応用するゲーミフィケーション[8][9]等がある。理想としては、“そっと後押しする”という意味を持つナッジで当人に認知させることなく自然に行動を促すことが望ましい。

一方で、ナッジは万能でないことも知られている。実際、“ナッジは効果があるもののその効果は小さい”と考える経済学者もあり[10]、単純なナッジの効果だけには限界があることが報告されている[11][12]。そこで、より行動変容を促すためにナッジに効用の高いインセンティブを組み合わせる手段が考えられ、実際にその組み合わせによる介入の有効性が報告されている[13]。

ナッジとインセンティブの組み合わせは様々な場面で応

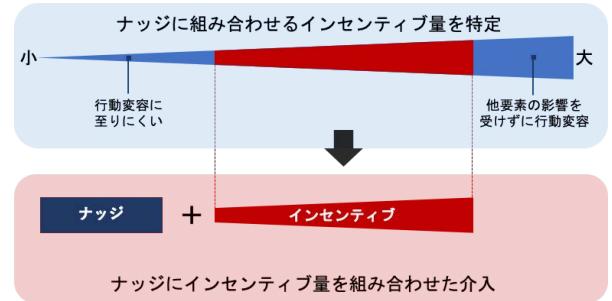


図1 インセンティブを組み合わせたナッジの効果測定のための実証設計イメージ

用可能な一方で、組み合わせる際のインセンティブ設計には以下の3点を考慮する必要がある。まず1点目は、インセンティブの効用が高くその影響が大きいという点である。具体的には、インセンティブが無い場合や極めて少ない場合は高い行動変容効果は期待できず、また、極めて多い場合はナッジとは関係なく高い確率で行動変容に至るためナッジの効果測定が困難となる。2点目は、サービス観点においてインセンティブ原資は有限である点である。そして3点目は、アンカリング効果[14]が働き、最初に大きいインセンティブを提示することでインセンティブを小さくした際の行動変容効果が減少してしまう。これら3点を考慮した上でナッジの効果検証を行えることは有意義であると言える。

そこで本論文では、インセンティブを組み合わせることを想定したナッジの効果検証を行うまでの実験設計方法を提案する。具体的には、インセンティブのみで効果測定を行うフェーズとナッジにインセンティブを組み合わせて効果測定を行うフェーズの2段階の実験設計である(図1)。これにより、ナッジによる行動変容効果向上の余地を残しつつ行動変容し得るインセンティブ量を特定した上で、ナッジの効果測定を可能とする。

本論文は次のように構成される。第二章では、ナッジとインセンティブについて述べ、第三章でナッジとインセンティブを組み合わせた実験設計方法について論じる。第四

Manuscript Format for Kickoff Symposium of IPSJ Behavior Transformation by IoT (BTI)

章でその実験設計に基づき実施した実証実験について概説し、最後に第五章で本論文についてまとめる。

2. 理論背景

2.1. ナッジ

2.1.1 ナッジの概要

Richard H Thaler らによって提唱されたナッジ (Nudge) とは、肘で軽くつくという意味を持ち、行動の意思決定の際に強制することなく自発的に当人や社会にとって望ましい行動を促す仕組みや仕掛けを表す [5]。

ナッジ理論は、人間は必ずしも合理的には行動しないが故に伝統的な経済学では説明困難な人間の経済活動を実証的に観測することで明らかにするという行動経済学の枠組みで提唱されたものである。その行動経済学の基礎となるのが、Daniel Kerneman らの提唱した二重過程理論 [15] であり、この理論によると人の思考、情報処理、意思決定は 2 つのシステムに分類される。1 つは、直感的で感情的、かつ自動的に処理するシステム 1 であり、もう一つは理論的かつ意識的にゆっくり処理されるシステム 2 である。つまり、本来このシステム 2 で物事の判断を行うことができれば、人は合理的に動くことができる。

しかし、人間は日常の中で様々な意思決定と行動をする必要があり、すべての物事を論理的に思考することは困難である。そのため、過去の経験や周囲の状況から半ば無意識的にシステム 1 の情報処理になることが多く、合理的ではない行動に至ってしまう。システム 1 は日常における意思決定の効率化に重要な一方で、時として損失や過失に繋がってしまう場合がある。例えば、衝動買いや SNS における浅慮な発言による炎上などである。

このように、半ば無意識の意思決定で社会的に好ましくない行動を行ってしまうことがあるため、外発的に当人や社会にとって望ましい方向へ行動変容を促すナッジが有効であるとされている。ここで留意したいのは、ナッジはあくまで個人の選択を強制することなくより良い意思決定を促す手法のことであり、個人が不利益を被ることや向社会的行動を困難にすることを促すスラッジ (Sludge) とは区別される。

2.1.2 ナッジの例

ナッジは通常、人の認知バイアスに基づき設計される。認知バイアスとは、周囲の環境やこれまでの経験、直感等の先入観による認知の偏りのことを表す。人はこの認知の偏りにより非合理的な意思決定をしてしまうが、ナッジは認知の歪みを認識するきっかけを与えたり、あるいは、この偏りを逆手に取りより良い行動を促すことを目的としている。

こうした認知バイアスを活用したナッジは様々あるが、本

論文では認知バイアスに基づくメッセージング（フレーミング効果）を扱う。フレーミング効果とは、情報の提示の仕方で受け取り方に変化が生じるバイアスである。言い換えると、内容が論理的に等価であっても、その言い回しにより印象が異なる。例えば、以下の 2 つの文章は等価であるためどちらを見ても同様の選択率であると考えられるが、実験によりポジティブな表現である前者を見た群の方が選択率が高い [16]。

- 術後 1 ヶ月後の生存率は 90 %ですが手術しますか？
- 術後 1 ヶ月の死亡率は 10 %ですが手術しますか？

2.1.3 ナッジの課題

より良い意思決定を促すために、日本では環境省が日本版ナッジ・ユニット (BEST:Behavioral Sciences Team) を立ち上げ、家庭における省エネや節電を推進する他、新型コロナウイルス対策におけるナッジの活用も進められている。このように、関係府省庁や地方公共団体、産業界で環境や健康、教育等の社会課題解決に向けてナッジの取り組みが推し進められる一方で、ナッジは万能ではなく効果が小さいと考える経済学者も少なくない [10]。人の意思決定は、置かれている状況や文化、習慣に大きく依存するため、“そっと後押しする” ナッジには限界がある。

こうした課題に対し、ナッジの効果を増強ないし持続可能なとすべく次世代のナッジ (Beyond Nudge) が世界的に検討されているが、その定義は提唱者ごとに異なり、まだ発展途上と言える。

2.2. インセンティブ

インセンティブとは、人の意思決定や行動を変容させる外発的な動機付けのための報酬を表す。インセンティブには、金銭的な物質的インセンティブと非金銭的な評価的インセンティブ、人的インセンティブ、理念的インセンティブ、自己実現的インセンティブがある [17]。本論文では、経済的な報酬を与える金銭的インセンティブを扱う。

インセンティブは動機付けのためのメジャーな手法の 1 つであり、禁煙や禁酒、運動行動の促進など、幅広い領域で適用可能であるが、行動変容を促すに足るインセンティブ量やその種類は行動の目的に依存することは明らかである。すぐに取り組める行動と手間と時間を要する行動で必要なインセンティブ量は異なり、目的に応じたインセンティブ設計が必要である。また、インセンティブ原資は有限であり、限られた原資で最大の行動変容効果を得られるように最適化していくことが重要である。

3. 実験設計

3.1. ナッジとインセンティブを組み合わせる際の課題

ナッジとインセンティブの組み合わせにおける相乗効果により高い行動変容率が期待される。実際、組み合わせた介入による行動変容効果を検証した研究がいくつか存在する。

例えば、エクササイズプログラムへの参加を促す研究において、参加回数が群内で上位 10 % の人に 20 ドルの報酬が与えられるコントロール群とそれにナッジ要素を組み合わせた群で比較検証をしている [13]。この検証においては、ナッジの有無で有意差を観測できていたものの、インセンティブ量の妥当性については検証されていない。つまり、インセンティブ量が 20 ドルではなかった場合、ナッジの有無で有意差を確認できなかった可能性がある。

ここでの課題としては、インセンティブ量が妥当でない場合は、行動変容効果がインセンティブによるものかナッジによるものか解釈が困難となる点である。具体的には、インセンティブ量が極めて多い場合、ナッジの有無によらずに高い確率で行動変容に至り、逆に極めて少ない場合、インセンティブによる効果向上が期待できなくなる。つまり、極めて多い量と少ない量を排除した妥当な量のインセンティブを組み合わせたナッジの効果検証が有意義である。

本来このようなインセンティブとナッジを組み合わせた検証は、ランダム化比較実験により行なうことが好ましい。しかし、インセンティブの種類の数を N_1 、ナッジの種類の数を N_2 とすると、検証群数は $N_1 N_2$ となり、一群あたりのサンプルサイズが小さくなり検証が困難になってしまう。言い換えると、妥当なインセンティブが事前に把握できており、その種類の数を N_3 とすると、 $(N_1 - N_3)N_2$ だけ検証群数を削減することが可能である。

3.2. 実験設計方法

本実験設計方法は、インセンティブ毎の効果を踏まえた効率的なナッジの効果検証を可能とする。具体的には図 1 の通り、インセンティブのみで効果測定を行い目的に応じたインセンティブ量を特定すること及び、インセンティブとナッジを組み合わせた効果測定の 2 つのフェーズで成り立つ。

最初のフェーズでは、目的の行動に対する妥当なインセンティブ量を特定するための効果測定を行う。そのために非常に少ない量からある程度多い量まで様々なインセンティブ量で効果検証を行う。その際、最初に与える情報が後の意思決定に影響を及ぼすアンカリング効果 [14] が働く可能性がある。つまり、最初に多いインセンティブを提示した場合、最初よりも少ないインセンティブの介入は行動変容効果が低下する可能性がある。そこで本提案では、はじめに少ないインセンティブを提示し、徐々にインセンティブ

量を増加させる。

インセンティブのみによる介入結果をもとに妥当なインセンティブ量を選択する。例えば、シンプルに妥当なインセンティブ量でナッジの効果検証を行いたい場合、インセンティブ全体の行動変容率を閾値とし、閾値に近い行動変容率のインセンティブ量を選択することが考えられる。もう一つ具体例を挙げると、ナッジを組み合わせることで、閾値に設定したインセンティブのみの介入と同等以上の行動変容効果が発揮可能であるという仮説の検証を行いたいとする。この場合、インセンティブ全体の行動変容率を閾値とし、閾値以下の行動変容率のインセンティブ量を選択することで検証を可能とする。

次のフェーズでは、選択したインセンティブ量にナッジを組み合わせた介入を行う。これにより、インセンティブによる行動変容効果を加味した上で効率的なナッジの効果検証を可能とする。

4. 実験

本章では、第 3 章で述べた実験設計方法に基づき実施した実証実験内容について提示する。

4.1. 実証実験概要

NTT ドコモと九州大学は、コロナ禍における混雑緩和を目的とした行動変容を促す実証実験を九州大学伊都キャンパスで 2021 年 4 月から 2022 年 1 月にかけて実施した。具体的にはランチタイムにおける混雑回避を想定し、キャンパス内にある複数の飲食店でランチタイム前（11:00～12:00）または後（13:00～14:00）に訪問を促すタイムシフト、ランチタイム（12:00～13:00）にイートインではなくテイクアウトを促すキャンペーン情報の配信を行なった。なお、被験者は実証実験への参加を希望した九州大学の学生約 1,500 名である。

実証実験専用アプリを通じてキャンペーン情報を配信し、被験者はタイムシフトまたはテイクアウトを行うことで d ポイントを受け取れる。行動に至ったかの定義は図 2 の通り、アプリのキャンペーン画面の“参加する”ボタンを押下後、指定の時間帯に対象店舗に設置された QR コードをスキャンし、イートインまたはテイクアウトをすることである。

本実証実験では、インセンティブとして d ポイント、ナッジとして PUSH 通知による認知バイアスを考慮したメッセージの効果を検証した。その際、ナッジの効果測定のために複数の通知文言を用意し、事前に被験者を群分けした。

なお、本実証実験は九州大学倫理委員会の承認（受承認番号：シス情認 2021-08-1）を受け、実施したものである。



図 2 実証実験における行動変容までの流れ

4.2. 検証方法

本検証において、インセンティブのみの効果測定を行う第一フェーズを 2021 年 4 月～2021 年 8 月、ナッジを組み合わせた効果測定を行う第二フェーズを 2021 年 10 月～2021 年 11 月に実施した。

第一フェーズにおける配布ポイント数は [1, 10, 50, 100, 150, 200] であり、週次でポイント数を増加させた。なお、週ごとのポイント数は [1, 10, 50] のように 3 つないし 2 つのレンジとなっており、そのレンジ内で日ごとの配布ポイントをランダムに選択した。

第一フェーズの検証の結果、1 ポイントではほとんど行動変容を促せないことが分かった。そのため、第二フェーズにおけるポイント数は [10, 50, 100, 150, 200] とし、ナッジを組み合わせた効果検証を行った（図 3）。

ナッジは認知バイアスを考慮したメッセージングによる介入と定義し、10 月、11 月それぞれで 3 つの検証群を設け、表 1 の通りの PUSH 通知本文を検証群ごとに出し分けた。なお、本フェーズにおいても、はじめのうちは少ないポイント数で配布し、徐々に増加させた。

表 1 PUSH 通知文言

PUSH 通知タイトル (共通)	
ランチタイムがお得になるキャンペーン！	
検証群 (10 月)	PUSH 通知本文
コントロール	{ 店舗 A } ^{*1} に行ってみよう！
ナッジ 1	今日は { 店舗 A } か { 店舗 B } を選ぶとお得
ナッジ 2	{ 店舗 A } で〇〇ポイントゲット
検証群 (11 月)	PUSH 通知本文
コントロール	{ 店舗 A } に行ってみよう！
ナッジ 1	【期間限定】ポイントを { 店舗 A } でゲットしよう
ナッジ 2	みんなで { 店舗 A } でご飯を食べてポイントを貯めよう

*1 {} 内の店舗名は今回の実証実験の対象となった 4 店舗の名称が表示される。

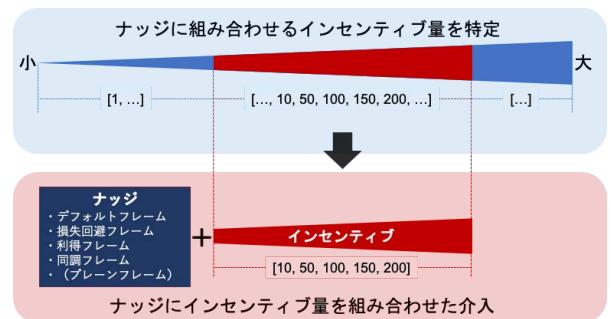


図 3 本検証におけるインセンティブ設計

4.3. 結果

まず、本検証の第一フェーズの結果を踏まえ第二フェーズにおけるインセンティブパターン数を 6 から 5 に削減したため、10 月、11 月それぞれでインセンティブとナッジの組み合わせが 18 パターンあったところを 15 パターンに削減できた。

本実験におけるコントロール群の平均行動変容率を基準とした行動変容受容率 [18] を表 2, 3 の通りまとめる。ナッジの効果検証結果については井手らの研究 [18] で報告した通りであり、結果の概要についてまとめると下記の通りである。

- コントロール群と比較して、デフォルトと呼ばれるナッジの介入群の平均行動変容受容率が約 20% 向上
- コントロール群と比較して、損失回避と呼ばれるナッジの介入群の平均行動変容受容率が約 33% 向上
- 検定の結果、ナッジ群（デフォルトナッジ、損失回避

表 2 10月における各週の行動変容受容率の平均

検証群	week1	week2	week3	week4	平均
コントロール	0.77	1.03	1.11	1.08	1.00
ナッジ 1	0.89	1.04	1.38	1.45	1.19
ナッジ 2	0.74	0.94	1.20	1.13	1.00

表 3 11月における各週の行動変容受容率の平均

検証群	week1	week2	week3	week4	平均
コントロール	0.87	0.83	1.22	1.11	1.00
ナッジ 1	0.90	1.02	1.55	1.80	1.33
ナッジ 2	0.73	0.63	1.20	1.30	1.00

ナッジ) はコントロール群よりも優位に行動変容率が高い

- ナッジを組み合わせる場合もインセンティブ量と行動変容率は相関がある

以上のことから、本提案の実証実験方法はインセンティブを組み合わせることを前提としたナッジ効果の検証において意義のあるものであると言える。

5.まとめ

本稿では、ナッジとインセンティブを組み合わせることを想定したナッジの効果検証のための実験設計方法について述べた。具体的には、目的に対して行動を変え得る妥当なインセンティブ量を測定し、それに基づきナッジを組み合わせた検証を行うという 2 段階の実験設計方法を提案した。本提案に基づき実験設計を行ったコロナ禍における混雑回避を促す実証実験において、ナッジの効果検証を効率的に行えたことと実際に有意なナッジの効果を確認できしたことから、本提案による実験設計方法が有効であるという示唆を得られた。

今後はインセンティブとナッジを組み合わせた効果検証だけではなく、インセンティブ量ごとのナッジへの及ぼす影響を明らかにし、より行動変容を促せる最適化手法を検討していきたい。

謝辞 本実証実験に参加した九州大学の学生および実証実験の運用に関わった株式会社 イマーゴ iQLab プロジェクト関係者に感謝の意を表す。

参考文献

- [1] 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, pp. D-K28_1 (2020).
- [2] 田中重人：「3 密」概念の誕生と変遷: 日本の COVID-19 対策とコミュニケーションの問題, 東北大学文学研究科研究年報= The Annual Reports of Graduate School of Arts and Letters Tohoku University, Vol. 70, pp. 140–116 (2021).
- [3] Keiichi, O., Hiroyuki, S., Hiroshi, K., Takahiro, I., Toru, O., Akira, Y., Tatsuya, Y., Hiroki, O., Takuya, S. and Yutaka, A.: Encouraging Crowd Avoidance Behavior using Dynamic Pricing Framework Towards Preventing the Spread of COVID-19, *2022 IEEE International Conference on Big Data (IEEE BigData 2022)* (2022).
- [4] 陳美怡, 幡井皓介, 西山勇毅, 濱崎薫ほか: 感染症予防行動を促進させるインセンティブモデルに関する一検討, 研究報告モバイルコンピューティングと新社会システム (MBL), Vol. 2021, No. 3, pp. 1–7 (2021).
- [5] Leonard, T. C.: Richard H. Thaler, Cass R. Sunstein, Nudge: Improving decisions about health, wealth, and happiness: Yale University Press, New Haven, CT, 2008, 293 pp, \$26.00 (2008).
- [6] Sugden, R.: On nudging: A review of nudge: Improving decisions about health, wealth and happiness by Richard H. Thaler and Cass R. Sunstein (2009).
- [7] 板谷祥奈, 竹内穂波, 松村真宏：「ひじでつく」ナッジ、「そそる」仕掛け (2018).
- [8] Hamari, J., Koivisto, J. and Sarsa, H.: Does gamification work?—a literature review of empirical studies on gamification, *2014 47th Hawaii international conference on system sciences*, Ieee, pp. 3025–3034 (2014).
- [9] 根本啓一, 高橋正道, 林直樹, 水谷美由起, 堀田竜士, 井上明人ほか: ゲーミフィケーションを活用した自発的・持続的行動支援プラットフォームの試作と実践, 情報処理学会論文誌, Vol. 55, No. 6, pp. 1600–1613 (2014).
- [10] Gneezy, U., Meier, S. and Rey-Biel, P.: When and why incentives (don't) work to modify behavior, *Journal of economic perspectives*, Vol. 25, No. 4, pp. 191–210 (2011).
- [11] Ito, K., Ida, T. and Tanaka, M.: Moral suasion and economic incentives: Field experimental evidence

- from energy demand, *American Economic Journal: Economic Policy*, Vol. 10, No. 1, pp. 240–67 (2018).
- [12] 依田高典, 石原卓典: 金銭的インセンティブとナッジが健康増進に及ぼす効果: フィールド実験によるエビデンス, 行動経済学, Vol. 11, pp. 132–142 (2019).
- [13] Zhang, J., Brackbill, D., Yang, S., Becker, J., Herbert, N. and Centola, D.: Support or competition? How online social networks increase physical activity: A randomized controlled trial, *Preventive medicine reports*, Vol. 4, pp. 453–458 (2016).
- [14] Tversky, A. and Kahneman, D.: Judgment under Uncertainty: Heuristics and Biases: Biases in judgments reveal some heuristics of thinking under uncertainty., *science*, Vol. 185, No. 4157, pp. 1124–1131 (1974).
- [15] Stanovich, K. E. and West, R. F.: Advancing the rationality debate, *Behavioral and brain sciences*, Vol. 23, No. 5, pp. 701–717 (2000).
- [16] 井田政則, 小橋眞理子: BIS/BAS が意思決定フレーミング効果に及ぼす影響, 立正大学心理学研究年報= The Journal of Psychology Risho University, Vol. 11, pp. 1–12 (2020).
- [17] 石塚浩ほか: 非営利ゆえの強さ: 理念的インセンティブ・システムとしての NPO, 経営論集, Vol. 2, pp. 1–14 (2016).
- [18] 井手崇博, 大滝亨, 山田暁, 佐藤弘之, 落合桂一, 川上博, 矢野達也, 大川博生, 和久井圭祐, 白井拓也, 荒川豊: 新型コロナウィルス感染防止に向けた混雑回避行動を促進する行動変容技術の検証, 情報処理学会 IoT 行動変容学研究グループキックオフシンポジウム (2022).

製造業における DX 推進に向けたリテラシー醸成の取り組み

磯田祐世^{†1}

株式会社 UACJ

浅田勝義^{†2}

株式会社 UACJ

山本佑樹^{†3}

株式会社 UACJ

1. はじめに

(株)UACJは、アルミニウムの軽圧メーカーである。当社主力設備である圧延機は、第三次産業革命(third industrial revolution, 3IR)の圧延機自動制御化により生産能力を飛躍的に向上させてきた。現在は、情報技術が主役の第四次産業革命(forth industrial revolution, 4IR)と呼ばれる時代であり、データ・デジタル技術を活用した産業構造の変化が起きつつある。このような時代において企業が競争上の優位性を確立するには、変化する社会や顧客の情報を常に捉え、デジタルトランスフォーメーション(以下、DX)と呼ばれる変革が企業に求められている。当社においても、産業 IoT の活用を含む DX を推進している。

国内外企業の DX への取り組み状況は以下の通りである。DX 白書 2023[1]によると、日本での取り組み割合は 2021 年度調査の 55.8%から 2022 年度調査は 69.3%に増加しており、米国の 77.9%に近づいている。しかしながら、成果創出の割合については、米国の 89.0%に対して日本は 58.0%と依然としてその差は大きくおくれをとっている。具体的には、日本企業はデジタイゼーションやデジタライゼーションといった領域では成果をあげてはいるものの、価値創出やビジネスモデルの変革といったトランクスフォーメーションは不十分であり、「X=変革」に向けた更なる取り組みの深化が必要であると報告されている。

ところで、当社では DX を行動変容の一つの形態と捉えているが、どのような変容を目指すべきかが大きな議題となっている。その要因として、変容を想像・創造する力(=リテラシー)がないことが挙げられる。企業が DX を実現するには、経営層を含め従業員一人ひとりが DX を自分事と捉え、全てのビジネスパーソンに DX に関するリテラシーを身につけた人材の確保・育成が必要であると考えられる。

様々な製造業において人材育成の取り組みがなされており、当社と同じ金属メーカーである日本製鉄では、総合職を対象とした育成プログラムや高度なデータ分析・活用スキルを備えるデータサイエンティストの育成に着手している[2]。当社においては、総合職だけでなく現場作業者も含めた全従業員を対象に、職位や役割に応じて必要なリテラシーを身につけるという位置づけで教育方法を設計し、実践している。その取り組みの一部を紹介する。

2. DX 推進の現状と課題を踏まえた方策

DX リテラシー教育実践にあたって、現状把握として当社従業員の DX に関する意識を調査した。これをもとに教育に必要な条件を明確にし、その方策を立案した。

2.1 従業員の DX に関する意識

当社従業員の DX に関する意識調査結果を表 1 に示す。これは、当社のこれまでの DX 活動の振り返りの中で得られた情報をリテラシー／マインドの観点で整理したものである。当社との比較対象として、一般企業における意識調査結果についてもあわせて列挙する。リテラシーについては帝国データバンク[3]、マインドについては RECEPTIONITST 社の調査結果[4]である。

表 1 より、当社従業員の DX リテラシーおよび DX に対するマインドが当社における課題であることがわかる。当社の DX リテラシーは従業員それぞれが業務内容に応じて身につけ「digital-ready(デジタル技術を理解しレベルに応じた活用ができる状態)」となる必要がある。しかしながら、現状は DX に関するスキル不足やそれにより向かうべきゴールを描くのが困難な状態である。DX に対するマインドについては、新人は現状維持バイアスがない状態であるものの中堅／管理職以上の従業員となると現状を維持したい・自分には無関係といったマインドであることがわかった。これらは一般企業の意識調査結果と同様の傾向であった。

表 1 UACJ における DX 推進に関する意識調査結果

Table1 Perceptions of DX promotion in UACJ.

対象	リテラシー	マインド
UACJ従業員	<ul style="list-style-type: none"> ・スキルがないとゴールを設定できない ・DXのゴールがわからない ・DXと言わても何から始めたらよいかわからない 	<p>＜新人＞</p> <ul style="list-style-type: none"> ・DXに対してニュートラルな思考 <p>＜中堅／管理職以上＞</p> <ul style="list-style-type: none"> ・DXの重要性は理解するが、困っていないため手間をかけて変化したくない ・DXリテラシーがある人がやってしまえばいい、誰かがやってくれる
一般企業	<ul style="list-style-type: none"> ・必要なスキルやノウハウがない ・対応する時間が確保できない ・どこから手をつけてよいかわからない ・データ活用の方針がない 	<ul style="list-style-type: none"> ・変化を受け入れられない ・組織風土が変わるので恐れている ・仕事を増やしたくない ・必要性を感じていない

Measures to foster literacy for DX promotion in the manufacturing industry

^{†1} YUSEI ISODA, UACJ Corporation^{†2} KATSUYOSHI ASADA, UACJ Corporation^{†3} YUKI YAMAMOTO, UACJ Corporation

2.2 DX リテラシー教育が成功する条件および方策

DX に対して現状維持バイアスのない当社総合職一年目社員を教育対象とした。表 2 に DX リテラシー教育が成功する条件および方策を示す。③④の自発性／心理的安全性については、一般的な新入社員の特徴[5]をもとに条件および方策を模索した。

各条件に対する方策の詳細は以下の通りである。

表 2 DX リテラシー教育が成功する条件および方策

Table2 Prerequisites and measures for DX literacy education.

教育が成功する条件	方策
①人数効率	集合研修形式で教育を実施
②学習効果	・ハンズオンによる学習 ・グループ討議の実施 ・デュアルディスプレイの使用
③受講者の自発性を促す	・チューターの配置 ・チューターによるアサーティブコミュニケーションの実施 ・返報性の利用
④心理的安全性 活発な質問、良い質問を引き出す	

① 人数効率

局所的な育成による人海戦術が通用しないこと、全社員に教育が必要なため、集合研修の形式で中長期的かつ体系的に DX 人材を育成することとした。2022 年度は全社実施に先だって行われたトライアルの位置づけである。

② 学習効果

単なる座学ではなく、「理論×実践」のハンズオン学習により表 3 に示すデータ活用スキルを把握できるようにすることで学習効果を高めるようにした。またハンズオン学習に加え、図 1 に示すような実施形態でグループ討議することで、学習内容の理解促進を促すようにした。快適性という観点では、PC とモバイルモニタを用いてデュアルディスプレイにすることで、コーディングや練習問題を効率良く解けるようにした。

表 3 DX リテラシー教育のコンテンツ内容

Table3 Contents of the DX literacy education.

内容	
基礎編	1回目 Pythonとは、Pythonにおける変数、データ構造
	2回目 繰り返し文(for文、while文)、制御文(if文)
	3回目 関数
	4回目 データの前処理(Numpy,Pandas①)
	5回目 データの前処理(Pandas②)、データの可視化(matplotlib)、その他標準ライブラリ(os, pickle, read, random, 正規表現)
	6回目 探索、ソート、数理最適化
応用編	7回目 統計学①
	8回目 統計学②
	9回目 機械学習
	10回目 機械学習
	11回目 機械学習
	12回目 SQLAlchemy (データベース)

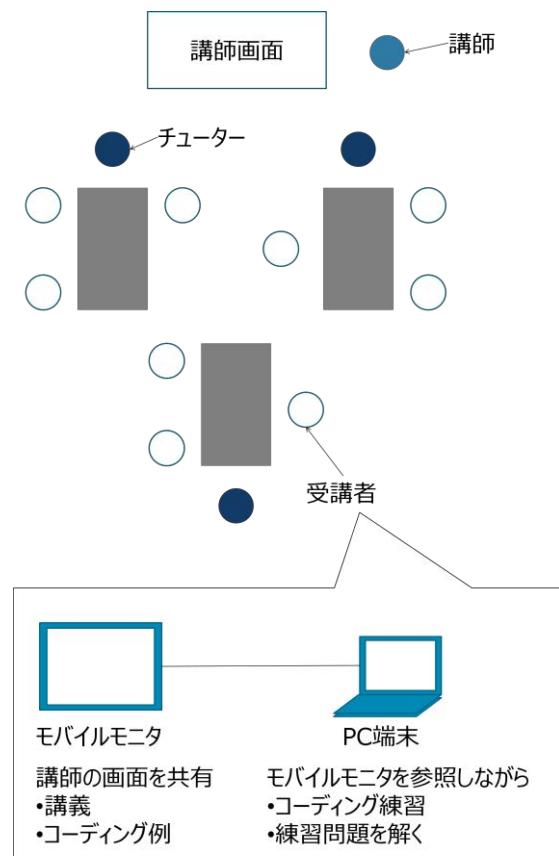


図 1 DX リテラシー教育の実施形態

Figure1 Setup of the DX literacy education.

③ 受講者の自発性を促す

受講者の自発性を促すため、各グループに受講者と歳の近いチューターを配置し、アサーティブコミュニケーションを導入した。アサーティブコミュニケーションとは、以下の特徴を持った相互尊重コミュニケーションスタイルのことである[6]。

- 場に沿った適切な表現を選択できる
- 相手の気持ちや意見を受け止め大切にする
- 自分の気持ちや意見をはっきりと主張できる
- 相手と意見が対立してもお互いが納得できる結論を導くことができる

図 2 にアサーティブコミュニケーションの例を示す。

近年の新入社員は、指示待ちで主体性が低いと言われている。そこで、受講者の様子やその場の雰囲気・状況に応じて誠実に指導・支援するといったアサーティブコミュニケーションを用いる。その結果、受講者の返報性に作用して自発性を促すことができると考えている。他者から何らかの恩恵を受け、お返しをしなければいけないという感情を抱くことを返報性の原理といい、デニス・リーガンによって提唱されたものである[7]。

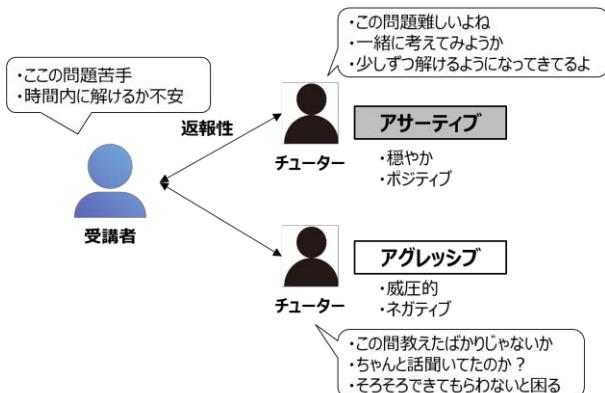


図 2 アサーティブコミュニケーションの例

Figure2 Examples of assertive communication.

④ 心理的安全性

自分の考えや気持ちを誰に対してでも安心して発言できる状態のことを心理的安全性という[8,9]。心理的安全性に対する方策として、③の自発性と同様に各グループにチューターを配置し、チューターにはアサーティブコミュニケーションを意識させた。教育の場において、受講者から質問行動を引き出すのがチューターの役割である。教育の場における質問行動は、学習内容の理解促進・思考の深化ができるため重要である。しかし、近年の新入社員は質問行動に消極的な傾向がある。秋田らは、大学生の質問行動に寄与する要因について調査しており、「つまらない質問で馬鹿だと思われたくない」、「目立ちたくない」といった理由が、質問への抵抗感と関連することを報告している[10]。これと同様に、自分の行動に対して上司・先輩がネガティブな反応をするのではないかという懼れから質問行動に消極的になると思われる。よって、図 2 に示すアサーティブコミュニケーションを用いることで安心して発言できる状態をつくりだし、質問行動を引き出すこととした。

上記を盛り込んだ教育効果の確認として、受講者の追跡調査および事後アンケートを実施した。

3. 結果と考察

3.1 受講者の自発性に関する追跡調査

DX リテラシー教育で学んだデータ活用スキルを自発的に業務で活用しているか確認するため、受講者へヒアリングを実施した。その結果を図 3 に示す。

全受講者の 40% にデータ活用スキルを用いて自職場課題を解決するといった DX に対する自発性という点で積極的な変容が認められた。このことから、今回開発した教育は DX リテラシー醸成においてある一定の効果が期待できると考えられる。行動変容のなかった受講者に対しては、さ

らなる追跡調査をするなどして別途施策を検討・実施するのが今後の課題である。

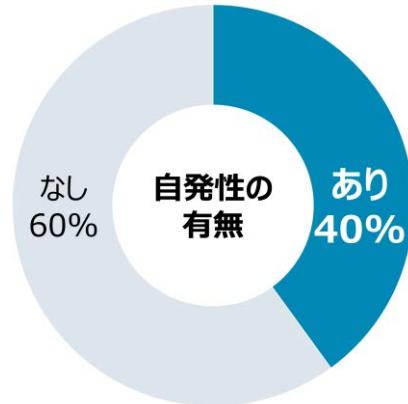


図 3 データ活用スキルを自発的に業務で活用している受講者の割合

Figure3 Percentage of participants who voluntarily use data utilization skills in their work.

3.2 満足度に及ぼすチューターの影響

DX リテラシー教育の受講者満足度に及ぼすチューターの影響を把握するため、教育終了後にアンケートを収集した。アンケート結果をもとに満足度とチューターに関する相関分析を実施した。その結果を図 4 に示す。

図 4 より、受講者満足度に対するチューターの相関係数は 0.72 となっており、全アンケート項目の中で最も相関が強いことがわかった。これより、DX リテラシー教育の良し悪しにチューターとのコミュニケーションが大きく作用していることがわかった。

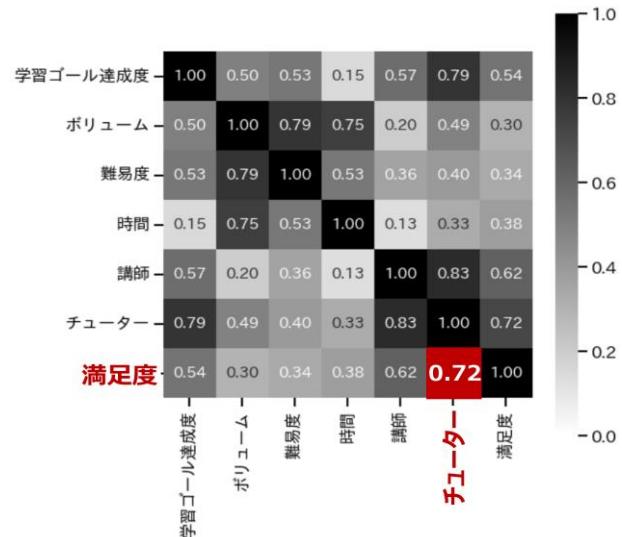


図 4 満足度に及ぼすチューターの影響

Figure4 Influence of tutors on satisfaction.

ところで、企画段階で想定できなかった新たな気づきが研修を通して得られた。具体的には、以下のようなシチュエーションである：特に機械学習の回においてであるが、受講者の半数が原理原則を隅々まで理解しなければ先に進めず、足踏み状態に陥ってしまうような状況である。受講者に対して、機械学習には多様かつ有用なロジックが膨大に存在するが、その原理原則を隅々まで覚えなければ応用できないということは一般にない(機械学習は問題解決するのが目的で原理起点ではない)という意識付けをすべきであった。

4. まとめ

当社総合職一年目社員を対象とした DX リテラシー教育を開発し、集合研修形式で教育を実践した。その結果、全受講者の 40% にデータ活用スキルを用いて自職場課題を解決するといった DX に対する自発性という点で積極的な変容が認められた。

本教育では、受講者の自発性を促すためにチューターによるアサーティブコミュニケーションをとりいれており、以下の知見を得ることができた。

- 受講者満足度に及ぼすチューターの影響を相関分析したところ、DX リテラシー教育の良し悪しにチューターが大きく作用していることがわかった。
- 教育の心構えとして、機械学習には多様かつ有用なロジックが膨大に存在するが、その原理原則を隅々まで覚えなければ応用できないということは一般にないという意識付けを受講者にすべきであった。

このような新人教育を継続していくことで IoT を活用するリテラシーが身につき、どのような変容を目指すべきか想像・創造できるようになる。また今後は、教育に IoT を導入して受講者のリアクションや生体反応をセンシングし、更なる学習効果の向上や心理的安全性確保の合理化等を検討していきたい。

参考文献

- [1] IPA 独立行政法人 情報処理推進機構.“DX 白書 2023”. IPA 独立行政法人 情報処理推進機構ホームページ. 2023.
<https://www.ipa.go.jp/files/000108048.pdf>, (参照 2023-2-16)
- [2] 日鉄 DX 具体的な取り組み | デジタルトランスフォーメーション (DX) 戦略の推進 | 企業情報 | 日本製鉄. 日本製鉄. 2022.
<https://www.nipponsteel.com/company/dx/specific.html>, (参照 2023-2-16)
- [3] 株式会社帝国データバンク 情報統括部.“DX 推進に関する企業の意識調査（2022 年 9 月）”. TDB 景気動向オンラインホームページ. 2022. <https://www.tdb-di.com/2022/10/sp20221028.pdf>, (参照 2023-2-16)
- [4] 株式会社 RECEPTIONIST.“調査レポート】DX 化が遅れる大企業の共通点 若手社員の 7 割が課題意識薄い | 株式会社 RECEPTIONIST”. 株式会社 RECEPTIONIST - ビジネスコミュニケーションをアップデートする。. 2022.
<https://receptionist.co.jp/release/report-office-dx/>, (参照 2023-2-16)
- [5] 武石 美有紀.“2022 年新入社員意識調査 | 定着・早期立ち上がりに向けて”. 人材育成・研修・組織開発のリクルートマネジメントソリューションズ. 2022,
https://www.recruit-ms.co.jp/issue/inquiry_report/0000001077?theme=starter, (参照 2023-2-16)
- [6] 平木 典子. アサーションの心. 朝日新聞出版, 2015
- [7] ロバート・B・チャルディニー. 影響力の武器. 誠信書房, 2007
- [8] Edmondson, A. (1999). Psychological safety and learning behavior in work teams. *Administrative Science Quarterly*, 44(2), 350-383.
- [9] Edmondson, A. C. (2019). The fearless organization: Creating psychological safety in the workplace for learning, innovation, and growth. John Wiley & Sons.
- [10] 秋田 喜代美. 心理学に対する授業観と質問行動:一般教育課程と心理学専攻の比較検討. 立教大学心理学科研究年報. 1995, 38, 25-38.

継続的な学習機能を有する WiFi ポータルシステムの設計と実装

松尾周汰 ^{†1} 中村優吾 ^{†2} 荒川豊 ^{†3}
九州大学 九州大学 九州大学

1. はじめに

スマートフォンの普及やデジタル活用の推進に伴い、インターネット利用が拡大しているが、安全にインターネットを利用するには、情報セキュリティへのリテラシーが求められる。総務省による調査^{*1}によると、個人のインターネット利用率は 83.4% であり、利用目的は「電子メールの送受信」が全体では最も多く、「ソーシャルネットワーキングサービスの利用」については、全ての年齢階層で利用率が上昇している。ICT サービスの利用が増えている一方で、インターネットを利用して「不安を感じる」又は「どちらかと言えば不安を感じる」との回答は合わせて 74.2% となっている。また、インターネット利用に伴う個人の被害経験として、「迷惑メール・架空請求メールを受信」が 61.4% と最も多くなっており、「何らかの被害を受けた」企業の割合は 51.4% と半数以上である。インターネット利用に対する不安感や被害の原因として、学校におけるセキュリティリテラシー学習の機会や、企業におけるセキュリティ対策の未熟さが挙げられる。文部科学省により実施された学校における教育の情報化の実態等の調査^{*2}によると、無線 LAN の設置やインターネット接続率などの ICT 環境の整備状況や、ICT の活用や情報セキュリティの基本的な知識などを指導する教員の能力には都道府県により格差があり、教育方法や教育体制が学区現場に十分に行き渡っていないとの指摘もある[1]。また、企業における情報セキュリティ対策^{*3}について、「何らかの対策を実施している」企業は 98.2% であるが、「社員教育」を行なっている企業は 49.4% と半数を下回っている。

これらを解決するための継続学習支援として、マイクロラーニングとナッジの手法に着目した。短時間で手軽に取り組むことができる学習方法であるマイクロラーニングは、

学習の習慣化や行動変容が期待されている[2]。しかしながら、言語学習以外のトピックの取り扱いが少なく、ラーニングシステムの設計とその影響については十分に理解されていない。また、プラウザ拡張機能やモバイルアプリの PUSH 通知を活用した学習提案では、限られたフィールドでしか学習ができないことや、回数を重ねるごとに無視されやすくなるという課題もある。行動経済学を発端とするナッジ理論では、選択肢を禁じたり、金銭的インセンティブを与えることなく、他人の行動を予測的に変化させる[3]。社会や健康といった観点からはじまり、プライバシー保護やセキュアな選択を促進するための手法としてもナッジが活用されている。しかし、これらの実験は、実験のための仮想環境にとどまった実証が多く実世界へのフィードバックが無いものや、文脈によってはナッジの効果がない場合もあり、まだ議論の余地がある。

そこで本研究では、学習システムの設計と実証運用を行った。学習の機会として、フリー WiFi 接続時に出現するような、ダイアログやプラウザとして表示されるランディングポータルを活用する。以下のようなフローで学習を支援する。

- (1) アクセスポイントの選択時に、認証と共に情報セキュリティに関するテストが出題される。
- (2) 回答後画面のインターフェース情報にナッジ要素を取り入れ、学習に対する動機付けを狙う。
- (3) 回答すると正誤結果と共に、コミュニティのスコアがフィードバックされる。
- (4) スコアボードには、その日のコミュニティにおける回答数、正答率、回答人数が表示される。
- (5) また、当日の回答数と目標数がプログレスバーと共に示され、回答が一定数未満であると、再認証までの利用時間が 1 時間に制限される旨の文言を表示される。

これらのスコアボードの下には再挑戦ボタンを設置し、新たにテストに回答することができる設計とした。これらの文章やシステム情報の設計には、ナッジ理論のフレームワークである EAST^{*4}を活用している。テストの回答を 1 問のみに限定し、一回のアクションを小さくすることで Easy,

Design and implementation of a WiFi portal system with continuous learning capabilities

^{†1} SHUTA MATSUO, Kyushu University

^{†2} YUGO NAKAMURA, Kyushu University

^{†3} YUTAKA ARAKAWA, Kyushu University

^{*1} <https://www.soumu.go.jp/johotsusintokei/statistics/statistics05.html>

^{*2}

^{*3} <https://www.soumu.go.jp/johotsusintokei/statistics/statistics05.html>

^{*4} <https://www.bi.team/publications/east-four-simple-ways-to-apply-behavioural-insights/>

コミュニティの他メンバーの回答情報を示していることから Social, インターネット利用行動の動線上でのテストであることから Timely を実現した。

アンケートにより学習システムの受容性と、システム利用ログからインターフェース情報による回答動機付け効果の評価を行う。学習システム運用前後での情報セキュリティテストのスコアの比較を行う。これらの調査の結果、スコアボードやプログレスバーの受容性が高いことがわかった。また、目標回答数の可視化が、繰り返し学習へ取り組むことへの動機付けとなることがわかった。さらに、学習システム前に比べ、後のテストでは有意にスコアが上昇したことが確認された。

2. 関連研究

本研究は、マイクロラーニングやナッジ手法に関する既存の研究に基づいている。これらの関連研究を示し、課題の整理と本研究の位置づけを行う。

2.1. マイクロラーニングを活用した関連研究

マイクロラーニングとは、学習タスクを小さな単位に分解し、短い時間で取り組むことを目的としている [2]。短く小さな取り組みに分解することで、日常的な活動の中のわずかな隙間時間に、簡単かつ柔軟に学習を組み込むことができる [4]。マイクロラーニングを取り入れるアプローチとして、コンピューターのスクリーンセーバーを活用したもの [2] や、Web ブラウザの拡張機能を活用したもの [5] がある。これらの学習システムは、ユーザーに受容され自発的な学習を促した一方で、インタラクティブ性が小さく、徐々に学習への取り組みが弱まる事も指摘されている。また、Web ブラウザを利用する場合、ブラウザ内の限られたフィールドしか活用できない拡張性の限界や、対応するブラウザに依存するという制約がある。スマートフォンのセンサーから退屈状態を判定して PUSH 通知による割り込みを行う研究 [6] も行われたが、通常状態の学習効果との有意な差は見られなかった。ユーザーの動作や状態の判定では、提供する情報が多く、特に位置情報は端末バッテリーの負荷となったり、情報の種類によってはプライバシーの侵害にあたる可能性があるため、ユーザーの受容性が課題として挙げられる。

2.2. ナッジを活用した関連研究

ナッジとは選択の自由を奪ったり、金銭的報酬を変えることなく、人々が望ましい選択をするよう影響を与える手法である [8]。オンライン環境におけるプライバシー関連のナッジデザインの研究 [7] では、Slack 上でのプライバシーに関

する意思決定をするインターフェースに対し、ユーザーの選好調査を行った。その結果、ユーザーは視覚的なデザインや赤や緑といった色に基づくナッジを好む一方で、情報提供やフィードバックのようなテキスト要素に基づくナッジはあまり好まれないことが明らかになった。色の表示とフレーミングによるナッジは理解がしやすく、プライバシー関連の情報を公開するかどうかの意思決定を迅速に行うことには役に立つことが分かった。セキュリティ選択とナッジの組み合わせの研究 [8] では、公衆 WiFi 選択、クラウドサービスの選択、スマートフォンの暗号化、パスワード生成というサイバーセキュリティに関するシナリオにおいてナッジの効果検証を行っている。色や図表を用いた直感に訴えかける変更を行うものや、選択に関する説明による情報提供を行うものをナッジとして活用した。実験の結果、全てのシナリオにおいて、ナッジを用いるとよりセキュアな選択をすることがあることが分かった。既存の研究では、アンケートフォーム上の仮想的な環境や一部の限られた環境で実験が行われている。その環境では、SNS の投稿やセキュリティの選択などの自分の行動が実際に現実世界には反映されないため、実際の選択や行動と異なる可能性があることが課題として挙げられている。

2.3. 本研究の位置づけ

上記関連研究より、さまざまな媒体や環境でマイクロラーニングシステムの設計と開発が行われ、ナッジのテキストやインターフェースへの有効な活用方法が明らかになっているものの、システムの設計や実証シナリオにおいては課題や議論の余地がある。既存研究の考えられる課題を解決しつつ、効果的な学習システムを設計するために、マイクロラーニングとナッジの観点から以下の視点を取り入れるべきであると考える。

- 学習システムには、ブラウザやアプリケーションに依存しない環境や、ユーザーにとって負担を感じさせない介入方法、回答データを活用したインタラクション性への工夫が必要である。
- セキュリティ文脈のナッジ実験においては、実環境への影響がある状態で、ナッジ介入の受容性や行動変容の効果測定が必要である。

これまでの研究の課題や以上の考察を踏まえ、本研究ではマイクロラーニングやナッジの概念に基づき、情報セキュリティリテラシーの向上支援をする学習システムの設計と実装を行うことを目的とする。その上で、以下 3 つのリサーチエクスプローラー (RQ) を立て、調査を行う。

- (1) WiFi 接続時のランディングポータルが、マイクロラーニングの学習環境として受容されるか。



- (2) 学習システム上のインターフェースにインタラクティブなナッジ情報を追加することは、回答数に影響を与えるか。
- (3) 学習システムの運用が、情報セキュリティリテラシーの向上に繋がるか。

3. 実験方法

この実験では、情報セキュリティリテラシー向上を支援する学習システムを設計、実装し、研究室内を実験環境として実証運用することで、RQ を検証する。RQ1 つ目に対して、一定期間、実験環境で学習システムを運用した後、アンケートによってシステム受容性の調査を行う。RQ2 つ目に対して、実験期間中の回答実績の分析を行う。RQ3 つ目に対して、実験期間の前後のテスト結果でスコアに比較を行う。

3.1. 学習システム設計

学習システムの構成図と学習フローを図 1 に示す。アクセスポイントとなるルータの設定により、ゲストネットワークを使用してインターネットにアクセスするときに、接続端末にゲスト向けのランディングポータルを出現させることができ。このランディングポータルの仕組みは主に、コーヒーショップやコンビニ、観光地で提供されているフリー Wi-Fi に活用されている。ランディングポータルとは、クライアントがユーザー認証を行うまで、ネットワークへアクセス出来ない環境を作る機能である。この機能を使用すると、端末がネットワークに接続したときに、表示されたポータル上で、指定した Web ページにリダイレクトして誘導することができる。また、このゲストネットワークは接続セッションの制限時間を設定することも可能である。制限時間が過ぎると、再びランディングポータル上の学習システムが起動され、ポータルの表示後にインターネット接続が可能となる。本実験のため、デフォルトの接続制限時間を 3 時間とした。

はじめに、被験者は WiFi 接続のために学習システムに紐づいた SSID を選択する。アクセスポイントの WiFi ネットワークに接続されると、PC 画面上へのダイアログやスマートフォンのブラウザ上でランディングポータルが出現する。ポータル上では、Web アプリケーションとして実装された学習システムにリダイレクトされ、情報セキュリティ

Question ID: 37
あなたは、業務で利用するためのソフトウェアを購入しようと考えています。色々と調べてみたところ、良さそうなソフトウェアがあったのですが、ソフトウェアの価格が予想していたよりも高額です。このような場合に、適切な対応を全て選択してください。

SNSでの状況を拡散したところ、当該ソフトウェアを無料で提供してくれるという人が現れたので、ありがたくソフトウェアを頂きインストールしてみる。

他に代用できる低価格のソフトウェアがないか探してみる。

予算を確認して、ソフトウェアの必要性を再検討してみる。

公式サイトに、A大学の学生であれば無料で当該ソフトウェアを利用できると書いてあったので、知り合いからアカウント情報を聞き、A大学の学生ということで登録して、ソフトウェアをインストールしてみる。

submit

図 2 回答画面の例

に関する問題が 1 問、ランダムに表示される。回答後の画面では、正誤や問題解説のフィードバックのほか、その他の被験者達の回答状況に関する情報が表示される。再挑戦ボタンも設置しているので、同一ポータル上から繰り返し問題に取り組むことができる。インターネット接続の制限時間が切れると、再び学習ステップが繰り返される。なお、離席などにより一度 WiFi 接続が切れ、手動または自動的に再接続される場合であっても、ポータルによる学習システムが起動される。

本学習システムでは、インターネット接続制限を活用する。被験者による一日あたりの問題回答数の合計が一定数以上に達しない場合、次の日の接続制限時間をデフォルトの 3 時間から 1 時間に短縮するというメッセージを表示する。また、回答データを活用したインタラクティブな学習システムのインターフェース設計を説明する。学習システムの回答画面の例を図 2 に、回答後の画面のうち、ナッジ要素を含んだ部分の例を図 3 に示す。

回答結果画面は主に、4 つの部分に分かれている。研究室全体のスコアには、当日の合計回答数、正答率、累計回答者数が表示される。社会規範ナッジ [9] を参考に、身近な人たちの情報を付与することで、周りがやっているから自分もやるという動機付けにより、回答数の増加を狙う。目標回答数と現在の回答数、それに基づく回答率がプログレスバーとともに視覚的に表示される。これは、目標達成が目に見える形で近づくとモチベーションが高まる心理を活用する [10]。また、目標回答数を 30 とした。この値は、一日あたりの研究室への出席人数が約 10 人、一人当たり 2 回程度の学習システムの起動を仮定して算出した。回答数が目標に達しない場合、接続時間がより短く制限されることをの説明が表示される。時間を強調するために、「1 時間」という箇所を赤色にした。なお、目標回答数を超えると、「今日の

情報セキュリティテスト Humanophilic System Lab. English

Question ID: 37

おめでとうございます 正解です！

研究室全体のスコア

回答数 正答率 回答者

7 件 29 % 1 人

回答数: 7 / 30 23%

目標に達しない場合、再認証までのインターネット接続が 1時間 に制限されます。

もう1回挑戦する

あなたは、業務で利用するためのソフトウェアを購入しようと考えています。色々と調べてみたところ、良さそうなソフトウェアがあったのですが、ソフトウェアの価格が予想していたよりも高額です。このような場合に、適切な対応を全て選択してください。

SNSでこの状況を拡散したところ、当該ソフトウェアを無料で提供してくれるという人が現れたので、ありがとうございました。ソフトウェアを頂きインストールしてみる。

✓ 他に代用できる低価格のソフトウェアがないか探してみる。

✓ 予算を確認して、ソフトウェアの必要性を再検討してみる。

公式サイトに、A大学の学生であれば無料で当該ソフトウェアを利用できると書いてあったので、知り合いからアカウント情報を聞き、A大学の学生ということことで登録して、ソフトウェアをインストールしてみる。

図 3 回答後の画面の例

目標回答数に達しました！」との表示に変更される。「もう1回挑戦する」という文字付きのボタンが表示される。ボタンの色は、クリックへの動機付けの効果が高める赤色とした[11]。また、これらの要素に加え、問題文と正誤情報もフィードバックされる。

3.2. 実験条件

実験対象として、研究室メンバーである20代～40代の男女、学生と教員合わせて52人を被験者とした。ただし、研究室への出席は任意であるため、実際に学習システムとインタラクティブな交流を持った被験者は、対象期間に出席した人に限られる。実験期間は3日間とした。また、実験前に情報セキュリティテストの受講を案内し、実験後にアンケートによるシステムの受容性調査とセキュリティテストの再実施を行った。セキュリティテストの内容は、九州大学の教職員向けに配信された「九州大学情報セキュリティ教育及び自己点検」の内容を参考に、一部改変したものを探用した。設問内容は、ナッジに対する受容性と、介入タイミングに対する受容性に関するものである。ナッジ受容性の設問は、健康や持続可能性に関する行動[12]や政策におけるオプトイン/オプトアウト[13]、フレーミングナッジに対するナッジ[14]を調査した研究に基づいて作成された。介入タイミング受容性の設問は、スクリーンタイムや、モバイルデバイスによる待機時間や外出時間を活用したマイク

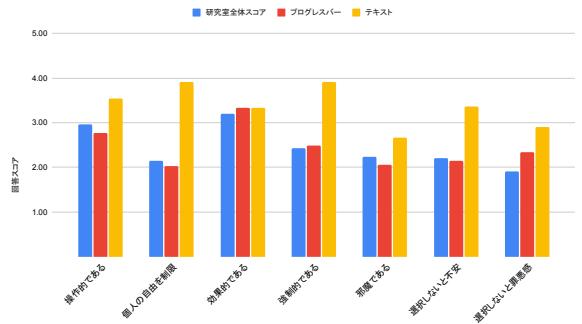


図 4 ナッジに対する受容性調査の結果

ローラーニングの研究[2][6][15]における受容性調査に基づいて作成された。

4. 実験結果

4.1. 学習システムの受容性

本節ではRQ1つ目の「WiFi接続時のランディングポータルが、マイクロラーニングの学習環境として受容されるか」に対して結果の確認と考察を行う。

4.1.1 インターフェース情報の受容性

インターフェース情報に関しては、研究室全体のスコア、目標達成率を表すプログレスバー、目標未達成の場合のインターネット接続制限を説明するテキストに対してそれぞれ受容性の調査を行った。各箇所に対する5件法による調査の集計結果を図4に示す。なお、設問内容は以下の通りである。

- (1) 「もう1回挑戦する」を選択することに対して、判断を迷わせ、操作的であると思う
- (2) 「もう1回挑戦する」を選択することに対して、個人の自由を制限していると思う
- (3) 「もう1回挑戦する」を選択することに対して、効果的であると感じる
- (4) 「もう1回挑戦する」を選択することを、強制的にさせるように感じる
- (5) 「もう1回挑戦する」を選択することの判断に、邪魔であると感じる
- (6) 「もう1回挑戦する」の選択をしない場合に、不安を感じる
- (7) 「もう1回挑戦する」の選択をしない場合に、罪悪感を感じる

各設問ごとに、2つのインターフェース情報に対して対応ありt検定を行った結果、インターフェース情報によって受容性が異なることが確認された。研究室全体のスコアとP

ログレスバーは、『「もう一回挑戦する」の選択をしない場合に、罪悪感を感じる』の項目において $p < 0.05$ で有意差となったが、その他の全ての項目において、有意差は見られなかった。接続制限を説明するテキストとログレスバーは『「もう一回挑戦する」の選択することに対して、効果的である』の項目において、有意差は見られなかったが、それ以外の項目においては全て $p < 0.05$ で有意差ありとなった。研究室全体のスコアとテキストにおいては、すべての項目に対して、 $p < 0.05$ で有意差ありとなった。

これらの結果から、目標回答数によって接続時間を制限することは、受容性の評価に大きな影響を与えることがわかった。特に、「個人の自由を制限している」「強制されるように感じる」といった項目に対するスコアが比較的高く、ユーザーによっては、再び問題へ回答することを作為的に促されているように感じる恐れがある。この説明テキストはもう一度回答することに対し、直接的に操作をしているわけではないが、ナッジの活用においては選択肢を強制してはならないため、慎重に取り扱うべきである。一方で、研究室全体のスコアにおいては、「個人の自由を制限している」「邪魔であると感じる」の項目のスコアが低いことから、ユーザーにとって束縛感がなく、自然に受け入れられやすい要素であると考えられる。「効果的である」の項目のスコアが、接続制限説明テキストと同様に高いことから、ユーザーに負担をかけずに問題へのもう一度回答することを促す手段としては適していると言える。また、ログレスバーにおいては、「罪悪感を感じる」「個人の自由を制限している」の項目のスコアが低いことから、研究室全体のスコアと同様、ユーザーへの負担が少なく、再度回答を促す手段として有効であると言える。

4.1.2 システム起動タイミングに対する受容性

学習システム起動タイミングの受容性に対する調査を行った。5件法による調査の集計結果を図5に示す。なお、設問内容は以下の通りである。

- (1) 学習システムが起動するタイミングは適していた
- (2) 学習システムによって、進行中の活動が妨げられた。
- (3) 学習システムの起動に対して、苛立ちを感じた
- (4) 学習システムでの回答後、実行しようとしていたタスクにすぐに取り掛かることができた
- (5) 学習システムによって、集中力が失われたと感じた

1日のうち初回のシステム起動時と2回目以降のシステム起動時の結果に対して、対応あり t 検定を行った。検定の結果、初回と2回目以降のシステム起動の受容性の間に有意差は見られなかった。1日のうち初回の学習システムの起動タイミングは、その日初めて学習システム対応 WiFi に接続された時である。このタイミングはタスクへ取り掛かると同時にタイミングであるため、スクリーンセーバー

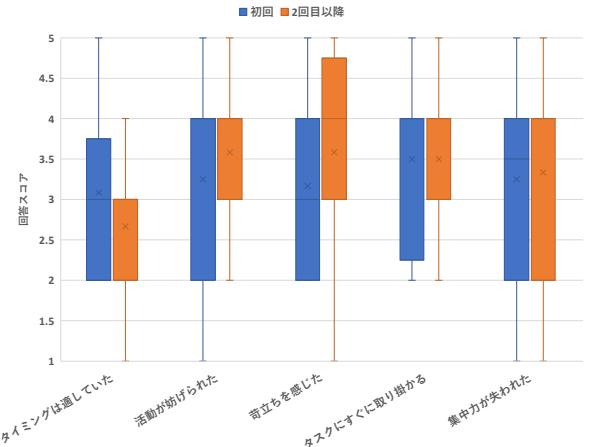


図5 学習システム起動に対する受容性調査の結果

によるマイクロラーニング [2] と同様に、タスクの途中での介入とはならず、精神的負荷への影響はない想定していた。一方で、2回目以降の時間経過による学習システム起動は二次タスクの割り込みにあたり、精神的負荷への影響が考えられた。しかしながら、これらには有意な差が見られなかった。特に、「進行中の活動が妨げられた」「集中力が失われたと感じた」の項目においては、初回起動時の回答でも、平均値が3を超えており、中央値がどちらも4という結果であった。これは、初回起動のタイミングはタスクへの介入とはならないという予想とは反する結果である。初回起動に対しても精神的負荷を感じた理由として、実際にブラウザの立ち上げや、インターネットを使用するアプリケーションを使用する直前まで行動を移していたタイミングで介入が行われたからであると考えられる。スクリーンセーバーによる介入は、PC起動直後であり、ブラウザやアプリケーション立ち上げよりも前に学習タイミングが発生する。このタイミングの僅かな差の間に、ユーザーの注意力が取り掛かろうとしているタスクへ向いている状態になるため、学習システムの起動が、一次タスクへの割り込みであると感じられたのではないかと考えられる。また、「実行しようとしていたタスクにすぐに取り掛かることができた」の項目においては、初回起動と2回目以降どちらも平均値が3.5、中央値が4という同等の結果となり、ある程度容易に一次タスクへの切り替えができたと考えられる。WiFiポータルの起動後は、ウィンドウの「完了」ボタンでポータル自体を閉じることが可能である。問題に対する回答は選択式であり、回答後に続けて回答するかどうかはユーザーに委ねられるため、一度の学習にかかる時間は最短で1分未満であると想定される。学習システムの終了の容易さや、学習完了までの時間が短いことは、元のタスクへの復帰には重要な観点であると考えられる。

以上の結果から、RQ1 つ目の「WiFi 接続時のランディングポータルが、マイクロラーニングの学習環境として受容されるか」への回答を述べる。インターフェース情報の観点からは、学習システムを共有しているコミュニティユーザーの回答状況や、目標回答数を可視化することは受容されやすく、繰り返しの回答動機付けの効果があると言える。一方で、インターネット接続制限を表すことに対しては、操作的、強制的な印象から受容されにくく、システム設計やインターフェース情報への採用は慎重に行うべきである。最後に、システム起動タイミングの観点からは、どのタイミングでの起動でも一次タスクの妨げになる可能性が大きく、頻繁な介入は受容性を低下させる恐れがある。そのため、学習環境として、起動回数の考慮が必要である。また、学習システムから迅速に元のタスクに復帰できることも受容性で重要な観点であり、容易に学習システムを閉じられることや、一回の学習時間を短くする工夫が必要であるといえる。

4.2. 学習システムの利用頻度

本節では RQ2 つ目の「学習システム上のインターフェースにインタラクティブなナッジ情報を追加することは、回答数に影響を与えるか」に対して、学習システムの利用ログから、日毎と全期間の回答結果と時間帯別の回答数についての結果を示し、考察を行う。

実証運用期間における、学習システムの利用結果を図 6 に示す。日毎の回答結果から、日を追うごとに回答数と正答率が増えていることが確認できる。回答目標数を 30 件としており、2 日目と 3 日目はこの目標数を超える回答数となった。2 日目と 3 日目の時系列に沿った回答記録を図 7 と図 8 に示す。同じ色の箇所は、同一 ID で連続して回答していた記録を指す。2 日目と 3 日目の間で色と ID の対応はない。この結果から、目標回答数の 30 件に達する前では短時間による繰り返しの学習が行われており、目標数を達成した後は繰り返しの学習ほとんど見られなかったことがわかった。

以上の結果をもとに、回答数について考察を行う。回答数の推移を見ると、目標の 30 件を達成する前後で連続回答の数に変化がある。目標回答数を達成する前は、連続回答が続いていることが確認される。これは、回答数を増やすプログレスバーを進めることで、インターネット接続制限を解除しようとする動機付けが作用したと考えられる。反対に、目標回答数を達成すると、プログレスバーが埋まり、インターネット接続制限を説明するテキストが「今日の目標回答数に達しました！」という内容に変わるために、再度回答することへの動機付けの効果が無くなったと考えられる。これらより、RQ の 2 つ目である「学習システム上のインターフェースにインタラクティブなナッジ情報を追加する

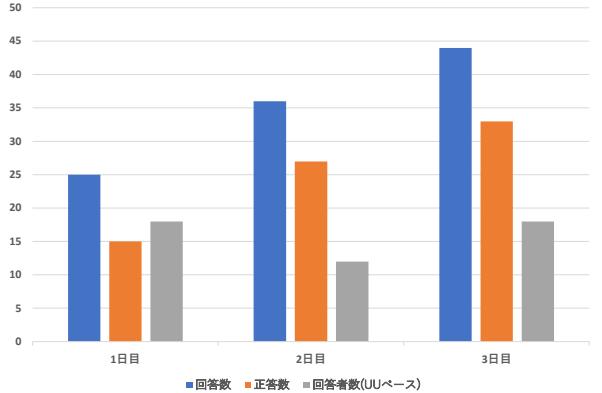


図 6 日毎の回答結果

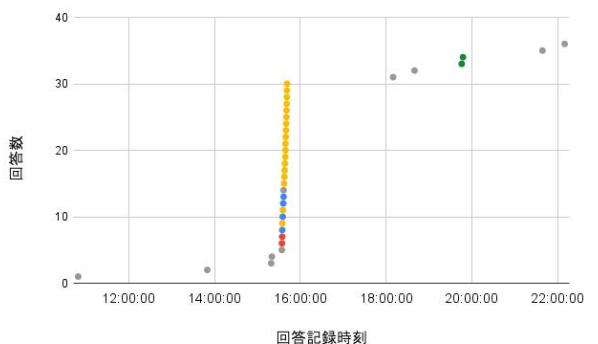


図 7 2 日目の時刻ごとの回答記録

ことは、回答数に影響を与えるか」に対して、目標回答数を達成する前は、回答数を増やす効果があるが、目標回答数を達成した後は、回答数を増やす効果は見られなくなると結論づけられる。

4.3. セキュリティテストのスコア比較

本節では RQ3 つ目の「学習システムの運用が、情報セキュリティリテラシーの向上に繋がるか」に対し、情報セキュリティテストの回答スコアの結果を示し、考察を行う。実験前、実験後全ての回答データを集計した結果を図 9 に、実験前、実験後どちらにも回答があったデータを集計した結果を図 10 に示す。これらのテストはおよそ 3 週間ほど間隔を空けて実施された。はじめに、実験前と実験後に収集した全ての回答データを比較した。実験前と実験後の回答人数がそれぞれ異なるため、対応なしの検定を行った。F 値を算出したところ、0.438 であり、分散は等しいことが確認されたので、対応なし片側 t 検定を行った。検定の結果、 $p < 0.05$ で有意差ありとなった。つぎに、実験前と実験後のどちらにも回答があったデータに対して検定を行った。対応あり片側 t 検定を行った結果、 $p < 0.05$ で有意差ありと

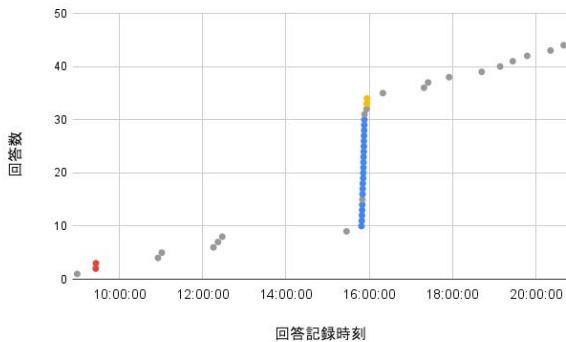


図 8 3 日目の時刻ごとの回答記録

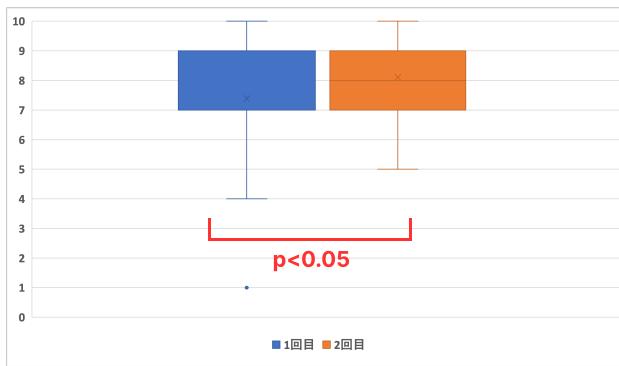


図 9 すべての回答データの集計結果

なった。これらより、対応なしの全データに対しても、対応ありのデータに対しても、学習システムの実証運用前に比べ、後の方がテストのスコアが有意に高い結果となった。

学習システムの問題には、テストに使用した問題と同じものが含まれており、実証運用期間中に同一問題に取り組んだことで正答率が高まったと考えられる。また、テストには使用されていない問題であっても、学習システムを通じて情報セキュリティの題材に触れることで、セキュリティリテラシーへの意識が高まり、テスト回答時の影響があったと考えられる。これらの結果より、RQ の 3 つ目である「学習システムの運用が、情報セキュリティリテラシーの向上に繋がるか」に対して、学習システムはリテラシー向上を促すと結論づけられる。

5.まとめ

5.1. 本研究の成果

本研究では、継続的な学習機能を有する WiFi ポータルシステムの設計と実装を行い、マイクロラーニングとナッジの観点から介入の効果や受容性の評価を行った。各実験結果から得られた結果は以下の通りである。

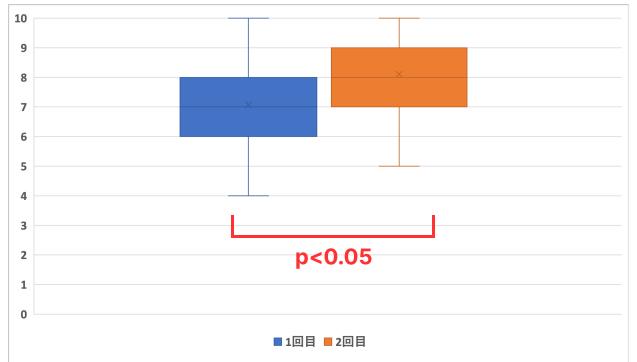


図 10 実験前、実験後のどちらにも回答があったデータの集計結果

- ・ インターフェース情報に研究室全体のスコアや目標回答数を表示し、可視化することは受容され、繰り返しの学習行動への動機付け効果がある
- ・ 介入タイミングはインターネット接続時間の制限やタスク割り込みによる心理的負担を考慮すべきであり、短時間で 1 単位の学習を終えることは受容性の観点で重要である
- ・ 学習システムの運用前後で、情報セキュリティテストのスコアが有意に上昇する

5.2. 研究上の制限

実験結果からの考察とともに各 RQ への結論を述べたが、実験条件や考察に関して、課題となるいくつかの考慮すべき点が存在する。一つ目は、学習システムの運用期間である。本実験は、受容性の調査も目的としていたため、運用した期間が 3 日間と短いものであった。この学習システムをより長期的に運用することで、回答数の推移やセキュリティテストのスコアが変動する可能性がある。

二つ目は、2 回の情報セキュリティテストの回答期間と学習システムの問題への取り組みが与える、テストスコアへの影響である。テストに使用した問題は実験前と実験後で同じものであったので、実験前のテスト結果による復習効果が持続している可能性がある。また、本学習システムには、セキュリティテストの問題以外の題材も採用しており、問題はランダムで出題されるため、実証運用期間中にテストで出題される問題に取り組んでいない可能性がある。そのため、実証期間中に学習システム上でテストに出題される問題には取り組んでいない被験者がいることも考慮すべきである。

最後に、被験者の基礎的な情報セキュリティリテラシーである。本実験では、情報系の大学生、大学院生が在籍する研究室を実験環境とした。そのため、基礎的な情報リテラシーが一般的な値よりも高く、学習前後のスコアの変動

や、学習への取り組みに影響を与えた可能性を考慮すべきである。

5.3. 今後の展望

今後の研究において、実験環境やシステム設計に対しての改善点や工夫点が挙げられる。実験環境として、被験者の属性の偏りをなくすことや、長期の学習システム運用を行うことで、より一般的な議論を行うことができる。システムの設計として、目標回答数を出席者数や前日の回答結果をもとに変動させることで、目標達成までの繰り返しの回答をより多く誘発できる可能性がある。また、インターネット接続制限時間に対して、個人のテストスコアや回答状況を反映させることで、学習の個別最適化も考えられるさらに、ローカルデータとして管理していた問題をサーバー上で管理することで、コミュニティ全体や個人の回答結果によって出す回数や頻度を操作し、効率の良い学習の促進ができると考えられる。このように回答状況を細かく分析することで、より受容されやすく、学習への動機付けの効果があるシステムの設計ができると考えられる。

謝辞 本研究の一部は、科学研究費補助金 (JP19KT0020) および JST さきがけ (JPMJPR21P7) の助成を受けたものである。

参考文献

- [1] 堀田龍也, 佐藤和紀. 日本の初等中等教育における情報リテラシーに関する教育の動向と課題. 電子情報通信学会 通信ソサイエティマガジン, Vol. 13, No. 2, pp. 117–125, 2019.
- [2] Gerhard Gassler, Theo Hug, and Christian Glahn. Integrated micro learning—an outline of the basic method and first results. *Interactive computer aided learning*, Vol. 4, pp. 1–7, 2004.
- [3] Thomas C Leonard. Richard h. thaler, cass r. sunstein, nudge: Improving decisions about health, wealth, and happiness, 2008.
- [4] 徳田博行, 高橋雄太, 松田裕貴, 荒川豊, 安本慶一. 辞書検索履歴の自動収集に基づく英単語マイクロラーニング支援手法. 情報処理学会モバイルコンピューティングとパーソナルシステム (MBL) 研究会, No. 44, pp. 1–8, 2019.
- [5] Andrew Trusty and Khai N Truong. Augmenting the web for second language vocabulary learning. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 3179–3188, 2011.
- [6] Carrie J Cai, Anji Ren, and Robert C Miller. Wait-suite: Productive use of diverse waiting moments. *ACM Transactions on Computer-Human Interaction (TOCHI)*, Vol. 24, No. 1, pp. 1–41, 2017.
- [7] Sofia Schöbel, Torben Jan Barev, Andreas Janson, Felix Hupfeld, and Jan Marco Leimeister. Understanding user preferences of digital privacy nudges—a best-worst scaling approach. In *Hawaii International Conference on System Sciences (HICSS)*, 2020.
- [8] Verena Zimmermann and Karen Renaud. The nudge puzzle: matching nudge interventions to cybersecurity decisions. *ACM Transactions on Computer-Human Interaction (TOCHI)*, Vol. 28, No. 1, pp. 1–45, 2021.
- [9] Michael Hallsworth, John A List, Robert D Metcalfe, and Ivo Vlaev. The behavioralist as tax collector: Using natural field experiments to enhance tax compliance. *Journal of public economics*, Vol. 148, pp. 14–31, 2017.
- [10] Joseph C Nunes and Xavier Dreze. The endowed progress effect: How artificial advancement increases effort. *Journal of Consumer Research*, Vol. 32, No. 4, pp. 504–512, 2006.
- [11] 松尾周汰, 中村優吾, 荒川豊. ナッジフレームを適用した captive portal によるセキュリティリテラシ向上支援システムの検討と評価. 情報処理学会 IoT 行動変容学研究グループ 第 2 回研究会 (IPSJ BTI2).
- [12] Laurens C van Gestel, Marieke A Adriaanse, and Denise TD de Ridder. Who accepts nudges? nudge acceptability from a self-regulation perspective. *PloS one*, Vol. 16, No. 12, p. e0260531, 2021.
- [13] Haoyang Yan and J Frank Yates. Improving acceptability of nudges: Learning from attitudes towards opt-in and opt-out policies. *Judgment and Decision Making*, Vol. 14, No. 1, pp. 26–39, 2019.
- [14] Ismaël Rafaï, Arthur Ribailleur, and Dorian Jullien. The impact on nudge acceptability judgements of framing and consultation of the targeted population. *Behavioural Public Policy*, pp. 1–16, 2021.
- [15] Tilman Dingler, Dominik Weber, Martin Pielot, Jennifer Cooper, Chung-Cheng Chang, and Niels Henze. Language learning on-the-go: opportune moments and design of mobile microlearning sessions. In *Proceedings of the 19th international conference on human-computer interaction with mobile devices and services*, pp. 1–12, 2017.

ハラスメント被害時の 主観的ストレス差異による心拍変動特徴分析

上野貴弘^{1,a)} 大橋正良²

概要：いじめやパワーハラスメントなどのハラスメント行為の解決に向けて、早期発見・対応が重要な要素である。一方で、ハラスメントとなるか否かは被害者の主観的な感情に基づいているため、目視による判断が難しい。本研究では、被害者の心拍情報を基に機械学習手法を用いた主観的高ストレスの検出を目的としている。実験では、VR動画視聴による疑似的な体験から、被験者のハラスメント被害時とリラックス時における心拍情報の取得と主観ストレス評価を実施した。また検出モデルの評価において、これまでの研究では、線形モデルでの分類が難しく、多変量の組み合わせによる特徴選択が必要であることが判明している。そこで本報告では、全被験者（16名）と高ストレス被験者（9名）、低ストレス被験者（7名）データを用いた各決定木モデルにおける有効な特徴量を詳細に比較するため、しらみつぶし探索によって特徴量の組み合わせの分類精度統計を算出した。そして比較結果を基に、主観的高ストレスの検出に有効な特徴量を分析した。

キーワード：主観的ストレス検出、心拍変動、しらみつぶし探索、特徴選択

1. はじめに

いじめやパワーハラスメントなどの、関係性のアンバランスを利用して個人の心身を傷つける事象（以下、ハラスメント）は、被害者にメンタルヘルスの不調を引き起こす。その結果、うつ病などの重篤な精神傷害や自殺につながる要因となるため、我が国では大きな社会問題となっている。文部科学省の「令和三年度児童生徒の問題行動・不登校等生徒指導上の諸課題に関する調査結果」によると、いじめの認知件数は615,351件、重大事態件数は705件といずれも平成二十五年度以降では過去最高を記録している[1]。また、厚生労働省の「令和三年度個別労働紛争解決制度の施行状況」によると、民事上の個別労働紛争相談件数は284,139件と平成二十四年度以降で過去最高を記録しており、特に相談項目は「いじめ・嫌がらせ」が最多で、件数が86,034件と24.4%を占めている[2]。

ハラスメントに対しては、早期発見・対応を行うことが被害者の重大事態を防ぐ重要な要素である。一方で、第三者ではハラスメントか否かの判断及び発生状況の把握が難しいことが課題として挙げられている[3]。これは、文部科学省がいじめを「当該児童が、一定関係にある者からの行為に対して心身の苦痛を感じるもの」という内容で取り扱っている[4]ように、ハラスメントが被害者の主観的な感情に基づいていることが要因である。

以上の背景から、近年はセンサデバイスが安価に入手できるようになったことに伴い、クラウドサービスとセンシングデータ、機械学習を活用した、ハラスメント行為を検出するためのデザインが検討されている[5][6]。またハラス

メント検出手法の面では、市販のデバイスを用いて収集したセンシングデータから、加害者が行う「叩く」「押し倒す」のような暴力的な事象と非暴力的な事象を分類する研究が行われている[7][8]。しかし、ハラスメントは被害者自身がアクションを起こさない限り、客観的に断定することは難しい。加えて、ハラスメントの態様の多くは精神的攻撃であり[1][9]、身体的暴力の検出だけでは多くのハラスメントに対応できない。

そこで、本研究では被害者の心拍情報を基にハラスメントを検出する機械学習モデルの構築を目的としている。心拍情報は市販のウェアラブルセンサで取得できるため、日常のストレス推定で使用されている。さらに、ストレスは被害者自身の精神的苦痛を伴う視点でハラスメントを捉えるため、網羅的な態様の検出が期待できる。また、検出結果は第三者の介入及び被害者自身が異常を認識するための判断材料になり得る。ただし、現在多く行われている心拍変動データを用いたストレス認識に関する研究は、緊張状態をストレスと捉えた実験課題が主である。そのため、被害者の絶対主観的な事象であるハラスメントに適用できない可能性があることから、ハラスメント特有の主観的ストレスについてモデルの分類精度及び検出に有効な特徴量を調査する必要がある。

本報告では、精神的攻撃に起因したハラスメントを疑似的に受けるVR動画を視聴した被験者のECG（Electrocardiogram）を基に、ハラスメント被害の検出に有効な心拍変動の特徴量を分析する。ただし、ハラスメントか否かの判断は、先述の通り被害者の主観的な感情が基準となっている。そのため、ハラスメントを受ける実験課題

1 福岡大学大学院工学研究科情報・制御システム工学専攻
Department of Information and Control Systems, Fukuoka University

2 福岡大学工学部電子情報工学科
Faculty of Electronics and Computer Science, Fukuoka University
a) td216501@cis.fukuoka-u.ac.jp

のデータだけでは被験者の主観評価が考慮されておらず、ハラスメント被害特有の心拍変動を正確に評価できない。そこで本実験では、被験者に主観ストレスアンケートを実施し、全被験者データ及び、高ストレス被験者データ、低ストレス被験者データの3種類を作成している。また特微量分析では、各作成データを用いて構築した機械学習モデルに対し、しらみつぶし探索で算出した分類精度統計を基に、ハラスメントの分類に寄与する特微量の組み合わせを比較評価する。しらみつぶし探索は、これまでの研究成果[10][11]において、単変量では分類精度に寄与しない特微量が多変量では寄与する可能性が考えられることから、包括的に特微量の組み合わせを調査するために採用している。

2. 関連研究

2.1 被害者視点のハラスメント検出に関する研究

被害者視点でのハラスメント検出には、被害者に関するデータが必要であるため、倫理的問題から検討された研究は少ない。その中で、Ferdinando らは、暴力的な事象（押しつけ、タックル、揺さぶり、前からの押し付け、肩をぶつけるなど）と非暴力的な事象（走る、歩く、動きを伴う遊び、ジャンプなど）を小学生にシミュレートさせることで取得した ECG を基に、6 秒未満のサンプルで暴力を検出する可能性を検討している[7]。これにより、短時間で暴力に起因したハラスメント事象の検出が期待できる。また、Vasileios らは、傍観者の EEG(Electroencephalography)からいじめの脳内刺激を特定する試みを行っている[12]。実験では、いじめの事例に関する動画視聴により被害者の共感を誘発しているため、傍観者の感情に沿った多くの態様に関するハラスメント事象が検出できる。一方で、これらの研究は被害者の主観性が含まれておらず、ハラスメントを断定する要素が少ない。そのため本研究では、被害者の精神的苦痛が市販のデバイスで評価できる心拍情報に焦点を当て、主観性の含まれたデータを主眼として、ハラスメントの検出に有効な特微量を分析する。また、主観的ストレスに応じた特微量分析では、二値分類の結果から選択された特微量を基に、ストレス反応の生理的特徴との関連性を確認する。

2.2 感情認識に関する研究

感情認識は音声処理、自然言語処理、画像処理など様々な分野で研究が行われており、ECG を用いた研究も多く存在する。データ収集の面では、怒りや幸せなどの基本感情や、Russell の円環モデルに基づく感情次元を、音や画像、動画での誘発や[13][14]、自然の会話からラベル付けする[15]ことで実施している。

一方で、ストレスを感情として捉えたストレス認識の分野では、就職面接時や自動車運転時、暗算タスク時の状況を作り出すことで、ストレス状態を誘発している[16]。そし

て、収集データを基にして、クラスごとの分類精度によって認識精度が評価される。

本研究は動画視聴により、ハラスメント被害におけるストレスを誘発している。これは、前述のストレス認識の実験課題とは異なり、感情認識と同様の手法である。そのため、ハラスメント被害特有の感情である可能性が考えられる。さらに、ハラスメントは主観的なストレスが検出基準であり、被験者の主観的なストレスが低いと学習用データとして用いることができない。そのため、主観的なストレス差異による特微量を比較分析することは、ハラスメント検出の有効性を示す要素となり得る。

3. ECG 測定実験

ハラスメント又はリラックスのコンセプトに沿った動画を視聴している被験者から、ECG データを取得する。実験実施にあたり、福岡大学医に関する倫理委員会の承認を得た。本実験手法はこれまでの研究[10][11]と同様であるが、被験者数及び使用データの選択が異なっている。

3.1 被験者

被験者は 16 名（男性 12 名、女性 4 名、 21.8 ± 1.6 歳）の健常な大学生及び大学院生である。実験にあたり、動画内容は伏せた状態で、事前に健康被害のリスクと個人情報の扱い、実験の同意撤回について十分な説明を行い、書面で参加の同意を得た。そして、測定前の 6 時間はカフェイン及びアルコールの摂取を禁止した。各被験者には S1~16 の番号を付与している。

3.2 測定方法

ECG の取得には市販の胸部型心拍センサ Polar H10 を使用した。動画(映像媒体、動画時間)は、ハラスメント (VR, 166 秒)、リラックス(VR, 601 秒)の 2 種類を用意した。各動画は状況に没入できるように 360 度見渡せる VR 映像を採用した。ハラスメント動画は、厚生労働省が提供しているハラスメント被害体験動画[17]である。仕事ミスに対して怒られる被害者の、一人称視点でのパワーハラスメント被害を体験できる。また、リラックスは浜辺の海を眺める動画である。視聴機器として、VR ヘッドセット Oculus Rift S 及び Meta Quest 2 を使用した。実験にあたり、気温の変化による心拍への刺激を避けるため、実験室の空調を 25°C に設定した上で測定を実施した。被験者は心拍センサを装着し、モニタと VR ヘッドセットが設置されたデスクに向かった状態でオフィスチェアに座った。測定の様子を図 1 に示す。

動画視聴時の主観ストレスのアンケート項目を表 1 に示す。本アンケートでは、ハラスメント被害における感情を尺度として扱うために、主観的なストレスの度合に置き換えて質問している。アンケートによる評価は、1.全くスト



図 1 VR 動画視聴実験の様子

表 1 ハラスメント被害体験動画に対する 4 段階評価

得点	ストレス評価
1	全くストレスを感じなかった
2	あまりストレスを感じなかった
3	少しストレスを感じた
4	かなりストレスを感じた

レスを感じなかった, 2.あまりストレスを感じなかった, 3.少しストレスを感じた, 4.かなりストレスを感じた, の 4 段階である。

被験者は、ハラスメント→リラックスの順番で動画を視聴した。次の動画を観るまでには、実験による精神的影響をリセットするため 1 時間以上の間隔を空けた。ハラスメント動画視聴後は、主観ストレスのアンケートを実施した。

4. 特徴量抽出

4.1 ECG データの切り出し

ECG データの開始位置～終了位置は、ハラスメント動画の場合「加害者役の暴言発話開始～再生終了 60 秒後」、リラックス動画の場合「動画の再生開始～再生終了時」である。各動画の ECG データ時間を以下に示す。

- ・ハラスメント : 192 秒
- ・リラックス : 601 秒

また、各動画の ECG データに対して解析用に、1 サンプルにつき窓幅 60 秒、スライド幅 5 秒に切り出した。

4.2 特徴量

解析用データである ECG に対して、ピーク検出を用いて RRI (隣接する R 波の時間間隔) データを抽出する。前処理として、RRI データの 400 以下 2000 ミリ秒以上の外れ値を除去した後、平均 0 分散 1 の標準化を行った。

表 2 RRI データから抽出する特徴

特徴	説明
RRI	R-R 間隔[ms]
HR	$RRI * 60 / 1000$
LF	交感神経成分
HF	副交感神経成分
LF/HF	交感神経系の作用成分
CVNN	RRI の変動係数
CVLF	LF の変動係数
CVHF	HF の変動係数

表 3 特徴量を算出するための統計量

統計量	説明
mean	平均
min	最小値
max	最大値
range	$max - min$
var	分散

本実験では、RRI データから下記の 8 種類の特徴（表 2）と 5 種類の統計量（表 3）を組み合わせた 40 種類の特徴量を算出した。なお、周波数データである LF と HF は、RRI データに対して、スプライン補間後 2Hz でリサンプリングし、AR 法により算出したパワースペクトルである。LF は 0.04~0.15Hz, HF は 0.15~0.4Hz の周波数領域である。さらに、下記の 3 種類を追加し、合計 43 種類の特徴量を 1 サンプルにつき使用する。

- ・NN50
隣接する RRI の差が 50 ミリ秒以上の総数
- ・pNN50
RRI の総数に対する NN50 の割合
- ・RMSSD
隣接する RRI の差の 2 乗平均の平方根

5. 評価方法

5.1 主観的なストレスの評価について

表 4 に被験者データの主観ストレス評価得点と各評価カテゴリの人数内訳を示す。主観ストレス評価得点が 3 以上を高ストレス、得点が 2 以下を低ストレスと定義している。高ストレスと低ストレスの被験者数の差が 2 名であることから、本実験は被験者ごとにストレスの感じ方に差が生じていることが判明している。全被験者の評価得点平均 2.75 は、低ストレスの基準を上回っているが、高ストレスの基準未満であるため、主観的なストレスの度合いでは中性と言える。また、本実験において低ストレスと評価した女性は 0 名であった。

5.2 機械学習モデルの構築と分類精度評価方法

表 4 被験者の主観ストレス評価得点と人数内訳

	評価得点	被験者数	
		男性	女性
全被験者	2.75 ± 0.75	16	
		12	4
高ストレス	3.33 ± 0.47	9	
		5	4
低ストレス	2 ± 0	7	
		7	0

機械学習モデルは、得られた特微量に対して、「ハラスメント」と「リラックス」の2値分類を行う。モデルの構築に使用する機械学習手法は、決定木を採用した。主な理由として、これまでの研究[10][11]で得られた成果の中で、本実験データは線形モデルと比較して、非線形モデルである決定木の分類精度が高い結果であったことが挙げられる。また、5.3章で後述する特徴選択手法において、同じ決定木系のランダムフォレストと比較して計算量の問題を軽減できる。なお、決定木のパラメータである深さは4に設定している。

モデルの組み合わせは、Harassment-Relax(All Subject)とHarassment-Relax(High Stress), Harassment-Relax(Low Stress)の3種類である。Harassment-Relax(All Subject)は全被験者(16名)のデータを用いており、全般的なハラスメント刺激を分類する。Harassment-Relax(High Stress)は本アンケートでハラスメント動画を高ストレスと評価した被験者(9名)のデータを用いており、主観的な高ストレスを分類する。また、Harassment-Relax(Low Stress)は本アンケートでハラスメント動画を低ストレスと評価した被験者(7名)のデータを用いており、主観的な低ストレスを分類する。全モデルの分類精度の評価指標はF1scoreである。ただし、Harassment-Relax(Low Stress)についてはRelaxの分類精度にも焦点を当てるため、NegativeF1scoreを追加する。NegativeF1scoreは(3)式で表す。

$$\text{Specificity} = \frac{TN}{FP + TN} \quad (1)$$

$$\text{NegativePrecision} = \frac{TN}{FN + TN} \quad (2)$$

$$\text{NegativeF1score} = \frac{2\text{Specificity} * \text{NegativePrecision}}{\text{Specificity} + \text{NegativePrecision}} \quad (3)$$

分類精度の各評価指標はLeave-One-Subject-Out(LOSO)による交差検証で算出している。LOSOはデータセットにおける被験者1名のデータをテスト用とし、残りの被験者データを訓練用として使用する交差検証法である。

5.3 特徴量選択手法

本報告では、しらみつぶし探索で算出した分類精度統計を基に、各モデルの比較結果から特徴量を選択する。しらみつぶし探索は、全組み合わせを検証することで最適な特徴量を探索する方法である。ただし、43種類の特徴量における全組み合わせの分類精度を算出することは計算量が多く難しい。そのため、1~4種類の組み合わせに限定して分類精度統計を算出している。便宜上1~4種類の組み合わせを1~4kindと表記する。また特徴量の評価・選択方法について、特徴量の組み合わせによる分類精度を、各kindに分けて降順にソートした統計から、各モデルに有効な特徴量及び共通項を比較することで実施する。

分類精度統計から各モデルに有効な特徴量の評価にあたり、比較対象である特徴量同士の多重共線性を検証する必要がある。そのため、偏相関係数を用いた分散拡大係数(VIF)を多重共線性の評価指標とした。VIFは各特徴量同士の偏相関係数をRとした場合、(4)の式に表す。

$$VIF = \frac{1}{1 - R^2} \quad (4)$$

特徴量同士の多重共線性の問題を認める基準はVIF > 10とした。

6. 評価結果

6.1 各モデルの分類精度評価

各モデルの分類精度統計を表5~8に示す。全kindのHarassment-Relaxモデルにおいて、第一位のF1scoreがHigh Stress > All Subject > Low Stressであった。この序列は、ハラスメント被害に対して、主観的なストレスが高いほど分類精度に影響を与えていることを示唆している。また、Harassment-Relax(Low Stress)のNegativeF1scoreは全kindにおいて90%以上であり、高精度にRelaxを分類できることが判明した。ただし、女性の被験者データが入っていないため、性差が含まれていない。

表5~8によると、kind数の増加に伴い第一位の分類精度が向上している。他方で、ひとつ前のkindで第一位であつた組み合わせをベースとして、1種類追加した組み合わせが分類精度統計の上位を占めているkindがある。加えて、これらの組み合わせは分類精度に1%以上の向上が見られず、ベースよりも減少する組み合わせも見られた。この結果は、ベースの組み合わせに対して、個体差のある特徴量の追加が分類精度に影響を与えている可能性がある。

6.2 各モデルの分類に有効な特徴量について

主観的なストレス差異によって分類精度が大きく変化するような特徴量は見られなかった。ただし、全特徴量が

表 5 1kind における各 Harassment-Relax モデルの分類精度統計

	All Subject		High Stress		Low Stress			
	F1score	特微量	F1score	特微量	F1score	特微量	NegativeF1score	特微量
第一位	67.63	pNN50	77.83	pNN50	50.85	pNN50	91.81	pNN50
第二位	66.86	HR_mean	76.78	NN50	43.66	RRI_mean	91.17	RRI_mean
第三位	66.26	RRI_mean	76.56	RMSSD	43.14	HR_mean	90.12	HR_mean

表 6 2kind における各 Harassment-Relax モデルの分類精度統計

	All Subject		High Stress		Low Stress			
	F1score	特微量	F1score	特微量	F1score	特微量	NegativeF1score	特微量
第一位	73.49	HR_max	84.09	HR_max	57.31	HR_max	92.42	HR_range
		NN50		NN50		NN50		pNN50
第二位	72.71	NN50	83.97	RRI_min	56.96	RRI_min	92.16	HR_range
		pNN50		NN50		NN50		NN50
第三位	71.08	LF/HF_var	81.97	RRI_min	55.76	RRI_min	91.89	RRI_mean
		NN50		HR_mean		pNN50		CVLF_min

表 7 3kind における各 Harassment-Relax モデルの分類精度統計

	All Subject		High Stress		Low Stress			
	F1score	特微量	F1score	特微量	F1score	特微量	NegativeF1score	特微量
第一位	73.72	HR_var	87.57	RRI_min	62.72	RRI_min	93.36	RRI_min
		CVLF_min		HF_var		LF/HF_var		LF/HF_var
		pNN50		NN50		NN50		NN50
第二位	73.49	HR_max	87.11	HR_max	60.96	HR_mean	93.22	RRI_mean
		HF_min		HF_var		HR_max		HR_max
		NN50		NN50		LF/HF_var		LF/HF_var
第三位	73.49	HR_max	85.94	HR_max	60.32	RRI_min	92.98	HR_mean
		HF_range		HR_range		HR_mean		HR_max
		NN50		NN50		LF/HF_var		LF/HF_var

表 8 4kind における各 Harassment-Relax モデルの分類精度統計

	All Subject		High Stress		Low Stress			
	F1score	特微量	F1score	特微量	F1score	特微量	NegativeF1score	特微量
第一位	74.04	HF_range	88.43	RRI_min	63.60	HR_max	93.51	HR_max
		LF/HF_max		HF_var		LF/HF_var		LF/HF_var
		NN50		LF/HF_min		CVLF_var		CVHF_var
		pNN50		NN50		NN50		pNN50
第二位	73.86	HF_range	88.43	RRI_min	60.96	RRI_min	93.43	RRI_min
		LF/HF_range		HF_var		LF/HF_var		LF/HF_var
		NN50		CVHF_max		CVHF_var		NN50
		pNN50		NN50		pNN50		pNN50
第三位	73.86	LF/HF_max	88.26	RRI_min	60.32	HR_max	93.42	HR_max
		CVHF_max		LF_var		LF_range		HR_range
		NN50		HF_var		LF/HF_var		LF/HF_var
		pNN50		NN50		NN50		NN50

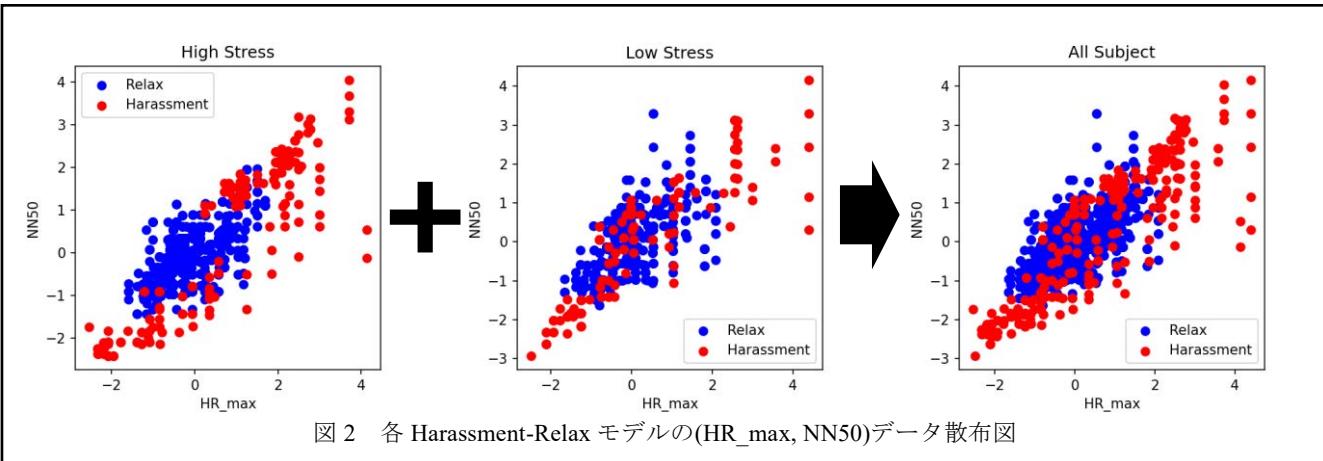


図 2 各 Harassment-Relax モデルの(HR_max, NN50)データ散布図

前節の分類精度順位に沿う傾向にあった。また、前節の考察を踏まえて、各モデルの統計を参照すると、Harassment-Relax(All Subject)では(HR_max, NN50)と(HR_var, CVLF_min, pNN50), Harassment-Relax(High Stress)では(RRI_min or HR_max, HF_var, NN50), Harassment-Relax(Low Stress)のf1score及びNegative-F1scoreでは(RRI_min or HR_max, LF/HF_var, NN50 or pNN50)がベースの組み合わせとして考えられる。ただし、(RRI_min, HR_max), (NN50, pNN50)については各モデルの分類精度統計において組み合わせが競合していた。

2kindでは、(HR_max, NN50)の組み合わせが全モデルにおいてF1scoreが最も高かった。この結果は、(HR_max, NN50)がハラスメントを分類するための共通した特徴量の組み合わせであることを示唆している。図2に、各モデルデータにおける(HR_max, NN50)の散布図を示す。Harassment-Relax(High Stress)はHarassmentデータが分散している一方、Harassment-Relax(Low Stress)では分散が見られなかった。このことから、(HR_max, NN50)の分散差が分類精度に表れている。

さらに、各モデルにおける上記の共通した組み合わせとその競合の特徴量を除くと、Harassment-Relax(High Stress)ではHF_var, Harassment-Relax(Low Stress)ではLF/HF_varが各主観的ストレスを分類する上で有効な特徴量であることが考えられる。また、ハラスメント検出モデルの評価対象である3kindのHarassment-Relax(High Stress)では、(RRI_min or HR_max, HF_var, NN50)が第三位のf1scoreと1.2%以上差があることから、他の特徴量と比較してHF_varの有意性を補強していると言える。ただし、1kindにおいてHF_varは3.96%と低精度であり、単変量での有意差が少ない。

ストレスの一般反応として交感神経が賦活し、副交感神経が抑制されるが、本実験ではこの特徴が確認できなかつた。よって、様々な実験課題におけるストレス状態でだけでなく、他の感情との比較が必要である。

表 9 VIF10以上の特徴量の組み合わせ

モデル	VIF	組み合わせ	
All Subject	11.7	RRI_range	CVNN_range
	12.2	HF_min	CVHF_min
	5945	LF/HF_max	LF/HF_range
	31.5	NN50	pNN50
High Stress	12.3	RRI_min	HR_max
	10.8	HF_min	CVHF_min
	5547.4	LF/HF_max	LF/HF_range
	18.7	NN50	pNN50
Low Stress	13.4	RRI_max	HR_min
	20	RRI_range	CVNN_range
	13.4	RRI_var	CVNN_var
	12	LF_max	LF_range
	11.2	HF_min	CVHF_min
	12000	LF/HF_max	LF/HF_range
	53.3	NN50	pNN50

6.3 VIFを含めた特徴量の組み合わせ評価

特徴量において多重共線性の問題がある組み合わせを表9に示す。全モデルデータで、(LF/HF_max, LF/HF_range), (NN50, pNN50), (HF_min, CVLF_min)の組み合わせが多重共線性の問題に該当していた。また、Harassment-Relax(High Stress)のデータでは(RRI_min, HR_max)も該当していたことから、前節で述べた分類精度統計において競合した特徴量同士が同質であることを示唆する結果となった。

7. まとめ

本報告では、VR動画視聴によってハラスメント疑似体験をした被験者のECGを基に、ハラスメント事象における被害者的心拍変動の特徴量を算出した。加えて、被験者の主観的なストレスが異なる3種類のデータを作成し、それぞれの決定木モデルを構築した。そして、しらみつぶし探索を用いて、1~4種類の特徴量の組み合わせに限定した

分類精度統計を算出した。その結果、主観的なストレスが高いほど分類精度に影響を与えることが示唆された。また、(HR_max, NN50) がハラスメントを有効に分類できる共通した特徴量の組み合わせであることが示唆された上、各モデルのデータ散布図を比較した結果、ストレスが低いデータほど分散度合いが低い傾向にあった。そして、ハラスメント検出モデルの対象である、主観的な高ストレスの被験者データを用いたモデルでは、(RRI_min or HR_max, HF_var, NN50) の特徴量の組み合わせが有効であることが示唆された。ただし、単変量での有効性が低い特徴量も含まれており、ストレスの一般反応における交感神経系の特徴も確認できなかった。そのため、今後の課題としては、様々な実験課題におけるストレス状態でだけでなく、他の基本感情との生体反応比較を行う必要がある。また、有効な特徴量を正確に評価するためには定量的な特徴選択手法による検証が必要となる。

参考文献

- [1] “文部科学省 令和3年度 児童生徒の問題行動・不登校等生徒指導上の諸課題に関する調査結果,” https://www.mext.go.jp/content/20221021-mxt_jidou02-100002753_1.pdf, (参照 2023-02-15).
- [2] “厚生労働省 令和3年度個別労働紛争解決制度の施工状況要,” <https://www.mhlw.go.jp/content/11909000/000959370.pdf>, (参照 2023-02-15).
- [3] “厚生労働省 令和2年度職場のハラスメントに関する実態調査報告書,” https://www.soumu.go.jp/main_content/000538672.pdf, (参照 2023-02-15).
- [4] “文部科学省 いじめ防止対策推進法（概要）,” https://www.mext.go.jp/a_menu/shotou/seitoshidou/1337288.htm, (参照 2023-02-15).
- [5] Kim, J., Ho, T. D., Kim, J., et al.: Can Bullying Detection Systems Help in School Violence Scenarios?: A Teachers' Perspective, In Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems, pp.1-7(2020).
- [6] Brahnam, S., Roberts, J.J., Nanni, L., et al.: Design of a bullying detection/alert system for school-wide intervention, Proc. Human-Computer Interaction: Interaction Technologies, pp.695-705 (2015).
- [7] Ferdinando, H., Ye, L., Han, T., et al.: Violence detection from ECG signals: a preliminary study, Journal of Pattern Recognition Research, Vol.12, No.1, pp.7-18 (2017).
- [8] Ye, L., Wang, P., Wang, L., et al.: A combined motion-audio school bullying detection algorithm, International Journal of Pattern Recognition and Artificial Intelligence, Vol.32, No.12, 1850046 (2018).
- [9] 入江正洋: 職場のパワーハラスメント : 現状と対応, 健康科学, Vol.37, pp.23-35 (2015).
- [10] 上野貴弘, 大橋正良: 心拍情報に基づくいじめ検出手法の一検, 情報処理学会研究報告, Vol.2022-HCI-199, No.25, pp.1-6 (2022).
- [11] 上野貴弘, 大橋正良: 心拍情報に基づくいじめ検出手法の一検, 電子情報通信学会総合大会, H-2-3, 2023年3月 (発表予定) .
- [12] Baltatzis, V., Bintsi, K. M., Apostolidis, G. K., et al.: Bullying incidences identification within an immersive environment using HD EEG-based analysis: A Swarm Decomposition and Deep Learning approach, Scientific reports, vol.7, pp.1-8(2017).
- [13] Minhad, K. N., Sawal, H. M. A., and Mamun, B. I. R.: Happy-anger emotions classifications from electrocardiogram signal for automobile driving safety and awareness, Journal of Transport & Health, Vol.7, pp. 75-89 (2017).
- [14] Mavridou, I., Seiss, E., Kostoulas, T., Nduka, C., and Balaguer-Ballester, E.: Proceedings of the Workshop on Human-Habitat for Health (H3), No.4, pp. 1-6 (2018).
- [15] Park, C. Y., et al.: K-EmoCon, a multimodal sensor dataset for continuous emotion recognition in naturalistic conversations." Scientific Data, Vol.7.1, No.293 (2020).
- [16] Schmidt, P., Reiss, A., Dürichen, R., Laerhoven, K. V.: Wearable-Based Affect Recognition—A Review, Sensors. Vol.19, No.19, pp. 4079(2019).
- [17] “厚生労働省 動画で学ぶハラスメント”, <https://www.mhlw.go.jp/movie/>, (参照 2023-02-15)

資源循環促進に向けた住民の行動変容

江島直也^{†1} 日室聰仁^{†1} 笹鹿祐司^{†1} 後藤晶^{†2}
NECソリューションイノベータ株式会社^{†1} 明治大学^{†2}

1. 背景

近年、大量生産、大量消費、大量廃棄型社会から資源循環型社会への転換が求められている。宮城県南三陸町では地域資源循環の取組として、可燃ごみ回収とは別に生ごみを回収している。町内約250か所のごみ集積所に生ごみ回収用のバケツが常設されており、住民はいつでもバケツに生ごみを捨てることができるようになっている。回収した生ごみはバイオガス施設で電力や液肥に変換し地域内で再利用している[1]。生ごみ回収の課題として、生ごみの回収量を増やすことと生ごみの中の異物を減らすことがある。課題解決のためには住民のごみ分別行動を変容する必要があり、筆者らは行動経済学や心理学の知見にもとづいて介入をデザインし、実験を通して効果を検証した。本稿ではこれまでに実施した実験とその結果、そこから得た学びについて報告する。2章では生ごみの回収状況を把握するためには作成したシステムについて述べる。3章では1つ目の実験である感謝をフィードバックすることで住民の向社会的行動を促す2つの実験について述べる。4章では損失情報をフィードバックすることで住民の損失回避行動を促す1つの実験について述べる。

2. 生ごみ回収状況把握環境の整備

施策の効果を検証するためには施策前後のデータ比較が必要となるが、生ごみ回収のデータ計測方法には2つの問題があった。

1つは計測データがデジタル化されていない点である。元々は資源回収業者が紙にデータを記録し、それを図1のように手動で集計していた。そのため、分析に手間がかかる状況であった。

もう1つは分析に必要なデータが不足している点である。元々の生ごみ回収に関する計測データにはごみ集積所に生ごみが「ある／ない」の情報しか記録されていなかった。また、分別品質に関しては、複数のバケツのうち1つでも異物が混入している場合にすべてのバケツに異物ありと記録されており、情報量が少なくなっていた。

前者の問題に対して、筆者らはタブレットPCを使いデジタルデータとして計測データを保存するシステムを開発す

ることで解決を図った。システム開発の際には、タブレットPCを使い慣れていない回収業者がシステムを利用する想定し、インターフェースを単純な設計にした。また、ユーザーである作業員にシステムの使用感を確認しながら開発を行った。システムの回収状況入力画面を図2に示す。

後者の問題に対しては、計測データに記録する内容を見直した。生ごみの回収量については、バケツ毎に0(空)～5(超)の6段階とし、図3のような基準で記録することとした。分別品質については、バケツ毎に混入した異物を記録できるようにし、評価の際に異物無し2点・異物あり1点・資源なし0点として扱った。

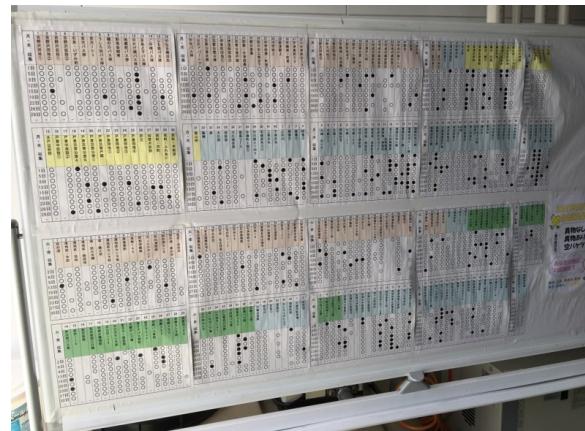


図1 生ごみ回収データの集計結果(手動)



図2 生ごみ回収状況入力システム画面

Change of residents' behavior to promote resource circulation

†1 NAOYA EJIMA, NEC Solution Innovators, Ltd

†1 AKIHITO HIMURO, NEC Solution Innovators, Ltd

†1 YUJI SASAKA, NEC Solution Innovators, Ltd

†2 AKIRA GOTO, Meiji University

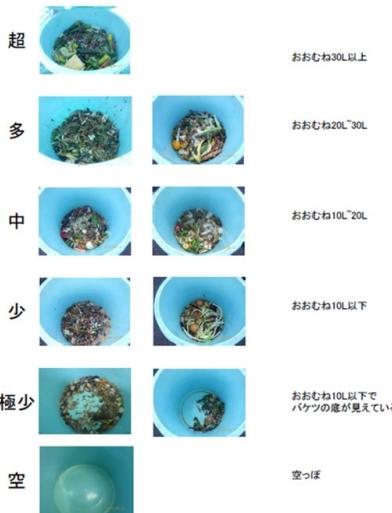


図 3 生ごみ回収量の記録基準



図 4 集積所に掲示した感謝状



図 5 看板を実際に設置している様子

3. 施策 1：感謝フィードバックによる施策

地域資源循環の取り組みにおける町民の行動は、一種の向社会的行動として理解できる。向社会的行動とは、社会のために無私の心で行う行為を指し、寄付やボランティア活動を支える心性のひとつである。地域資源循環の取り組みは金銭などの外的な報酬がなく、町民の地域をより良くしたいという善意によって成り立つ活動のため、筆者らは向社会的行動であるととらえた。また、藏永らの研究によると、感謝を受取るなどの恩恵の受領評価が向社会的行動に影響を及ぼすことが示唆されている[2]。住民の分別行動に対して感謝をフィードバックすることで、生ごみ回収量や分別品質の向上に繋がるのではないかと考え、以下の実験を実施した[3][4]。

3.1 生ごみ集積所への感謝状の設置実験

3.1.1 実験内容

感謝フィードバック前後の生ごみ回収量と分別品質の変化を確認するために、ランダムに抽出したごみ集積所に図4の感謝状を貼り付けた看板を設置し、感謝を町民へフィードバックした。看板には感謝状の他に、過去のごみ出し傾向、ワンポイントアドバイスを記載した紙を張り付けて設置した。1週間あたり6か所に感謝状を掲示し7週間で計42か所の生ごみ集積所に掲示した。実際に看板を設置している様子を図5に示す。

3.1.2 結果と考察

感謝状を掲示した42か所の集積所のうち、有効なデータを取得することができた31か所を掲示群として分析した。掲示群と非掲示群ごとに、掲示前後1カ月の生ごみ回収量の平均変化量を図6に、分別品質の平均変化量を図7に示す。生ごみ回収量・分別品質ともに、掲示群において非掲示群に対して向上が確認された。t検定による仮説検証では、分別品質にのみ有意差がみられた。集積所に感謝状を設置するという物理的制約のため、掲示群のサンプルサイズを十分に確保できず、生ごみ回収量が有意に向上するという証左までは得ることができなかったと考えられる。

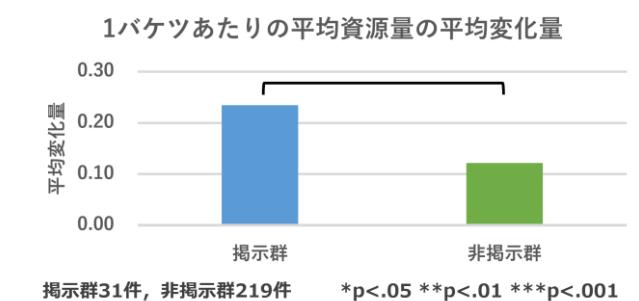


図 6 掲示群と非掲示群の生ごみ回収量の変化量の比較

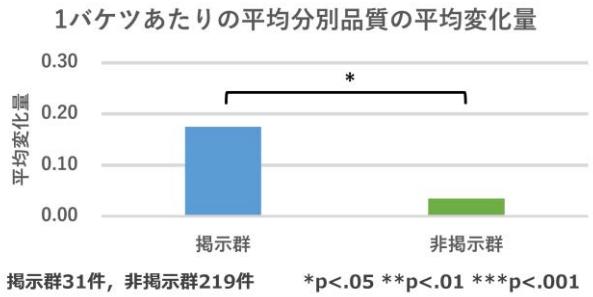


図 7 掲示群と非掲示群の生ごみ分別品質の変化量の比較

3.2 住民に対する感謝状の配布実験

3.2.1 実験内容

3.1.1 の感謝状設置の実験では、感謝状を設置した生ごみ集積所が 42 か所に対して、感謝状を設置しなかった生ごみ集積所が 219 か所と群分けのバランスが悪くなっていた。また、感謝状を設置したタイミングが集積所ごとに異なってしまい、季節性の影響を受け正しく分析できない可能性があるという問題点があった。本実験では上記の問題を解消するような実験設計として、感謝状を南三陸町の広報誌に挟み配布した。全 251 か所の集積所のうち、125 か所の集積所を利用する住民に感謝状を配布した。

3.2.2 結果と考察

感謝状を配布した住民が利用する 125 か所の集積所のうち、有効なデータを取得することができた 124 か所を介入群、感謝状を配布していない住民が利用する 126 か所の集積所のうち有効なデータを取得することができた 122 か所を待機群として分析を行った。介入群と待機群ごとに、介入群に感謝状を配布した前後一ヶ月の生ごみ回収量の変化量を図 8 に、異物混入率の変化量を図 9 に示す。生ごみ回収量は介入群の方が増加していたが異物混入率は介入群の方が悪化していた。t 検定による仮説検証では、異物混入率にのみ有意差がみられた。感謝状を広報誌に挟み配布したことによる、ごみの分別行動の促進は見られなかった。実験前に群ごとに回収傾向に大きな差がないことを確認していたが、群ごとに異物の種類ごとの混入率を確認したところ、介入群の方が待機群と比べ月ごとの値のブレが大きかった。この偏りが影響して介入群の異物混入率が悪化したように見えた可能性が考えられる。

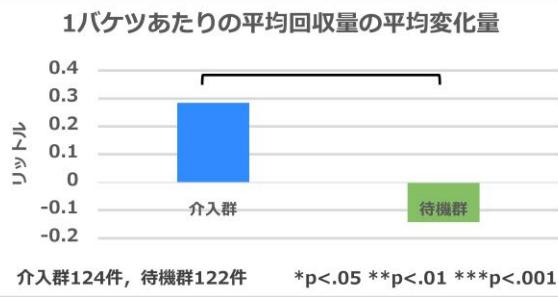


図 8 介入群と待機群の生ごみ回収量の変化量の比較

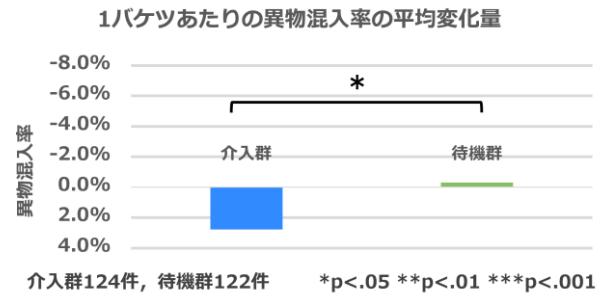


図 9 介入群と待機群の異物混入率の変化量の比較

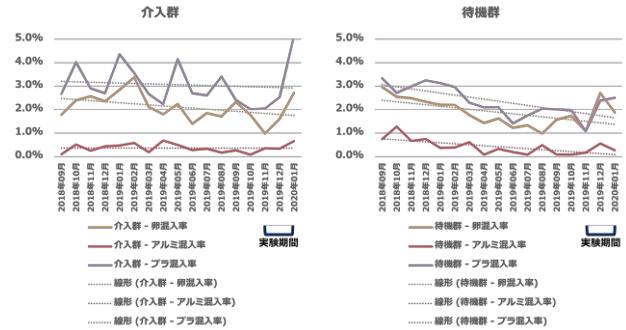


図 10 群別異物ごとの混入率推移

4. 施策 2：損失情報フィードバックによる施策

行動経済学で近年盛んに研究されているプロスペクト理論では、損失回避性というものがある。これは、人は利益を得るときよりも損をするときの方がより感情的に反応し、その結果、人は損失を回避するような行動を取るというものである[5]。

生ごみを分別せず可燃ごみとして廃棄することによる損失の情報をフィードバックすることで、損失回避行動を促し、生ごみ回収量の増加に繋がるのではないかと考え、以下の実験を実施した。

4.1 可燃ごみ袋売り場へのポスター掲示実験

4.1.1 実験内容

南三陸町には可燃ごみ処理施設がなく、隣町である気仙沼市に年間 4,200 万円を支払い処理している。4,200 万円の損失を町が負担しているということを記載したポスターをデザインし、住民に提示することで損失回避行動を促した。また、ポスターには生ごみ回収に協力することで可燃ごみ量が減少し、焼却費や CO₂ の排出量が削減できることを記載した。生ごみ回収という具体的な行動を記載することで、情報を受け取った人が行動に移しやすくなる狙いがある。町内で可燃ごみ袋を販売している 20 店舗のうち、11 店舗の可燃ごみ袋を販売している棚にこのポスターを掲示した[6]。可燃ごみ袋売り場に掲示することで、住民がごみについて考えているタイミングで損失の情報を伝えることができるようとした。掲示したポスターを図 11 に示す。

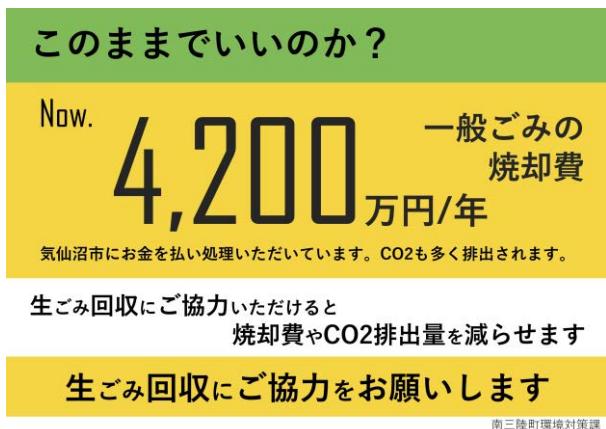


図 11 掲示したポスター

4.1.2 結果と考察

スーパー・マーケット等でポスターを見た住民がどの集積所を利用しているかを特定することができないため、本実験では 3.1.2 や 3.2.2 のように介入群と非介入群を比較し効果を確認することができなかった。そのため、過去の生ごみ回収量からポスター掲示後の生ごみ回収量を予測し、その値と実際に回収された生ごみ量を比較することでポスター掲示による介入効果を確認した。予測には時系列因果推論フレームワークである CausalImpact[7]を利用し、2016 年から 2020 年の生ごみ回収量を学習データとして本実験を実施した 2021 年の生ごみ回収量を予測した。予測値と実際の回収量を比較した結果を図 12 に示す。ポスターを掲示した 37 週目以前は予測値と実際の回収量が同じような値となっているが、ポスター掲示以降 12 週間は実際の回収量が予測値より 14.68% 高くなっていた。この結果から、ポスター掲示による住民への損失回避行動の促進には一定の効果があったと考えられる。理由としては、損失額を見せてることで住民の損失回避行動の促進に繋がったことや、ごみ袋売り場にポスターを掲示したことで住民がごみのことを考えているタイミングでの介入となり行動変容に繋がった可能性が考えられる。



図 12 生ごみ回収量と予測値

5. まとめ

本研究では、宮城県南三陸町で取り組んでいる生ごみ回収における、回収量の増加と分別品質の向上という 2 つの課題について行動経済学や心理学の手法を用いて 3 つの実験を実施した。

生ごみ集積所に感謝状を掲示することにより住民に感謝をフィードバックし、生ごみ分別行動を促す実験では回収量、分別品質ともに感謝状を設置した集積所の方が向上し分別品質は有意差がみられた。

感謝状を広報誌に挟み配布することで感謝をフィードバックし、生ごみ分別行動を促す実験では、回収量は感謝状を配布した住民が利用する集積所の方が向上していたが異物混入率は悪化しており有意差がみられた。

南三陸町の焼却費を記載したポスターをデザインし可燃ごみ袋売り場に掲示することで、損失回避として生ごみ分別行動を促す実験では、ポスター掲示によって生ごみ回収量が予測値より約 14.68% 高くなかった。

これらの結果から、生ごみ分別という行動に対して、行動経済学や心理学の手法を用いてポジティブに変容できる可能性が示されたが、一方で情報の提示方法次第で効果に差があることが認められた。一定の効果が確認されたケースはごみ捨てや買い物など既存の行動の導線上に情報を掲示し、かつ繰り返し接触する機会が提供できており、介入をデザインする上で重要な要素であると考えられる。実際の運用においては利用できるメディアの制約など、必ずしも効果的な導線上に介入を設計できないことも考えられ、そのような場合には感性に訴えかけるレイアウトを検討するなどの工夫が必要だろう。

参考文献

- [1] アミタホールディングス株式会社:南三陸 BIO(ビオ)資源循環の基盤づくり事例
https://www.aise.jp/case/circulation/minamisanriku_bio.html(参照 2023-02-13)
- [2] 蔵永瞳,樋口匡貴:感謝生起状況における状況評価と感情体験が対人行動に及ぼす影響,心理学研究 84.4(2013):376-385
- [3] 日室聰仁,笹鹿祐司,福井知宏,河又恒久:資源回収状況のデジタル化および感謝フィードバックによって向社会的行動が促進されることの検証,情報処理学会第 81 回全国大会 2K-03(2019)
- [4] 日室聰仁,笹鹿祐司,江島直也,福井知宏,後藤晶:IT を活用した感謝フィードバックによる資源循環促進の検証,情報処理学会全国大会第 82 回全国大会 4G-05(2020)
- [5] Kahneman, Daniel, and Amos Tversky. "Prospect Theory: An Analysis of Decision under Risk.",Econometrica 47.2 (1979): 263-292.
- [6] 行動経済学会第 16 回大会ベストナッジ賞コンテスト
http://www.abef.jp/conf/2022/common/doc/program_best-nudge.pdf(参照 2023-02-17)
- [7] KAY H. BRODERSEN, FABIAN GALLUSSER, JIM KOEHLER, NICOLAS REMY AND STEVEN L. SCOTT: INFERRING CAUSAL IMPACT USING BAYESIAN STRUCTURAL TIME-SERIES MODELS, The Annals of Applied Statistics 2015, Vol. 9, No. 1, 247–274

健康メディアデザインによる 大学生の生活習慣の改善と集中力の向上

千種 康民¹ 保坂 浩太郎¹ 熊谷 華菜¹ 落合 育海¹ 今野 駿太¹

概要：本報告は筆者らが 2015 年から発案して実践してきた健康メディアデザインに基づく成果をまとめたものである。メディア学部の多くの大学生は生活スタイルが夜型でかつ PC 作業やゲームなど液晶ディスプレイの使用時間が長く、朝食欠食も運動不足も長時間デスクワークも多い傾向にあり、結果として日中に実施される授業や学修に深刻な悪影響をもたらしていた。これらを踏まえて、最初は学生本人の睡眠不足や眼精疲労など健康状態を低下させている要因分析を行ない、悪影響への寄与の高い、あるいは現状の早急な改善が必要であるなどの改善対象となる健康テーマを決定する(PLAN)。そして健康テーマに関連すると考えられる行動や健康状態を学生本人が記録し、現状分析から改善に至るための有効な対策を必要最小限だけ実施し(DO)、その経過観察をデータサイエンス的な視点から可視化と分析し(CHECK)、改善手法の追加・変更(ACTION)という、健康 PDCA サイクル着想して実践し有効な結果を残してきた。これらの活動を卒業研究として位置づけるため、前述の過程を基に自身の健康テーマ改善に役立つ健康アプリをデザインすることも併行して実施し、情報系への就職活動に役立ててきた。本報告ではそれらの成果を報告する。

キーワード：健康メディアデザイン、健康 PDCA サイクル、健康アプリ、低負担な改善手法、睡眠改善、集中力向上

Improving lifestyle-Habits of Students by Health Media Design Method

YASUTAMI CHIGUSA¹ KOTAROUH HOSAKA¹
HANA KUMAGAYA¹ IKUMI OCHIAI¹ SHUNTA KONNO¹

Abstract: This report summarizes the results based on the health media design that the authors have proposed and practiced since 2015. Many university students in the Faculty of Media Studies have a night-time lifestyle and spend a lot of time using liquid crystal displays such as PC work and games, and tend to skip breakfast, lack exercise, and work at a desk for long periods of time. It had a serious negative impact on teaching and learning. Based on these, we will first analyze the factors that are deteriorating the student's health condition, such as lack of sleep and eye strain. Decide on a health theme (PLAN). Students themselves record behaviors and health conditions that are considered to be related to health themes, analyze the current situation, implement the minimum necessary effective measures for improvement (DO), and observe the progress from the perspective of data science. We have visualized and analyzed (CHECK), added and changed improvement methods (ACTION), conceived and practiced the health PDCA cycle, and left effective results. In order to connect these activities to graduation research, based on the above-mentioned process, I also designed a health app that would help me improve my health theme, which has been useful for my job hunting in the information field. This report reports those results.

1. 健康メディアデザインとは

研究者らは、2014 年に研究室に配属になった大学生の 1 名との対話の中で「卒業研究に必要なプログラミングに取り組みたいけれど、日中が眠くてほとんどプログラミングが学修できない」という発言を発端として、人間の生活習慣の状況を「健康メディア」ととらえ、人間ひとりひとりの生活習慣の乱れ等の正常化や改善を目的とする方法論を「健康メディアデザイン」と定義し、その改善プロセスを「健康 PDCA サイクル」と呼び、週 1 回の授業やゼミにおいて、記録⇒分析⇒報告⇒議論⇒次週対策の策定、を毎週の授業時に実施するスタイルを確立し、対外発表によ



図 1 2018 年卒研生の 3 年次の前期から後期の GPA の変化

1 東京工科大学メディア学部

Tokyo University of Technology, Faculty of Media Science



図2 2019年卒研生の3年次の前期から後期のGPAの変化

卒研関連が半期×3科目（卒研生：5～14名）、プロジェクト演習関連（履修生：3～8名）が半期×2科目、プレミアム科目（履修生3名前後）、大学院の特論（履修生：約20名）が四半期×1科目、であり、大多数が一定の成果を得てきました。総数126名の3分の2程度95名の健康改善テーマが「睡眠改善」であり、これに関しては50%以上が下記のKPI(Key Performance Indicator)が3項目以上改善し、90%以上が1項目以上改善した。KPIとして、①入眠時間短縮、②2度寝時間の短縮、③睡眠時間の増加、④朝食摂取率の上昇、⑤午前中の集中度の改善、⑥午後の集中度の改善、を使用した。卒業研究生においては⑦GPAの0.2ポイント以上上昇も使用した。

図1に示す2018年に3年後期から配属された卒研生13名の健康テーマは13名中12名が睡眠改善であり、配属されてから半年後のGPAは平均0.28ポイントの上昇を示している。図2に示す2019年の卒研生5名は全員の健康テーマが睡眠改善であり、同じ半年後のGPAは平均0.9ポイントの大幅上昇を示している。

学年全体のGPAの変化と比べて優位であることの確認が必要であるが、GPAが一定以上上昇した学生に関しては、健康メディアデザインの方法論の確立や少人数教育の効果などの影響のほか、睡眠改善により日中の授業における集中力が改善され、授業中の学修がスムーズであったことなどが考えられる。

2. 健康PDCAサイクルとその基本原則

最初の1年間を経過し、この2年間の健康PDCAサイクルで実施した記録は、日々の睡眠についての諸情報の記録を行なった。ベッドや布団に入る時刻である入眠時刻、入床してから入眠するまでの入眠時間、朝起きる起床時刻、一旦起きてから完全に起床するまでの二度寝時間、を記録

した。改善が好ましい学生はアラーム無しでの自然な目覚めの起床回数を記録した場合もあった。

さらに学生たちとの対話を踏まえて、睡眠に影響のある項目も随時増やしていく。入床前の環境づくりとして、歩数、入浴時間や入浴後から就寝までの時間、入床前のスマートフォンやPC作業の終了時刻、起床後の活動状態として、二度寝時間、朝食摂取の有無、午前中の集中の度合い、午後の集中の度合い、夕食後の集中の度合い、なども随時記録していく。さらに実施する対策として、学生が記録してきた状況や実施した対策の効果を確認しながら、睡眠状況を改善する手法を学生とともに調査しながら、学生当人に適していると思われる手法を実践していく。

この時点では、従来11時前後に起床していた学生が、スタート時に8時～9時起床と急激な変化のある行動を設定した場合に、適応できる学生と適応できない学生の差が生じた。これは適応力の高い学生、自律心の強い学生とそうでない学生の差と考えられるが、実際に急激な変化は人体への負荷が大きく、急激なダイエットはリバウンドするケースが多いなど、決して好ましい状況に至らないことも明らかである。

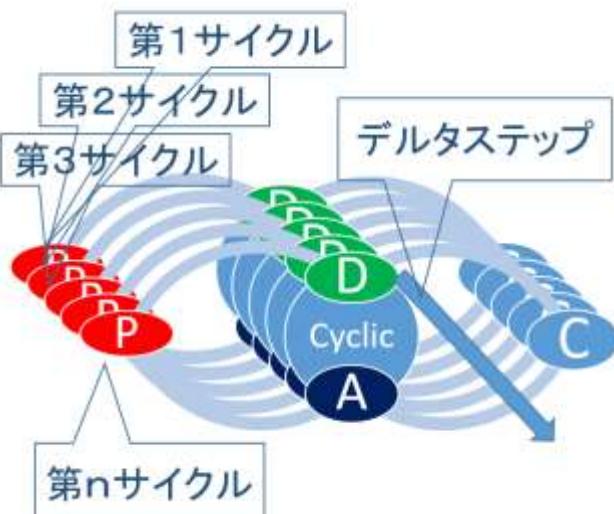
2015年から2年間でこの人体の適応力を考慮した健康PDCAサイクル(図3)の基本原則を以下のように設定した。

1. 健康PDCAサイクルは1週間1サイクル、半期で15サイクル前後である。
 2. 1～4週間後に改善意欲の強い健康テーマを選定し、その目標設定をし、記録する項目は必要最小限からスタートする。
 3. 様々な記録は、自動記録可能なスマートウォッチや睡眠記録アプリなどを積極利用し、途中から様々な記録ができる記録アプリを使用し、正確性と網羅性を担保する。
 4. 新たに導入する対策は週1回の授業中に、最も効果が期待できる、あるいは、もっとも導入が容易な対策の複数の中から学生本人がその選択理由とともに1つを選択し、その効果測定が可能なKPIを設定し、導入を決定し実施する。
 5. 新たに導入した対策の効果が確認できない場合でも1,2週間は経過観察する。3週以上経過して効果が確認できない場合は、対策を停止し、他の対策を1つだけ選択する。
 6. 週1回の授業前に、学生は各自の記録をまとめ分析し、その分析内容を報告できるようにしておく。
 7. 授業中は学生ひとり一人が1週間の記録とその分析結果を報告し、教員および他学生と議論を行ない、その後、次週の対策について検討する。
 8. 少しの体調不良でも自宅から参加でき、オンラインで習熟したzoomのチャットを用いて、他学生からのリアクションを得るとともに時短を図る。
- 以上の基本原則に則り、事前に提供するテンプレートとし

ての“行動記録シート”に毎日の状況や対策などを必要最小限の項目のみ記録することからスタートする。



(a) 1週間の健康 PDCA サイクル



(b) 繰り返し実施する健康 PDCA サイクル

図3 健康 PDCA サイクル

通常の一般企業などで実施される業務改善などの PDCA サイクルは、長い一定の期間をかけて改善量が数 10%と大幅な改善を目標設定し、失敗が許されなく厳格かつ多面的に同時進行させていく行動計画が設定される。それをそのまま人間の生活スタイルに適用すると人体への負荷が過大になる可能性が高く、より多くの学生が適応不可能となってしまうため、業務改善的な PDCA サイクルのアプローチは必ずしも適さない。

改善目標は容易に到達できるレベルより少しだけ上の位置を設定する。健康 PDCA サイクルは、毎週 1 サイクル、設定する対策とその効果測定のための KPI は 1 つもしくはごく少数としている。

そして実績の改善量は必ずしもプラスでなくともよく、例えば、風邪を引いたとか、冬休みに生活習慣が乱れて、改善した状態が悪化する場合など、改善しないケースも多々ありとして許容している。

例えば、起床時刻を早めるという健康テーマを設定した場合の最終的な目標は、例えば通常の社会人が通勤時間 1 時間かけて 9 時から 18 時まで勤務するような一般的な

社会人の生活スタイルから派生する「平日」 + 「起床時刻」 + 「7 時以前」であるが、いきなり来週からの目標が 7 時起床では、持続する前に破綻する可能性が高い。したがって、スタート時の現状が「平日」 + 「起床時刻」 + 「10 時以前」であるなら、1 ~ 4 週間後の目標設定を例えば「平日」 + 「起床時刻」 + 「9 時半以前」と緩めにであるが平日 5 日間と設定する。実現が困難であった場合は「平日」を「平日の 3 日以上」とより緩い目標設定にする場合もある。この目標設定が安定して達成された場合の次の目標設定は「9 時以前」とするのでなく、「平日」を「毎日」と変更し、毎朝の起床時刻を揃えて、人体の 24 時間時計の働きを有効活用する方向に誘導していく。

他にも多様な問題がある可能性があるが、起床時刻に問題がある場合は、就寝時刻や入眠時間にも問題がある場合が多いため、起床時刻だけを早期化しても、睡眠時間が短縮して悪化するケースもあり、そのままだと破綻してしまう。その場合は、一定以上の睡眠時間を確保するために、就寝時刻あるいは入眠時間の改善余地の多い、あるいは改善可能性が高い項目を選定し、少しだけ改善された起床時刻を固定したまま、どちらかの項目を改善するための健康 PDCA サイクルを実施する。前述の取り組みをしても、一定以上の睡眠時間が確保できない場合は、昼食後の眠気が深刻な場合が多く、午後の授業に対する集中度が著しく悪化するため、対症療法として、昼食後に 20 から 30 分の仮眠を取り入れるケースもある。

以上のように起床時刻、就寝時刻、入眠時間、睡眠時間、二度寝、昼食後の眠気、など睡眠問題に関しては多種多様で実績のある対策が多く蓄積されてきた。眼精疲労や朝食欠食に対する実績も地道に蓄積されてきている。

KPI は「」当面の起床時刻など、とすることになるが、1 ~ 4 週間後の目標値は本人の現状と目標から換算するにとって到達可能

数年間の実績により、事前に学生に提供する健康テーマのテンプレートは、適用人数の多いものから、”睡眠”、“朝食欠食”、“眼精疲労”、“首コリ・肩コリ”、“運動不足”、“ストレス”、“集中力”、である。コロナ禍の影響もあるが、ディスプレイ作業の多いメディア学部ならではの傾向があるともいえる。

3. 反転授業としての健康メディアデザイン

能動的学修が期待される、グループワークなどを取り入れたアクティブラーニングと自宅での学修を組み合わせた反転授業では、学修者（ここでは大学生）が受け身ではなく、自ら能動的に学びに向かうよう設計された教授・学習法により学修をしていくことにより学習効果を高めている。学修者の認知的、倫理的、社会的能力、教養、知識、経験といった能力を育むことが期待されている。

図 4 に示す 7 階層のラーニングピラミッド[9]において、従来の講義スタイルの学修定着率は 5%と言われており、

読書や視聴覚資料を用いた授業では学修定着率が 20%まで上昇する。さらに実習的なデモンストレーションにより、



図4 反転授業におけるラーニングピラミッドと学修定着率



(a) 健康メディアデザイン授業中（発表学生）



(b) 健康メディアデザイン授業中（他学生）



(c) 予習中（発表準備）

図5 健康メディアデザインの授業中と授業外

学修定着率は 30%まで上昇すると言われている[9]。またラーニングピラミッドの第 5 層から第 7 層に該当するアクティブラーニングの実施法と第 1 層から第 3 層に該当する事前予習法との適合性に応じて差異が生じることになる[10]。

一方、1週間ごとの健康メディアデザインの授業において、図 5 (a)に示す発表学生と教師のインタラクションは、アクティブラーニングとして、第 5 層の教師との討論を中心に、第 4 層の可視化された自身の記録データ分析に基づくデモンストレーション発表と 1 週間記録された第 6 層の効果実感を体験している。

また、図 5 (b)に示す、他学生が発表学生や教師との討

論などから受けるインタラクションは、第 3 層の可視化されたデータについての第 1 層の解説、それらを聞いた上で第 5 層の発表学生へのコメント、コメントは場合によっては発表学生へのアドバイスという第 7 層の教授、も発生する場合がある。

図 5 (c)の発表準備において、自分自身の 7 日分の行動記録において第 6 層の効果実感を得て、第 3 層の直前のデータ可視化と、その分析結果に関する第 2 層の調査が必要となり、第 4 層のレポート作成と第 5 層の考察により発表準備を実施する。

1 コマの授業においては、全受講生数の発表があり、学生は 1 回の発表学生と、受講生数 - 1 の回数の他学生を演じることになる。これらにより全学生の多様な健康テーマに関する情報も多様に共有され、学生の知識の獲得にも該当テーマだけに関わらず多種多様な効果が得られる。

4. ケーススタディ

4.1 夜型生活から朝型生活化による集中力改善例[5]

2020 年前期のプレミアム科目履修学生が夜型生活であり、夜間のながら作業による作業効率悪化と、その影響で入眠時間が 2020 年 5 月の段階で、平均睡眠時間 5 時間、平均入眠時間 60 分という睡眠不足問題があった。

5 月から当面は、入眠時間 60 分の短縮と、就寝時刻の早期化、定刻 8:30 の起床時刻化により 6 時間以上の睡眠時間の確保を当面の目標と設定した。約 1 か月後には就寝時刻の早期化と定刻起床が安定的に実現でき、6 時間半以上の睡眠時間および午前中 3 時間半が確保できた。その後、夜間にながら作業で作業効率が悪化していた創造的作業を安定的に確保できた午前中の 3 時間半内へシフトした。

6 月中盤からは、苦手科目のだらだら作業の時短をポモドーロ法により試行した。並行して前述の夜間作業の午前中への移行と、定刻 8:00 のより早期の起床時刻により、午前中の作業時間が 30 分増え 4 時間が確保でき、入眠時間の短縮にも効果が出始め 30 分程度まで改善できた。これらにより時短できた余剰時間は半分程度を睡眠時間に組み入れて増やし、さらに午前中作業も増やし、日中気分と集中力の向上に成功した。余剰した自由時間も確保できた。

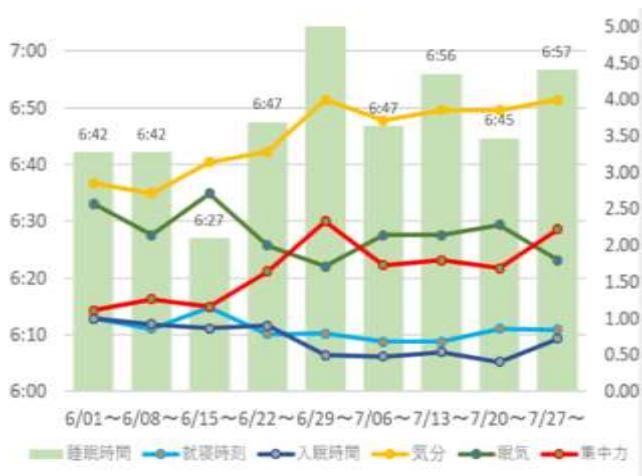
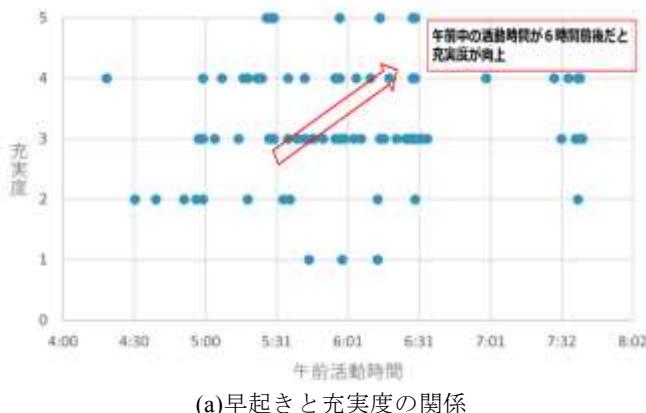


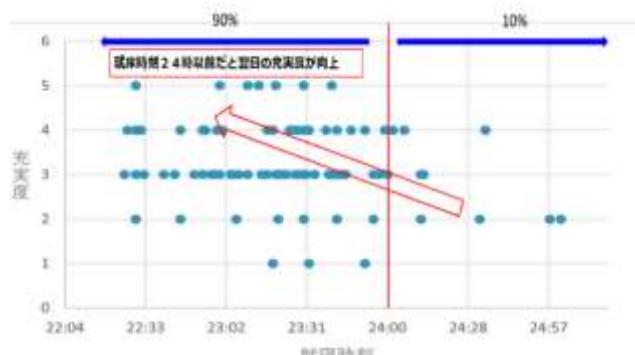
図 6 作業の朝型化と入眠時間の短縮による効果

4.2 早寝・早起きの効果測定（充実度として）

2022 年後期のプレミアム科目の履修生が朝活を実践し、その充実度を観測した。6 時くらいに起床する学生であったが、5 時半～6 時半の早起きをすると午前中の予定が順調にはかかるため、1 日の充実度が高くなった。また 23 時前後に就寝すると翌日の充実度が高くなかった。これは 23 時前後の就寝により、早朝起床になるためと推察される。



(a)早起きと充実度の関係



10/14 入床 1 時間前にスマホを置く

11/3 就寝前に紙の語学勉強開始

11/8 就寝前にお香開始

11/17 就寝前に塗り絵開始

悪い

良い

(b)早寝と充実度の関係

図 7 早寝と早起きの効果測定

4.3 入眠時間短縮に効果的なルーチン

2022 年後期のプレミアム科目の履修生が時々 60 分を超える入眠時間があり、翌朝に悪影響があり問題であった。この現状に対して、10/13～11/11 まで入眠時間と対策を記録した。①就寝 60 分前にスマホ利用を停止、などを実践し、現在は安定的に 10～20 分程度に安定している。

図 8 様々なルーチンを取り入れた入眠時間の改善

4.4 ストレスコーピングによる大学生活の充実化

2022 年後期のプレミアム科目の履修生が、苦手科目によりモチベーション(気分)が低下している問題を改善するためにストレスコーピングを用いて対策した。図 9 に示すように、開始前の気分は最悪であったがストレスコーピングを開始してから着実に効果が表れ、ジム利用開始なども効果がみられて、後半には、通常以上に改善できた。

ストレスコーピング開始



(a)改善の過程(1週ごとのサンプル)

番号	ジャンル	項目	効果	どういうときにやったか
1	音楽	カウセリングに行く	○	いつでも
2	音楽	いいことを声に出す	△	
3	音楽	よくなかったことを声に出す	○	
4	音楽	がんばれ私！と高い声で叫ぶ	○	バイト直前
5	音楽	歌謡曲	○	何もやる気が起きないととき
6	音楽	歌と話す	○	いつでも
7	音楽	友人と話す	○	ある程度元気な時
8	運動	ウォーキング	○	いつでも
9	運動	ジョギング50分	○	いつでも
10	運動	ジョギング20分	△	
11	運動	ジム	○	いつでも
12	運動	プール	○	いつでも
13	運動	練習	○	
14	音	自分とマッチした曲を聴く	○	モチベーションを上げたい時
15	音	お気に入りの曲を聴く	○	モチベーションが下がった時
16	音	サウンドトラックを聴く	○	モチベーションが下がった時
17	音	水の流れる音を聞く	○	
18	外出	コンビニに行く	○	
19	解放	カラオケに行く	○	いつでも

(b)ストレスコーピングリストと効果のある対策

図 9 ストレスコーピングによる苦手科目の克服

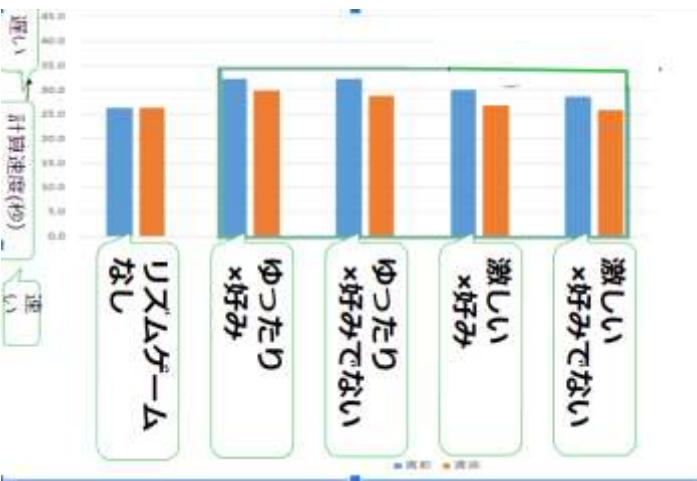
4.5 作業開始前のリズムゲームによる集中力向上法

卒業研究生が、昼食後の眠気のために午後の集中力が低下する問題があり、その問題解決のための対症療法として、本人が好んでいるリズムゲームにより短期的に集中力向上効果を確認した。



実験では、効果確認のために、四則演算計算をするアプリを使用して、一定の問題数を解く計算時間を測定した。計算時間が短いほど、計算間違いが少ないほど集中力が高いと仮定する。

測定実験では、リズムゲームをしない場合、リズムゲームのジャンルを分類し、好き一好きでない、ゆったり一激しい、の 2×2 の 4 パターンの曲群に分類して、リズムゲームの前後に計算アプリでその瞬間の集中力を測定することにより、リズムゲームの効果を測定した。



とにより、リズムゲームの効果を測定した。



(a) 曲の分類によるリズムゲームの集中力向上効果
(b) 時間経過後の集中力向上の持続性

図 10 作業前のリズムゲームによる集中力向上効果

図 10(a)に、リズムゲーム無し、4 つのジャンルの曲群ごとのリズムゲームの直前(左側)と直後(右側)の計算時間をグラフにしている。全体として、リズムゲームを実施すると集中力が平均 11% 向上している。曲群の違いによる有意な差はみられなかったが、リズムゲームのプレイした場合は総じて集中力が向上した。

また図 10(b)にリズムゲーム実施の直前、直後、30 分後、60 分後、の集中力を計測した。直後から 30 分後にかけていったん集中力が向上した。直前から 1 時間後までの遷移に違いがみられたが有意な差は見られなかった。

5. まとめと今後の課題

本報告では、これまで実施してきた健康メディアデザイ

ンによる学生個人の健康改善の効果を確認できた。非常に興味深い結果が観測できた場合があるが、エビデンスと言える統計結果は得られていないため、今後の追実験が必要である。

最後に今回、個別にピックアップして報告した事例では、すべてのケースにおいて、実験で得られた成果・知見をアプリ化またはポスター化することにより、自分自身を取り組んだ健康テーマの改善の成果を活かし、本人のポートフォリオの充実に役立てており、より達成感の得られた課題設定であったとの感想を得た。

参考文献

- [1] 千種他,"健康意識の改善を促進するストレス度のチェックテストと可視化法: メディアと健康における可視化の取り組み",信学技報 115(479), p.81-85 (2016).
- [2] 千種, 山崎,"PDCA サイクルに基づく健康メディアに対するセルフケア法の提案", vol.116, no.EMM-501, pp.67-71 (2017).
- [3] 千種, 岩,"健康メディアにおける集中力をはじめとしたセルフケアの分析と改善アプリの開発",信学技報 IE2017-139, pp.123-127,(2018).
- [4] 佐藤, 千種, "環境音と音楽をポモドーロテクニックに適用した集中力向上法", 映像情報メディア学会, 映像情報メディア学会技術報告, ITE-42 (34), pp.13-17 (2018)
- [5] 千種, 井上,"健康メディアデザインに基づく座位姿勢の悪化要因分析と改善法に関する研究", FIT2019, vol.4 pp. O-036-037 (2019).
- [6] 小林, 千種,"健康メディアデザインによる睡眠力と集中力の改善",大学コンソーシアム八王子講演予稿集, pp.122-123 (2020).
- [7] 矢野, 千種,"健康メディアデザインに基づく集中力向上を目指した新しいポモドーロ法",情処大会第 84 回全国大会講演論文集,vol.1,pp.69-70 (2022).
- [8] 井上, 千種,"健康 PDCA サイクルにより運動不足解消と免疫力改善を両立する運動力向上支援アプリの開発", 情処大会第 84 回全国大会講演論文集,vol.1, pp.107-108 (2022).
- [9] 千種, 黄,"健康メディアデザインに基づく夜勤アルバイト学生のための日勤夜勤混合タイムスケジュール法の一提案",信学技報 vol.ET-212, ET2022-28, pp.39-42(ET)(2022).
- [10] 安達一寿,"ブレンディッドラーニングでの学習活動の類型化に関する分析",日本教育工学会論文誌,vol.31,no.1, p.29-40 (2007).
- [11] 三保 紀裕, 本田 周二, 森 朋子, 溝上 慎一,"反転授業における予習の仕方とアクティブラーニングの関連",日本教育工学会論文誌, vol.40, p.161-164 (2016).

地磁気センサと BLE ビーコンを用いた低成本屋内混雑度推定システム

友浦一樹^{†1}立命館大学 情報理工学研究科^{†1}吉田政望^{†2}立命館大学 情報理工学研究科^{†2}野口拓^{†3}立命館大学 情報理工学部^{†3}

1. はじめに

近年、コロナ禍の影響により、感染症対策の観点からユーザーと商業施設の両方から混雑度推定の需要が増している。混雑度推定とは、一定空間内の群衆密度を推定することである。例えば、ショッピングモールを任意のエリアに分割し、エリア別の混雑度推定ができると3密（密閉、密集、密接）の回避ができ、感染症対策に活用できる。

混雑度推定には様々な手法が存在する。特に GPS (Global Positioning System) で得られる位置情報を利用した混雑度推定は手軽でかつ正確に混雑度推定が可能であるが、屋内や地下では衛星からの電波が十分に届かず、機能しない。そのため、屋内ではカメラや Wi-Fi, Bluetooth 等を用いたものが主流となっている。しかし、カメラや Wi-Fi 等を利用するとコストが高くなる傾向がある。

そこで本稿では、安価な BLE (Bluetooth Low Energy) ビーコンとスマートフォンに標準搭載されている地磁気センサを用いて既存システムと比較して低成本な屋内混雑度推定システムを提案する。本提案システムは、屋内を任意の数にエリア分割し、エリアごとの人数を把握することで混雑度推定を行う。この手法では、エリアの境界をまたぐユーザの移動に着目し、BLE ビーコンの配置と地磁気の収集場所を限定することで、既存システムと比較して導入コストを削減している。

本提案システムの有効性を示すために、エリア推定精度について、実証実験を行い BLE のみを用いたシステムとの比較を行う。続いて、コスト削減効果を検証するため、実在する大規模商業施設におけるデータ収集コストの評価を行う。

2. 屋内端末位置推定技術

2.1. 地磁気測位

地磁気とは、地球の磁場のことであり、屋内では地球の磁場の影響だけでなく、建物の鉄筋などの金属が地磁気に影響を与える。そのため、位置によって磁気の特徴が異なる。地磁気測位は、前述のように位置によってわずかに違いのある地磁気を収集し、大きさと方向を位置ごとにデータベースに保存する。地磁気センサは、 x, y の2軸または、 x, y, z の3軸のタイプがあり、それぞれの方向の磁力の値を取得する。地磁気測位は、事前に収集されてデータベースに登録されている地磁気とスマートフォン等が搭載している地磁気センサが収集した地磁気の特徴を用いてマッチングを行い、その合致度合によって位置を推定する。マッチングは機械学習を用いて行う。スマートフォンを用いた地磁気測位の既存研究 [1] では、各場所の地磁気の特徴のみではなく、ユーザが通過した経路の一連の磁気特性を特徴として利用する。まず、ビル内を一定の速度で歩行し、データを収集する。この経路で得たデータと事前に収集したデータを機械学習を用いてマッチングを行い、ユーザの経路を推定する。この手法は実験場が細い廊下のみであり、広いエリアで測位した際に精度が低下すると考えられる。

2.2. Bluetooth を用いた位置推定

2.2.1. Bluetooth を用いた位置測位

Bluetooth を用いた位置測位は、固定された Bluetooth ビーコンとスマートフォンが通信を行うときに確認できる RSSI (Received Signal Strength Indicator) を用いてユーザの位置を推定する手法である。

RSSI とは、Bluetooth や Wi-Fi などの無線通信機器が受信する信号の強度を示す指標である。RSSI は数値が大きいほど信号強度が強く、無線通信機器同士の距離が近いほど大きくなる。この関係を式 1 に示す。 R は RSSI(受信信号強度)、 T は送信出力、 D は発信機と受信機の距離を表す。図 1 は式 1 をグラフにしたものである。図 1 では T を -4dBm に設定している。 R と T が分かれれば、式 1 を用いて距離 D が推定できる。

Low-cost Indoor Congestion Estimation System Using Geomagnetic Sensors and BLE Beacons

^{†1} KAZUKI TOMOURA, Ritsumeikan University Graduate School of Information Science and Engineering

^{†2} MASAMI YOSHIDA, Ritsumeikan University Graduate School of Information Science and Engineering

^{†3} TAKU NOGUCHI, Ritsumeikan University College of Information Science and Engineering

$$R = T - \log_{10} D \quad (1)$$

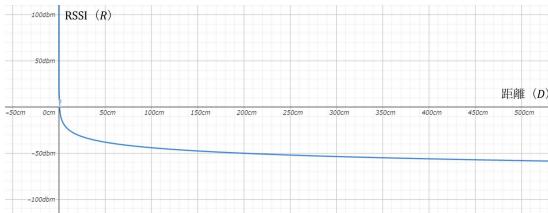


図 1 RSSI の距離特性

3つ以上のビーコンを利用し、それぞれのビーコンからユーザの距離が推定できること、それらの情報を元にユーザの位置を1つに推定できる。このような手法を3点測位と呼ぶ。自由空間においては、RSSIは距離と反比例して小さくなるため、距離の測定が簡単に行えるが、実際の空間では地面や壁などによる反射波の影響や障害物による電波の減衰など様々な要因によって、距離とRSSIが反比例しない場合がある。そのため、事前にRSSIを収集し、地磁気測位と同様に、各場所で収集されたRSSIをデータベースに記録する必要がある。また、この手法では、位置測位を行うすべての場所において複数のビーコンからの電波が届くようにビーコンを配置する必要がある。

2.2.2 BLE ビーコンを用いた混雑度推定システム

既存研究[2]では、複数のBLEビーコンを用いて大規模イベント会場での混雑度を推定している。BLEビーコンは安価かつ通信時の消費電力も少ない。この研究では、BLEビーコンを会場内に複数配置し、各ユーザのスマートフォンからビーコンの電波受信情報を収集することでビーコン周辺にいるユーザ数を把握している。しかし、既存のビーコンを用いた手法では、位置測位と同様に混雑度を推定したい範囲すべてをカバーできるようにビーコンを配置する必要がある。そのため、建物や施設が大きくなればなるほどコストが増加する。

3. 地磁気と受信信号強度を用いた低成本屋内混雑度推定システム

3.1. 概要

広い範囲をカバーできる屋内混雑度推定は、Wi-FiやBluetoothなど、システムの導入の際に、既存インフラや新たな機器の設置が必要である。しかし、広大な範囲を電波でカバーする必要があることから、必要な機器数も膨大となり導入コストが高くなる傾向がある。そのため、既存システムと比較して導入コストの低いシステムの需要が増している。本稿では、地磁気センサとBLEビーコンを用いた混雑度推定システムを提案する。本提案システムは、BLE

ビーコンの設置場所と地磁気を収集するエリアを限定することで、既存システムと比較してBLEビーコンの設置数を減らし、任意のエリアごとの混雑度を推定できる。

3.2. システム構成

本提案システムの構成図を図2に示す。まず、ユーザはスマートフォンに搭載されている地磁気センサとBLE機能を用いて地磁気データとRSSIを収集し、サーバへ送信する。サーバは、スマートフォンから受信したデータが事前にデータを収集し作成したデータセットとの類似度を機械学習を用いて算出し、比較することでユーザの位置を推定し、推定した位置情報を元に混雑度を推定する。



図 2 本提案システムの構成

3.3. 混雑度推定システム

3.3.1 推定手法の概要

本提案システムは、導入コストを削減するために屋内を任意の数のエリアに分割し、そのエリアの境界にのみBLEビーコンを設置し、境界内でのみ地磁気の事前収集を行う。BLEビーコンは1つで境界を1つ以上カバーできるように設置する。境界とエリアはそれぞれ番号ラベルを付与し、管理する。例を図3に示す。提案システムは、ユーザの所在エリアを推定し、エリア内のユーザ数を把握することでエリア別の混雑度推定を行う。ユーザの所在エリアは、ユーザのスマートフォンから収集された地磁気とRSSIのデータを用いて推定する。エリアを推定するためにユーザの境界通過を検知する。まず、BLEビーコンの電波範囲内で取得した地磁気とRSSIを用いてユーザの現在地がエリアの境界内か境界外なのかを推定する。ユーザの現在地が境界内であれば、機械学習を用いてユーザの所在エリアを推定する。また、ユーザの現在地が境界外であれば、推定の履歴から境界内で最後に推定されたエリアをユーザの所在エリアとする。

混雑度推定までの手順を以下に示す。

1. 事前に各境界に、境界全体をカバーできるようにBLEビーコンを設置する。地磁気センサを用いて境界内の地磁気を収集する。
2. 境界ごとに収集した地磁気データをデータセットにする。

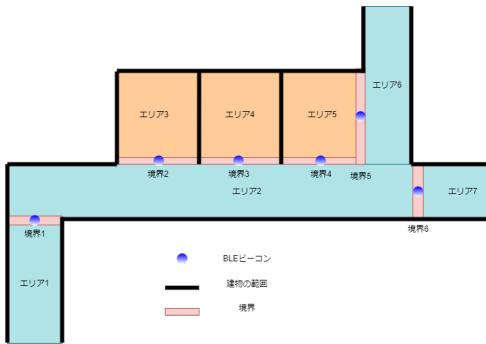


図 3 BLE ビーコンの配置例

3. ユーザはスマートフォンを用いて地磁気と RSSI を常に収集し、サーバに送信する。
4. サーバは、スマートフォンから受信したデータとデータセットを用いて機械学習を行い、ユーザの所在エリアを推定する。
5. エリア別のユーザ数を元にエリアごとの混雑度を算出する。

3.3.2 事前データ収集

本提案システムでは、事前に境界内の地磁気を収集し、データセットを作成する。データの収集はスマートフォンの地磁気センサを用いて行い、3 軸 (x, y, z) の地磁気を収集する。境界内での位置推定を行うために、境界をさらに分割しグリッドを作成する。地磁気収集はグリッド単位で行う。図 4 は境界内部の詳細である。グリッドのサイズは小さいほど高い密度でデータセット作成を行うことができるため、精度が高くなる。今回は、地磁気を用いた既存研究 [3] を参考に境界は 1 辺 1m のグリッドで構成する。また、それぞれのグリッドは所属エリアとグリッド番号が定められる。図 4 の例ではエリア 1 とエリア 2 に分類されるグリッドが存在する。スマートフォンを用いてグリッドごとに地磁気の事前データ収集を行う。収集は、それぞれのグリッドの中心で行う。また、地磁気の値はスマートフォンの姿勢に依存するため、1 グリッドにつき 24 方向のデータを収集する。スマートフォンの姿勢を 3 軸 (x, y, z) で示す場合、初期の姿勢を $(0, 0, 0)$ とする。この状態からそれぞれの軸を 90 度ずつ回転させ、スマートフォンの姿勢が重複しないように $(0, 0, 0), (0, 90, 0), (0, 180, 0), (0, 270, 0), (0, 0, 90), (0, 90, 90), (0, 180, 90), (0, 270, 90), (0, 0, 180), (0, 90, 180), (0, 180, 180), (0, 270, 180), (0, 0, 270), (0, 90, 270), (0, 180, 270), (0, 270, 270), (90, 0, 0), (90, 90, 0), (90, 180, 0), (90, 270, 0), (180, 0, 0), (180, 90, 0), (180, 180, 0), (180, 270, 0)$ の 24 方向とする。データ収集は 0.5 秒間隔で、1 方向につき 5 秒間行う。

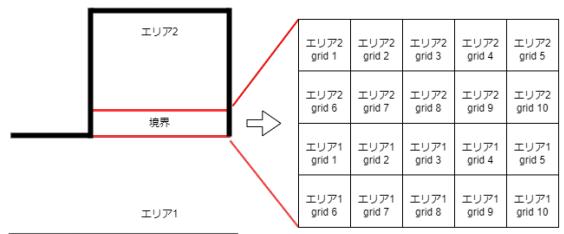


図 4 境界の詳細

3.3.3 データセット作成

収集した地磁気データからエリア推定に利用する機械学習用のデータセットを作成する。データセット作成の手順を以下に示す。

1. 24 方向の地磁気データをそれぞれ x, y, z 軸ごとに平均化し、特徴量とする。
2. 平均化した 3 軸の地磁気を利用して地磁気のノルムを算出し、新たな特徴量として追加する。これは式 2 で算出される。 N はノルム、 x_a, y_a, z_a はそれぞれ x, y, z の地磁気の平均を示す。

$$N = \sqrt{x_a^2 + y_a^2 + z_a^2} \quad (2)$$

3. 平均化した地磁気データにラベル付けを行う。ラベルは、データを収集したグリッドの番号、所属するエリア、および境界番号で構成される。
4. 1, 2, 3 と同様の手順で、1 グリッドにつき 24 方向分のデータセットが作成できる。

3.3.4 ユーザの動作と推定例

図 5 にユーザのエリア移動例を示す。ユーザは矢印に沿って点 1, 2, 3 を順番に通過すると仮定する。ユーザは初期状態ではエリア推定履歴を持っていない。本提案システムは、ユーザが境界を通過した場合には、ユーザの所在グリッドを推定し、そのグリッドに対応したエリアをユーザの所在エリアとする。ユーザは、入口の境界の右半分を通過した際にエリア 1 と推定される。その後、点 1 に到達するまでの境界外では直前の所在エリア履歴からエリア 1 と推定される。さらに点 1 の境界の下半分にユーザが存在する場合はエリア 1、上半分に存在する場合はエリア 2 と推定される。次に点 1 から点 2 までは境界外のためユーザの所在エリア履歴の中で最新の所在エリアであるエリア 2 と推定される。その後、点 2 で境界内の下半分を通過したとき、エリア 1 に更新される。その後、点 2 から点 3 までは境界外のため履歴からエリア 1 と推定される。

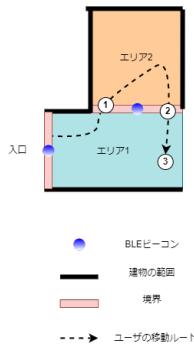


図 5 ユーザの移動例

3.3.5 最近隣境界の判定

本提案システムでは、境界の通過を検知してエリアを推定する。そこで、ユーザがどの境界を通過しようとしているのか判定する必要がある。これは、3.3.1 項でラベル付けした境界の番号ラベルを用いてデータセットの中から最も近い境界の番号ラベルが付与されているデータセットを選択するために行う。ユーザに最も近い境界は BLE ビーコンを用いて判定する。ユーザが BLE ビーコンからの電波が届かない場所にいる場合は、3.3.4 項で述べた境界外での処理と同様に直前の所在エリア履歴によるエリア推定が行われる。BLE ビーコンと境界は紐づけられており、ユーザの現在地に最も近い BLE ビーコンが判明すると最も近い境界も判明する。このユーザの現在地に最も近い BLE ビーコンは BLE ビーコンから発信されている電波の RSSI を用いて判定する。2.2 節で述べたように RSSI は発信元と距離が近いほど大きくなる特性がある。これを利用し、最も RSSI が大きい BLE ビーコンに紐づけされている境界が最も近い境界と判定される。

図 6 はユーザに最も近い境界を判定する例である。ユーザは現在地で境界 1 と境界 2 の BLE ビーコンの電波を受信している。境界 1 の BLE ビーコンによる電波の RSSI は -50.3dBm であり、境界 2 の BLE ビーコンによる電波の RSSI は -69.1dBm である。RSSI は前述のように距離が近いほど強く、大きな値になるため RSSI が -50.3dBm である境界 1 が最も近い境界だと判定できる。なお、RSSI が同じ値になった場合は境界に割り当てられた番号の小さい境界と判定される。

3.3.6 境界の内外判定

サーバはユーザが収集したデータと事前に収集したデータセットを利用してエリア推定を行う。その際にユーザが境界内に居るか境界外に居るのかを判定する必要がある。この境界の内外判定を、3 軸 (x, y, z) の地磁気およびノルムで構成される 4 次元ユークリッド空間上のユークリッド距離を利用して行う。まず、ユーザの収集したデータと

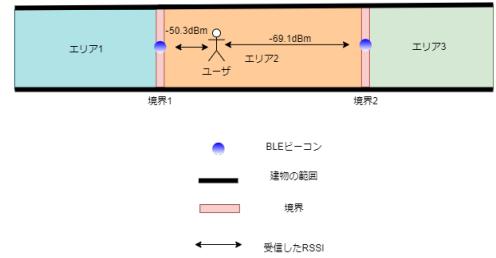


図 6 最も近い境界の例

3.3.5 項で判定したユーザに最も近い境界の番号ラベルが付与されているデータセットのユークリッド距離を算出する。ユークリッド距離の閾値を設定し、算出したユークリッド距離が閾値未満の場合は境界内、閾値以上の場合は境界外と判定する。ユークリッド距離は式 3 で算出される。 E は算出されるユークリッド距離を示す。 $x_m, x_n, y_m, y_n, z_m, z_n$ は地磁気データ、 N_m, N_n はノルムを示す。 m は事前に収集したデータ、 n はユーザが収集したデータを示す添字である。

$$E = \sqrt{(x_m - x_n)^2 + (y_m - y_n)^2 + (z_m - z_n)^2 + (N_m - N_n)^2} \quad (3)$$

3.3.7 エリアの推定

3.3.6 項でユーザが境界内であると判定された場合、エリア推定を行う。エリア推定には機械学習の k 近傍法 (k-nearest neighbor algorithm, k-NN) を用いる。このアルゴリズムは、データを分類する際に予測データに近いデータ k 個の多数決によってクラスを推測するものである。位置ごとの特徴を収集し、類似度によって位置推定を行う、フィンガープリント法の既存研究 [4][5] では、他のアルゴリズムと比較して k 近傍法が最も精度が良いとされている。

3.3.5–3.3.7 項までのエリア推定の手順を以下に示す。

1. RSSI を用いてユーザに最も近い境界を判定する。
2. 最も近い境界のデータセットを用いて境界の内外判定を行う。
3. 境界外と判定された場合、履歴から前回のエリアと推定される。境界内と判定された場合、k 近傍法を用いて境界内のデータセットの内、最も収集データに近いデータセットを算出する。
4. 算出された最も収集データに近いデータセットのエリアラベルを参照し、ユーザの所在エリアを推定する。

3.3.8 閾値の決定法

3.3.6 項で述べたように、境界の内外判定を行う際に、ユークリッド距離の閾値が必要となる。この閾値の決定は 2 ステップで行う。はじめに、スマートフォンの姿勢を固定した実験を行う。実験で取得した地磁気データは、実験シナリオと整合する必要があることを利用して、閾値の適切な

範囲を決定する。続いて、スマートフォンの姿勢を固定しない実験を行い、得られた閾値の範囲内でエリア推定精度が最も高くなる閾値を決定する。

スマートフォンの姿勢を固定した実験によるデータ収集

適切なユークリッド距離の閾値範囲を決定するため、はじめにスマートフォンの姿勢を固定した実験を行う。図7に実験シナリオを示す。この実験では、ユーザの移動経路とスマートフォンの姿勢を固定し、実験の誤差要因を可能な限り取り除く。ユーザは0.7–1.0m/sで歩行する。

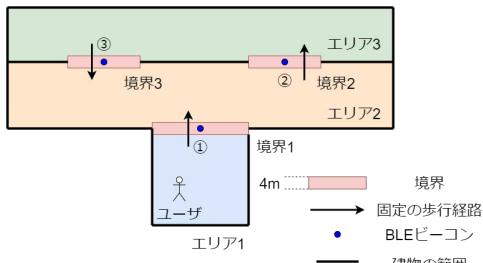


図7 スマートフォンの姿勢を固定した実験シナリオ

この実験の手順を以下に示す。

- スマートフォンを用いて境界内の地磁気を収集する。地磁気はスマートフォンの姿勢を固定した1方向(x , y , z)を収集する。
- 境界ごとに収集した地磁気データをデータセットにする。
- スマートフォンを用いて図7の3つの境界を矢印の番号と方向に従い通過しつつ地磁気データとRSSIを収集する。まず、エリア1から開始し、境界1を通過しエリア2に入る。次にエリア2から境界2を通過し、エリア3に入る。最後にエリア3から境界3を通過し、エリア2に戻る。境界の通過は任意のタイミングで行い、すべての境界の通過時間が合計1分間になるようを行う。また、スマートフォンの姿勢は一定姿勢で固定して実験を行う。データの収集間隔は3.3.2項と同様に0.5秒である。
- 収集した合計120個のデータをサーバに送信する。
- 1–4を1回として合計3回行う。

閾値の範囲決定方法

収集した1回120個のデータの内、閾値内に収まるべきデータ数を求める。閾値内に収まるべきデータ数は、実験においてユーザが境界内にいたときに収集されたデータ数である。これをユーザが通過した境界内の距離とデータ収集間隔、ユーザの歩行速度を用いて式4で求めることができる。

x は閾値内に収まるべきデータ数、 d は通過する境界内の距離、 v は歩行速度、 t はデータ収集間隔を表す。

$$x = \frac{d}{v \times t} \quad (4)$$

この実験では歩行速度は0.7–1.0m/sである。また、通過する境界内の距離は合計3つの境界を通過するため図7より12mとなる。データ収集間隔は0.5秒である。式4に上記の数値を代入した結果、この実験1回での理想的な閾値内のデータ数は24–34個となる。

図8はこの実験3回の平均データ120個の閾値ごとのデータ分布である。また、閾値は0.1刻みで変化させる。この結果より、実験シナリオと同じエリア推定結果となっている閾値は3.8–4.2である。これにより、境界の内外判定を行うために利用するべきユークリッド距離の閾値の範囲は3.8–4.2であると判明した。

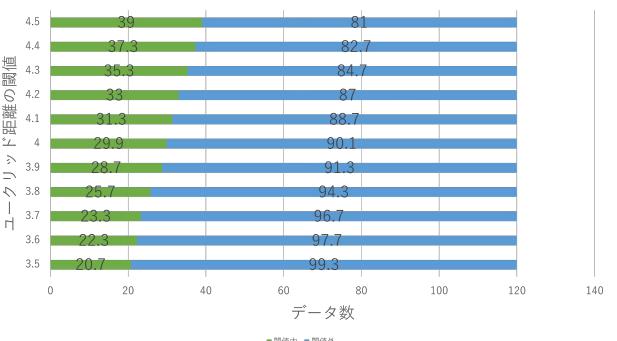


図8 実際のデータ分布

閾値の決定

図9は閾値の範囲3.8–4.2でのスマートフォンの姿勢を固定した実験におけるエリア推定の精度である。精度は閾値ごとに式5で算出される。この範囲内では閾値ごとの差は小さく、最大の差は閾値4.2と閾値4, 4.1のときで2%である。範囲内で最も推定精度が高いのは閾値4.2の85.3%である。

$$\text{エリア推定精度} = \frac{\text{エリア推定の正解データ数}}{\text{収集データ数}} \quad (5)$$

閾値3.8–4.2の範囲内で本提案システムのエリア推定の精度が最大になるようにユークリッド距離の閾値を1つに決定する必要がある。そのため、スマートフォンの姿勢を固定しないエリア推定実験を行い、その結果を利用する。図10はこの実験の様子である。実験の手順を以下に示す。データ収集やデータセットは、3.3.2項、3.3.3項の手順に従って作成する。

- 境界1, 2, 3で事前にデータ収集を行う。

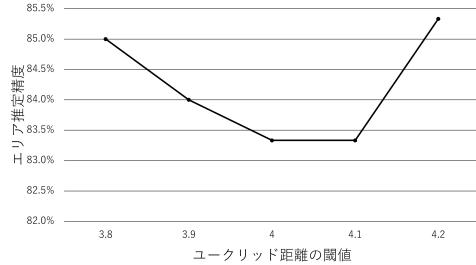


図 9 閾値ごとのエリア推定精度

2. 収集した地磁気データを用いてデータセットを作成する。
3. ユーザはエリア内を 1 分間スマートフォンの姿勢、移動経路、歩行速度を自由に歩行しつつ、スマートフォンでデータ収集を行う。
4. データセットと 1 分間歩行したデータを利用し、エリア推定を行う。
5. 実際にユーザが居たエリアと比較し、精度を算出する。
6. 1-5 を 1 回として合計 6 回行う。

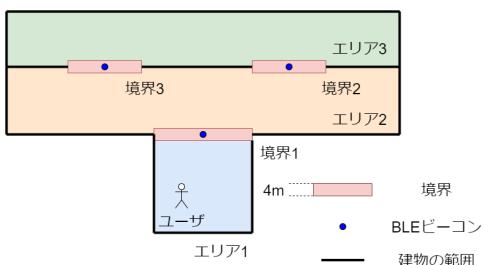


図 10 エリア推定の実験

図 11 はこの実験のエリア推定の精度である。data1-6 は、合計 6 回行った実験それぞれの精度である。また、図 12 は図 11 の data1-6 の平均精度である。精度は、方向を固定した実験と同様に式 5 で算出される。平均の精度は、閾値 3.8 が 72.1% と最も低くなっている。一方、閾値 4.2 は 78% と最も精度が高い。この結果からユーフリッド距離の閾値は、端末の姿勢を固定した実験結果から得られた範囲である 3.8-4.2 の中から、端末の姿勢を固定しない実験結果で最もエリア推定精度の良い閾値 4.2 を選択すべきであることが示された。

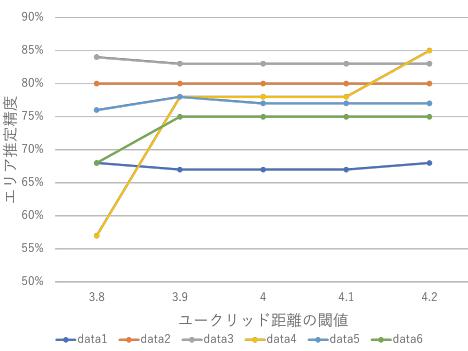


図 11 エリア推定精度

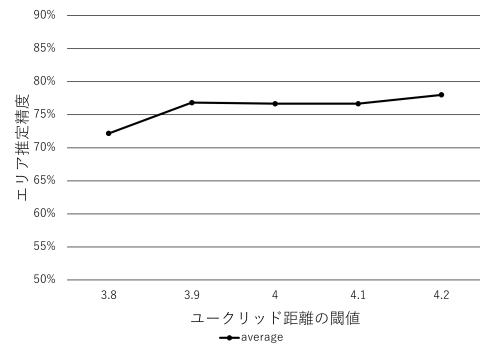


図 12 エリア推定精度の平均

4. 性能評価

4.1. エリア推定精度評価

本稿では、大学内の建物を利用した評価実験によるエリア推定の精度と実際の利用を想定した導入コストの評価を行った。また、本提案システムを BLE ビーコンのみを用いた手法と比較し、精度と導入コストの差を評価した。精度の評価は、評価実験を大学内の建物であらかじめ決めたエリア内でエリア推定を行い、結果を評価する。評価実験での BLE ビーコンはサンワサプライ社の MM-BLEBC5 を利用する。また、スマートフォンは Google 社の Android スマートフォンである、Nexus5 を利用する。コストの評価は、実在する施設を参考に BLE ビーコンの価格と事前データ収集における導入コストを算出することで行う。本提案システムの評価実験は、3.3.8 項の閾値の決定で述べた実験と同様に合計 6 回行う。エリア推定精度は、3.3.8 項の実験と同様に式 5 で算出する。

4.1.1 比較システムの概要

本提案システムに対する比較システムとして BLE ビーコンのみを用いた混雑度推定システム [2] を利用する。このシステムは、BLE ビーコンからの電波をユーザのスマ

トフォンで受信した際の RSSI のみを利用する。評価実験を行う各エリアの中心に 1 つずつ BLE ビーコンを設置する。ユーザのエリアは、データ収集の際に最も RSSI の強い BLE ビーコンの配置されているエリアと推定される。この比較システムで評価実験と同様の場所でエリア推定を行い、精度を比較する。

4.1.2 比較システムの実験内容

本提案システムと同様の場所でこの比較システムの評価実験を行う。この実験は精度に焦点を当てたものである。ビーコンの個数や境界の距離については、4.2 項で述べる。図 13 は実験の様子である。手順を以下に示す。

1. エリアの中心に BLE ビーコンを設置する。
2. エリア内をスマートフォンでデータ収集しつつ、1 分間スマートフォンの姿勢、移動経路、歩行速度を自由に歩行する。
3. 収集したデータと実際にユーザが居たエリアを比較し、精度を算出する。
4. 1-3 を 1 回として合計 6 回行う。

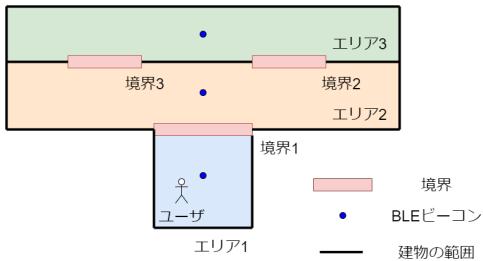


図 13 BLE ビーコンのみを用いた比較システムの実験

4.1.3 実験結果

表 1 にエリア推定精度の実験結果を示す。実験結果としては、本提案システムのエリア推定精度の最低値は 68%，最高値は 85% となり、平均値は 78.1% となった。比較システムの実験結果としては、最低値は 72%，最高値は 85% となり、平均値は 79.3% となった。

表 1 提案システムと比較システムの精度

システム	最低値	最高値	平均値
提案システム	68%	85%	78.1%
比較システム	72%	85%	79.3%

4.2. コスト評価

実際の施設を想定して BLE ビーコンと地磁気データの収集コストを算出する。BLE ビーコンの価格は今回使用したもの参考に、1 個 5000 円と仮定する。また、BLE ビーコンの通信半径は 15m とする。事前データ収集は 3.3.2 項と

同様にデータ収集を行うと仮定すると、1 グリッドにつき 2 分、境界の長さ 1m (4 グリッド) につき 8 分の時間がかかる。さらに、データ収集のための移動時間も考慮し、今回は 1m につき 9 分の収集時間が必要と仮定する。これをコストに換算するために日本の平均時給 [6] を参考にする。日本の平均時給は 961 円である。計算の簡略化のために時給 1000 円と仮定すると、データ収集 1m につき、150 円となる。これらの前提を元に本提案システムと BLE ビーコンのみを用いた比較システムにおける導入コストを算出する。

評価の場所は滋賀県草津市のイオンモール草津と滋賀県大津市のフォレオ大津一里山を想定する。イオンモール草津は店舗面積が約 78,000m² の大型ショッピングモールである。フォレオ大津一里山は店舗面積が約 41,000m² のショッピングモールである。これらを店舗ごとのエリア分割とジャンルごとのエリア分割の 2 つの手法で一定空間ごとにエリア分割する。ショッピングモールでは、店舗のジャンルによって店舗の配置が決められている。図 14 にジャンルによってエリア分割したイオンモール草津 1 階のフロアマップ [7] を示す。黄色が飲食、緑色がファッション、青色が雑貨、紫色がサービス、桃色がイオスタイル (スーパー・マーケット) となっている。店舗ごとのエリア分割では、通路をファッションエリア通路と飲食エリア通路の 2 種類に分割した。ジャンルごとのエリア分割手法では粒度が大きく、店舗ごとのエリア分割手法では粒度が小さくなる。この分割手法を各フロアに適用し、イオンモール草津ではジャンルごとのエリア数は 15、店舗ごとのエリア数は 178 となった。フォレオ大津一里山も同様の手法でフロアマップ [8] に基づいてエリア分割を行い、エリア数はジャンルごとの分割が 9、店舗ごとの分割が 98 となった。エリア数と境界数は比例しており、境界数が少ない程、地磁気収集コストと BLE ビーコン数が減少し、コストが低下する。

表 2 はイオンモール草津をエリア分割した場合のコストである。この場合、本提案システムは、BLE ビーコンのみのシステムと比較して約 13.1% のコスト削減ができる。また、表 3 はフォレオ大津一里山を店舗ごとにエリア分割した場合のコストである。この場合、本提案システムは、BLE ビーコンのみのシステムと比較して約 19.4% のコスト削減ができる。表 4 はこの方法でイオンモール草津をジャンルごとエリア分割した場合のコストである。この場合、本提案システムは、約 56.4% のコスト削減ができる。また、表 5 はフォレオ大津一里山をジャンルごとにエリア分割した場合のコストである。この場合、本提案システムは、BLE ビーコンのみのシステムと比較して約 53.4% のコスト削減ができる。つまり、本提案システムは、BLE ビーコンのみのシステムと比較してエリア分割の粒度が粗いほどコスト削減できることが示された。



図 14 イオンモール草津 1 階 [7]

表2 イオンモール草津における店舗ごとのコスト表

Digitized by srujanika@gmail.com

表3 フォレオ大津一里山における店舗ごとのコスト表

For more information about the study, please contact Dr. [REDACTED] at [REDACTED].

表4 イオンモール草津におけるジャンルごとのコスト表

For more information about the study, please contact Dr. [REDACTED] at [REDACTED].

5. おわりに

本稿では、安価な BLE ビーコンとスマートフォンに標準搭載されている地磁気センサを用いて既存システムと比較して低コストかつ同等の精度を持つ屋内混雑度推定システムを提案した。評価の結果、本提案システムは、BLE ビーコンのみのシステムと比較して同程度のエリア推定精度を保ちながらコストを削減できることを示した。また、エリア分割の粒度が粗いほど、コストを削減することができた。

今後の展望としては、実際に大規模な商業施設で実験を行い、周囲の環境による影響も考慮する。また、今回は考慮していなかった閾値の決定コストや効率的な閾値の決定手法についても考慮する。

表5 フォレオ大津一里山におけるジャンルごとのコスト表

© 2009 by the author

参考文献

- [1] Shervin, S. and Shahrokh, V.: GIPSY: Geomagnetic Indoor Positioning System for Smartphones, *IPIN* (2015).
 - [2] Sato, D., Mihara, Y., Sato, Y., Tanaka, Y., Miyamoto, M. and Sakuma, S.: BLE ビーコンを利用した混雑度可視化サービス, 情報処理学会論文誌コンシューマ・デバイス&システム, Vol. 8, No. 1 (2018).
 - [3] Azuma, K. and Arai, I.: 相互補完型 Wi-Fi, 地磁気 フィンガープリンティング手法の評価, 情報処理学会論文誌, Vol. 58, No. 2 (2017).
 - [4] Ranimol, J. and Swapna, B, S.: Indoor Positioning Using WiFi Fingerprint, *ICCSDET* (2018).
 - [5] Devin, Ryan, R., Marcel, Bonar, K. and Maria, Irmina, P.: Multi-Floor Indoor Positioning Mobile Application using Earth's Magnetic Field, *IEEE* (2017).
 - [6] ”日本労働組合連合会 | 労働・賃金・雇用 最低賃金労働・賃金・雇用”, <https://www.jtuc-rengo.or.jp/activity/roudou/chingin/>, (参照 2022).
 - [7] ”イオンモール草津 | フロアガイド”, <https://kusatsu.aeonmall.com/floorguide/>, (参照 2022-12-01).
 - [8] ”ショップガイド (フロア 1F) — フォレオ大津一里山公式ホームページ”, <https://otsu.foleo.jp/shop/?search=floor-1>, (参照 2022-12-01).

混雑緩和やユーザ体験価値向上に向けた 行動変容施策の設計および社会実装について

勝間田優樹^{†1}

株式会社 NTT ドコモ

吉川裕木子^{†2}

株式会社 NTT ドコモ

鈴木喬^{†3}

株式会社 NTT ドコモ

山田暁^{†4}

株式会社 NTT ドコモ

1. はじめに

「自分が好きな店に立ち寄っていたら自然と混雑を回避していた。」このような交通最適化された世界の実現を目指し様々な検討が進められている。交通最適化された世界では、人は無理なく混雑を回避することができ、その効果が他者にも波及することでメリットを社会全体で享受することができる。渋滞を回避する手段として、例に挙げたような立ち寄りに限らず、ピークシフトや交通手段の変更など人によって妥当性が異なる様々な手段が採用される。

交通最適化された世界の実現に向けては、本来の取るはずだった行動を、交通手段の変更や立ち寄りなどの行動へと変容させることが必要であり、そのためには行動変容をするための十分な動機付けが必要である。多くの人にとって、混雑を認識してもそれを回避することは容易ではない。例えば、多くの高速道路サービスエリアではリアルタイムの混雑情報が提示されており、それを見たドライバーによる緩和効果には一定の成果はあるものの、それだけで混雑を緩和することは難しい。

動機付けのアプローチの一つとして行動経済学に基づくナッジ^[1, 2]が着目されている。ナッジとは、人間の価値観や心理的な傾向を利用し、対象者にとってより望ましい選択をするように促す行動変容に関する一手法である。ナッジを活用した研究はこれまで数多く実施されており、例えば、Push 通知によって配信されるメッセージにナッジを適用することで配信されたメッセージの開封率を向上させる検討などが行われている^[3, 4]。

筆者らは、の中でもナッジのパーソナライズ化について検討している。ナッジのパーソナライズ化は、個人がそれぞれ持つ価値観・心理傾向の違いを考慮し、それに応じてナッジを含むメッセージ文言（以下、本稿ではナッジ文言と呼ぶ）の出し分けを行うことを意味する。例えば、損失回避

的な心理傾向を持つ人には「このサービスを利用しないと損してしまいます」と伝え、同調的な心理傾向を持つ人には「多くの人がこのサービスを利用しています」と伝える。対象者をナッジ文言を配信する群としない群とに分け、ナッジ文言を配信する群に対して一律に同じメッセージを配信するものとは異なり、個別にナッジ文言の出し分けを行うことで、行動変容施策や検討の効果を最大化させることを目指す。

本稿では、ナッジのパーソナライズ化の検討の中で実施した多数のフィールド実証で得られた知見や課題を述べる。

2. ナッジのパーソナライズ化

ナッジのパーソナライズ化を実現するためには様々なデータを収集する必要がある。扱うデータは想定されるサービスによって異なり、例えばユーザの性別・年代のようなデモグラフィックデータの他にも、アプリを使う目的や好きな食べ物のジャンルといった情報もアンケートなどを活用して取得する場合がある。また、ユーザのコンテキストの推定のためにユーザの位置情報などのリアルタイム情報を取得する場合もある。データはユーザに限らず、交通手段の状況や気象状況といった環境情報や立ち寄りを促す場合には立ち寄り先の店舗などの情報を扱う場合もある。

収集されたデータの一部からユーザの特徴量を抽出し、そのユーザに対してより行動変容効果の高いナッジ文言の推定を行う。推定する手法として、機械学習モデルの分類器を用いる。フィールドによっては収集できるデータが大きなサンプルサイズとならないことも考えられるため、十分に学習を進められるよう、フィールド実証開始前に広くアンケート調査を行い別途正解データを収集する場合もある。出力されるナッジ文言としては、例えば

- 「周辺の魅力的なお店にお出かけしませんか？せっかくのお得なクーポンを使わないのはもったいないですよ！」
- 「カフェで一休みはいかがですか？多くのお客様がお得なクーポンを利用しています！」

などがある。

Design and Implementation of Behavior Transformation Measures for Congestion Reduction and User Experience Improvement

^{†1} Yuki KATSUMATA, NTT DOCOMO, INC.

^{†2} Yukiko YOSHIKAWA, NTT DOCOMO, INC.

^{†3} Takashi SUZUKI, NTT DOCOMO, INC.

^{†4} Akira YAMADA, NTT DOCOMO, INC.

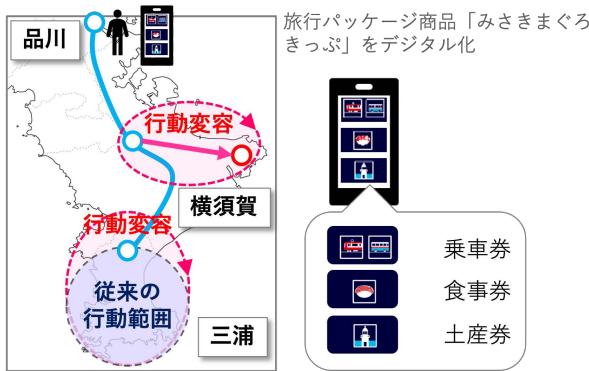


図 1 みうらよこすか MaaS の概要

Figure 1 Overview of Miura Yokosuka MaaS

ナッジ文言による介入手段にも様々なものがあり、その一つに端末アプリへの Push 通知がある。ナッジ文言による介入に対するユーザの行動有無から、効果の有無を判定し再度学習に用い、推定精度を向上させる。目的に応じてユーザの行動有無の判定の定義が異なり、例えば開封率の向上が目的であれば端末アプリ内のメッセージ開封有無のログを収集し、訪問率の向上が目的であれば店舗への訪問有無のログを収集し、それらをそれぞれ正解データとして学習に用いる。

3. 実証実験および社会実装

この章では、これまで取り組んできた社会実装に向けたフィールド実証の概要を述べる。検討してきたナッジのパーソナライズ化技術の適用先は大きく 2 種類ある。1 つは、混雑や渋滞の緩和を目的とした行動変容施策である。目的は、発生しているもしくは発生するであろう渋滞という損失を減少させることにある。もう 1 つは、街全体の周遊性向上を目的とした行動変容施策である。目的は新たな体験価値を提供し便宜を増加させることにある。

3.1. みうらよこすか MaaS [5]

本フィールド実証の概要を図 1 に示す。みさきまぐろきっぷ [6] という電車・バスの周遊券を含む旅行パッケージ商品がある。みさきまぐろきっぷを電子化し端末アプリで扱い、その端末アプリの中でナッジ文言を Push 通知することによる行動変容施策を実施した。観光動線が概ね固定化されているという課題に対し、自治体が持つ多様な観光資源への立ち寄りを促し周遊性を向上させることを目的としている。多様な観光資源の情報およびそこへ至る交通手段の情報をナッジ文言と併せて提示するという手段により、地域の様々な場所に周遊することで新たな体験価値を提供することを目指した。

行動変容施策として、いくつかのパターンで介入を行った。具体的には、これから旅先へ向かおうとしているユーザに対する旅前の介入、周遊中に食事をしたユーザに対する旅中の介入、旅行を終え帰路につこうとしているユーザに対する旅後の介入に分類される 3 パターンの介入を行った。本フィールド実証では、観光という非日常の中でのナッジ文言の効果検証を目的としており、ナッジ文言の出し分けを行うためのデータや知見がない状態での施策実施を行ったため、ユーザに対してランダムにナッジ文言を Push 通知し、ナッジ文言自体の行動変容効果有無の検証および出し分けの実現可能性についての事後検証を行った。

通知対象ユーザのうち、11 % が旅後の介入で立ち寄りを行うなどナッジ文言自体の行動変容効果が見られることができた他、メッセージ文言の種類によっては最大で約 25 % の行動変容率^{*1}を達成することができ、ナッジのパーソナライズ化の有効性について示唆が得られた。

3.2. バス快適乗車案内 [7]

本フィールド実証の概要を図 2 に示す。通勤時間など混雑時間帯のバスでは、バス停での混雑および乗り降りにかかる時間によるダイヤ乱れが発生し、それによりバス車内の混雑度の偏りが加速する場合があり、オペレーション煩雑化や利用者満足度の低下が生じる。加えて、コロナ禍における密回避の必要性も顕在化しており、バス毎の混雑度の平準化が喫緊の課題となっている。本フィールド実証では、混雑時間帯のバス車内の混雑度の平準化を目指し、混雑度が低い後続のバスが数分で到着すると見込まれる場合に、そのバスの混雑情報およびナッジ文言を Push 通知し混雑バスの見送りを促進するとともに、見送り実施したユーザに対してポイントによるインセンティブを付与することで動機付けを行った。

本フィールド実証では、日常の中でのナッジ文言の効果検証を目的としており、主な対象者が日常的に通勤や通学などでバスを使うユーザである。そのため、毎日同じような内容の Push 通知がなされることから、行動変容施策の効果が摩耗することが懸念される。ナッジ文言の有無に限らず、ユーザ端末へのメッセージ配信頻度による効果の摩耗 [8] や Web マーケティングにおいても同じ広告を配信することによる効果の摩耗 [9] が存在することが報告されており、同様の事象がナッジ文言の Push 通知にどの程度起こるのかを検証した。特定のナッジ文言で行動変容を行ったユーザを事後分析すると、特定の心理傾向が強い傾向であることが分かるなど、本フィールドのようなユースケースにおいても

^{*1} 本フィールド実装では、端末アプリの中で事前に設定できる旅程に対して、介入をすることで外れた行動をした場合に行動変容したと定義した。



図 2 バス快適乗車案内の概要

Figure 2 Overview of Comfortable Bus Ride

ナッジのパーソナライズ化が有効であることが示唆された。

3.3. ゆるり GO!TEMBA [10]

自家用車で行楽地に行った場合に、帰りの高速道路の渋滞に巻き込まれるという事象が発生する。大型連休などにおける帰省に際して発生する渋滞は U ターンラッシュとも呼ばれ、幅広く認知されている課題である。本フィールド実証では、行楽地からの帰路の高速道路で発生する渋滞を緩和およびそれによる CO₂ 排出を削減する目的で、これから帰ろうとしているユーザに対する行動変容施策を実施した。

具体的には、御殿場プレミアム・アウトレットに東京方面から自家用車で訪れたユーザに対して、東名高速道路の上り方向の混雑予測情報およびナッジ文言の Push 配信を行い、滞在時間延長または周辺施設への立ち寄りを促した。本フィールド実証では、観光という非日常の中での混雑緩和を目的とした検証を行ったため、過去の実証で得られたデータの活用による行動変容効果向上が必ずしも見込まれなかった。そのため、事前に広くアンケート調査を行いそのアンケート結果を正解データとする機械学習モデルを構築し、ナッジ文言の出し分けによる行動変容効果の検証を行った。検証を通して、高速道路の渋滞緩和にナッジのパーソナライズ化が一定の効果をもたらすことが示唆された。

3.4. 大学キャンパスにおける店舗内の密回避 [11]

本フィールド実証の概要を図 3 に示す。大学キャンパスのような閉空間においては、ランチタイム時に食堂などの飲食店に利用者が集中し局地的な混雑が発生する。とりわけコロナ禍においては密回避を行うことが重要であり、このような日常的に発生する混雑を緩和することが求められる。本フィールド実証では、局地的な混雑が日常的に起こるようなケースに対してピークシフトを実現することを目的として、混雑時間帯を避けて来店を促すタイムシフトおよび持ち帰りを促すテイクアウトの 2 種類のキャンペーン

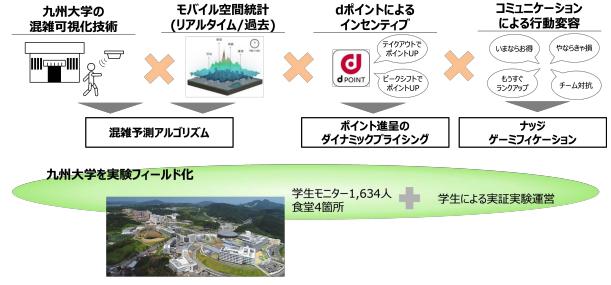


図 3 九州大学におけるフィールド実証の概要

Figure 3 Overview of Field Experiment at Kyushu Univ.



図 4 通勤電車の混雑緩和施策の概要

Figure 4 Overview of Behavior Transformation Measures for Easing Commuter Rush

施策を行った。

具体的には、上記キャンペーンの中で店舗の情報配信とポイントによるインセンティブ付与を行うとともに、行動変容施策として、ナッジ文言の Push 配信およびゲーミフィケーションによる行動変容効果の向上が見込まれるかを検証した [12, 13]。結果として、ナッジ文言による行動変容受容率 *2 は 20 % から 33 % 向上、ゲーミフィケーションによる行動変容受容率は 30 % から 41 % 向上させることが確認され、これらの行動変容施策が混雑緩和に対して一定の効果をもたらすことが示唆された。

3.5. 通勤電車の混雑緩和施策 [14]

本フィールド実証の概要を図 4 に示す。通勤時間帯の鉄道は慢性的に混雑が発生しており、3.2 と同様にダイヤ乱れによるオペレーション煩雑化や利用者の満足度低下が生じるだけでなく、コロナ禍においては密回避の必要性の高まりもあり、混雑の緩和が課題である。本フィールド実証では、通勤時間帯の鉄道利用者であるモニタに対してオフピークを促すナッジ文言の Push 通知を行った。具体的には、通勤時間帯を後ろ倒しにするタイムシフトもしくは別の交通手段を用いるモーダルシフトを促すメッセージの Push 通知を行った。

*2 本フィールド実装では、Push 通知を受けた全ユーザに対する、Push 通知を受けた後にタイムシフトもしくはテイクアウトを行ったユーザの割合を行動変容受容率として定義した。



図 5 まくはり MaaS の概要

Figure 5 Overview of Makuhari MaaS

具体的には、モニタに限定せずに収集したアンケートの結果を正解データとしユーザの心理傾向を推定する機械学習モデルを生成し、その機械学習モデルをモニタに適用することでモニタの心理傾向を推定した。モニタを2群に分離し、1つの群には推定した心理傾向に応じてナッジ文言の出し分けを行い、ランダムにナッジ文言を出し分けた群との比較を行った。結果として、推定した心理傾向に応じてナッジ文言の出し分けを行った群は有意にオフピークを達成することが分かり、行動変容効果の高いナッジ文言をユーザの特徴量から直接推定するナッジのパーソナライズ化手法だけでなく、推定したユーザの心理傾向を用いるナッジのパーソナライズ化手法にも一定の行動変容効果が見込まれることが示唆された。

3.6. まくはり MaaS [15]

本フィールド実証の概要を図5に示す。幕張新都心エリアにおいて、回遊性向上や賑わいの創出が課題となっている。次世代モビリティの提供とともにユーザ個人に寄り添ったレコメンドを提供することにより、回遊性向上による地域経済の活性化や賑わいの創出をするとともに、様々な施設に分散誘導することによるコロナ禍における安心・安全を意識した施策実施を行った。

本フィールド実証では、日常の中での周遊性促進を目的としたナッジ文言の効果検証を目的としている。具体的には、住民向けの端末アプリに対し、おすすめの店舗情報およびクーポン情報とともにナッジ文言のPush通知を行った。本実証においては、回遊性向上に向けたマイルストーンとして、Push通知の開封による認知率向上を目的としたナッジ文言の出し分けを行いその効果を検証した[16]。結果として、ナッジを含まないメッセージをPush通知した群のメッセージ開封率に対する、ナッジ文言を出し分けてPush通知した群のメッセージ開封率の増加率が20.3%であった。

4. 得られた知見と今後の課題

これまでの実フィールド実証の結果を受け、ナッジ文言を個人ごとに出し分けることによる行動変容には一定の効

果が見込まれることが示唆された。一方で、社会実装においても効果を見込むためにはいくつかの課題が存在する。

(1) 出し分けに資するユーザ情報の取得の難しさ：個人ごとにメッセージ文言を出し分け、最適化された状態を目指すものの、最適化に必要なユーザ情報を取得することは難しい。既存の端末アプリやWebアプリなどを用いる場合は取得可能な情報に制約がある上、既にサービス上同意取得している範囲を逸脱してデータ取得をすることはサービス規約の変更を伴うため一定のハードルを超える必要がある。一方で、新規に端末アプリなどを新規に作成する場合は、サンプルサイズを大きくするためにユーザへの認知や周知を十分に行う必要がある。情報の種類と量のいずれも担保するためには、社会実装した際の状態を精緻に予測しながら計画的に検証を進める必要がある。

(2) 実データの測定の難しさ：行動変容施策の最終的な目的は、ユーザの行動を変えることにある。例えば、移動行動に関する行動変容で店舗への送客を目的とした場合においては、送客先の店舗への訪問や送客先の店舗での決済を検知する必要がある。訪問の検知については、GPSを用いた手法により端末アプリでの実装だけで実現が可能で最も低成本で実現することが可能である。しかしながら、大型のショッピングモールや地下などGPSでは十分に精度が出ない環境においては、BluetoothビーコンやWi-Fiなどを用いた代替の検知手法が必要となるケースがあり、これらには信号の発信機や受信機などの機器の設置によるコストがかかる。決済の検知については、キャッシュレス決済システムなどを店舗ごとに導入することで実現が可能であるものの、システム間の連携や機器導入自体のコストやハードルが存在し、容易には実現できない。行動変容施策を実施する範囲は数個の店舗ではなく数十や百以上の店舗を対象として実施することが多く、全ての店舗に対してこれらの訪問や決済を検知する方法を実装することは難しい。

(3) 因果関係の明確化の難しさ：ここまで述べたような現実世界のデータを収集することが可能であったとしても、施策による行動変容効果を実計測することは難しい。サービス提供者や施策実施の意思決定者が期待する行動変容の効果は、「もともとそのつもりがなかった人が施策により行動を変えた」ことによる効果である。施策実施の対象者と施策実施後の行動を観察・計測できた場合でも、「もともとそのつもりがあったのか否か」を計測することは難しい。そのため、行動変容により何を実現したいのか、実現に向けてどのようにKPIを設定するのか、は施策設計者と綿密にすり合わせる必要がある。

(4) ナッジによる行動変容効果の限界：ナッジというのは付加的なものであり、それ単体で劇的な効果を發揮することはない。いかに素晴らしいメッセージ文言による介入を行ったとしても、限られたデータ取得の中でナッジによる

効果を明確に提示することは難しい。そのため、コンテンツの種類を増やす、ユーザの数を増やす、といったより根本的な部分での施策効果向上のための取り組みと並行して実施していくことが重要である。

5. おわりに

本稿では、筆者らが検討を進めているナッジのパーソナライズ化の概要を実フィールド実証の例とともに述べた。引き続き大規模かつ長期に渡る実フィールドでの実証および社会実装を進め、ナッジのパーソナライズ化による効果検証を推進していく。

参考文献

- [1] Thaler, R. H. and Sunstein, C. R.: *Nudge: Improving decisions about health, wealth, and happiness*, Penguin (2009).
- [2] Sunstein, C.: Nudging: A Very Short Guide, *Journal of Consumer Policy*, Vol. 37, No. 4, pp. 583–588 (2014).
- [3] Kim, H. B., Iwamatsu, T., ichiro Nishio, K., Komatsu, H., Mukai, T., Odate, Y. and Sasaki, M.: Field experiment of smartphone-based energy efficiency services for households: Impact of advice through push notifications, *Energy and Buildings*, Vol. 223, pp. 110–151 (2020).
- [4] Valle, C. G., Nezami, B. T. and Tate, D. F.: Designing in-app messages to nudge behavior change: Lessons learned from a weight management app for young adults, *Organizational Behavior and Human Decision Processes*, Vol. 161, pp. 95–101 (2020).
- [5] NTT ドコモ: 横須賀・三浦エリアで「観光型 MaaS」の実証実験を実施, https://www.ntt-docomo.ne.jp/binary/pdf/info/news_release/topics_210621_01.pdf.
- [6] 京浜急行電鉄: みさきまぐろきっぷ, https://www.keikyu.co.jp/visit/otoku/otoku_maguro/.
- [7] NTT ドコモ: バス車内の混雑平準化に向けた実証実験用アプリ「バス快適乗車案内」を提供, https://www.ntt-docomo.ne.jp/binary/pdf/info/news_release/topics_210916_00.pdf.
- [8] Pop-Eleches, C., Thirumurthy, H., Habyarimana, J. P., Zivin, J. G., Goldstein, M. P., de Walque, D., MacKeen, L., Haberer, J., Kimaiyo, S., Sidle, J., Ngare, D. and Bangsberg, D. R.: Mobile phone technologies improve adherence to antiretroviral treatment in a resource-limited setting: a randomized controlled trial of text message reminders, *AIDS*, Vol. 25, No. 6, pp. 825–834 (2011).
- [9] Moriwaki, D., Fujita, K., Yasui, S. and Hoshino, T.: Fatigue-Aware Ad Creative Selection, Workshop SUM' 20: State-based User Modelling at The 13th ACM International WSDM Conference (WSDM'20), p. 8 (2020).
- [10] NTT ドコモ: 行動変容で「CO₂ の削減」をめざす実証実験を御殿場プレミアム・アウトレット付近で開始, https://www.ntt-docomo.ne.jp/binary/pdf/info/news_release/topics_211105_00.pdf.
- [11] NTT ドコモ: 店舗の混雑状況で、d ポイントの進呈率を変動! 店舗内の「密」を防ぐ共同研究開始, https://www.ntt-docomo.ne.jp/binary/pdf/info/news_release/topics_201119_00.pdf.
- [12] 井手崇博, 大滝亨, 山田暁, 佐藤弘之, 落合桂一, 川上博, 矢野達也, 大川博生, 和久井圭祐, 白井拓也, 荒川豊: 新型コロナウイルス感染防止に向けた混雑回避行動を促進する行動変容技術の検証, 情報処理学会 SIG-BTI シンポジウム, pp. 52–55 (2022).
- [13] 酒井亮勢, 落合桂一, 井手崇博, 大滝亨, 山田暁, 鈴木喬, 佐藤弘之, 川上博, 矢野達也, 大川博生, 和久井圭祐, 白井拓也, 荒川豊, 星野崇宏: インセンティブとの組み合わせを前提としたナッジ効果測定のための実験設計方法, 情報処理学会 SIG-BTI シンポジウム (2023).
- [14] NTT ドコモ: 鉄道の混雑解消に向け、行動変容を促す共同実験を実施, https://www.ntt-docomo.ne.jp/binary/pdf/corporate/technology/rd/topics/2021/topics_220121_00.pdf.
- [15] NTT ドコモ: 幕張新都心エリアにおける「千葉市幕張 MaaS 実証事業」を推進, https://www.ntt-docomo.ne.jp/binary/pdf/info/news_release/topics_220125_00.pdf.
- [16] 吉川裕木子, 勝間田優樹, 鈴木喬, 山田暁: アプリログによる属性情報の拡張を用いたナッジ最適化レコメンデーションシステムの提案と評価, 情報処理学会 SIG-BTI シンポジウム (2023).

オンライン会議中の参加状態推定と 音声変換提示による参加促進手法の提案

城谷知葵 ^{†1}

九州大学大学院システム情報科学研究院

中村優吾 ^{†3}

九州大学大学院システム情報科学研究院

難波洸也 ^{†2}

九州大学大学院システム情報科学研究院

荒川豊 ^{†4}

九州大学大学院システム情報科学研究院

1. はじめに

会議においては発言やリアクションなど場への参加度が高いことが望ましい。発言やリアクションなどの情報によって相手の状況を理解することができ、会議を円滑に進めることができる。これは近年盛んであるオンライン会議においても同じである。オンライン会議は場所を気にすることなく参加できるため、従来の対面会議に比べて会議への参加が容易になった。オンライン会議において参加度を上げるために効果的な方法はカメラとマイクを ON にすることである。カメラを ON にすることで表情や頷きなどの仕草を確認することができ、相手の反応を理解することが出来る。つまり参加者の情報を対面会議同様に得ることができる。

しかしオンライン会議中にカメラとマイクを ON にすることは参加者に測定されている・監視されているなどの印象を与えててしまう。そのため対面会議よりも心理的な負荷が大きくなる。さらに話し手・聞き手ともに無意識のうちに自然な振る舞いが阻害されることや心理的な緊張感による疲れも懸念される。会議などのコミュニケーションにおいて心理的な安全性を確保することは会議を活発にする上で会議の質を高める上でも大切である。

上記の問題を解決するにはカメラとマイクを OFF にすることを考えられる。カメラを OFF にすることによって自分と相手の顔が見れなくなり、心理的負担を減らすことができる。しかし、カメラやマイクを OFF にするとノンバーバル情報(身振り・手振り、息遣い、視線、布ズレ音、椅子に座り直す音など)が欠落し、他の参加者の反応を理解することが困難になる。その結果会議を円滑に進めることができなくなり、会議の質が低下してしまう。

カメラやマイクを OFF でもリアクションを伝えるために

Teams, Zoom, Discord, Slack, hangout, WebEx などの各種コミュニケーションツールはチャットやエモート、他者のチャットへのいいねなどの機能が存在している。しかしこれらの機能は参加者自らが用いる必要があり、対面会議時のリアクションのように自然に発生するものとは異なる。そのため参加者に心理的負担を与えることになる。よって参加者のリアクションを自動でフィードバックするシステムがあれば、情報開示のような負荷やリスクの影響なく、会議を円滑に進めることができる。

しかしカメラ OFF でのオンライン会議においてはリアクションを伝えることよりもオンライン会議特有の問題を先に解決すべきである。1つ目はカメラを OFF にすることによって参加者の緊張感が減ることである。対面会議ではオンライン会議のカメラほどではないが、他の参加者に見られているかもという緊張感を少なからず与えることができる。しかし、オンライン会議でカメラを OFF にすると他の視線が完全にカットされるため、会議への集中力が減ってしまうことがある。2つ目は認知資源が他のことに奪われやすいことである。オンライン会議に使用しているデバイス(PC やタブレット、スマートフォンなど)を前に置いていることによってポップアップ通知やその他の機能に気が散ってしまう。また個々の作業空間によってはデバイスの外にも気が散るもののが存在し、注意が別の方向に向いてしまうことがある。

以上のことからリアクションを自動で伝えるシステムの開発より先にオンライン会議中に他所にそれる注意を引き戻すシステムが必要である。

Arakawa らは、ビデオベースの学習において非集中状態を音声変化によって再び集中状態に戻す手法を提案している[1]。Mindless Attractor は実験参加者の視線によってビデオへの集中力を測定しており、PC から目線が外れている(スマホなどを見ている)場合には、ビデオ音声のピッチを変化させることで集中状態に引き戻している。しかし上記の手法ではオンライン会議への参加と PC 上での内職(ネットサーフィンや動画視聴)を識別することが困難である。ま

^{†1} TOMOKI JOYA, Kyushu University

^{†2} KOYA NAMBA, Kyushu University

^{†3} YUGO NAKAMURA, Kyushu University

^{†4} YUTAKA ARAKAWA, Kyushu University

た音声変化に用いているのはピッチだけであり、音声変化が1種類だけであると馴化してしまう可能性がある。

そこで本研究では、オンライン会議において参加者のPC情報からオンライン会議への集中状態を推定し、その状態に応じて音声を変化させ、参加者をオンライン会議へと注意を引き戻す手法を提案する。オンライン会議への集中状態を推定するために用いるのはPCのプロセス情報だけであるため、カメラやマイクなどによる測定感を軽減することができ、参加者に不快感を与えることがなくなる。さらにプロセス情報によってPC上で内職を識別することもできる。また、非集中状態と推定された参加者に対しては、オンライン会議の音声に対し、ピッチやリバーブといった音変化をさせることにより、参加者に強制感を与えずに会議への注意を引き戻すことができる。

オンライン会議中、非集中状態と推定された参加者に対し、ピッチやリバーブといった音声変換提示をすることで、参加者の注意を会議へ引き戻し、集中状態へ戻る効果があるかを明らかにするための評価実験を実施した。7名の実験協力者に対し、普段オンライン授業に参加するときと同じ状況下において、ビデオ講演を聞いてもらった。音声変換ありとなしの場合の集中状態時間を比較した結果、音声変換を体験した実験参加者5名中4名に対し、ビデオ講演への集中時間割合が増加した。また、実験参加者は、ピッチやリバーブといった音声変換機能の詳細まで意識的に気づかなくとも、集中状態への復帰行動を取ることが確認できた。

2. 関連研究

本章ではユーザの状態をセンシング技術によって取得する研究と、聞き手を集中状態に引き戻す研究について述べる。

2.1. ユーザの状態を認識する研究

ユーザの集中状態を外部センサを用いて取得する研究が多く行われている[2, 3]。文献[2]では、被験者が装着したウェアラブルデバイスから生理学的なデータを取得することで被験者の参加状態をセンシングしている。また文献[3]では外部センサをPCに装着することで視線情報を取得している。しかし、外部センサを使用するためにはコストがかかる上に、会議のためにセンサを装着することは現実的ではない。

オンライン会議で用いるPCに付属しているカメラでユーザの状態をセンシングする研究も行われている[4-7]。文献[4]ではオンライン講義を受講している学生の表情をPCのカメラで撮影し感情を分析している。文献[5]では表情に加えて頷きや眉の動きをセンシングしている。PCのカ

メラを用いる手法は外部センサを新しく導入する必要がないため手軽に実施できるが、カメラの画角に顔を収める必要がある。また、オンライン会議ではカメラをオフにすることもあり、その場合は自分がカメラの画角内に収まっているか確認することができない。

そこで本研究ではCPU使用率やマウスの挙動などのユーザのPC情報に着目した。PC情報を用いることで画角を気にすることなくデータを収集することが可能となる。本研究ではActivityWatch[8]を参考にユーザのPC情報を取得し参加度を推定する。

2.2. 非集中状態のユーザを集中状態に戻す研究

次に参加度が低い聞き手に対してのフィードバック手法について述べる。文献[9]では視線情報から参加度を推定し、参加度が低い聞き手に対して試聴している動画を止めて画面上にアラートを表示する。文献[10]ではプレゼン資料上に四角形の枠を表示して聞き手に注意を促している。このようなPCの画面に提示する手法では聞き手がPCの画面を見ている必要があるが、別の画面を見ている場合には有効ではない。

そこで本研究では音による提示によって聞き手へフィードバックする。文献[1]では、聞き手が聞く音声のピッチを変換することで無意識的に会議に集中するよう促す手法を提案している。本研究でも音声変換をすることにより参加度が低い聞き手の参加度を向上させることを目標とする。

3. 提案システム

本章では、オンライン会議中において、注意が会議に向かっていない非集中状態にいる参加者を、会議へと注意が向いている集中状態に引き戻すシステムの概要について記す。まずははじめに、3.1節でオンライン会議における参加者の状態や遷移について定義し、ターゲットとなる参加者の状態と目指すべき参加者の状態について述べる。次に、3.2節でターゲットとなる参加者の状態のセンシング推定手法について述べ、3.3節でセンシング結果に基づいてシステムが参加者へどういった介入をすべきかについて述べる。最後に、3.4節で上記要件を満たしたプロトタイプ開発の詳細について述べる。

また、本研究で提案するシステムの適応対象となるオンライン会議の想定環境について下記に記す。

- ・会議形式：先生（話し手）と生徒（聞き手）といったような役割が分かれている会議
- ・発表資料：あり。話し手は画面共有をしている
- ・ハードウェア：PCやスマートフォンといった視覚と聴覚機能を有するデバイス
- ・ソフトウェア：TeamsやZoomといったオンライン会

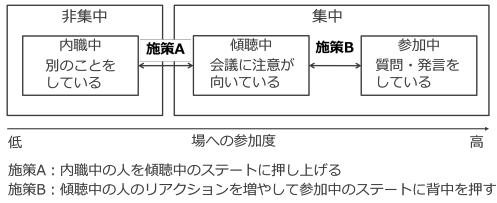


図 1 オンライン会議における参加者の状態遷移

議システム

- ・参加者のマイクの状態：問わない
- ・参加者のカメラプレビュー：OFF

3.1. オンライン会議における参加者の状態遷移

図 1 に示すように、オンライン会議における参加者の状態は、非集中状態と集中状態に大別できる。

非集中状態とは、参加者の注意がオンライン会議から外れており、心身的な理由から集中度が低い・乱れている状態、または会議システムを実行している機器の別のアプリケーションに認知資源を割いている状態、会議システムを実行している機器の外（参加者が置かれている環境）に認知資源を割いている状態を示す。代表的な行動としては、会議には直接関係のないアプリケーションを使用する（メールチェックや Web 検索）、会議に参加せずにスマートフォンを操作するなどが挙げられる。

集中状態とは、オンライン会議に注意を向け、高いパフォーマンスが發揮されている状態である。集中状態は参加者の状態を傾聴中・参加中の二つに分類することができる。傾聴中はその名の通り、話に耳を傾けている状態を示す。参加中の代表的な行動としては、発話をする・チャットに書き込む・エモートを押すなどが挙げられる。

図 1 に示したように、左側に行くほど場への参加度が低くなり会議への貢献が低下し、反対に右側へ行くほど場への参加度が高まり、会議への貢献が向上すると見える。1章でも触れた通り、我々は会議の参加者の場への参加度を高めるため、会議の参加者をこの一番右側の参加中の状態へと促すことを目指したい。しかし先に挙げた様々な要因から参加者の集中が保たれず、参加者の注意がオンライン会議から逸れて、非集中状態に落ちてしまうことがある。

そこで、オンライン会議の参加者の場への参加度を高めるためには、図 2 のようにまず参加者の状態が集中・非集中状態のどちらにあるのかを検出し、参加者の状態に応じて、”施策 A：内職中の人在傾聴中のステートに押し上げる”と、”施策 B：傾聴中の人のリアクションを増やして参加中の状態に背中を押す”を使い分ける必要がある。これらの施策のうち、本研究ではオンライン会議中に注意が逸れてしまう参加者の注意を引きもどすシステムを提案する。また本論

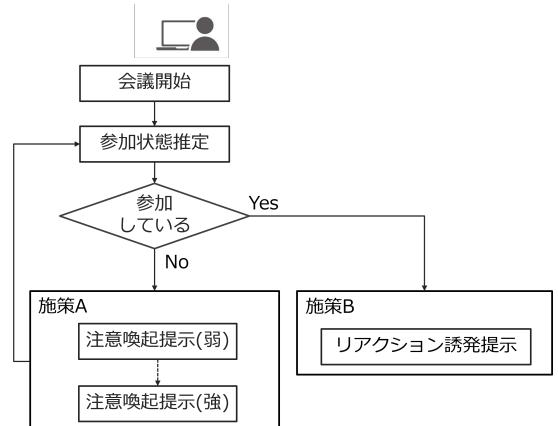


図 2 参加度を高めるための全体フロー

文では、デバイスは普段使用しているノート PC を想定し、代表的な非集中状態の行動として、内職作業を取り上げる。

3.2. 参加状態推定プロック

参加状態推定プロックではユーザーの参加状態を推定する。図 3 に参加状態推定プロックの概要を示す。

まずユーザーの PC から PC 情報を取得する。PC 情報は CPU 使用率・Active Window (PC 上で作業中や選択中の画面)・AFK 時間 (マウスやキーボードを触っていない秒数) で構成されている。PC 情報を取得するプログラムは python で記述しユーザの PC で動作させる。プログラムは 1 秒毎に CPU 情報を取得し csv ファイルに書き込む。

次に取得したデータをそれぞれ前処理プロックに入力する。CPU 使用率データに対しては、全て値が 0 である CPU プロセスを削除した後に使用した Web 会議システムの CPU 使用率で正規化する。Active Window データに対しては、使用した Web 会議システムかそれ以外かの 2 値に分類する。AFK データに対しては、閾値を設定して 2 値化する。本研究では閾値は 3 秒とする。

特微量抽出プロックでは前処理したデータを窓で区切り、それぞれ特微量を抽出する。本研究では窓サイズは 4 サンプル、スライド幅は 2 サンプルとする。CPU 使用率データからは最大値と最小値の差・標準偏差を計算する。Active

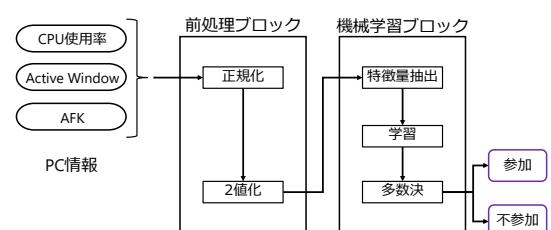


図 3 参加状態推定プロックの全体像

Window と AFK データからは合計値を計算する。

機械学習ブロックでは教師あり学習を用いて特徴量からユーザーの参加状態を推定する。正解ラベルは参加状態と不参加状態の 2 値である。機械学習モデルは限定しないが、本研究では lightGBM を用いた。

最後に推定されたラベルについて数サンプル毎に多数決をとり、最も個数が多かった状態を最終状態として出力する。本研究では 5 サンプル毎に多数決をとる。

3.3. 施策 A: どのように参加者の注意を引き戻すのか

施策 A の目的は、オンライン会議の非集中状態における参加者に対し、再びオンライン会議へ注意を向けさせることである。この際、我々が重要視するのは、参加者に対して強制感を与えることなく、感情や心理に訴えかけてさりげなく行動を促したいということである。よって、システムの介入によって参加者が作業を中断する必要がある・不快な思いをするようなことは避けたい。また、オンライン会議では、非集中状態にある参加者は常に PC 画面を見ているとは限らず、PC 画面を見ていたとしても打ち合わせアプリケーションではない他のアプリケーションを閲覧している状況も考えられる。こういった状況下で、参加者に対して有効な提示をするには、視覚的な提示よりも聴覚提示の方が適していると考えた。先にも述べたが、文献 [1] では、聞き手が聞く音声のピッチを変換することで無意識的に会議に集中するよう促す手法を提案している。しかし、ピッチ変換は音声変換の開始と終了時に注意を惹くことはできるが、変換中は馴化してしまい、効果としては不十分であると予測されたため、音声のピッチ変換だけでなく、音声変換中ずっと効果の持続性が期待できると考えられるリバーブ変換に着目した。

3.4. 施策 A：プロトタイプ実装

図 4 に提案手法の一連の流れを示す。我々は、非集中状態と判定された場合には、会議の音声を変換するプロトタイプ実装を行った。音声変換は 3 種類（ピッチ、リバーブ弱、リバーブ強）用意した。非集中時間の継続時間に応じて、ピッチ→リバーブ弱→リバーブ強と段階的に刺激の強度が変化するようにし、参加者が非集中状態になるごとに、繰り返されるようにした。オンライン会議システムに入力された話し手の音声を、仮想マイクを介することで聞き手の PC にインストールされたプログラムに入力して音声変換を施す。

使用 OS は macOS のみを対象としており、音声変換は pedalboard、CPU 情報取得には psutil を用いた。PC 情報は 1 秒間に 1 回取得し、音声変換は非集中状態の場合は 15 秒に 1 回施した。

4. 実験

4.1. 参加度推定性能

参加度推定の機械学習モデルを作成するために、教師データ取得実験と評価用データ取得実験を行った。いずれの実験でも、実験参加者のパソコンに「CPU 使用率」・「Active Window」・「AFK」を取得するソースコードをインストールし、コードを実行した状態で、Microsoft Teams を使用したオンライン会議を行った。

まず教師データ取得実験のため、オンライン会議に入った 5 人の実験参加者に指定した行動をとってもらった。参加している状態では「Teams の画面を真面目に見る」・「議事録をとる」・「Teams のチャットに質問を書く」の行動をしてもらった。不参加の状態では「ネットサーフィンする」・「資料を作成する」・「YouTube 動画を見る」などの行動をしてもらった。なお、パソコンを使用する以外の不参加状態はここで取得していない。

次に評価用データ取得実験を行った。2 名の実験参加者を参加している状態と不参加の状態に分け、オンライン会議室に入らせてもらった。教師データ取得実験とは異なり、実験参加者には指定した行動ではなく自由な行動をとってもらった。実際は、参加状態では「Teams の画面を真面目に見る」、不参加状態では「ネットサーフィンする」の行動があった。

取得した教師データを用いて 3.2 章の手順で機械学習モデルを作成した。前処理とアンダーサンプリングの結果、機械学習モデルに使用したデータは参加状態と不参加状態いずれも 1770 サンプルであった。作成した機械学習モデルを評価データで検証したところ正解率は 0.776 であった。

4.2. 施策 A：予備実験

ピッチやリバーブ音声変換によって、参加者の注意をオンライン会議に引き戻す効果があるかを検証するにあたり、まず参加者がピッチやリバーブ音声変換に気づくかどうか、またその機能はシステムの挙動として受け入れられる（受容性）のか、ピッチとリバーブ変換はどちらが効果的なのか

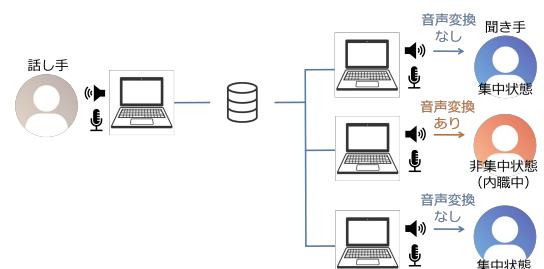


図 4 提案手法の一連の流れ

といった知見を得るための予備実験を行った。

4.2.1 実験概要

音声変換データを3種類（ピッチ、リバーブ弱、リバーブ強）用意した。14名の実験参加者に対し、3種類の音声変換データのうち1種類を体験するようにグループ分けした。内訳はピッチ変換（5名）、リバーブ弱変換（5名）、リバーブ強変換（4名）である。テスト中、実験参加者は変換された音声を聞きながら、手元のPCで文章を読む作業を行ってもらった。テスト終了後に音声内容理解の確認テストと作業負荷や音声に対する印象についてのアンケート調査を行った。

4.2.2 実験結果

アンケート調査によって、ピッチよりもリバーブ音声変換の方が気づきやすいという結果が得られた。また、テスト中に変換された音声を聞くことによる認知負荷については、変換されていない音声を聞くときと比較して大きな変化は見られなかった。変換されていない音声と変換された音声を聞く場合を比較して、音声内容の理解度や参加者の手元PC画面でのタスクパフォーマンスに大きな違いは見られなかった。

4.2.3 まとめと本実験に向けて

アンケート調査から明らかになった、ピッチよりもリバーブ変換の方が気づきやすいという結果から、音声変換した際に参加者に与える刺激の強度としてピッチくりバーブの関係があると考えられる。また、テスト中に変換された音声を聞くことによる認知負荷については、変換されていない音声を聞くときと比較して変化はなかったことから、ピッチやリバーブによる音声変換提示が参加者の心理や感情を阻害することはないと言える。ピッチ変換は音変換の開始と終了時に注意を惹けるが、変換中は馴化てしまい、また元々の話し手の音声に依存する部分も重なって効果が不十分であることが明らかになった。一方でリバーブ変換にはその懸念がなく、音変換開始、終了時に注意を惹けるのはもちろんのこと、音声変換中ずっと効果の持続性が期待できると考えられる。よって、本実験では、非集中時間の継続時間に応じて、ピッチ→リバーブ弱→リバーブ強と段階的に刺激の強度が変化するようにし、参加者が非集中状態になるごとに、繰り返されるようにした。

4.3. 施策 A：本実験

4.3.1 目的

予備実験の結果および得られた知見を踏まえ、オンライン会議において参加者が別アプリケーションに認知資源を割いている状態において、音声変換提示が、参加者の注意を引き付けて別アプリケーション作業を中断させる効果があるか否かを明らかにするための実験を行った。

4.3.2 タスク

実験参加者は、各自が普段使用しているノートPC内の打ち合わせアプリケーション（Teams）を通して、用意された音声ビデオを学習するように指示された。

4.3.3 素材

15分程度のUIST2019のプレゼンテーション発表動画を用意し、実験参加者に学習状況を再現させた。

4.3.4 実験参加者

実験はファシリテータ1名と実験参加者7名で行った。ファシリテータ1名と実験参加者7名は、共に九州大学の学生である。

4.3.5 実験手順

実験参加者は、椅子に座り、机の上にPCを置き、有線のイヤホンまたはヘッドフォンを装着してもらい、普段オンライン授業に参加するときと同じ状況下で実験を行った。音声変換機能がある場合とない場合での、違いを比較するため音声ビデオ学習を2回実施した。その際、順序効果をなくすため、最初に音声変換機能を体験するグループと2回目で体験するグループに実験参加者を分けた。図5に実験の流れを示す。まずははじめに、実験参加者に、これから音声ビデオ学習を行ってもらう旨を伝えた（事前説明）。音声ビデオ学習は2回に分けて行い、ビデオ学習後に作業負荷や音声に対する印象についてのアンケート調査を行った。すべてのテスト終了後、事後インタビューを行った。

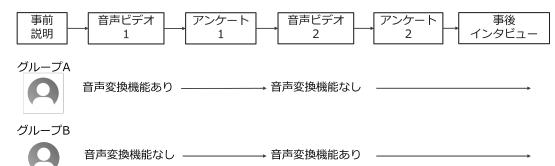


図5 実験の流れ

4.3.6 測定データ

音声ビデオ学習中は、参加者のPCのCPU使用率、ActiveWindow、AFKのログを採取した。学習後に行ったアンケートについては、ポジティブ/ネガティブ感情の変化や音声変換に気づいたか否かを調査した。すべてのテスト後の事後インタビューによって、参加者の感想を集めた。

4.3.7 実験結果

図6に、音声変換機能なしとありの場合の集中状態時間割合の変化を示す。7名の実験参加者のうち、複数回にわたって音声変換を体験したのは5人（ID:3,4,5,6,7）であった。図6のグレー色で示した2名（ID:1,2）はテスト中ずっと集中していたため、音声変換機能をほとんど体験しなかった。音声変換を体験した5人中4人が、音声変換提示をすることで、音声ビデオへの集中時間の割合が上がった。また、音声ビデオへの集中状態へ戻ったのは、リバーブ変換発動時

が最も多く、ピッチ変換だけでは音声ビデオへの集中状態に戻らなかった。また、テスト後の事後インタビューから、実験参加者 7 人中 2 人が、音声変換に気づき、残りの 5 名は気づかなかった。音声変換を体験した 5 人中、リバーブ変換機能に気づいた人は 1 人しかいなかった。実験参加者 7 人中 2 人が、音が変化した際に Teams 画面を確認したとコメントした。音声変換による、ポジティブ/ネガティブ感情の大きな変化は見られなかった。

参加者ID	音声変換なし(%)	音声変換あり(%)	変化率(%)
3	21.1	27.6	130.8
4	65.1	69.6	106.9
5	29.9	41.0	137.1
6	64.1	45.6	71.1
7	25.2	56.2	223.0
平均	56.1	62.3	111.0

図 6 実験参加者の集中時間割合の変化

4.3.8 まとめと今後の課題

音声変換を体験した 5 名中 4 名に対し、音声ビデオへの集中時間割合が増加する結果が得られたことから、本システムによる音声変換提示は、参加者の注意を惹き戻す効果があると言える。音声ビデオへの集中状態へ戻ったのは、リバーブ時が多くかった理由としては、実験参加者が、そもそもピッチ変換に気づいていない可能性が高い。本実験で使用したプロトタイプは、ピッチ変換時の音声の変化量が小さかったために参加者に気づかれにくかったことや、実装の都合上、変換時にノイズ（ぶつぶつ途切れる音）が混ざってしまったことが原因と考えられる。音声変換を体験した 5 人中、リバーブに気づいた人は 1 人だけだったが、意識的に気づいてなくとも実際に集中状態への復帰行動は複数の実験参加者の間で観測されていた。最後に、音声変換によるポジティブ/ネガティブ感情の大きな変化は見られなかったことから、本システムによる音声変換に対する参加者の受容度（受け入れやすさ）に関しては問題ないレベルであったといえる。

今後の課題としては、まず今回の本実験におけるピッチ音声変換時にノイズが混ざってしまったことで、純粋にピッチ変換だけの効果ではなく、ピッチ+ノイズの効果になってしまった点が挙げられる。また、音声変換提示は、話し手の声の性質や聞き手との関係性（聞き手にとって面識のない人が発表する際はより気づきにくくなる可能性）、参加者が使用しているヘッドフォンの性能、出力方法（ヘッドフォン or スピーカー）といった外部環境に注意する必要がある。

5. おわりに

本研究ではオンライン会議においてカメラを ON にすることなく参加者を集中状態にする手法を提案した。そのための第一歩として内職中の人の傾聴中のステートに押し上げる手法（施策 A）の有効性を検証した。提案手法では、まず参加者の PC から取得した CPU 使用率、Active Window, AFK 情報を機械学習モデルに入力し参加状態を推定する。次いで非集中状態と判定された参加者の音声に対してピッチシフト、リバーブの加工を施す。実験の結果より、音声変換を体験した実験参加者 5 名中 4 名に対し、ビデオ講演への集中時間割合が増加したこと、音声変換を施しても無意識的に実験参加者を集中状態へ復帰できることを確認した。今後は、傾聴中の参加者のアクションを増やして参加中の状態に背中を押す手法（施策 B）について取り組む予定である。

謝辞

本稿の研究を行うに際し、ご助言いただいたソニーグループ株式会社川名氏、森氏、朽木氏に感謝の意を表する。

参考文献

- [1] Arakawa, R. and Yakura, H.: Mindless Attractor: A False-Positive Resistant Intervention for Drawing Attention Using Auditory Perturbation, *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pp. 1–15 (2021).
- [2] Di Lascio, E., Gashi, S. and Santini, S.: Unobtrusive Assessment of Students' Emotional Engagement during Lectures Using Electrodermal Activity Sensors, Vol. 2, No. 3 (online), 10.1145/3264913 (2018).
- [3] Yao, N., Brewer, J., D'Angelo, S., Horn, M. and Geragle, D.: Visualizing Gaze Information from Multiple Students to Support Remote Instruction, *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI EA '18, New York, NY, USA, Association for Computing Machinery, p. 1–6 (online), 10.1145/3170427.3188453 (2018).
- [4] Yao, N., Brewer, J., D'Angelo, S., Horn, M. and Geragle, D.: Visualizing Gaze Information from Multiple Students to Support Remote Instruction, *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI EA '18, New York, NY, USA, Association for Computing Machinery, p. 1–6 (online), 10.1145/3170427.3188453 (2018).
- [5] Murali, P., Hernandez, J., McDuff, D., Rowan, K., Suh, J. and Czerwinski, M.: AffectiveSpotlight: Facilitating the Communication of Affective Responses from Audience Members during Online Presentations, *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, CHI '21, New York, NY,

- USA, Association for Computing Machinery, (online), 10.1145/3411764.3445235 (2021).
- [6] Watanabe, K., Soneda, Y., Matsuda, Y., Nakamura, Y., Arakawa, Y., Dengel, A. and Ishimaru, S.: Discaas: Micro behavior analysis on discussion by camera as a sensor, *Sensors*, Vol. 21, No. 17, p. 5719 (2021).
- [7] Chen, C., Arakawa, Y., Watanabe, K. and Ishimaru, S.: Quantitative Evaluation System for Online Meetings Based on Multimodal Microbehavior Analysis, *Sensors and Materials*, Vol. 34, No. 8, pp. 3017–3027 (2022).
- [8] : ActivityWatch, <https://github.com/ActivityWatch/activitywatch>.
- [9] D'Mello, S., Olney, A., Williams, C. and Hays, P.: Gaze tutor: A gaze-reactive intelligent tutoring system, *International Journal of human-computer studies*, Vol. 70, No. 5, pp. 377–398 (2012).
- [10] Sharma, K., Alavi, H. S., Jermann, P. and Dillenbourg, P.: A Gaze-Based Learning Analytics Model: In-Video Visual Feedback to Improve Learner's Attention in MOOCs, *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge*, LAK '16, New York, NY, USA, Association for Computing Machinery, p. 417–421 (online), 10.1145/2883851.2883902 (2016).

スマートウォッチによる煩わしさを低減する通知タイミングの推定方法に関する検討

神田琢也[†] 高橋公海[†] 宮島麻美[†]
NTT 社会情報研究所[†]

概要

健康行動を促すことを目的として、多くのヘルスケアアプリケーションで Push 通知が活用されている。Push 通知はユーザに行動を促す上で有用である一方、そのわずらわしさによってアプリの利用をやめてしまうケースも多く見られる。こういった通知の煩わしさを低減させる方法の一つとして、行動の切り替わりのタイミング (breakpoint) に通知をするという方法が挙げられる。そこで本研究では、breakpoint を検出しそのタイミングで Push 通知を送ることで煩わしさを低減させることを目指し、スマートウォッチのセンサ情報を用いて機械学習で breakpoint を推定する方法を検討し、予備実験でその有効性を検証した。

1. はじめに

高齢化社会や COVID-19 の影響により、人々のヘルスケアへの注目は近年ますます向上している。スマートホンや、スマートウォッチなどウェアラブルデバイスの普及に伴い、それらのデバイスを活用し、個人の健康管理を行う Fitbit (Fitbit Inc., San Francisco, CA, USA)¹ や Apple Health (Apple Inc., Cupertino, CA, USA)² , Google fit (Google Inc., Mountain View, CA, USA)³ などに代表される健康行動を促すモバイルヘルスケアアプリケーションが多く登場している。

IQUIVIA の 2021 年の調査 [1]によれば、現在 350,000 を超えるヘルスケアアプリが利用可能であり、毎日約 250 個の新しいヘルスケアアプリがリリースされている。こういったヘルスケアアプリはアプリユーザの健康管理に対する手間を削減し、健康の維持や疾病の予防に対して大きく寄与すると考えられている。

また、個人の健康管理という目的だけでなく、新しい疾患の治療手段として Digital Therapeutics (DTx) が注目されている。医薬品のように薬事承認を必要とし、医師の管理下で疾病患者自身が使用する治療目的のプログラムである。日本でもニコチン依存症改善アプリの CureApp SC (CureApp Inc., Tokyo, Japan)⁴ や高血圧治療アプリの

CureApp HT (CureApp Inc., Tokyo, Japan)⁵ など医療機器認定されたヘルスケアアプリが登場し注目を集めている。これらの DTx は主にスマートホンやスマートウォッチなどのデバイスを用いたアプリを活用し、疾病患者の生活習慣や行動に変化を生じさせるよう促すことにより、治療効果をもたらすことを目的としている。このように、ヘルスケアアプリで個人の生活習慣や行動に変化をもたらすことが、今後、健康状態の改善・維持においてさらに大きな役割を果たすことが予想される。

しかし、ヘルスケアアプリの課題としてユーザの利用が継続しないことやエンゲージメントの低さが指摘されている [2]。継続率を低下させる可能性のある主要な要因の一つとして、具体的にはヘルスケアアプリからの Push 通知が挙げられる。ヘルスケアアプリにおける Push 通知は、内容とタイミングが適切である場合、アプリユーザの満足度の向上に寄与すると言われている [3, 4]。一方で、内容やタイミングが適切でなかった場合や、頻度が高すぎる Push 通知はユーザの満足度を低下させ、アプリを削除してしまう可能性があると指摘されている [4, 5]。ヘルスケアアプリでは健康に関わる情報提供や健康行動のリマインダする際に Push 通知が使われることが多く、健康行動を促す際に重要な機能の一つである。そのため、その煩わしさを低減させることは、多くのヘルスケアアプリの改善につながることが期待できると我々は考えている。また、ヘルスケアアプリが対象としている健康行動の内容に関わらず、広い範囲のヘルスケアアプリへの応用が可能であることから通知の内容ではなく、タイミングに着目して研究を進める。

本研究では、ヘルスケアアプリの Push 通知のタイミングに着目し、スマートウォッチのセンサデータを活用して Push 通知の適切なタイミングを自動的に推定する手法を検討とする。スマートウォッチでセンシングしたデータを用いることで、従来のスマートホンなどのデバイスを活用した手法と比較して、適切なタイミングをより高頻度で検出できる可能性がある。本稿では適切な通知タイミングを推定することを目的とし、スマートウォッチのセンサ情報を用いた機械学習による推定手法の有効性の検証を行った。

Study on Estimation Method of Notification Timing by Smartwatch to Reduce Annoyance

^{†1} TAKUYA KANDA, NTT Social Informatics Laboratories.

^{†2} MASAMI TAKAHASHI, NTT Social Informatics Laboratories.

^{†3} ASAMI MIYAJIMA, NTT Social Informatics Laboratories.

¹ <https://www.fitbit.com/>

² <https://www.apple.com/jp/ios/health>

³ <https://developers.google.com/fit/>

⁴ <https://sc.cureapp.com/d/>

⁵ <https://cureapp.co.jp/productsite/ht/>

予備実験として実験参加者 11 名の通常の業務時のスマートウォッチのセンサ情報と適切な通知タイミングと考えられる breakpoint のアノテーションデータを収集し、Long Short-Term Memory (LSTM) ネットワークを用いた機械学習モデルを構築し推定の精度の検証を行った。その結果機械学習による breakpoint の推定に対する一定の有効性を確認した。

2. 関連研究

2.1 Breakpoint を活用した通知の煩わしさの低減手法

Push 通知の煩わしさを低減させる手法の研究として、時間帯や場所、一週間のどの日か、Bluetooth で検出したユーザ間の距離、電話・SMS のログに着目した手法など多くの先行研究が存在する [6]。先行研究の中でも主要なアプローチの一つとして breakpoint に着目した手法が挙げられる。breakpoint は心理学分野における概念であり、「意味がある自然な行動が終わり、新たな行動が始まったタイミング」と定義されている [7]。この breakpoint に通知を送ることで、通知によって起きる悪影響が軽減されているという結果がいくつかの先行研究で示されている [8~15]。

Iqbal ら[8~11]はオフィス作業をさせた際にパソコン上のクリックなどのイベント情報をもとに breakpoint を推定するミドルウェアの Oasis を提案し、Oasis による breakpoint の検出とその通知の煩わしさを低減させる効果を検証した。Oasis では実験参加者の作業中の動画をもとに、その作業に関する専門家がアノテーションをすることによって breakpoint の機械学習を行っている。

Okoshi ら [12~15] はモバイル端末を対象に、breakpoint を検出する手法を提案している。これらの研究ではスマートホン操作（アプリの起動、スワイプなど）やセンサ情報をもとにした行動認識 API の認識結果をもとに breakpoint を検出する手法の提案を行っている。

Breakpoint を活用した Push 通知の煩わしさを低減させる手法は、Push 通知を breakpoint が検出されるまで遅延させることによって、Push 通知の煩わしさを低減させている。そのため、検出できる breakpoint の頻度が低いと、必要なタイミングに Push 通知を送ることができない可能性がある。

パソコンやスマートホンなど多くのデバイスで breakpoint を推定する手法が提案されている [8~15] が、これらのタイミングの推定は実験参加者がデバイスを扱っている、もしくは身に着けている状態を前提とした手法である。スマートウォッチは体に密着しているデバイスであることから、breakpoint をより高頻度で検出できる可能性がある。

スマートウォッチを活用して breakpoint の検出を行った研究としてスマートウォッチとスマートホンを併用して使

用した際の breakpoint の検出手法の検討を行った [13] が挙げられる。[13] における実験の一部としてスマートウォッチ画面のタッチなどの UI 操作と、スマートウォッチに搭載されているセンサ情報をもとにした行動認識 API を活用し「立つ」、「座る」などの行動を認識し、認識された行動の切り替わりから breakpoint を推定する手法を提案した。

2.2 先行研究と本研究の違い

本研究では breakpoint を検出しそのタイミングで Push 通知を送ることで煩わしさを低減することを目指し、スマートウォッチなどのリスト型デバイスに一般的に搭載されている加速度センサとジャイロセンサの情報をもとに、高頻度で breakpoint を検出する手法の検討を行う。

行動認識を活用した breakpoint の推定手法は「立つ」、「座る」といった行動認識 API などのシステム側が検出する対象として事前に設定された行動のカテゴリ間の切り替えタイミングで breakpoint を検出することができる。しかし、事前に設定された行動のカテゴリだけでなく、ほかの種類の行動を切り替えたタイミングでも breakpoint が発生することが想定され、これらの breakpoint の検出が可能になれば、Push 通知のタイミングを breakpoint まで遅延させた煩わしさ低減の手法が適用しやすくなると考えられる。

そこで、加速度センサやジャイロセンサから得られる情報に対して、行動認識を活用した breakpoint の検出ではなく、センサ情報に対して行動認識を活用せず、直接 breakpoint を推定する方法の検討を行う。本稿ではまず基礎検討として、機械学習による breakpoint の推定手法を構築し、予備実験でその有効性の検証を行う。

3. 実験手順

3.1 データ収集

機械学習による breakpoint の推定に用いるデータ収集のために、研究業務に従事している社員 11 名を募集し、業務中のスマートウォッチのデータと業務の様子を撮影した。実験参加者は実験の説明を受けたのち、実験参加に関する同意書にサインした。その後、利き手ではない方の手首にスマートウォッチの Fitbit Sense (Fitbit Inc., San Francisco, CA, USA) を装着し、Fitbit Sense 上に Fitbit SDK (Fitbit Inc., San Francisco, CA, USA) を用いて実装された加速度と角速度を記録するアプリを起動した。データ収集実験時の構成を図 1 に示す。

アプリの起動後参加者の作業の様子がわかる角度で動画の撮影をはじめ、通常の業務を 1 時間実施するよう依頼した。データ収集時の Fitbit に搭載された加速度と角速度のデータは Bluetooth を経由し、スマートホンの Google Pixel 5a (Google Inc., Mountain View, CA, USA) 内のサーバに送信

された。また、実験参加者が作業をしている様子の撮影には Insta 360 ONE RS (Arashi Vision Inc., Shenzhen, China) を用いた。実験中にデータ収集システムに不具合が生じた場合は、実験参加者に作業を中断してもらいデータ収集システムの再起動の実施を依頼した。

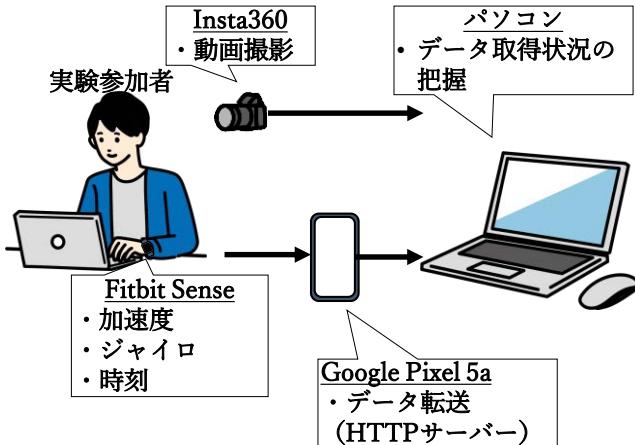


図 1 データ収集実験時の構成

3.2 Breakpoint アノテーションデータの作成

各実験参加者の作業動画に対して、アノテーター2名にアノテーションデータの作成を依頼した。アノテーションデータの作成には音声・映像アノテーションツールである ELAN 5.8 [16] を用いた。アノテーターはデータ収集実験に参加しておらず、なおかつ ELAN 5.8 の操作方法に精通しているものに依頼した。アノテーターはデータ収集実験参加者の作業動画を視聴しながら、一定の間隔で分割された動画の期間（タイムステップ）において breakpoint の発生を記録した。Iqbal ら[8~11]の実験では breakpoint の有無だけでなく、発生した breakpoint を3段階の粒度 (Course, Medium, Fine) に分類したアノテーション方法を実施したが、本稿では breakpoint の発生自体を機械学習で推定可能か検証を行うために breakpoint の有無のみを記録した。

本実験ではタイムステップの間隔を2秒と設定した。アノテーターには下記の文面でアノテーションを依頼した。
「実験参加者が作業している動画を見ていただきながら Breakpoint をアノテーションして頂く作業になります。

（※Breakpoint：意味がある自然な行動が終わり、新たな行動が始まったタイミング）。具体的には動画を見ていて実験参加者が行動を切り替えたと感じたタイミングにフラグを入れてください。どう感じたかを知りたいので、正解や間違いはありません。」

アノテーション作業完了後、アノテーターごとに判断に偏りがある可能性を考慮して、2人のアノテーターがいずれも breakpoint があると記録したタイムステップもしくは、一つずれたタイムステップに breakpoint があると記録

されたタイムステップを Breakpoint クラス、無いと記録されたタイムステップを Not Breakpoint クラスとした。



図 2 ELAN によるアノテーションの画面例

4. 機械学習による breakpoint 推定モデルの構築

4.1 センサデータの前処理

スマートウォッチの Fitbit Sense から得られた3軸の加速度センサ及び3軸の角速度のデータには欠損値が存在したため、前後の平均値を算出して補完した。欠損値補完処理を行った後、実験参加者の体の動きの情報をより、学習データに反映するために、加速度センサのデータに対してローパスフィルタを用いることで重力加速度の影響を除去する処理を行った。また、重力加速度の影響を除去した加速度センサから合成加速度の算出を行った。

4.2 センサデータの特徴量選択

4.1 項までの前処理を施した加速度 (x, y, z), 合成加速度、角速度 (x, y, z) のデータに対して、学習に用いる特徴量として、タイムステップごとに以下の値を算出した。

- 平均値
- 中央値
- 最大値
- 最小値
- 標準偏差

スマートウォッチ上のデータ収集アプリの不具合により、スマートウォッチのセンサ情報が記録できていないタイムステップを排除した結果、各実験参加者のデータに含まれるタイムステップ数は表 1 のようになった。

表 1 Not Breakpoint クラス, Breakpoint クラスの
タイムステップの数

実験参加者 ID	Not Breakpoint	Breakpoint	合計
P1	1327	16	1343
P2	1701	11	1712
P3	1256	2	1258
P4	864	2	866
P5	1504	13	1517
P6	1524	20	1544
P7	1733	5	1738
P8	1444	6	1450
P9	1491	28	1519
P10	1747	2	1749
P11	1756	14	1770
合計	16347	119	16466

4.3 機械学習モデル

スマートウォッチのセンサデータをもとに各タイムステップの breakpoint の有無を推定するため、4.2 項までの処理を施したタイムステップごとの特徴量を入力とした、機械学習モデルを作成した。機械学習のモデルを選択する際にロジスティック回帰や Support Vector Machine (SVM), 勾配ブースティングなど複数の手法を検討したが、最終的にテスト検証で精度の良かった Long Short-Term Memory (LSTM) ネットワークを選択した。本実験で使った LSTM の構造を表 2 に示す。作成した機械学習モデルは入力とした特徴量の次のタイムステップに breakpoint が存在する (Breakpoint クラス), 存在しない (Not Breakpoint クラス) かの 2 値分類を予測結果として出力した。入力データと予測対象のタイムステップの関係を示したものを見図 3 に示す。また、Breakpoint クラスと Not Breakpoint クラスの比率に大きな偏りが生じているため、クラスごとの重みを考慮した係数をかけた学習を行った [17]。

表 2 モデルの構造

入力サイズ	7×5
入力タイムステップ	30
LSTM の次元数	10
活性化関数	Sigmoid 関数
全結合層	1
最適化手法	Adam
損失関数	Binary Cross Entropy

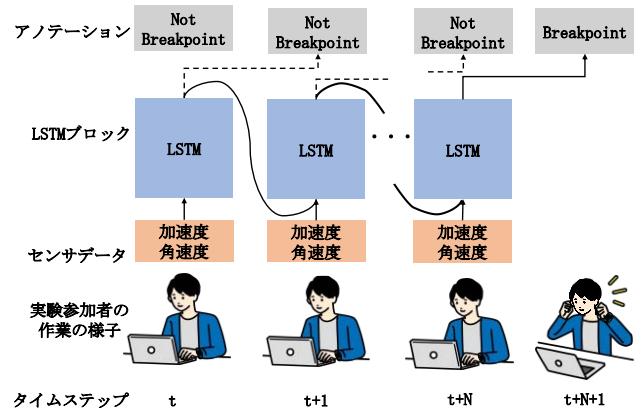


図 3 入力データと予測対象のタイムステップとの関係

4.4 評価

実験参加者 11 名のデータに対して Leave One Person Out (LOPO) 法を用いて学習を行い、モデルの推定精度を評価した。評価に用いたハイパーパラメータを表 3 に示す。

表 3 ハイパーパラメータ

バッチサイズ	32
エポック数	20
確信度の閾値	0.95

なお、加速度、角速度のデータが取得できていなかったタイムステップは学習時もテスト時も除外した。学習時のモデルのばらつきを考慮し、学習モデルを 10 回作成し、それぞれの予測結果の評価指標 (Accuracy, Precision, Recall, F1 値) の平均値を算出した。表 4 に各実験参加者のデータに対する各評価指標の平均値を示す。また、図 4 に各実験参加者のデータに対する予測結果の評価指標の箱ひげ図を示す。

表 4 評価指標の平均値

	Accuracy	Precision	Recall	F1 値
P1	1.00	0.99	0.86	0.88
P2	0.99	0.78	0.89	0.78
P3	1.00	1.00	1.00	1.00
P4	1.00	0.81	0.90	0.84
P5	0.99	0.77	0.85	0.76
P6	0.90	0.88	0.84	0.77
P7	1.00	0.64	0.78	0.65
P8	0.99	0.89	0.73	0.79
P9	0.99	0.77	0.84	0.77
P10	1.00	0.95	1.00	0.97
P11	1.00	0.90	0.86	0.87

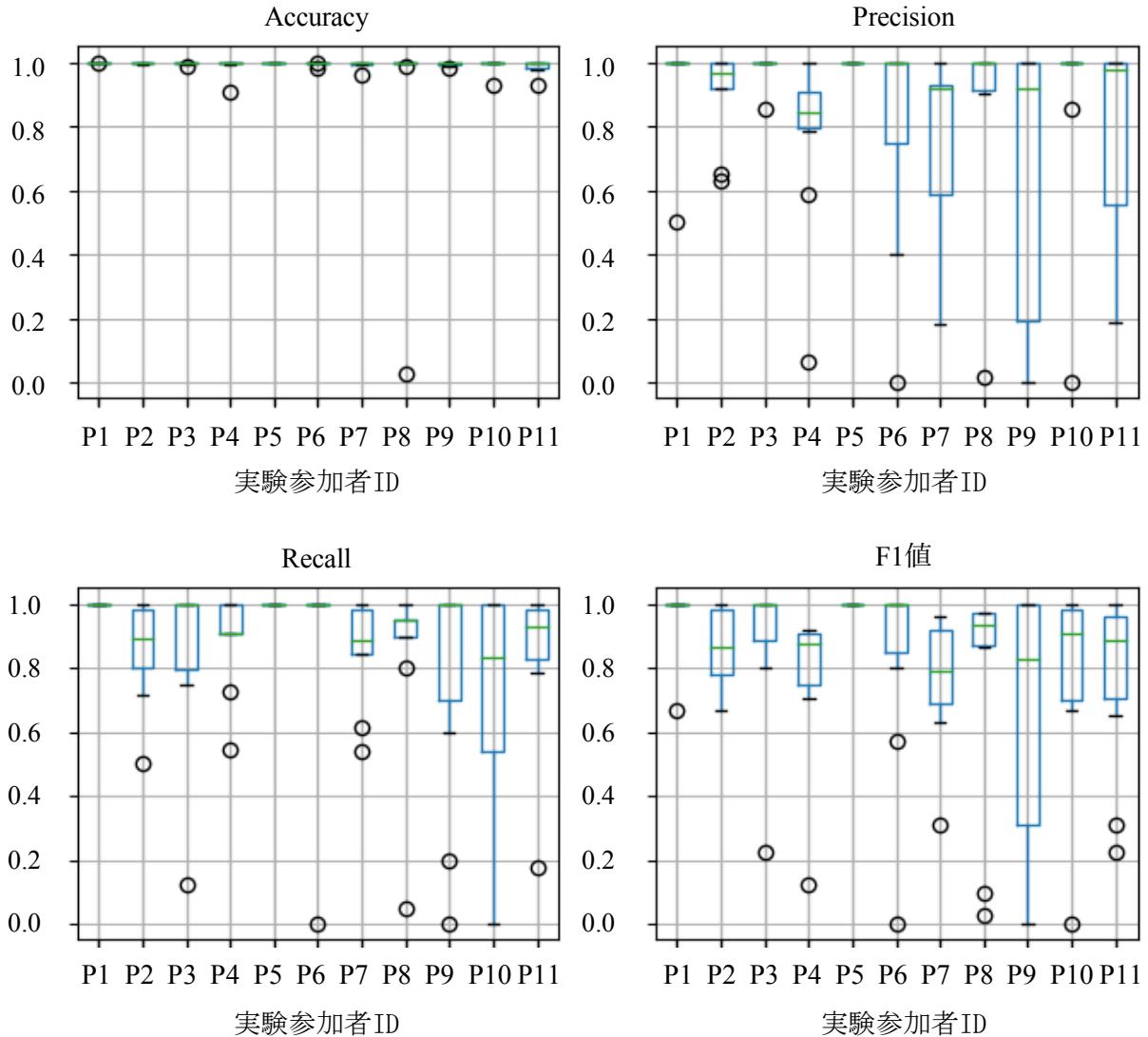


図4 各実験参加者のデータに対する予測結果 (Accuracy, Precision, Recall, F1 値)

5. 考察

5.1 Breakpoint の推定精度に関して

Accuracy は全体的に高かったが、これは Not Breakpoint クラスと Breakpoint クラスの比率に大きな偏りが生じていたためであると考えられる。Precision と Recall に関しては Accuracy ほど高い精度での予測はできなかったが、ある程度の精度で予測ができたためスマートウォッチのセンサ情報を用いた機械学習による推定手法に一定の有効性があることが示唆された。

Precision と Recall に関しては実験参加者のデータごとに大きな差が出ている。P9 のデータに対する予測精度が Precision, Recall ともにほかの実験参加者のデータと比較して低い結果が得られた。ほかの実験参加者はデータ収集中の作業がパソコン作業であったのに対して、P9 のデータ収集実験時の作業にはパソコン作業以外の作業も多く含まれていた。今回の実験では Breakpoint クラスの学習データ

が少なかったため、LOPO 法で評価するために学習データから排除した。そのため、作業内容がほかの実験参加者のデータと大きく異なる P9 のデータに対する予測結果の精度がほかの実験参加者のデータと比較して低くなったと考えられる。

本研究は Push 通知の煩わしさを低減させることを目的としている。煩わしくないタイミングを煩わしいタイミングと誤推定しても、アプリユーザは煩わしいと感じない。一方で、煩わしいタイミングを誤って煩わしくないタイミングであると誤推定し、Push 通知をするとアプリユーザに煩わしいと感じさせてしまうことが想定される。そのため、煩わしいタイミングに通知をしないという観点から、Recall の精度の向上よりも Precision の精度の向上がより重要であると考えられる。

5.2 Breakpoint をアノテーションする作業におけるアノテーター間の個人差

本実験ではデータ収集実験で撮影した実験参加者の作業動画に対して 2 人のアノテーターにアノテーションを実施してもらった。各アノテーターの Breakpoint クラスと Not Breakpoint クラスの数を比較したものを表 5 に示す。アノテーターをそれぞれ A1, A2 と記載する。同じ実験参加者の作業動画であっても、P2, P3, P4 のアノテーションデータは各クラスのアノテーションの数が大きく異なっていた。一方で、Breakpoint クラスであるとアノテーションしているタイムステップは同一、もしくは隣り合った前後のタイムステップであった。アノテーター間で Breakpoint の数に差が生じた理由は Iqbal らの研究 [13~16] で指摘されている breakpoint の粒度による影響であると考えている。アノテーションデータ作成後にアノテーションを付与した基準など感想をアノテーターに尋ねたところ、A2 のアノテーターからは「視線が大きく作業から外れた場合は breakpoint と判断した」というコメントがあり、A1 のアノテーターからは「視線が作業から外れた場合もその後作業が継続していた場合 breakpoint とは判断しなかった」とコメントがあった。これは、アノテーターごとに、breakpoint だと認識する場面に差異があるのではなく、アノテーションした際に粒度が小さいものでも Breakpoint クラスであるとアノテーションを付与したかどうかの差異であるように見受けられる。

表 5 アノテーターごとの各クラスのアノテーション数の比較

	Not Breakpoint		Breakpoint	
	A1	A2	A1	A2
P1	1318	1319	25	24
P2	1701	1667	11	45
P3	1256	1230	2	28
P4	862	829	4	37
P5	1502	1499	15	18
P6	1514	1524	30	20
P7	1733	1724	5	14
P8	1442	1431	8	19
P9	1491	1504	28	15
P10	1747	1743	2	6
P11	1756	1747	14	23
合計	16322	16217	144	249

5.3 本研究の適用範囲

本稿の実験ではオフィスのデスクワーク時の作業に関するデータを収集した検証であったため、今後はさらに生活の様々なシーンに適用範囲を広げていきたいと考えてい

る。また、本実験の参加者は普段からデスクワークを行っている研究者であり、作業の内容や実験参加者の属性として偏りが生じていたため、全てのオフィスワーカーで本手法が有効であるかどうかは実験の範囲外である点は留意する必要がある。

6. 結論・今後の展望

本稿では、ヘルスケアアプリの Push 通知の煩わしさを低減させることを目指し、適切な通知タイミングの推定を取り組んだ。Push 通知の煩わしさを低減させる手法の中でも breakpoint を活用した手法に着目し、當時アプリユーザに装着されているスマートウォッチのセンサデータを用いて LSTM ネットワークを活用した機械学習モデルで breakpoint を推定する手法を構築し、予備実験で検証を行った。

予備実験では、デスクワーク時のスマートウォッチのセンサデータの収集と breakpoint のアノテーションを行い、収集したセンサデータとアノテーションデータから LSTM ネットワークを活用したモデルで推定が可能かという検証を行った。その結果、今回の収集データでは比較的高い精度で breakpoint を推定することができたことから、スマートホンやスマートウォッチの行動認識 API などを活用せず、スマートウォッチに搭載された加速度センサやジャイロセンサの情報から直接 breakpoint を推定する手法が有効である可能性が示唆された。

今後は、さらに多く、多様な実験参加者からの作業時のデータ収集を行い検証する必要がある。また、本実験で実施した参加者は通常からデスクワークを行っている研究者であり、参加者の属性としては偏りがある。そのため、参加者の属性をより広げた実験をしたいと考えている。実験で対象とした行動に関しても、デスクワークと限られたシチュエーションであることから、さらに多様なシチュエーションでデータ収集を実施したいと考えている。また、これまでの先行研究との組み合わせを行うことで、Push 通知によってより効果的に発生する煩わしさを低減することが可能になるのではないかと考えている。

アノテーションに関しては、今回 2 人のアノテーターがアノテーションを行ったが、アノテーター間でアノテーションを付ける数に大きな差が発生していた。これは個人の主観に基づいてアノテーションを付けていることから、breakpoint だと判断するアノテーターの基準に差があったためであると考えられる。今後、大規模な実験や実用化を考える上では、さらに多くのアノテーションデータが必要になると考えられ、一人のアノテーターが大量の breakpoint のアノテーションデータを作成することは現実的ではないことから、複数人のアノテーターがアノテーションデータを作成することが考えられる。そのため、アノ

データー間のアノテーションに差が生じないようなアノテーション方法を検討していきたい。

将来的には、機械学習モデルを実際にスマートウォッチ上に実装し、breakpointを検出した際にPush通知することによって、その煩わしさが低下するかを実際に検証することを検討している。また、煩わしさを低減させることによる、ヘルスケアアプリの継続利用率の変化や、健康行動を促すリマインダのPush通知をアプリ利用者が受け取った際の健康行動の実施率を調査することで、行動の変化にどの程度繋がるかを検証する必要があると考えている。

7. 参考文献

- [1] IQVIA INSTITUTE FOR HUMAN DATA SCIENCE, Digital Health Trends 2021: Innovation, Evidence, Regulation, and Adoption (online), available from <<https://www.iqvia.com/-/media/iqvia/pdfs/institute-reports/digital-health-trends-2021/iqvia-institute-digital-health-trends-2021.pdf>>, (accessed 2023-02-13).
- [2] Amagai, S., Pila, S., Kaat, A., Nowinski C. and Gershon, R.: Challenges in Participant Engagement and Retention Using Mobile Health Apps: Literature Review J Med Internet Res, doi: 10.2196/35120 (2022).
- [3] Bidargaddi, N., Almirall, D., Murphy, S., Nahum-Shani I, Kovalcik, M., Pituch, T., Maaieh, H. and Strecher, V.: To Prompt or Not to Prompt? A Microrandomized Trial of Time-Varying Push Notifications to Increase Proximal Engagement With a Mobile Health App, JMIR Mhealth Uhealth, doi: 10.2196/10123 (2018).
- [4] Haggag, O., Grundy, J., Abdelrazek, M. and Haggag, S.: A large scale analysis of mHealth app user reviews. Empir Softw Eng, doi: 10.1007/s10664-022-10222-6 (2022).
- [5] Morrison, L.G., Hargood, C., Pejovic ,V.: Geraghty AWA, Lloyd S, et al., The Effect of Timing and Frequency of Push Notifications on Usage of a Smartphone-Based Stress Management Intervention: An Exploratory Trial., PLOS ONE, doi: <https://doi.org/10.1371/journal.pone.0169162>, (2017).
- [6] Abhinav, M., Musolesi, M.: Intelligent notification systems: A survey of the state of the art and research challenges, arXiv preprint arXiv:1711.10171, (2017).
- [7] Newson, D. and Engquist, G.: The perceptual organization of ongoing behavior, Journal of Experimental Social Psychology, vol. 12, no. 5, pp. 436–450 (1976).
- [8] Iqbal, S. T. and Horvitz, E.: Notifications and awareness: A field study of alert usage and preferences, Proc. 2010 ACM Conference on Computer Supported Cooperative Work (CSCW '10), pp. 27–30, (2010).
- [9] Iqbal, S. T. and Bailey, B. P.: Understanding and developing models for detecting and differentiating breakpoints during interactive tasks, Proc. SIGCHI Conference on Human Factors in Computing Systems (CHI '07), pp. 697–706 (2007).
- [10] Iqbal, S. T. and Bailey, B. P.: Investigating the effectiveness of mental workload as a predictor of opportune moments for interruption, Proc. 05 Extended Abstracts on Human Factors in Computing Systems (CHI EA '05), pp. 1489–1492 (2005).
- [11] Iqbal, S. T. and Bailey, B. P.: Leveraging characteristics of task structure to predict the cost of interruption, Proc. SIGCHI Conference on Human Factors in Computing Systems (CHI '06), pp. 741–750 (2006).
- [12] Okoshi, T., Ramos, J., Nozaki, H., Nakazawa, J., Dey, A. K. and Tokuda, H.: Attelia: Reducing user's cognitive load due to interruptive notifications on smart phones, Proc. 2015 IEEE International Conference on Pervasive Computing and Communications (PerCom), IEEE, pp. 96–104 (2015).
- [13] Okoshi, T., Ramos, J., Nozaki, H., Nakazawa, J., Dey, A. K. and Tokuda, H.: Reducing users' perceived mental effort due to interruptive notifications in multi-device mobile environments, Proc. 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp), pp. 475–486 (2015).
- [14] Okoshi, T., Tsubouchi, K., Taji, M., Ichikawa, T. and Tokuda, H.: Attention and engagement-awareness in the wild: A large-scale study with adaptive notifications, Proc. 2017 IEEE International Conference on Pervasive Computing and Communications (PerCom), IEEE, pp. 100–110 (2017).
- [15] Okoshi, T., Tsubouchi, K. and Tokuda, H.: Realworld product deployment of adaptive push notification scheduling on smartphones, Proc. the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data mining (KDD), pp. 2792–2800 (2019).
- [16] Brugman, H., Russel, A.: Annotating multi-media/multimodal resources with ELAN, Proc. the Fourth International Conference on Language Resources and Evaluation (LREC'04) (2004).
- [17] King, G., Zeng, L.: Logistic regression in rare events data. Political analysis, 9(2), pp.137-163 (2001).

行動変容を定着化させるための介入制御方式の検討

原田将治^{†1}

富士通株式会社

鳥谷彰^{†2}

富士通株式会社

1. はじめに

より良い社会に向かうために、人々を行動変容させなければ達成できない社会課題がある。例えば、健康推進やセキュリティ向上といった目標を達成するには、より多くの人の行動変容の定着化が必要となる。しかし、これらの行動変容の定着化に向けたプロセスにおいては、数回行動するだけでは成果が得られず、むしろ、すぐに行動による身体的な負担や心理的な負担を感じることが多いため、目標達成に至ることが容易ではない。

さらに、目標を達成するための行動は、義務感や強制ではなく主体的な取り組みであることが望まれることが多い。なぜなら、運動習慣やセキュリティ対策などは、ただ決められたことをやればよいというものではないからである。例えば、運動習慣なら、どれくらいの負荷や頻度であれば適切に継続できるのか各自のやり方は異なる。また、セキュリティ対策では、規制やルールの想定外の事態に直面することが多々あり、目的を理解して自ら判断して適切に行動することも求められる。

そこで本稿では、ICTをどのように活用したら、より多くの人の行動変容が定着化し、目的達成のための主体的な活動へと昇華させられるのかについて検討する。

2. 関連研究

行動変容を促す働きかけや支援となる介入は多くの種類や方法が試みられており、Behavior Change Technique Taxonomy(v1)[1]では、93種類もの介入方法が定義されている。ICTを活用した介入も多く検討がなされている。例えば、ライフログを活用したチーム内の「競争」と「協力」による介入の有効性の検証[2]や仲間と比較した電力消費量ランキングと報酬の有効性の検証[3]といったように、ICTでは計測した対象者の結果を提示するだけでなく、他者との合計、他者との比較といった見せ方を工夫して介入することができる。他にも、スマートクッション[4][5]において、対象者が良い姿勢を保てるように触覚刺激を与えるといったように常時介入し続ける仕組みはICTの得意とするところである。それ以外にも、対象者の状態に応じた介入、例えば、行動変容ステージモデルに基づきながらも、その間に生じる対象者の内面状態に応じたメッセージを提供したり[6][7]、対象者の忙しさや鼓動パターンに応じた提案を

提供したり[8][9][10]と、対象者に合わせた画一的でない介入も提案されている。

また、ICTによる機械的な介入が嫌がられないようとする試みもある。否定的なフィードバックであっても、優しく間接的に対象者に働きかける諭しの有効性を検証する試み[11][12]や、情報の閲覧を強制するが自ら選んで閲覧したと感じさせる手法の有効性を検証する試み[13]といったように、介入自体が対象者に不快感を与えない仕組みについての検討がなされている。

3. 行動変容の定着化

行動変容を定着化させるには長期的な戦略が必要と考えた。行動を主体的な取り組みとさせるには、一時的な行動変容ではなく、介入者側が戦略を持ちながら、少しづつ時間を掛けて、対象者の同意や共感を得ていくことが大切と考えたからである。例えば、6ヶ月の間に継続的に行動させることを想定してみても、対象者はその間、様々な心理状態になるし、様々な状況にも直面することが想像できる。これに対し、同じ介入、あるいは、最良の介入をただ繰り返していても、やがては飽きられることが容易に想像できる。例えば、意欲がなくなる度に仲間の状況を見せて奮起させる、ゴール達成者の感想で勇気付けるといった同じことを繰り返していても徐々に刺激が弱まり、いつまでも行動変容を持続できるとは限らない。つまり、1回の介入で、対象とする行動をすぐにさせることだけを単調に目指していくても、定着化するまでを支援し続けることは困難である。したがって、定着化を目指すのであれば、時には行動変容に価値があることを論理的に納得させること、行動は体感的に意外と簡単だと気付かせること、行動変容にどのような意味があるかを考えさせること、といった直接行動を促すことを目的としない介入も必要であり、行動変容の定着化に向けて徐々に進めていくことが重要である。

そのような長期的な戦略の一つとして、認知を変容させることに着目するとよいのではないかと考えた。最終的なゴールとして、主体的に行動に取り組ませるには、本人が行動やその目的の必要性や効果をどのように捉えているかという認知が内発的動機を高めることにつながると考えたからである。また、その認知の状態によって、介入自体の受け入れやすさや、介入によって生じる効果が変わることが予想されるので、認知に着目することは有用と考えた。例えば、単純なリマインダー通知であっても、行動変容の重要性を認知した状態であれば、状況やタイミングが悪く

Proposal of intervention control method to habituate behavior change

^{†1} SHOJI HARADA, Fujitsu Ltd.

^{†2} AKIRA KARASUDANI, Fujitsu Ltd.

行動に至らなかつたとしても、通知されることに対する不快感は強くないであろう。逆に、それほど行動変容に好感を持っていない認知状態の場合、少しでも状況やタイミングが悪いときに通知されてしまうと、通知されたという介入自体を邪魔に感じ不快感が生じる。したがって、どんなに単純な介入であっても、対象者の認知状態で受け止め方が変わるので、認知は介入の効果を左右する大事な着眼点になると考える。

また、介入を続ける際には、その不快感に対しても過度に与えたり、蓄積させたりしない仕組みが必要である。介入は必ずしも対象者に受け入れられるものではなく、タイミングが良くても介入内容が合わず、どうしても不快感を与えてしまうことがある。にもかかわらず、介入により不快感を与え続けることで、行動変容自体を挫折させてしまう可能性があるからである。

介入により不快感を与える要因としては、認知のギャップやタイミングの悪さ以外にも、次のようなものが考えられる。

- 現状維持バイアス
- 心理的リアクタンス
- ネガティブな感情
- 行動への阻害要因
- 介入内容が不明瞭

「現状維持バイアス」により、介入が対象者に生じさせる変化の度合いが高いほど受け入れにくくなり、不快感を生じやすい。さらに、介入が日常をどれくらい妨害するのか、つまり、どれだけ自由を奪われたのかによっても「心理的リアクタンス」が生じて不快感が増す。時には「ネガティブな感情」が強かつたり、「行動への阻害要因」が生じ

たりしていると、合理的には受容していた介入内容に対しても不快な気持ちを抱くことがある。また、対象者にとって「介入内容が不明瞭」であった場合、介入の意味が伝わらないことで、意図したつもりはなくとも不快感を生じさせることがあるだろう。このように様々な要因があるため、対象者への介入において不快感を生じさせることは避けにくい。むしろ、全く不快感を生じさせない介入は対象者に何も刺激を与えられていないともいえるので、完全な回避は目指すことではないと考える。

4. 提案方式

4.1 認知変容ステージモデルと介入制御

認知に着目した戦略とするため、行動変容ステージモデルを参考に表1のような認知変容ステージモデルを提案する。この認知変容ステージモデルに基づいて、対象者は認知の状態によって6種類に分けられ、そのステージに応じて認知に変化を与えられるように介入を制御する。したがって、1回1回の介入の主目的は対象者に行動を起こさせることではなく、認知に刺激を与えることにある。なお、認知も無理に1回の介入で変容させることは目指さず、段階を追いながら行動に対して対象者が自身にとって納得のいく捉え方を見付けさせることで、主体的な行動変容の定着化を目指す。次に各ステージにおける認知の状態と介入制御について説明する。

0) 反対期は、何らかの理由で目的や行動に反対しており、価値や意義がないと認知している状態である。そのため、ほとんどの介入は受容することが困難になっている。

表1 認知変容ステージモデル

Table 1 Attitude change stage model

ステージ	行動の捉え方	介入の方向性	状況	ステージにおける傾向			
				介入の受容	意図	行動	意欲
0) 反対期	意義や価値がない。	反対する理由の解消。	行動や目的に反対しているか、賛同していない。	困難	やらない	なし	なし
1) 無関心期 ▼	関心がない。	興味の喚起。	関心がないので、行動の目的も価値も分からない。	やや難	ない	なし	なし
2) 奨励期 ▼	関心はあるが、価値は分からない。	目的への同意、理解の深化。	関心はあるが目的には不同意、行動の価値は分からない。	ある程度	ない >ある	なし >あり	低い
3) 期待期 ▼	価値はありそう。	価値の実感。	行動に価値はありそうだが、効果は実感していない。	ほぼ受容 ≈ある	ない ≈ある	なし ≈あり	中
4) 実感期 ▼	価値はあるが、意義は分からない。	意義の会得。	価値は実感したが、本人にとって取り組む意義は分からない。	ほぼ受容 やや不要 <ある	ない やや不要 <ある	あり	高い
5) 会得期	自分にとって行動する意義を会得した。	介入しない。	本人にとって行動する意義を感じ、主体的に取り組んでいる。	ほぼ不要	ある	あり	高い

したがって、介入制御としては、反対する理由を解消し、行動に価値がありそうと思わせ、期待期への移行を目指す介入を選出する。反対期は認知の段階としては、行動に挫折した状態でもあり、どのステージからも残念ながら反対期に移行してしまうことがある。

1) 無関心期は、行動変容ステージモデルと同様に関心を持っていないため、行動の目的や価値も知らないか、理解していない状態である。したがって、介入制御としては、興味を喚起し、関心を持ってもらう介入を選出することになる。反対期と同様に介入によって無理に行動させることは逆効果になる恐れがある。

2) 興味期は、関心は持ったものの、行動の目的や価値を十分に理解できていない、あるいは、同意には至っていない状態である。傾向としては、関心はあるので、意欲や意図がある時もあるが、概ね行動を起こしておらず、介入もあまり好まれない。この興味期は、行動を起こしたときに、行動変容によって得たい効果をすぐに得られないにも関わらず、行動コストはすぐに実感できてしまうため、マイナス面を認知しやすい時期もある。したがって、この時期も無理に行動を促すことよりも、行動変容に対する理解を深めて、行動コスト以上に価値がありそうだと認知させる介入を選出することが重要となる。

3) 期待期は、行動に価値がありそうと思っているが、実感としては認知していない状態である。傾向としては、価値があるとは感じているので、意欲も意図も中程度で、行動もそれなりに見られる。介入も受け入れられやすく、行動を促せば比較的行動させやすいが、積み重なる行動コストに目を向けてしまうと止めやすい時期もある。したがって、少しでも効果に気付かせるとともに、行動コストを低く感じさせる介入を選出することになる。どうしたら行動を継続できるのか自分に合った正解を見付けさせる介入も有効になると考える。

4) 実感期は、行動に価値があること、つまり、効果があることは体感しているが、自分にとって行動する意義については認知していない状態である。例えば、歩行習慣といった行動で、心身が健康になると実感していても、良い散歩コースや散歩仲間がなくなったら止めてしまうことがある。これは、対象者が行動の目的である健康の価値、つまり、自分にとって健康であることの意義を十分に認知していないかったことが原因と考えられる。この実感期は、行動の価値は実感しているため、行動を継続する傾向ではあるので、事例によってはここをゴールと定めても構わない。しかし、確固とした定着化を目指すのであれば、対象者にとって行動の目的を達成することの意義をしっかりと認知させるような介入を選出する必要がある。

5) 会得期は、自分にとって行動する意義を認知している状態である。そのため、取り組み方は主体的となっており、どんな環境や状況の変化が生じても、継続できるよう努力を惜しまないであろう。ここは理想的なゴールであり、介入も不要となる。もちろん、認知の状態はいつでも変わるので、ここで終わりとは限らない。また、行動変容を促してきた介入者側にも終わりではなく、そのままでよいケースは少なく常に改善が求められるであろう。ここまででは、介入される側の変容に着目してきたが、介入する側も、何をゴールとするのかを明確に定め、介入者自身も目標に向けて継続的に取り組むことが重要である。

以上のように、行動変容ステージモデルと違い、対象者を意図や行動の様子で分類するのではなく、対象者が行動やその目的をどのように捉えているかに基づいて分類し、介入方法や介入内容を選出することで、1回の介入で大きな変化を求めず、長期的に対象者が定着化するまで導くことを目指す介入制御を提案した。1回1回の介入においては、感情的に気分が乗らなかったり、行動する場所や時間がなかつたりして、行動を起こさないこともあるだろう。

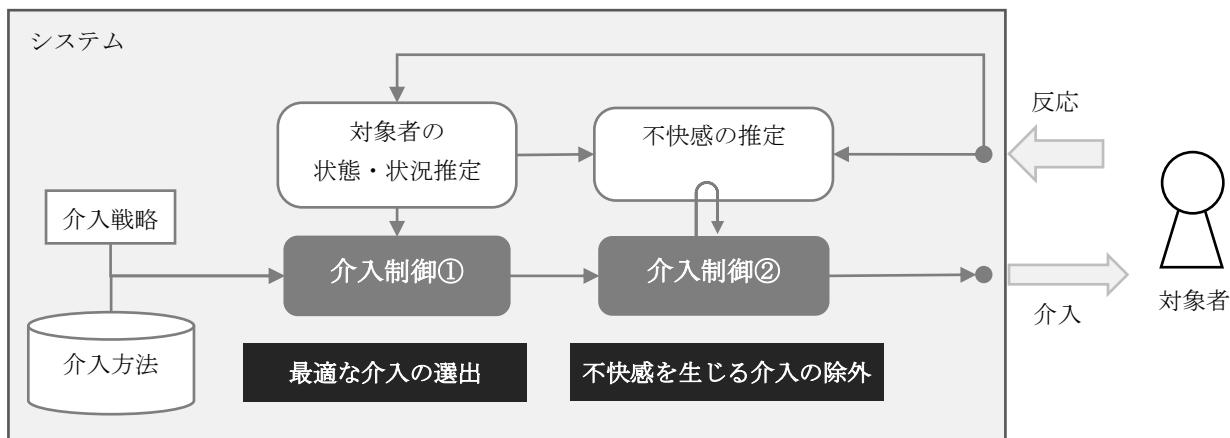


図 1 介入制御方式の構成図
Figure 1 system configuration diagram of intervention control method

そのようなときは、無理に行動を喚起するのではなく、少なくとも余計な介入は避け、例えば、改めて何のために行動するのか、どうして行動に嫌な気持ちを持つのか、といったことを考えさせる介入が必要である。他にも、意欲や感情面を考慮して、これまでの行動を賞賛したり、行動できないことに共感したりといったように、行動させることを主目的としない介入も、上手く活用すべきである。

4.2 不快度を考慮した介入制御

介入の不快度を考慮した介入制御を行うために、図1のような二段階の構成による介入制御方式を提案する。介入制御①においては、前述の通り、認知変容ステージモデルを用いて対象者の認知の状態に合わせ、次に目指すべき状態となるにはどのような介入が良いかという観点で選出する。これだけでも、重要性や価値観といった認知のギャップによる介入の不快感はかなり抑制される。

さらに、介入制御②では、選出された介入によって対象者に生じる不快感を推定して介入の可否を判断する。不快感は、介入頻度、介入への反応、介入方法や介入内容に起因する妨害度や嫌悪感、これまでの介入により与えてきた不快感を踏まえ推定する。特に介入内容に関しては、過去に実施した内容と同一、あるいは、類似している場合、介入が受容されていたとしても、それらの介入による不快感の蓄積が強烈なインパクトを与えることがあるため、最も考慮すべき点であると考える。このように介入の適切性や成功確率とは別に、介入が与える不快感によって、その介入を実行するか否かを決定し、時には何も介入しないという判断がなされる。したがって、いかに行動や認知の変容の成功確率が高い介入方法であっても、不快感を生じさせる介入は実行されなくなり、定着化の挫折、つまり、二度と介入を受け入れないような事態を回避することができる。逆に不快感を生じさせない介入であれば問題はなく、例え成功確率が低くても介入することは、無視されたとしても、その無視されたという反応を次の介入に活かすことができる。どのような介入に対して、どのような反応があったのか、例えば、メッセージを読んでももらえなかつたのか、読んだけど変化がなかつたのかによっても、対象者の状態や好む介入の傾向を見極めるヒントになるからである。

5. まとめ

ICTを活用して、行動変容を定着化させるため、挫折を回避して長期的に介入を継続するための介入制御方式を提案した。ひとつは、認知変容ステージモデルに基づく介入制御である。対象者が行動やその目的の重要性や効果をどのように捉えているかに応じ、段階的な変容を促し、定着化を目指す長期的な戦略である。これにより、単に行動を実行し続けるだけでなく、その価値を認識し、主体的な取り組みとして定着化させるところまで支援することを目指す。もうひとつは、継続的な介入を続けるために、不快感を考

慮した介入制御である。これにより、行動や介入自体を拒絶されることを回避し、長期的に対象者に寄り添った介入を継続することを目指す。

今後は、具体的な事例に適用し、提案方式によるシステムの有効性を検証していく。また、認知の状態だけでなく、気分といった感情面や自己効力感、モチベーションといった要因が長期的な介入に与える影響についても検討を進め、より効果的に統合していくことをを目指す。

参考文献

- [1] Michie, S., Richardson, M., Johnston, M. et al. "The Behavior Change Technique Taxonomy (v1) of 93 Hierarchically Clustered Techniques: Building an International Consensus for the Reporting of Behavior Change Interventions," Annals of Behavioral Medicine, vol. 46, pp. 81–95, 2013. <https://doi.org/10.1007/s12160-013-9486-6>.
- [2] 西山 勇毅, 大越 匠, 米澤 拓郎, 中澤 仁, 高沢 一紀, 徳田 英幸, “ライフログデータを用いたチームの行動変容促進,” 情報処理学会論文誌, Vol. 56, No. 1, pp. 349-361, 2015.
- [3] Ming-Chuan Chiu, Tsai-Chi Kuo, Hsin-Ting Liao, “Design for sustainable behavior strategies: Impact of persuasive technology on energy usage,” Journal of Cleaner Production, Vol. 248, 2020. <https://doi.org/10.1016/j.jclepro.2019.119214>.
- [4] Munir, F., Biddle, Stuart, Davies, MJ, Dunstan, D., Esliger, D., Gray, LJ, Jackson, BR, O'Connell, SE, Yates, T and Edwardson, CL , “Stand More AT Work (SMArT Work): using the behaviour change wheel to develop an intervention to reduce sitting time in the workplace,” BMC Public Health 18, Article No. 319, 2018. <https://doi.org/10.1186/s12889-018-5187-1>
- [5] 上岡 玲子, “デスクワーク中の良姿勢習慣化のための触覚提示インターフェースシステムの提案,” 電子情報通信学会技術研究報告, Vol. 120, No. 391, MVE2020-66, pp. 121-126, 2021.
- [6] 佐藤 妙, 青木 良輔, 小安 宗徳, 篠崎 一慶, 大島 直樹, 武川 直樹, 渡部 智樹, 犬童 拓也, “行動変容ステージモデルに着目したユーザ状態遷移の検討,” 電子情報通信学会技術研究報告, Vol. 119, No. 38, HCS2019-1, pp. 1-6, 2019.
- [7] 佐藤 妙, 青木 良輔, 小安 宗徳, 篠崎 一慶, 大島 直樹, 武川 直樹, 渡部 智樹, 犬童 拓也, “行動変容のための「認知的不協和」におけるメッセージ提示方法,” マルチメディア, 分散協調とモバイルシンポジウム 2019 論文集, pp. 670-675, 2019.
- [8] 森 拓也, 島川 博光, “行動変容をもたらすユーザに寄り添うヘルスケアシステムの提案,” 情報科学技術フォーラム講演論文集(FIT), Vol. 19, No. 4, pp. 205-206, 2020.
- [9] Takuya Mori, Fumiko Harada and Hiromitsu Shimakawa, “Effects of Nudge Reflecting Day-of-Week Pattern against Induction to Walk,” Advances in Social Sciences Research Journal, Vol. 8, No. 1, pp. 610-630, 2021. <https://doi.org/10.14738/assrj.81.9686>.
- [10] 湯浅 智也, 島川 博光, “忙しさと行動変容ステージを考慮した健康増進のための歩行量改善,” 情報科学技術フォーラム講演論文集(FIT), Vol. 20, No. 4, pp. 359-360, 2021.
- [11] 二田 悠史, 杉原 太郎, 佐藤 健治, 五福 明夫, “行動変容促進を目的とした諭しおよび警告メッセージの予備的検討,” 情報処理学会研究報告, Vol. 2019-AAC-10, No. 13, pp. 1-8. 2019.
- [12] 二田 悠史, 杉原 太郎, 佐藤 健治, 五福 明夫, “敵意を喚起しない行動変容促進を目的とした諭しおよび警告メッセージの検討,” 情報処理学会研究報告, Vol. 2021-HCI-192, No. 32, pp. 1-8, 2021.
- [13] 西條 涼平, 佐藤 妙, 永徳 真一郎, 渡辺 昌洋, “習慣化支援のための視線誘導に着目した情報表示方法の基礎検証,” 電子情報通信学会技術研究報告, Vol. 119, No. 477, LOIS2019-59, pp. 19-24, 2020.

行動多動性とプッシュ通知による行動変容容易性

田中宏昌 ^{†1}
株式会社 NTT ドコモ

實成優馬 ^{†2}
株式会社 NTT ドコモ

吉川 裕木子 ^{†3}
株式会社 NTT ドコモ

勝間田 優樹 ^{†4}
株式会社 NTT ドコモ

山田渉 ^{†5}
株式会社 NTT ドコモ

落合桂一 ^{†6}
株式会社 NTT ドコモ

概要

混雑緩和や地域活性化を目指した人の流動性向上などの様々な目的のため、人を特定の場所に誘導する方法の研究が行われている。その中でもスマートフォンのプッシュ通知を用いたアプローチはスマートフォンを持つユーザに対して広く実施できるため、他のアプローチに比べて比較簡単に実施することができる。その便利さの一方で、過剰なプッシュ通知はユーザにとってストレスを感じる原因となるだけでなく、プッシュ通知を出すアプリケーションをスマートフォン上から削除する行動につながってしまう [1]。この対応策として、ユーザの興味対象に沿った施設を推薦するプッシュ通知を出す方法が開発されているが、プッシュ通知によって行動が変容されにくい性質をもつユーザが存在する場合は、ユーザの興味に沿ったプッシュ通知を提供したとしてもユーザのストレスを増加させる。そこで本研究では、プッシュ通知によって誘導されやすい性質をもつ人と誘導されにくい特性を持つ人が存在することを示し、その性質を日々の行動ログから推定できる方法を提案する。具体的には、人々の日々の施設へのチェックインログから、どの程度多様な施設にチェックインしやすいかという行動多動性という指標を定義し、その行動多動性を特徴量として、プッシュ通知によって誘導されやすい性質を持つ人か、誘導されにくい性質を持つ人かを判定する機械学習モデル（分類器）を提案する。このモデルを用いることで、プッシュ通知を出す前に事前に目的地に誘導できるかを予測できるため、行動変容が期待できる人のみにプッシュ通知を出すことで、ユーザ体験の向上に寄与する。

1. はじめに

人々が地域内で回遊して賑わいを創出することは、その地域にとって、地域経済活性化や防犯を始めとした様々な利点がある。実際に日本においても、地域の回遊性向上や賑わいの創出を目的とした NTT ドコモによる実証実験（以下、この実験を「幕張新都心エリアでの実証実験」と呼ぶ）が行われている^{†1}。幕張新都心エリアでの実証実験では、人々の局所的な集中を防ぎつつも人を様々な場所に流動させるために、スマートフォンへのプッシュ通知を用いて、人々を様々な施設に分散誘導させることを目指した。

幕張新都心エリアでの実証実験のように、人を特定の場所に誘導するための研究には、都市や建物自体の設計を工夫して人が流れやすい場所を作るアプローチ [2] を始めとして様々なアプローチが存在する。このようなアプローチの中でも、プッシュ通知を用いたアプローチはスマートフォン利用者（以下、ユーザと呼ぶ）に対して比較的容易に実施できるという点で有効である。

プッシュ通知は比較的容易に実施できる一方で、プッシュ通知によって誘導されにくいユーザが存在する場合、そのようなユーザにとって通知は不快なものであり、プッシュ通知を可能にしているアプリケーションの削除につながる [1]。アプリケーションを通してサービスを提供する多くの企業にとって、アプリケーションの削除は顧客との接点を喪失することに繋がるため、プッシュ通知によるユーザ体験の低下は、企業にとってもユーザにとっても避けるべき事象である。実際に幕張新都心エリアでの実証実験においても、ユーザ体験向上のためにユーザが興味を持てるような施設の通知を行っていた。しかし、たとえ興味対象の施設を推薦していてもそこに行かないような、個人特性としてプッシュ通知によって誘導されにくい性質を持つユーザが存在する場合は、いくら推薦モデルを改善してもユーザ体験は低下する一方である。

そこで本研究では、プッシュ通知によって誘導されやすい性質を持つユーザと誘導されにくい性質を持つユーザが存

Analysis of Relationship between Hyperactivity of Behaviour and Ease of Induction

^{†1} Hiroaki Tanaka, NTT DOCOMO INC.

^{†2} Yuuma Jitsunari, NTT DOCOMO INC.

^{†3} Yukiko Yoshikawa, NTT DOCOMO INC.

^{†4} Yuki Katsumata, NTT DOCOMO INC.

^{†5} Wataru Yamada, NTT DOCOMO INC.

^{†6} Keiichi Ochiai, NTT DOCOMO INC.

^{*1} https://www.ntt-docomo.co.jp/binary/pdf/info/news_release/topics_220125_00.pdf

在する（この性質を、以下では行動変容容易性と呼び、誘導されやすいユーザを行動変容容易性が高く、誘導されにくいユーザを行動変容容易性が低いと呼ぶ）ことを示し、さらに行動変容容易性はユーザの日々の行動データから予測できることを示す。具体的には、幕張新都心エリアでの実証実験で取得したデータから、プッシュ通知によって施設を推薦した回数に対しての実際にチェックインした割合を個人ごとに算出し、チェックイン割合の分布が二極化することを確認する。そして、チェックイン割合の中央値によってユーザを2つの群に分割して、行動変容容易性が中央値よりも高い群のチェックイン割合分布と、行動変容容易性が中央値よりも低い群のチェックイン割合分布に差があることを統計的仮説検定によって示す。

さらに、行動変容容易性が低いユーザか高いユーザかをプッシュ通知を出す前に判定するために、ユーザのチェックイン行動に関する多動性を表す指標として「行動多動性」を提案し、行動多動性を特徴量として行動変容容易性の高低を目的変数とした機械学習モデル（分類器）を構築する。行動変容容易性は、プッシュ通知を出す2週間前までの1日あたりの平均チェックイン数によって定義され、ユーザが一日にいろいろな場所にチェックインするほど大きい値をとる。この定義方法は、多動性のある子どもが衝動的に行動を変容しやすい[3]という特徴から発想を得ており、様々な場所に他動的にチェックインするユーザほど、プッシュ通知によって行動を変容しやすいと考えられる。

構築した分類機の評価のため、幕張新都心エリアでの実証実験データを用いた実験を行った。その結果、行動変容容易性が高いユーザラベルをポジティブ、行動変容容易性が低いユーザラベルをネガティブとしたとき、構築した分類機は適合率77%、再現率88%でユーザの行動変容容易性を推定できることができた。まとめると、本研究の貢献は以下の2点である：

- プッシュ通知によって誘導されやすいユーザと誘導されにくいユーザが存在することをデータから示した
- ユーザの行動多動性という指標を定義し、行動多動性によって行動変容容易性は推定できることを示した

2. 関連研究

2.1. 特定のエリアへの人の誘導

人を特定の場所に誘導するアプローチには、街や建物の設計自体を工夫するアプローチ[4, 2]や人の目に入る標識などのデザインを工夫するアプローチ[5]、さらには音声を用いて人を誘導するもの[6]などの様々なアプローチが存在する。これらのアプローチには、対象エリアに存在するすべての人に対して介入を実施できるという利点があるが、

実施のしづらさという欠点も存在する。たとえば、街や建物の設計を工夫するアプローチは、エリアや建物の設計を既存のものから変更する必要がある。また、標識などのデザインを工夫するアプローチや音声を用いて人を誘導するアプローチにおいても、事前に標識を設置したり、音声を発生させるための機材を設置する必要がある。

これらのアプローチに対して、スマートフォンのプッシュ通知を用いて人を誘導するアプローチは、世帯保有率86.8%^{*2}を締めるスマートフォン所有者に対して事前のハードウェアの準備なしに実施できる。そのため、比較的容易に実施できるアプローチであるといえる。

2.2. スマートフォンにおけるプッシュ通知

スマートフォンにおけるプッシュ通知は、企業がスマートフォンユーザと積極的にコミュニケーションを取るための有効な手段であり[7]、様々な企業がプッシュ通知を利用している。一方で、過剰頻度のプッシュ通知はユーザにとって好ましいものではなく、アプリケーションのアンインストールにつながることが指摘されている[1]。そのため、ユーザに対してプッシュ通知をどのぐらいの頻度で送るのが良いかを決定する研究も行われている。たとえば、スマートニュース社は新規ユーザに対して自社のアプリケーションをアンインストールされる事なく継続的に利用してもらうためにはどの程度の頻度でプッシュ通知を送信すればよいかを明らかにするため、プッシュ通知のネガティブ影響を機械学習を用いて予測する研究を行っている^{*3}。

スマートニュース社の研究のように、個別のアプリケーションにおけるアプリケーション利用率を維持または向上するためのプッシュ通知に関する研究が行われている一方で、スマートフォン上のプッシュ通知が実在する場所への誘導にどの程度効果的なのかや、どのような人が実際に誘導されたのかという評価は行われていない。したがって本研究では、プッシュ通知によって推薦された場所に誘導されやすいかを示す行動変容容易性という概念を定義し、さらにユーザの日々の行動を表現する行動多動性とい概念を定義して、行動多動性によって行動変容容易性を推定できることを示す。

3. データセット

本研究では、幕張新都心エリアでの実証実験において、幕張新都心エリアの居住者1129人から事前同意を得て取得したデータを用いる（以下、データ取得対象者を被験者と呼

^{*2} 令和3年版度情報通信白書（総務省）<https://www.soumu.go.jp/johotsusintokei/whitepaper/ja/r03/html/nd111100.html> より

^{*3} <https://medium.com/smартnews-inc/optimizing-push-frequency-for-new-users-5db5e04d76df>

データ取得期間	2022年2月1日から2022年3月21日
被験者数	1129人
被験者あたりプッシュ通知送信数の平均 ± 標準偏差	4.97 ± 5.08
プッシュ通知誘導による被験者あたり施設チェックイン数の平均 ± 標準偏差	1.06 ± 2.18
被験者あたり施設チェックイン数の平均 ± 標準偏差	199.20 ± 499.11
プッシュ通知による誘導先施設数	399

表1 本研究で利用するデータセットの主要な統計値

Table 1 Statistics of Dataset

ぶ). 本データセットは2022年2月1日から2022年3月21日までの期間で収集されており、被験者が普段の生活の中でどのような施設にチェックインしているか、被験者に対してプッシュ通知を送信したタイムスタンプ、プッシュ通知による推薦施設が記録されている。データセットの主要な統計値は表1に示す。

4. 予備実験：行動変容容易性の存在

4.1. 実験目的と実験方法

本研究の有効性である、「行動変容容易性の低いユーザに対してプッシュ通知を減らすことで該当ユーザのストレスを軽減し、アプリのアンインストールを回避することができる」という主張が成立する前提として、行動変容容易性の低いユーザの存在がある。そこで本節では、行動変容容易性が高いユーザと低いユーザ、すなわちプッシュ通知によって誘導されやすいユーザと誘導されにくいユーザが存在することをデータから確認する。

データセットから、各被験者に対して

$$R_{bc} = \frac{\text{(実際に誘導先にチェックインした回数)}}{\text{(プッシュ通知を受け取った回数)}}$$

を算出し、ヒストグラムを確認する。 R_{bh} はプッシュ通知によって目的地へと誘導された割合を表しているため、 R_{bh} の値が1に近いほどプッシュ通知によって誘導されやすく、0に近いほど誘導されにくいことを表す。

4.2. 実験結果

データから算出した R_{bc} のヒストグラムを、カーネル密度推定した結果は図1のようになった。図1から分かるように、 R_{bc} の分布には $R_{bc} = 0$ と $R_{bc} = 1$ の2箇所にピークが存在しており、行動変容容易性が低い被験者群と高い被験者群が存在することが分かる。

行動変容容易性が低い群と高い群の間に統計的に有意な差があるかを確認するため、 R_{bc} がとりうる値の中央値である $R_{bc} = 0.5$ をしきい値として被験者を2つの群に分けて統計的仮説検定を実施した。具体的には、 $R_{bc} < 0.5$ の群

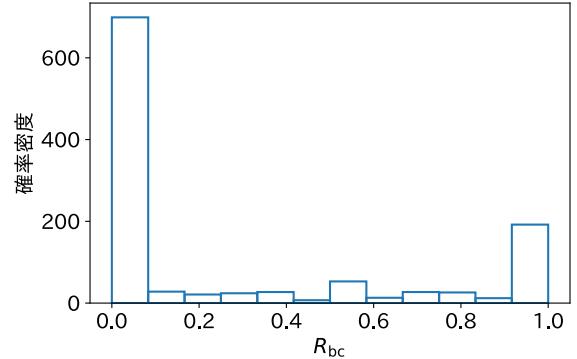


図1 R_{bc} のヒストグラムをカーネル密度推定した結果：0付近と1付近の2つの群にピークが立っていることが分かる。

Figure 1 Kernel Density Estimation of R_{bc} : We can find that the density achieve two peaks around 0 and 1.

を G_{low} , $R_{bc} \geq 0.5$ の群を G_{high} とし、 μ_{low}, μ_{high} をそれぞれ G_{low} と G_{high} の母平均として、帰無仮説 H_0 と対立仮説 H_1 をそれぞれ

$$H_0 : \mu_{low} \geq \mu_{high}, \quad H_1 : \mu_{low} < \mu_{high}$$

とした有意水準5%のWelchのt検定を行った。その結果、 p 値は 1.13×10^{-23} (t 統計量は 71.52) となり、帰無仮説 H_0 を棄却できた。すなわち、 G_{high} に属する被験者は G_{low} に属する被験者に比べて有意に行動変容容易性が高いといえる。以上のことから、プッシュ通知によって特定の施設に誘導されやすいユーザ群と、誘導されにくいユーザ群が存在することが示された。

5. 行動変容容易性判定モデル

5.1. 提案指標と提案モデル

4章で示したように行動変容容易性が低いユーザ群が存在するため、そのようなユーザ群を日々の行動ログから特定する行動変容容易性判定モデルを提案する。提案モデル

		Predicted Label	
		G_{high}	G_{low}
True Label	G_{high}	72	10
	G_{low}	81	182

表 2 提案モデルによるテストデータでの分類結果： G_{high} は行動変容容易性が高いクラスを表し、 G_{low} は行動変容容易性が低いクラスを表す。

Table 2 Result of classification by proposed model: The label G_{high} expresses the user class whose *human behavior change easiness* is high; on the other hand, G_{low} represents the user class whose *human behavior change easiness* is low.

は、プッシュ通知を受け取る日付 T における行動変容容易性が高いか低いかの 2 値ラベルを目的変数にとり、特徴ベクトルには以下で定義する行動多動性

$$x_d = \sum_{t=t_0}^{T-\Delta t} \frac{c_{td}}{\sum_{d=1}^D c_{td}}, \quad (1)$$

c_{td} = (日付 t における施設 d へのチェックイン回数)

を誘導先候補となる施設数分だけ並べたベクトル

$$\mathbf{x} = [x_1 \dots x_D] \quad (2)$$

を用いる。(1) は一日あたりに施設にチェックインする回数が増えるほど、すなわち他動的であるほど大きい値をとるため、行動変容容易性の判定対象となるユーザが普段からどの程度他動的な性質を持つのかを表現できる。

5.2. 評価方法と評価結果

提案モデルを評価するために、3 章で説明したデータセットを用いる。データセットに含まれる被験者 1129 人のうち、幕張新都心エリアでの実証実験期間において R_{bc} が 0.5 以上であった被験者 (G_{high} に属する被験者) は 323 人、 R_{bc} が 0.5 未満であった被験者 (G_{low} に属する被験者) は 806 含まれている。これらの被験者を、学習データ：開発用データ：テストデータに 565 : 281 : 283 に分けて、機械学習モデルの一つである LightGBM [8] の学習・ハイパーパラメータチューニング・評価を行う。ただし、分割されたそれぞれのデータセットには G_{high} と G_{low} それぞれのクラスに属する被験者が同程度の割合で含まれている。また、ハイパー パラメータのチューニングには Optuna [9] を用いる。

本研究では行動変容容易性が低いユーザに対するプッシュ通知を減らすことを目指しているため、行動変容容易性が低い人を誤って行動変容用意性が高いと予測している

$$\text{割合} = \frac{\text{(False Positive)}}{\text{(True Positive)} + \text{(False Positive)}}$$

が低いこと、すなわち適合率

$$\text{(正確度)} = 1 - \frac{\text{(False Positive)}}{\text{(True Positive)} + \text{(False Positive)}}$$

が高いことが望ましい。したがって、本実験の評価指標では正確度を最も重視し、次点で行動変容容易性が高い人に対してプッシュ通知を漏れなく送信することを重視して再現率で評価する。テストデータで評価した結果は表 2 のようになり、適合率は 0.77、再現率は 0.88 となった。すなわち、提案モデルは行動変容容易性が高いと分類した人のうち 77% の人は実際に行動変容用意性が高く、行動変容容易性が低い人の内 88% を実際に発見できていることが分かった。

6. おわりに

本研究では、人々を様々な施設に誘導するためのスマートフォンによるプッシュ通知を対象として、不要なプッシュ通知を無くしてユーザのストレスを減少させることを目的とし、プッシュ通知によってユーザの行動が変容しづらいユーザを特定する方法を提案した。具体的には、プッシュ通知による施設推薦に誘導されやすいユーザと誘導されにくいユーザが存在する、すなわち行動変容容易性が高いユーザと低いユーザが存在することを示し、行動変容容易性は行動多動性と呼ばれる日々の行動特徴によって推定できることを示した。行動多動性は、ユーザが日々の行動においてどの程度他動的に施設に移動を行っているかを表す指標であり、普段から色々な施設に移動する傾向がある人ほどプッシュ通知による施設推薦によって推薦施設に誘導されやすいと考えられる。

行動変容容易性を推定するために、行動多動性を特徴量とする機械学習モデルによる分類器を構築し、幕張新都心エリアでの実証実験において取得したデータセットを用いて分類機の学習及び評価を行った。その結果、行動変容容易性が高いユーザと低いユーザを適合率 77%、再現率 88% で発見できることを確認した。この結果から、行動変容用意性が高いと判断されたユーザのみに対してプッシュ通知を送った場合、すべてのユーザに対してプッシュ通知を送る場合に比べて、誤って行動変容用意性が低いユーザにプッシュ通知を送ってしまう割合を 23% 低減する事が可能になる。また、再現率も 88% であるため、行動変容容易性が高いユーザ群に対しても 88% は正しく行動変容容易性が高いと推定できている事がわかる。

評価実験の結果から、行動多動性は行動変容容易性を推定するために有効な指標の一つであると考えられる。その一方で、本稿では他の指標との比較を実施していないため、

考えうる様々な指標の中で行動多動性が行動変容容易性の推定に最も有効な特徴量ではない可能性がある。また、本研究の結果は幕張新都心エリアでの実証実験で取得したデータから導かれたものであるため、都市としての特性が著しく異なるようなエリア、たとえば公共交通機関が極端に発展していない山間部等には安易に一般化できないという制限がある。

参考文献

- [1] Atilla Wohllebe, Dirk-Siegfried Hübner, U. R. and Podruzsik, S.: Mobile apps in retail: Effect of push notification frequency on app user behavior, *Innovative Marketing*, Vol. 17, No. 2, pp. 102–111 (2021).
- [2] Vaez, S., Burke, M. and Alizadeh, T.: Urban form and wayfinding: Review of cognitive and spatial knowledge for individuals' navigation, *Australasian Transport Research Forum (ATRF)*, 38th, 2016, Melbourne, Victoria, Australia (2016).
- [3] Singh, A., Yeh, C. J., Verma, N. and Das, A. K.: Overview of Attention Deficit Hyperactivity Disorder in Young Children, *Health Psychology Research*, Vol. 3, p. 2115 (2015).
- [4] Raubal, M. and Egenhofer, M. J.: Comparing the complexity of wayfinding tasks in built environments, *Environment and Planning B: Planning and Design*, Vol. 25, No. 6, pp. 895–913 (1998).
- [5] Shi, Y., Zhang, Y., Wang, T., Li, C. and Yuan, S.: The Effects of Ambient Illumination, Color Combination, Sign Height, and Observation Angle on the Legibility of Wayfinding Signs in Metro Stations, *Sustainability*, Vol. 12, No. 10 (online), <https://www.mdpi.com/2071-1050/12/10/4133> (2020).
- [6] Milliman, R. E.: Using Background Music to Affect the Behavior of Supermarket Shoppers, *Journal of Marketing*, Vol. 46, No. 3, pp. 86–91 (1982).
- [7] Gavilan, D. and Avello, M.: Enabling smartphone push notifications: the effect of a framed opt-in request, *International Journal of Mobile Communications*, Vol. 21, No. 1, pp. 1–18 (2022).
- [8] Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q. and Liu, T.-Y.: LightGBM: A Highly Efficient Gradient Boosting Decision Tree, *Advances in Neural Information Processing Systems*, Vol. 30, Curran Associates, Inc. (2017).
- [9] Akiba, T., Sano, S., Yanase, T., Ohta, T. and Koyama, M.: Optuna: A Next-generation Hyperparameter Optimization Framework, *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, KDD '19, New York, NY, USA, Association for Computing Machinery, pp. 2623–2631 (2019).

案件割当における従業員の作業時間の均一化を図る割当手法

小林久也 ^{†1}富士通株式会社 ^{†1}尾形晋 ^{†2}富士通株式会社 ^{†2}鳥谷彰 ^{†3}富士通株式会社 ^{†3}

概要

少子高齢化に伴う労働力不足が顕在化しつつある中、多くの業務案件を適切に従業員へ割り振り、業務の効率化や生産性を向上したいとのニーズが高まっている。このようなニーズへ対応する方法の一つに、多数の業務案件を複数の従業員へ割り当て、分担するマッチングの仕組みがある。例えば、作業期日の遵守と従業員の作業負担の公平性の観点から、全案件の処理に要する時間は短く、従業員毎の担当案件処理に要する作業時間をなるべく均一にしたいようなケースでは、線形計画法を用いて割当ることが一般的である。しかし、従来の線形計画法では、割当する期間や対象となる業務案件内のみでの最適化を意図しているために、その中で従業員毎の作業時間にばらつきなどが生じた場合、その割当を繰り返すことで、作業時間のばらつきが積みあがって大きくなる問題があった。

これを解消するためには従業員毎の作業時間のばらつきを把握し、業務が追加されて割当を繰り返す際に、従業員毎の総作業時間が均一になるように調整を行うことが考えられる。

本稿では、これを実現するために、既存の従業員毎の作業時間のばらつきから、追加案件の割当により従業員毎の総作業時間が均一になる調整量を求め、線形計画問題における従業員毎の担当案件処理に要する作業時間の上限値の制約式に加えることで、全案件の処理時間を短くしつつ、かつ従業員毎の総作業時間を均一にする新たな割当手法を提案する。

1. はじめに

チームにおける各個人へ案件の割り振りを最適化し、工数等を最小化するマッチングサービスのニーズが高まっている。

例えば、池上(2005)は病院に属する複数の看護師に対し、日勤や夜勤などの勤務シフトを割り振る数理計画モデルを用いたマッチングサービスを報告している[1]。マッチング

する際には、各勤務シフトにおける看護師の必要人数の確保や、看護師による休みの希望や不規則な勤務シフトから生じる健康影響等の条件を考慮している。

また、マッチングサービスへのニーズ例として、黒坂(2007)は役所などの滞納整理業務について、職員の習熟度向上や困難な滞納整理案件を職員同士でカバーしながら、滞納整理業務の効率化を図るマネジメントの必要性を報告している[2]。この仕組みを実現するためには、滞納整理案件における未納や回収不能などの問題による案件の難易度、各職員の経験や能力から生じる習熟度を考慮する必要がある。

以下、滞納整理業務を題材に、マッチングサービスに必要な要件や課題について論じる。滞納整理業務について、職員の滞納整理業務における習熟度と滞納整理案件の難易度が既知のものとし、滞納整理案件の処理に要する時間は、担当職員の習熟度と案件の難易度より決まるものとする。このとき、職員毎の担当案件処理に要する作業時間をなるべく均一にするような職員への滞納整理案件の割当が、職員同士のカバーに繋がると考える。また、全ての滞納整理案件の処理に要する時間を短くすることが、滞納整理業務の効率化に繋がると推察する。しかし、人手でこのような職員への滞納整理案件の割当を行うのは、職員の習熟度や、案件の難易度の違いを考慮するため困難である。

本稿では、書類確認業務等で多数の案件を、複数の従業員に割当することで分担し、全案件の処理が完了した後、多数の追加案件を同じ従業員へ割当することを繰り返す業務を対象とする。このとき、多数の案件を複数の従業員へ割当する従来方式として線形計画問題を考える。前提として、案件には易しい、標準などの難易度、従業員には新任、中堅などの習熟度が付与されているものとし、一案件当たりの処理時間は、担当者の業務への習熟度と案件の難易度より決まるものと想定する。

以上の議論から、本稿での案件割当の目標として、作業期日の遵守の観点から全案件の処理に要する時間は短く、従業員間の作業負担の公平性の観点から従業員毎の総作業時間が均一になる仕組みの実現を目指す。この目標を実現するため、線形計画問題における目的関数を従業員毎の担当案件処理に要する作業時間の総和の最小化とする。そして、線形計画問題における制約条件として次の4点を設定する。

Manuscript Format for Kickoff Symposium of IPSJ Behavior Transformation by IoT (BTI)

^{†1} HISAYA KOAYASHI, Fujitsu Ltd.

^{†2} SUSUMU OGATA, Fujitsu Ltd.

^{†3} AKIRA KARASUDANI, Fujitsu Ltd.

1. 一案件につき、担当の従業員は一名とする。
2. 全ての従業員は、必ず一件以上の案件を担当する。
3. 従業員毎の担当案件処理に要する作業時間はなるべく案件割当担当者が指定する作業時間に収める。
4. 従業員毎の担当案件処理に要する作業時間をなるべく均一にする。

しかし、案件割当担当者が指定する作業時間や割当する案件内のみで最適化を行うため、必ずしも従業員毎の担当案件処理に要する作業時間を均一にすることはできず、ばらつきが生じる。この状況で繰り返し割当を行うと、従業員毎の総作業時間のばらつきは拡大し、従業員間の作業負担に大きな差が生じる。

本稿では、既存の従業員毎の作業時間のばらつきから、追加案件の割当により従業員毎の総作業時間が均一になる調整量を求め、制約式3に加えることで、全案件の処理時間を短くしつつ、かつ従業員毎の総作業時間を均一にする新たな割当手法を提案する。

2. 提案手法による従業員への案件割当

2.1. 従業員への案件割当における線形計画問題

はじめに、従来方式として線形計画問題による従業員への案件割当を説明する。1章の線形計画問題における前提として、従業員の習熟度を表す集合 $L_1 = \{1, 2, \dots, l\}$ ($l \in \mathbb{N}$)、案件の難易度を表す集合 $L_2 = \{1, 2, \dots, d\}$ ($d \in \mathbb{N}$) とし、習熟度が s ($s \in L_1$) の従業員が難易度 q ($q \in L_2$) の案件処理に要する時間を $e_{s,q}$ ($e_{s,q} > 0, e_{s,q} \in \mathbb{R}$) とする(表1)。

この前提をもとに、線形計画問題における案件構成と従業員構成を定める。従業員へ割当する案件数 n_t ($n_t \in \mathbb{N}$) とし、各案件の識別番号 $T = \{1, 2, \dots, n_t\}$ とする。また、難易度が q の案件について、案件数 n_{t_q} 、案件の識別番号の集合 T_q 、連番の集合 T'_q とする(表2)。同様に、従業員数 n_w ($n_w \in \mathbb{N}$) とし、従業員の識別番号 $W = \{1, 2, \dots, n_w\}$ とする。そして、習熟度 s の従業員について、従業員数を n_{w_s} 、識別番号の集合 W_q 、連番の集合を W'_q とする(表3)。このとき、識別番号が w ($w \in W$) の従業員に対し、識別番号が t ($t \in T$) の案件を割当することを $x_{w,t} = 1$ 、割当しないことを $x_{w,t} = 0$ で表す(表4)。最後に、案件割当担当者が指定する作業時間は、案件割当担当者が定める所要日数と案件処理における1日の作業時間の積 D とする。

これまでに定めた表1から表4における各変数を用いて、識別番号 t_q ($t_q \in T_q$) の案件を担当する従業員数 f_1 は(1)式のように得られる。

$$f_1 = \sum_{s \in L_1} \sum_{w'_s \in W'_s} x_{n_w(s-1)+t'_q, t_q} \quad \left(= \sum_{w \in W} x_{w, t_q}\right) \quad (1)$$

ただし、 $n_w(i) = \sum_{k=1}^i n_{w_k}$ とする。

表1 1件の案件調査に要する時間の見積値

Table 1 Estimated time for an employee with level s to finish a case with difficulty level q .

		L_2				
		1	...	q	...	d
L_1	1	$e_{1,1}$...	$e_{1,q}$...	$e_{1,d}$
	\vdots	\vdots	\ddots	\vdots	\ddots	\vdots
	s	$e_{s,1}$...	$e_{s,q}$...	$e_{s,d}$
	\vdots	\vdots	\ddots	\vdots	\ddots	\vdots
	l	$e_{l,1}$...	$e_{l,q}$...	$e_{l,d}$

表2 案件情報

Table 2 Cases list.

T'_q	T_q
1	$\sum_{k=1}^{q-1} n_{t_k} + 1$
2	$\sum_{k=1}^{q-1} n_{t_k} + 2$
\vdots	\vdots
n_{t_q}	$\sum_{k=1}^q n_{t_k}$

表3 従業員情報

Table 3 Employees list.

W'_s	W_s
1	$\sum_{k=1}^{s-1} n_{w_k} + 1$
2	$\sum_{k=1}^{s-1} n_{w_k} + 2$
\vdots	\vdots
n_{w_s}	$\sum_{k=1}^s n_{w_k}$

表4 従業員への案件割当テーブル

Table 4 Variables in the linear programming problem.

		T				
		1	...	t	...	n_t
W	1	$x_{1,1}$...	$x_{1,t}$...	x_{1,n_t}
	\vdots	\vdots	\ddots	\vdots	\ddots	\vdots
	w	$x_{w,1}$...	$x_{w,t}$...	x_{w,n_t}
	\vdots	\vdots	\ddots	\vdots	\ddots	\vdots
	n_w	$x_{n_w,1}$...	$x_{n_w,t}$...	x_{n_w,n_t}

そして、識別番号 w_s ($\in W_s$) の従業員へ割当する案件数 f_2 は(2)式のように得られる。

$$f_2 = \sum_{q \in L_2} \sum_{t'_q \in T'_q} x_{w_s, n_t(q-1)+t'_q} \quad \left(= \sum_{q \in T} x_{w_s, q}\right) \quad (2)$$

ただし、 $n_t(i) = \sum_{k=1}^i n_{t_k}$ とする。また、識別番号 w_s の従業員が担当案件処理に要する作業時間 $f_3(w_s)$ は(3)式のように得られる。

$$f_3(w_s) = \sum_{q \in L_2} e_{s,q} \sum_{t'_q \in T'_q} x_{w_s, n_t(q-1)+t'_q} \quad (3)$$

よって、従業員毎の担当案件処理に要する作業時間の総和 f_4 は(4)式のように得られる。

$$f_4 = \sum_{s \in L_1} \sum_{w_s \in W_s} \sum_{q \in L_2} e_{s,q} \sum_{t'_q \in T'_q} x_{w_s, n_t(q-1) + t'_q} \quad (4)$$

従って、従業員への案件の割当における線形計画問題は次のように定式化する。

$$\begin{aligned} \text{minimize : } & f_4 + \sum_{i \in W} a_i + \sum_{j=1}^2 \left(\sum_{s \in L_1} \sum_{i=1}^{n_{w_s}-1} b_{j,i} + \sum_{i=1}^{l-1} c_{j,i} \right) \\ \text{s.t. : } & \forall q \in L_2, \forall t_q \in T_q, \quad f_1 = 1 \end{aligned} \quad (5)$$

$$\forall s \in L_1, \forall w_s \in W_s, \quad f_2 \geq 1 \quad (6)$$

$$\forall s \in L_1, \forall w_s \in W_s, \forall i \in W, \quad f_3(w_s) - a_i \leq D \quad (7)$$

$$\forall s \in L_1, \forall w_s \in W_s, \forall i \in \{1, 2, \dots, n_{w_s} - 1\}, \quad f_3(w_s) \geq f_3(w_s + 1) - b_{1,i} \quad (8)$$

$$f_3(w_s) - b_{2,i} \leq f_3(w_s + 1) \quad (9)$$

$$\forall s \in L_1, \forall w_s \in W_s, \forall i \in \{1, 2, \dots, l-1\}, \quad f_3(w_s) \geq f_3(w_{s+1}) - c_{1,i} \quad (10)$$

$$f_3(w_s) - c_{2,i} \leq f_3(w_{s+1}) \quad (11)$$

ただし、 $a_i, b_{1,i}, b_{2,i}, c_{1,i}, c_{2,i}$ は非負の実数とし、 $b_{1,i}$ と $b_{2,i}$ 、そして $c_{1,i}$ と $c_{2,i}$ はなるべく等しい値とする。この線形計画問題から得られる最適解を用いて (3) 式から、識別番号 w の従業員の担当案件処理に要する作業時間 w_{h_w} ($w_{h_w} > 0$) が得られる。また、 w_{h_w} が D に対し、超過あるいは不足する時間 d_{h_w} は (12) 式のように得られる。

$$d_{h_w} = w_{h_w} - D \quad (12)$$

2.2. 提案手法による従業員の総作業時間を均一化する案件割当

提案手法による従業員への案件割当を説明する。提案手法におけるフローチャートは図 1 に記す。ここで、提案手法では既存の従業員毎の作業時間のばらつきを追加案件の割当により解消するため、過去における従業員の作業時間を w'_{h_w} 、案件割当担当者が指定した作業時間を D' とする。識別番号 w の従業員の過去における作業時間の超過あるいは不足時間 d'_{h_w} は、(12) 式を用いることで得られる。このとき、 $\forall w \in W, d'_{h_w} = 0$ の場合、過去における従業員毎の作業時間に超過あるいは不足時間がないため、過去における従業員毎の作業時間にばらつきがないと判断する。よって、既存の従業員毎の作業時間のばらつきを解消する必要はなく、前述の線形計画問題から得られる最適解が、従業員毎の担当案件処理に要する作業時間のばらつきをなるべく抑えた結果であることから処理を終了する。一方で、 $\exists w \in W, d'_{h_w} \neq 0$ の場合、過去における従業員毎の作業時間に超過あるいは不足時間があるため、過去における従業員毎の作業時間にばらつきがあると判断する。このばらつきを追加案件の割

当によって解消し、従業員毎の総作業時間を均一にするためには、前述の線形計画問題による最適解から得られる従業員毎の担当案件処理に要する作業時間を調整する必要がある。

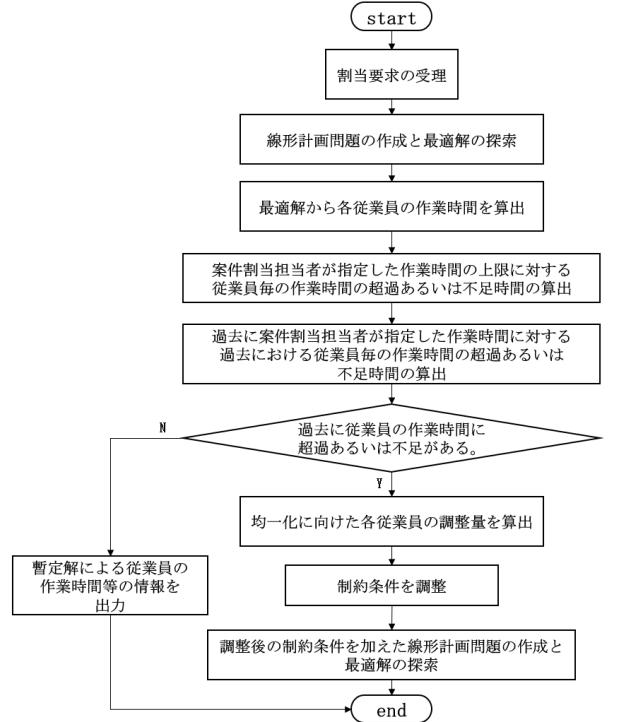


図 1 提案手法におけるフローチャート

Figure 1 Flowchart of the proposed method.

ここからは、過去における従業員毎の作業時間のばらつきを追加の案件の割当により従業員毎の総作業時間が均一になる調整量を求める。先ず、過去と前述の線形計画問題による最適解における全従業員の作業時間の超過あるいは不足時間 d_{h_w} と d'_{h_w} を平準化する。この平準化した値 eq は (13) 式のように得られる。

$$eq = \frac{1}{n_w} \sum_{w \in W} (d_{h_w} + d'_{h_w}) \quad (13)$$

従業員毎の総作業時間を均一にするために全従業員の作業時間の超過あるいは不足時間を平準化した値 eq を用いて、識別番号 $w \in W$ の従業員の担当案件処理に要する作業時間の調整量 ad_w は (14) 式のように得られる。

$$ad_w = eq - d'_{h_w}. \quad (14)$$

この調整量 ad_w を線形計画問題における制約式 (7) の右辺に加えると、(15) 式のように得られる。

$$\forall s \in L_1, \forall w_s \in W_s, \forall i \in W, f_3(w_s) - a_i \leq D + ad_w \quad (15)$$

前述の線形計画問題における制約式(7)を(15)に変えた新しい線形計画問題での最適解から得られる従業員毎の担当案件処理に要する作業時間により、既存の従業員の作業時間のばらつきを解消する。

3. 具体例を用いた従業員への案件割当

本章では、2章で定式化した従来方式と提案手法を用いて、具体例での、従業員毎の総作業時間を比較し、提案手法の有効性を示す。すなわち、従来方式による従業員への案件割当を繰り返すことで生じる職員毎の総作業時間のばらつきに対して、提案手法を用いることで総作業時間が均一になることを示す。(図2)。

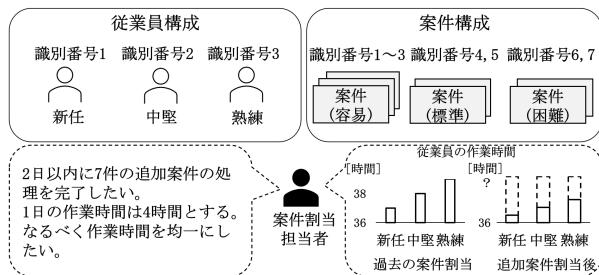


図2 従業員への案件割当(具体例)

Figure 2 Example of assigning seven cases to three employees.

はじめに、従来方式と提案手法で次の5点を共通の問題設定とする。

- ・従業員の習熟度は新任、中堅、熟練の3種類とする。
- ・従業員数は新任、中堅、熟練のそれぞれを1名とした合計3名とする。
- ・過去における従業員毎の作業時間にそれぞれの超過がある。
- ・案件の難易度は容易、標準、困難の3種類とする。
- ・案件数は容易3件、標準と困難を2件ずつとした合計7件とする。

上記の問題設定をもとに、2章における各変数を定義する。従業員の習熟度を新任、中堅、熟練の3種類、案件の難易度を容易、標準、困難の3種類とすれば、 $l = d = 3$ と定めることができ、 $L_1 = L_2 = \{1, 2, 3\}$ で与えられる。このとき、 $s(s \in L_1), q(q \in L_2)$ に対し、1案件当たりの処理時間 $e_{s,q}$ を表5とする。次に、具体例を簡単にするため、各難易度の案件数を容易3件、標準と困難をそれぞれ2件ずつとすれば、 $T = \{1, 2, \dots, 7\}, n_{t_1} = 3, n_{t_2} = n_{t_3} = 2$ で与えられる。また、各難易度の案件における識別番号の集合は $T_1 = \{1, 2, 3\}, T_2 = \{4, 5\}, T_3 = \{6, 7\}$ となり、連番

の集合は $T'_1 = \{1, 2, 3\}, T'_2 = T'_3 = \{1, 2\}$ となる。同様に従業員数を新任、中堅、熟練のそれぞれ1名ずつとすれば、 $W = \{1, 2, 3\}, n_{w_1} = n_{w_2} = n_{w_3} = 1$ となる。識別番号が $w(\in W)$ の従業員に対し、識別番号が $t(\in T)$ の案件を割当することを $x_{w,t} = 1$ 、割当しないことを $x_{w,t} = 0$ で表す。そして、案件割当担当者が定める所要日数を2日、1日における作業時間を4時間とすれば $D = 8$ となる。次に、既存の従業員毎の作業時間のばらつきとして、過去における識別番号 $w = 1$ の従業員の担当案件処理に要した作業時間 $w'_{h_1} = 37$ 時間、識別番号 $w = 2$ の従業員は $w'_{h_2} = 38$ 時間、識別番号 $w = 3$ の従業員は $w'_{h_3} = 39$ 時間とする。過去において案件割当担当者が指定した所要日数を12日間とし、案件処理における1日の作業時間を3時間とすると、案件割当担当者が指定した作業時間 D' は $D' = 36$ 時間となる。よって、過去における従業員の作業時間の超過は、識別番号 $w = 1$ の従業員は $d'_{h_1} = 1$ 時間、識別番号 $w = 2$ の従業員は $d'_{h_2} = 2$ 時間、識別番号 $w = 3$ の従業員は $d'_{h_3} = 3$ 時間として得られる。

表5 具体例における1件の案件調査に要する時間の見積

Table 5 Ex. Estimated time for an employee with proficiency level s to finish a case with difficulty level q .

		L_2		
		1	2	3
L_1	1	1	4	6
	2	2	3	5
	3	3	2	4

3.1. 従来方式による従業員への案件割当

2.1章における線形計画問題に対し、上記の具体例を適用した線形計画問題は次のように得られる。

$$\begin{aligned} & \text{minimize : } f_4 + \sum_{i \in W} a_i + \sum_{j=1}^{l-1} \sum_{i=1}^{l-1} c_{j,i} \\ & \text{s.t. : } \forall q \in L_2, \forall t_q \in T_q, \quad f_1 = 1 \end{aligned} \quad (16)$$

$$\forall s \in L_1, \forall w_s \in W_s, \quad f_2 \geq 1 \quad (17)$$

$$\forall s \in L_1, \forall w_s \in W_s, \forall i \in W, \quad f_3(w_s) - a_i \leq D \quad (18)$$

$$\forall s \in L_1, \forall w_s \in W_s, \forall i \in \{1, 2, \dots, l-1\}, \quad f_3(w_s) \geq f_3(w_{s+1}) - c_{1,i} \quad (19)$$

$$\begin{aligned} & f_3(w_s) - c_{2,i} \leq f_3(w_{s+1}) \end{aligned} \quad (20)$$

ただし、 $a_i, c_{1,i}, c_{2,i}$ は非負の実数とし、 $c_{1,i}$ と $c_{2,i}$ はなるべく等しい値とする。ここで、各習熟度において従業員は1名ずつであるため、(8)式と(9)式は消去される。

従って、上記の線形計画問題による従来方式から得られる最適解は表 6 となる。

表 6 従来手法による従業員への案件割当結果

Table 6 Assignment result by the conventional method.

		T						
		1	2	3	4	5	6	7
W	1	0	0	0	0	0	0	1
	2	0	0	0	1	1	0	0
	3	1	1	1	0	0	1	0

この表 6 を用いて (12) 式より、識別番号 $w = 1$ の従業員の担当案件処理に要する作業時間 $w_{h_1} = 7$ 時間、識別番号 $w = 2$ の従業員では $w_{h_2} = 10$ 時間、そして識別番号 $w = 3$ の従業員では $w_{h_3} = 11$ 時間となることがわかる。従って、表 6 による従業員の作業時間の超過あるいは不足時間 d_{h_w} は、識別番号 $w = 1$ の従業員では $d_{h_1} = -1$ 時間、識別番号 $w = 2$ では $d_{h_2} = 2$ 時間、識別番号 $w = 3$ では $d_{h_3} = 3$ 時間として得られる。これらを踏まえ、表 6 の結果による従業員毎の総作業時間は、識別番号 $w = 1$ の従業員は $w_{h_1} + w'_{h_1} = 44$ 時間、識別番号 $w = 2$ の従業員は $w_{h_2} + w'_{h_2} = 48$ 時間、識別番号 $w = 3$ の従業員は $w_{h_3} + w'_{h_3} = 50$ 時間のように得られる。

3.2. 提案手法による従業員への案件割当

はじめに、提案手法における従業員毎の総作業時間を均一化するための調整量を算出する。(13) 式より、全従業員の作業時間の超過あるいは不足時間を平準化した値 eq は $eq = \sum_{w=1}^3 (d_{h_w} + d'_{h_w}) / 3 = 3.3$ 時間として得られる。よって、式 (14) より、従業員毎の総作業時間を均一にするための調整量は識別番号 $w = 1$ の従業員は $ad_1 = 3.3 - 1 = 2.3$ 時間、識別番号 $w = 2$ の従業員は $ad_2 = 3.3 - 2 = 1.3$ 時間、識別番号 $w = 3$ の従業員は $ad_3 = 3.3 - 3 = 0.3$ 時間として得られる。この調整量 ad_1 から ad_3 を制約式 (18) に加えた新たな線形計画問題における最適解は表 7 として得られる。

表 7 提案手法による従業員への案件割当テーブル結果

Table 7 Assignment result by the proposed method.

		T						
		1	2	3	4	5	6	7
W	1	0	1	0	0	0	1	0
	2	0	0	0	1	1	0	0
	3	1	0	1	0	0	0	1

表 7 から従業員毎の担当案件処理に要する作業時間は、識別番号 $w = 1$ の従業員は $w_{h_1} = 11$ 時間、識別番号 $w = 2$

の従業員は $w_{h_2} = 10$ 時間、識別番号 $w = 3$ は $w_{h_3} = 9$ 時間として更新される。これら $\forall w \in \{1, 2, 3\}, w_{h_w}$ に対し、過去における従業員毎の担当案件処理に要した作業時間 $\forall w \in \{1, 2, 3\}, w'_{h_w}$ を加算すると、従業員毎の総作業時間 $\forall w \in \{1, 2, 3\}, w_{h_w} + w'_{h_w} = 48$ 時間のように得られる。

3.3. 結果と考察

3.1 節と 3.2 節より、従来方式と提案手法による従業員毎の総作業時間を比較した結果は図 3 となる。

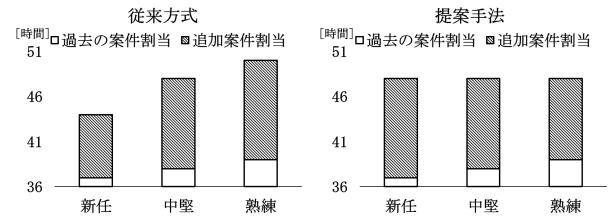


図 3 従来方式と提案手法による従業員の作業時間の比較

Figure 3 Comparison of total working hours for three employees by the conventional and proposed method.

過去における従業員毎の作業時間は新任の従業員が 37 時間、中堅は 38 時間、熟練は 39 時間であることから、新任と熟練の従業員毎の作業時間のばらつきは 2 時間であった。従来方式による追加案件の割当を行うと、従業員毎の総作業時間は新任の従業員で 44 時間、中堅は 48 時間、熟練は 50 時間の結果となった。従って、新任と熟練の従業員の総作業時間のばらつきは、2 時間から 6 時間へと拡大した。

一方で、提案手法による追加案件の割当を行うと、過去における新任と熟練の従業員の作業時間のばらつきを解消し、従業員毎の総作業時間が 48 時間として均一になる結果が得られた。

従来方式を用いて本具体例と類似の案件構成を従業員へ割当する場合、新任と熟練の従業員の総作業時間のばらつきは 4 時間ずつ拡大してしまう。この総作業時間のばらつきの調整を追加案件の割当時に行なうことは、案件割当担当者にとってかなりの負担となり得る。また、定期的に従業員毎の総作業時間を調整するため、総作業時間が少ない従業員に対し多くの案件を割当してしまうと、この従業員の負担となってしまう。

これに対し、提案手法を用いて本具体例と類似の案件構成を従業員へ割当する場合、既存の従業員毎の作業時間のばらつきから、追加案件の割当により従業員毎の総作業時間が均一になる調整量を使い、従業員毎の総作業時間の均一化を自動的に行なうため、案件割当担当者の負担を軽減する。そして、毎回の従業員への案件割当時に従業員毎の総作業時間が均一になる調整を行なうため、従業員に負担の大きな割当にはならないと考えられる。

4. おわりに

従業員への案件割当において、従業員の習熟度と案件の難易度による作業時間の差異を考慮し、作業期日の遵守と従業員作業負担の公平性の観点から、全案件の処理に要する時間は短時間にしつつ、従業員毎の総作業時間をなるべく均一にする新たな割当手法を提案した。さらに、小規模な具体例を用いて、従来方式と提案手法による従業員への案件割当結果による従業員毎の総作業時間のばらつき具合の比較を行い、提案手法において従業員毎の総作業時間が均一になることを確認し、本手法の有効性を確認した。今後、実業務への適用において、本手法で必ずしも従業員毎の総作業時間が均一にならないケースが生じた場合、現場の案件割当担当者による微調整の方法を確認し、提案手法を取り入れたい。

参考文献

- [1] 池上敦子 (2005). ナース・スケジューリング -調査・モデル化・アルゴリズム-, 統計数理, 第 53 卷 第 2 号, p.231-259
- [2] 黒坂昭一 (2007). 滞納整理における処理促進のための一考察 -リスク管理及び滞納事案の類型別アプローチ手法を中心に-, 税大論叢, 53 号

製造業における技術革新と行動変容

～アルミ圧延加工事業を例に～

浅田勝義^{†1}
株式会社 UACJ

磯田祐世^{†2}
株式会社 UACJ

山本佑樹^{†3}
株式会社 UACJ

1. はじめに

世界の産業は過去に三度の産業革命を経て、大きく姿を変えてきた。近年は、大量データの活用を主役とした第四次産業革命の時代に入ったと言われている。株式会社UACJ(以下、当社)においては主力のアルミ圧延工程において、1980年代から進んだコンピュータによるファクトリーオートメーション(Factory automation, FA)導入により、圧延加工における従業員のマニュアル機械設定および操作の機会が減少し、人による機械操作は熟練が必要な、例えば圧延頻度の少ない材料の通板などの場面に限られるようになった。更に近年ではIoT導入により、前述の様な熟練者の技術に頼らざるを得ない場面についても、一部自動化が可能となっている。上記のように、生産技術の発展に伴い、製品の生産性・品質は飛躍的に向上し、量産現場における従業員の業務の大部分はマニュアル機械設定・操作からデスクワークへと変容した。

一方、当社における技術発展(FA・IoT)の経緯を振り返ると、品質の良いものを大量に効率的に作るという観点では一定の成功を収めてきたが、従業員一人一人の幸福度という観点では必ずしもそうではない。また、エンゲージメントの低下、離職者増加等の問題は日増しに深刻となっている。IoT・デジタル技術は、生産性・品質を追求のみならず、従業員一人一人の幸福度向上への応用も期待される。

本報告では、当社におけるFA・IoT導入による製造現場の変容の具体的な事例を紹介するとともに、IoTおよびデジタル技術活用により、製造業が今後目指すべき働き方および企業の組織形態を対照しながら組織の在り方について分析・議論する。

2. 技術と働き方

当社ではFA(1980年代～)、IoT導入(2010年代～)によって、労働生産性向上を高める活動をしてきた。特に図1に示す熱間仕上圧延機は機械的性能や表面品質の作りこみに最も重要な工程であり優先的に技術導入を進めてきた。

2.1 当社のファクトリーオートメーション(FA)

FAとは一般に工場の自動化を示すが、ドメインにより自動化対象は異なる。当社の圧延加工部門では圧延機の制御を自動化対象としている。圧延とは回転している二本の圧延ロールの間に材料である金属を挿入し、所定の厚みに成形する加工法のことである。

高温に加熱されたスラブ(およそ厚さ 500 mm×幅 1500～2000mm×長さ 10000 mm程度の直方体、図2)は熱間粗圧延機にて往復圧延され、板厚 30 mm程度まで伸ばされる。その後、複数のスタンドが並んだ状態となっている熱間仕上圧延機に搬送される。熱間仕上圧延機では 1 スタンドあたり最大 50%程度の圧下率で出厚 2～10 mm程度まで圧延され、コイル状に巻き取られる[1]。代表的な製品である缶材の場合、顧客毎の要求に合わせて機械的性質、表面品質の作りこみを行う為、冷間圧延工程に移行し、製品厚 0.3mm 程度まで圧延されたコイル状の製品として出荷される。その後、当社の顧客である製缶メーカーで缶として成形され、飲料メーカーで中身を充填後、缶飲料として市販されている(図3)。

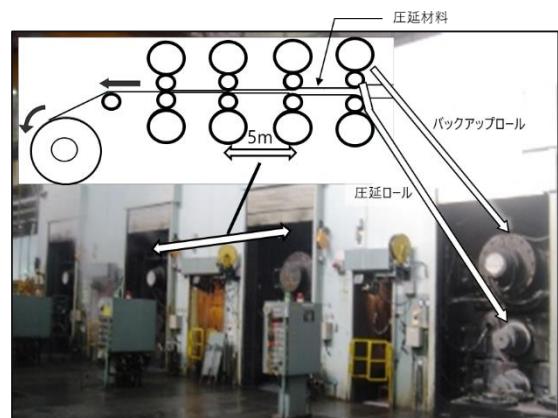


図1 热間仕上圧延機の概観
Figure1 Overview of hot finishing mills

Technological innovation and behavioral change in the manufacturing industry
～Taking the example of the aluminum rolling and processing business～
†1 KATSUYOSHI ASADA, UACJ Corporation
†2 YUSEI ISODA, UACJ Corporation
†3 YUKI YAMAMOTO, UACJ Corporation

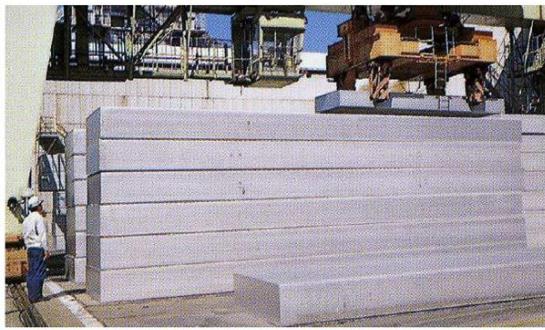


図2 スラブの例

Figure2 Examples of slabs



図3 圧延コイルと飲料缶

Figure3 Rolling coils and beverage cans

一方、圧延機で各製品の要求特性に応じて、板厚や温度を高精度に制御する為には、①圧延ロールの間隙設定(ロールギャップ設定)、②速度設定、③ロール冷却・潤滑スプレーのタイミング・圧力などの多くの条件設定が必要となる。

FA導入前、圧延機の多数の条件設定は、マニュアルでの打ち込み作業となっていた。そのため指示書や標準書にある多くの設定条件パターンを人が参照しながら入力するには限界があった。その結果、粗い条件設定となり、高精度制御は実現できていなかった。1980年代からコンピュータによる設定条件のプリセットが可能となり、品質、生産性が向

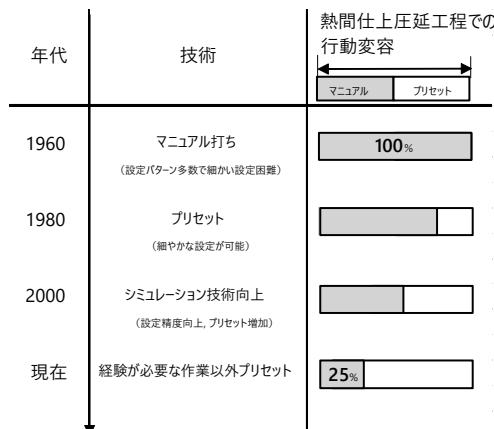


図4 技術と行動変容

Figure 4 Technology and behavior transformation

上すると共に、マニュアルでの打ち込み作業は大幅に減少した。その後、シミュレーション技術の向上により、前述の様な熟練者のノウハウに頼らざるを得ないケースを除いてマニュアル設定作業の多くは無くなっている(図4)。

2.2 IoT 導入について

近年の IoT 導入により、製造における確認・点検メンテナンスに必要なデータがリモートで収集されるようになった。その結果、以前は1日の就業時間の中で時間を要していた確認・点検などの設備メンテナンスの為の現地確認作業が減少し、会議/デスクワークの比率が増加した。代表的な技術者、の典型的な一日のタイムチャートを分析した結果、一日の行動は図5の様に変容している。現在、当社では IoT 導入を更に進めており、上記の行動変容は今後、更に進んでいくと予想される。

一方で、図6に示すように FA・IoT 導入後、満足度について従業員をランダムに選定しヒアリング(12名)した結果、大部分の人の仕事に対する満足度は変化していない結果となっており、IoT 導入で製造における労働生産性は上がっているもののそこに携わる従業員の総合的な満足度という観点ではまだ課題が残る。

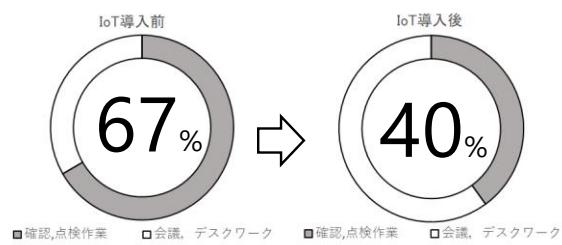


図5 IoT導入による行動変容
Figure 5 Behavioral changes due to IoT implementation

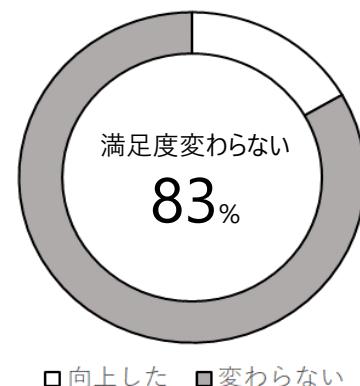


図6 FA, IoT導入後の仕事への満足度(n=12)
Figure 6 Satisfaction with work after introducing FA and IoT (n=12)

3. 製造業と組織形態

一般に規模の大きい企業ではすべての従業員に対してきめ細やかなマネジメントを行うことが難しくなり、上位者の管理が管理対象組織全体に行き届かなくなることも増えてくる。その結果、組織のどこで何が起きているのか、変化を把握できなくなり、課題やトラブルが生じてもすぐに気づけなくなるケースが増加する。従って組織の規模が大きくなるとともに、組織形態として、トップダウンで上長から下位へと指示や命令が下り、経営者以外は他者にマネジメントをされている状態をつくり出せるピラミッド型の組織形態を採用する企業が多い。

当社のような製造業では、課題やトラブルにすぐに対応できなければ、大量の品質不良や大きな設備トラブルに繋がりビジネスに影響する。その為当社もやはりピラミッド型組織となっている。またピラミッド型組織は指揮命令系統が1方向となるため、明確かつシンプルで、混乱、変化が起きにくいという点がメリットとして挙げられる。その為、当社の様にモノを大量生産する企業にとって適切であり、必然性を持った組織形態であるとも言える。

しかし、指揮命令系統が1方向のみでシンプルである反面で例外対応業務の規則を属人化してしまっていることが問題の一つになっている（業務の属人化とは、ある特定の社員のみが担当し、業務関連情報が担当者にしかわからない状態になることである）。実際、当社の様々な部門へヒアリングすると、真っ先に業務の属人化が挙げられるケースが多い。属人化の弊害として、(1) 業務の標準がないため業務プロセスや業務品質の改善が困難、(2) 担当者が退職等でいなくなったりした場合、重要な情報が失われ引継ぎも困難、(3) ノウハウが共用されず全体レベルが向上しない等が挙げられる。また、ピラミッド型組織は、責任の所在が明確で、下位の各担当が自発的な起こす前には、責任者である上長の承認が必要となる。上長へは多くの課題が集中するため、判断に時間を要するなど、トラブルリスクへの不安感から、上司が担当者へ非常に細かい説明を何度も要求する場合もある。その為、各担当の自発的な意思決定や行動は抑制されてしまい、結果として、自ら考えて動かない自律性の低い人材、つまり「指示待ち人間」が増えるという問題が増加している。統率されたマネジメントの中で、責任者からの指示で動くことを要求される中では、一人ひとり個性を持った人

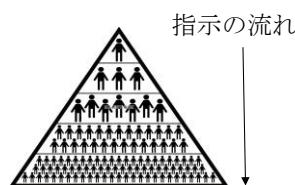


図7 ピラミッド型組織
Figure 7 Pyramidal organization

間の良さ/裁量が發揮されない部分もある。これらのこととは、FA・IoT導入後も、仕事への満足度が向上しないことと密接に関わっていると考えられる。

4. Well-being と組織形態論

ラファエル・A・カルヴォ、ドリアン・ピーターズ[2]によれば、Well-beingには大きく医学的、快楽心理的、そして人の潜在能力の発揮にかかるエウディモニア(ギリシャ語で幸福)の観点があるとされ、総じて「人生に意義を見出し、自分の潜在能力を最大限に発揮している状態」と定義されている。さらに、Well-beingを人間の能力発揮と捉える定義は、人口減少や働き手のニーズの多様化を背景として、厚生労働省が示す方向性「投資やイノベーションによる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作ること」[3]とも一致している。このように、働く中で Well-beingをどのように実現していくかということは社会的課題と言われる。

製造業は顧客の要望に基づくきめ細かい対応が必要で、典型的な労働集約的な産業である。製品を作る従業員の処理能力は限られており、個人個人が処理能力ギリギリの負荷を負う形で顧客要求が満たされていた。結果として従業員は重労働で精神的に負担が大きい仕事となっていた。その状況において、FA・IoT導入は、一人で対応できる作業の数を増やす、労働に伴う負担を減らすなど、労働生産性を上げる目的で導入してきた。その結果、「生産性」的な効率化には成功した。しかし一方で従業員は仕事において十分な満足感を得られているとは言えず、いまのところ Well-beingが達成されている状況とは言えない。

従業員が仕事を含めた Well-beingな状況となるためには仕事(企業)における組織形態を考えることは重要である。組織形態に関して、近年、生産性/ビジネス/Well-beingを両立し得る組織形態としてティール組織が提唱されている[4]。その中では図8に示すように組織モデルをその進化過程ごとに色で区別した5つの組織モデルがあり[5]、現状当社の組



図8 5つの組織モデル
Figure 8 Five organization models

組織形態はアンバー～オレンジの組織に近い。アンバー～オレンジ組織はピラミッド型組織であるが、ピラミッド型組織のトップダウン階層構造の中ではトップの意思は絶対となりその為下位からの反論が許容されにくい。結果として上司とのコミュニケーション不足に陥り、上司に対して不満を抱えた状態で、例外対応として業務を属人化し、自律的に働くことをあきらめて指示待ち人間になるケースが認められる。このことは、人間関係の問題として表面化することが多い。事実、一般的にも社員がモチベーションを低下させる因子として職場の人間関係が好ましくないことはよく挙げられる[6]。すなわちピラミッド組織は人が離れていく組織形態と考えられる。

対してティール組織は組織内に権力を集中させたリーダーは存在せず、現場においてメンバーが必要に応じて意思決定をおこなうことが特徴である。ティール組織の提唱者であるラルフ氏は、表1に示すような3つの重要な要素を唱えている[4]。これら3つの要素の共通項はコミュニケーションであり、フラットで活発なコミュニケーションをとれる環境が確保された組織であることで、人間の能力は最大限に發揮され、Well-beingの実現に繋がる[7]。ティール組織は表2に示すようにピラミッド型の問題を解決できる組織と考えられ、事実、製造業の枠組みにおいてもティール組織を採用して成功しているフランスの自動車部品製造メーカー(FAVI、従業員500名)の例も報告されている[4]。

ただし、暗黙知の多い、大規模製造業で従来機能でのアウトプットを劣化させることなくティール組織を成功させるにはセルフマネジメントにおける「助言プロセス」と呼ばれ

表1 ティール組織の3つの柱
Table 1 Three Pillars of Teal Organizations

柱	内容
セルフマネジメント	メンバーには責任に応じて大きな裁量が与えられるメンバーが意思決定する権利を持っている状態となるため、会社が保有している情報は、基本的にメンバーに対して開示され、適切な意思決定をするための助言を得られる仕組みも必要となる。
ホールネス	ティール組織におけるメンバーは、それぞれが対等(フラットな関係)であり、組織はメンバー全員のものであると考える「組織の社会的使命を果たすために自分ができること」と「自分自身の目標達成のための行動」が一致した組織である。
エボリューションナリ・バーバス	メンバーが組織の目的をはっきりと理解し、組織の使命を果たすための自律的な行動ができなくてはならない。

表2 ピラミッド組織、ティール組織の比較
Table 2 Comparison of Pyramid organization Teal organization

	評価軸			
	マネジメント	業務の属人化	社員の自律性	社員のWell-being
ピラミッド型組織	Good	Bad	Bad	Bad
ティール型組織	Good	Good	Good	Good

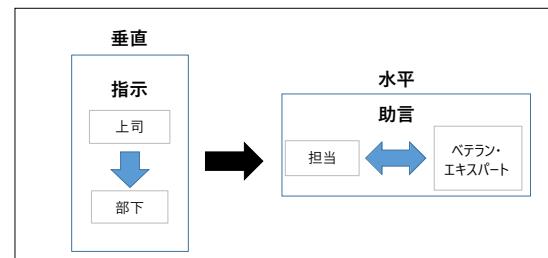


図9 指示から助言へ（垂直から水平へ）

Figure 9 Transformation from instructions to advices

る考え方がキーになる。「助言プロセス」とは適切な意思決定を助ける「助言を得られる仕組み」であり、従来のピラミッド組織において行われていた上から下への業務指示を代替するものである(図9)。このプロセスにより組織で動く人間はアクションを起こす前に適切な情報を得ることができ、従来の上司→部下でやり取りされていた情報伝達の内容が損なわれることはない。

5. 当社の向かうべき姿

筆者らは、旧来の組織にティール組織の仕組を組み込むことで高付加価値創出と Well-being を実現できる組織形態に進化(ある種の行動変容)していくと考えている。その変容において課題となるのはピラミッド型で実施していた上意下達業務フローをどのように仕事の質を損なわずに変えていくかである。4項で述べた「助言を得られる仕組み」を実現する手段として、図10に示すようなIoT/デジタル技術による知識ベースを包含した「集合知プラットフォーム」が必要となってくる[8]。従業員が集合知プラットフォームに接続された組織では、現在の上位者の指示通りに動く個の尊重の少ない行動から、個が自律して助言し合うような行動へ変容する。またそこでは個人の気づき、工夫を集合知プラットフォームにアップロードすることで知識ベースを強化したり、AIによる社内ジョブマッチングなども可能である。これにより、従業員は常に各人の性格、価値観、体調に適し

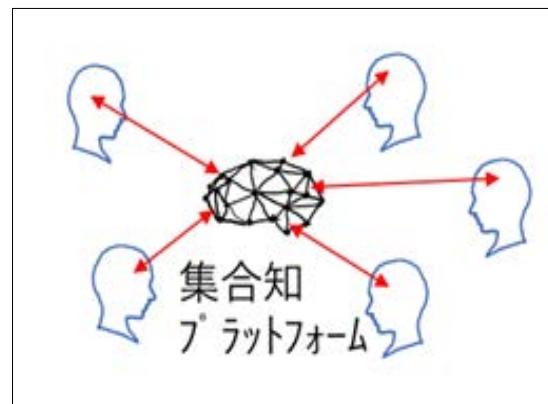


図10 集合知プラットフォーム

Figure 10 Collective knowledge platform

た仕事を得ることが可能となり、結果として全体(Well-being, 会社, 社会)の利益をも最大化していく。現在、筆者らはこのような将来向かうべき姿の実現に向けて技術開発に取り組んでいる。

参考文献

- [1] 藤森崇起：アルミニウム板製品の製造プロセスと品質管理, 第25回 表面探傷シンポジウム講演論文集, pp.1-8 (2022).
- [2] ラファエル・A・カルヴォ, ドリアン・ピーターズ.
Positive Computing : Technology for Well-being and Human Potential :
MIT press
(渡邊淳司, ドミニク・チェン監訳『ウェルビーイングの設計論』, BNN, 2017) (2014)
- [3] 厚生労働省. 『働き方改革』の実現に向けて,
<https://www.mhlw.go.jp/stf/seisakunitsuite/bunya/0000148322.html>
(2019年7月25日現在)
- [4] フレデリック・ラルー. Reinventing Organizations. 英治出版
(2014)
- [5] ティール組織とは? それぞれの組織モデルをわかりやすく解説
<https://hr-trend-lab.mynavi.jp/column/organizational-development/1509/> (2023年2月17日現在)
- [6] 厚生労働省. 平成30年若年者雇用実態調査の概況
- [7] テレワーク導入による Well-being の向上—個人と組織の Well-being—デジタルプラクティス Vol.10 No.4.(Oct. 2019)
<https://www.ipsj.or.jp/dp/contents/publication/40/S1004-S01.html> (2023年2月17日現在)
- [8] ティール組織の内部コミュニケーションはどうなっているか(5
分読書#46)
<https://hr-trend-lab.mynavi.jp/column/organizational-development/1509/> (2023年2月17日現在)

IoT 行動変容に向けた向社会的行動の他者推奨意向に関する検討

山崎悠大 ^{†1}
KDDI 総合研究所

曹蓮 ^{†2}
KDDI 総合研究所

柏本幸俊 ^{†3}
KDDI 総合研究所

上坂大輔 ^{†4}
KDDI 総合研究所

概要

近年、災害の激甚化やコロナ禍を受け、人と人との助け合いがより重要視されてきている。相手のことを思いやって、または誰かのために行う行動は向社会的行動と呼ばれ、Inducements(誘引) や Persuasion(説得) に代表される介入を行うことで向社会的行動を促進する行動変容研究が数多くなされている。しかしながら、既存研究の課題として直接介入されていない人に対する行動変容の実現が困難であることが挙げられる。限られた対象者への介入でより多くの人や社会に影響を及ぼすためには、介入対象者の向社会的行動を促進するだけでなく、他者への推奨を促す行動変容技術が必要である。そこで本研究では、行動変容のステージが説得受容（介入によって説得を受け入れ、対象行動を行う意欲がある状態）と他者推奨意向（人に対象行動を推奨する状態）の2段階で進行すると仮説を立て、介入対象者が他者推奨意向に至り、行動変容が他者に伝搬する要因が何であるかを明らかにする。向社会的行動の1つである「パトロールラン」を対象とし説得的介入を含む206名を対象にしたWEBアンケート調査結果の媒介分析(Bootstrap法)により、説得的介入を受けた実験群で2つの変数の媒介効果が有意となった（向社会性認知: $z = 3.43$, $p < .01$, 自発性: $z = 2.08$, $p < .05$ ）。よって説得受容から他者推奨意向に至る過程で、自発性と向社会的行動の向社会性認知に媒介効果があることを明らかにした。本プロセスを情報システムによる向社会的行動変容支援へ活用することで、その効果が広く波及することが期待される。

1. はじめに

一般に相手や誰かを思いやって行う利他行動や援助行動は向社会的行動と呼ばれる。向社会的行動は社会に利益をもたらすことから、その促進に関する様々な研究が行われている。Capraraらは中学生を対象に思春期の向社会的行動を促進する介入プログラムを実施し、身体および言葉に

より攻撃の減少と支援行動の増加を確認しており[1]、曹らは特定の向社会的行動の促進が日常的な向社会的行動の般化を及ぼすかどうかについて検証している[2]。

このような特定の行動を促進する介入の1つにPersuasion(説得)が存在する。Marjaらは説得を金銭の授受や行動の強制を伴わずに言語的および非言語的なシンボルを用いて人の行動の促進を行うことであると説明づけている[3]。説得はその効果が生じるまでに時間がかかる一方で効果の持続時間が長いという特徴があり、行動の習慣化に適している。よって現在までに座りすぎ行動の抑制[4]やオンライン学習活動に対する学生の関与促進[5]など幅広い行動を対象とした説得による介入が検討されている。また、同様に向社会的行動の促進においても説得の活用について様々な研究がなされている[6][7]。

ボランティア活動や寄付活動などの向社会的行動はより多くの人々によって実施されることで助け合いを生み、社会的に大きな効果を持つようになる。よって向社会的行動の促進介入においてはその影響を広く波及させることが重要である。説得介入は介入対象者の行動変容を促すことは自明であるが、介入対象者を媒介した他者への介入効果の伝搬については未検証である。本研究では向社会的行動の他者推奨の促進に向けて、向社会的行動を促す介入を受けた対象者は、まず向社会的行動を行う意欲がある状態（説得需要）に移行し、次に他者推奨意向を生じると仮定し、以下の課題に取り組む。

課題1 説得的介入によって向社会的行動の説得受容との関係性

課題2 向社会的行動の説得受容から他者推奨意向を導く要因の調査

課題1においては他者推奨意向について検証をする前に説得的介入によって介入対象者が説得を受け入れ、説得受容に移行することを確認する。そこで、本研究では社会的意義付けによる説得的介入を用いて対象行動の促進を行い説得受容度合を評価する。また、本研究における行動変容の対象行動には向社会的行動の中の1つである、日常の屋外での軽い運動（散歩・ジョギング）のついでに街の見回りを行う向社会的行動「パトロールラン」を選択する。

課題2では向社会的行動の他者推奨意向を生じる要因に

Manuscript Format for Kickoff Symposium of IPSJ Behavior Transformation by IoT (BTI)

^{†1} YAMAZAKI YUDAI, KDDI Research, Inc.

^{†2} CAO LIAN, KDDI Research, Inc.

^{†3} KASHIMOTO YUKITOSHI, KDDI Research, Inc.

^{†4} KAMISAKA DAISUKE, KDDI Research, Inc.

について検討する。特定の対象を他者へ推奨する態度や行動についてはレコメンド分野で広く研究がなされており、多くの研究が満足感と他者推奨の関係性について言及している。Bontis らは銀行業においての顧客満足度が推奨意向に影響を与えると述べており [8]、Al-Ansi らはハラル食品について信頼度や満足度が推奨意向に影響を与えると述べている [9]。これらの研究は概して購買行動における満足の度合が高まることで他者推奨につながることを示唆している。しかしながら既存の研究では向社会的行動の他者推奨に関する検討が十分になされていない。向社会的行動は購買行動と異なり必ずしも自己が利益を得る行動ではないため、異なる要因を用いて他者推奨を促進する必要があると考えられる。そこで、本研究では向社会的行動の促進を目的とした説得的介入によって介入対象者は初めに説得受容に移行し、その後に説得を受け入れた対象行動を人に推奨する状態（他者推奨意向）に移ると仮定する。本仮定に基づき説得受容と他者推奨意向を媒介する要因について検証を行う。ここで本研究は媒介変数の候補として考えられる次の 3 つの変数について検討する。まず 1 つめの検証対象は行動の向社会性認知である。向社会的行動は他者のために実施する行動であることから、その意義すなわち行動の目的やその価値を認知することで他者推奨意向を生じる可能性がある。2 つめの媒介変数の候補としては自発性が考えられる、自発性が高い人は自ら向社会的行動に取り組む傾向があると考えられるため、同様に自発的な他者推奨意向を示す可能性がある。3 つ目の媒介変数候補としては自分に対して対象行動が利益をもたらすと感じる度合（自己利益）が考えられる。向社会的行動においては行動が直接自身の利益になるとは限らないが、他者や社会への貢献を通して向社会的行動が自身にとっても有益であると感じる場合にはその他者推奨傾向が見られる可能性がある。以上より本研究では課題 2 の対処を目的として介入対象者の行動の向社会性認知、自発性、自己利益について検証する。また、本課題 2 は課題 1 の説得的介入の効果検証に基づいて実施するため、その対象行動についても課題 1 と同じ「パトロールラン」を用いる。

以上の方針を踏まえ、本研究では説得的介入を含む WEB アンケート調査を実施する。アンケートでは回答者を統制群と実験群に分け、実験群にのみ「パトロールラン」を促進する説得的介入を行う。本調査では課題 1 の説得的介入の効果検証を行うために説得受容を測定し、分析する。次に課題 2 の対処として行動の向社会性認知、自発性、自己利益、他者推奨意向について測定し、媒介分析によって説得受容から他者推奨意向の間の媒介要因を明らかにする。



図 1 社会的意義付けによる説得記事

以下に本稿の構成を述べる。第 2 章では本研究の課題を検証するために実施した WEB アンケート調査について述べ、第 3 章でその分析結果について説明する。第 4 章では本稿の結論を述べる。

2. WEB アンケート調査

2.1. アンケート設計

本研究では課題 1 の向社会的行動の促進を目的とした説得的介入の効果と課題 2 の向社会的行動の説得受容と他者推奨意向の媒介要因を検証するするためにアンケート調査を行った。本アンケートでは初めにその調査対象を対象行動「パトロールラン」が可能な人に絞るため、週 1 回以上屋外での軽い運動（散歩・ジョギング）を行っており、今後も引き続き継続しようとする意欲がある 24 歳-69 歳の回答者計 206 名を対象とした。次に説得的介入の効果を比較するため、被験者全体を統制群 103 名と実験群 103 名の 2 組にランダムに割り付け、実験群にのみ「パトロールラン」の社会的意義付け記事を提示して説得的介入を行った。最後に後述する複数の質問調査項目を用いてアンケートを実施し、課題 1 の説得的介入の効果および課題 2 の説得受容と他者推奨意向を媒介する要因を分析した。本アンケートは WEB アンケート調査会社に依頼し、2022 年 12 月 12 日（月）～2022 年 12 月 13 日（火）に実施した。

表 1 質問調査項目

尺度	質問調査項目
説得受容 (実験群)	今後、パトランを意識した屋外での軽い運動（散歩・ジョギング）を行おうと思いますか？
説得受容 (統制群)	今後、誰かのためになることを意識した屋外での軽い運動（散歩・ジョギング）を行おうと思いますか？
向社会性認知	あなたが現在行っている屋外での軽い運動（散歩・ジョギング）は誰かの役に立つと思いますか？
自発性	今までやったことのない新しい目標を立て進んで軽い運動（散歩・ジョギング）を行う 軽い運動（散歩・ジョギング）をしようと思えば、誰かに言われなくても進んで行う 誰かの役に立つ軽い運動（散歩・ジョギング）であれば、言われなくても進んで行う
自己利益	誰かのためになることを意識した屋外での軽い運動（散歩・ジョギング）は、 今後自分の利益にもなると思いますか？
他者推奨意向	普段行われている屋外での軽い運動（散歩・ジョギング）を誰かに勧めたいと思いますか？

2.2. 社会的意義付けによる説得記事

アンケート回答者のうち実験群には図 1 に示す記事を提示した。本記事は「パトロールラン」についての説明とその社会的意義を伝えるものとなっており、閲覧者に「パトロールラン」を意識して軽い運動を行うと街の安全安心に貢献できることを提示した。

2.3. 質問調査項目

本アンケートにおける尺度と質問調査項目を表 1 に示す。まず課題 1 の対処のために説得受容の質問調査を行った。次に行動の向社会性認知、自発性、自己利益、他者推奨意向に関して質問することで、どの変数が説得受容と他者推奨意向の因果関係を媒介するのかを明らかにした。以降は各調査項目について説明する。

説得受容

説得受容は実験群と統制群で質問項目が異なる。実験群については説得記事を踏まえて今後「パトロールラン」を意識した軽い運動を行う意思がどれくらいあるか測定することを目的として独自項目を作成した。また、統制群については本質問項目の回答を依頼することが説得的介入とならないように質問文の「パトロールラン」を「誰かのためになること」に置き換え、一般的な向社会性を意識した軽い運動を行う意思があるかどうかについて質問している。本項目により統制群では説得的介入が無い状況における向社会的行動の受容度合い測定する。いずれも 5 件法 (1: まったく当てはまらない—5: 非常に当てはまる) を用いて測定した。

向社会性認知

向社会性認知は自身の実施する軽い運動に向社会性をどれくらい感じるか測定することを目的とした独自作成項目である。5 件法 (1: まったく当てはまらない—5: 非常に当てはまる) を用いて測定した。統制群については説得記事による介入を実施していないため、本質問項目によって一般的な軽い運動に対する向社会性の認知度合を測定した。

表 2 井上らによる自発性尺度 [10]

質問調査項目
今までにやったことのない新しい行事などはすすんでやる 皆の役に立つことであれば進んでその仕事を引き受けれる 自分で発言しようと思えば誰に言われなくても進んで発表する

他者推奨意向

他者推奨意向は軽い運動に対する他者推奨意向を測定することを目的とした独自項目であり、5 件法 (1: まったく思わない—5: 非常にそう思う) で測定した。また、本質問項目について「非常にそう思う」「ややそう思う」と答えた回答者についてその理由を「運動が楽しいと思うから」、「健康のためになると思うから」、「一緒に運動する仲間が欲しいから」、「運動をすること自体に意義を感じるから」、「その他」の 5 つの選択肢それぞれについて 5 件法 (1: まったく当てはまらない—5: 非常に当てはまる) で質問した。

自発性

自発性は井上らが開発した中学校における自主性尺度項目 [10] の下位尺度である自発性尺度 (表 2) を参考に、本研究の対象行動（パトロールラン・軽い運動に）に対する回答者の自発性を測定する質問調査項目を作成した。各質問項目について 5 件法 (1: まったく当てはまらない—5: 非常に当てはまる) を用いて測定した。本アンケート調査において α 係数は .75 であった。

自己利益

自己利益は向社会的行動としての軽い運動が自身にとつてどれくらい利益をもたらすを感じるか測定することを目的とした独自項目である。5 件法 (1: まったく当てはまらない—5: 非常に当てはまる) を用いて測定した。

3. アンケート分析結果

3.1. 説得的介入の効果確認

はじめに課題 1 に対応するため「パトロールラン」に関する社会的意義付け記事を用いた説得的介入と説得受容との関係を確認した。まず実験群に対して「パトロールラン」の記事の理解を確認する質問を行い、正しく記事を理解できていない回答者については記事を教示通りに閲覧していないとみなして以降の分析から除外した。ここでは実験群のうち 19 名を除外し、実験群 84 名、統制群 103 名のデータを採用した。次に屋外での軽い運動を向社会的行動として行う意思があるかどうかを説得受容の項目で測定した。測定結果より実験群は「3: どちらともいえない」を基準としたとき有意に高い値 ($t_s > 2.27, p < .05$) が示されたのに対して、統制群は同基準値に対して有意な差は示されなかった。以上の結果から説得的介入を受けた実験群は介入によって説得を受け入れ対象行動を行う意欲がより高く、説得的介入の効果が確認された。続いて向社会的行動の他者推奨について検討する。

3.2. 向社会的行動の他者推奨意向

続いて本節では課題 2 の説得受容と他者推奨意向の因果関係を媒介する要因が何であるかを検証するため、向社会性認知、自発性、自己利益について Bootstrap 法を用いた媒介効果の検証を行った。また、本媒介分析には心理統計分析用のソフトウェア HAD を用いた [11]。

向社会性認知の媒介効果

まず、向社会性認知が向社会的行動の説得受容を他者推奨意向に導くかどうかを実験群と統制群で検証した。向社会的行動の他者推奨意向を従属変数、説得受容を独立変数、向社会性認知を媒介変数とおいて実験群の媒介分析を行った結果を図 2 に示した。分析結果より説得受容から他者推奨意向 ($\beta = .54, p < .01$)、向社会性認知 ($\beta = .61, p < .01$) にそれぞれ有意な正の効果が確認できた。また、向社会的行動の向社会性認知から他者推奨意向 ($\beta = .37, p < .01$) に有意な正の効果が示された。加えて、向社会的行動の説得受容が向社会性認知を媒介して他者推奨意向に有意な媒介効果がみられた ($z = 3.43, p < .01$)。また、この際媒介変数の投入後に $\beta = .54$ が $\beta = .31$ と減少しており、媒介変数による効果が確認できた。

続いて統制群についても実験群と同様の媒介分析を実施し、その結果を図 3 に示した。統制群では説得記事による介入は実施していないが、実験群と同様に行動の向社会性認知に関する質問調査によって測定された一般的な軽い運動に対する向社会性認知度合に基づいて分析を行った。分析結果より統制群についても説得受容から他者推奨意向

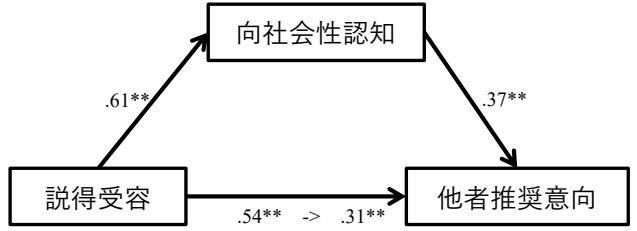


図 2 実験群における向社会性認知の媒介効果

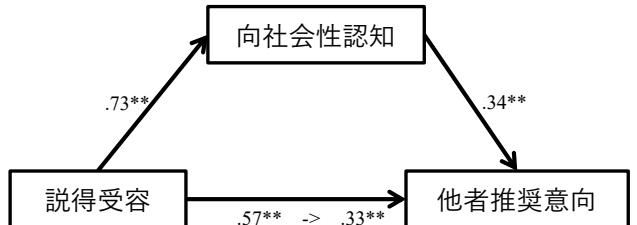


図 3 統制群における向社会性認知の媒介効果

($\beta = .57, p < .01$)、向社会性認知 ($\beta = .73, p < .01$) にそれぞれ有意な正の効果が確認でき、向社会的行動の向社会性認知から他者推奨意向 ($\beta = .34, p < .01$) に有意な正の効果が示された。さらに実験群と同様に媒介分析を行ったところ媒介効果が有意であることが示され ($z = 2.89, p < .01$)、媒介変数の投入前後で $\beta = .57$ が $\beta = .33$ に減少していることから媒介変数による効果が示された。

以上の結果から、説得記事による介入の有無にかかわらず向社会性を意識した軽い運動を行う意思を持っている人は、自身の軽い運動の社会的意義を感じているほど他者推奨意向が促進されることが示された。よって自身の軽い運動の社会的意義を感じるように介入を実施することで、より向社会的行動の他者推奨を促進させることができると考えられる。また、介入を行わずとも元より向社会的行動を行う傾向のある人は自身の軽い運動の社会的意義を感じることで自発的に他者推奨を行う傾向があると示唆している。

自発性の媒介効果

次に自発性が向社会的行動の説得受容と他者推奨意向の因果関係を媒介するかどうかについて実験群と統制群で検証した。向社会的行動の他者推奨意向を従属変数、説得受容を独立変数、自発性を媒介変数とおいて実験群の媒介分析を行った結果を図 4 に示した。分析結果より説得受容から他者推奨意向 ($\beta = .54, p < .01$)、自発性 ($\beta = .59, p < .01$)、にそれぞれ有意な正の効果が確認できた。また、自発性から他者推奨意向 ($\beta = .26, p < .05$) に有意な正の効果が示された。続いて説得受容と他者推奨意向を自発性が媒介することを検証する目的で媒介分析を行った。結果として有意な媒介効果が確認された ($z = 2.08, p < .05$)。さらに自

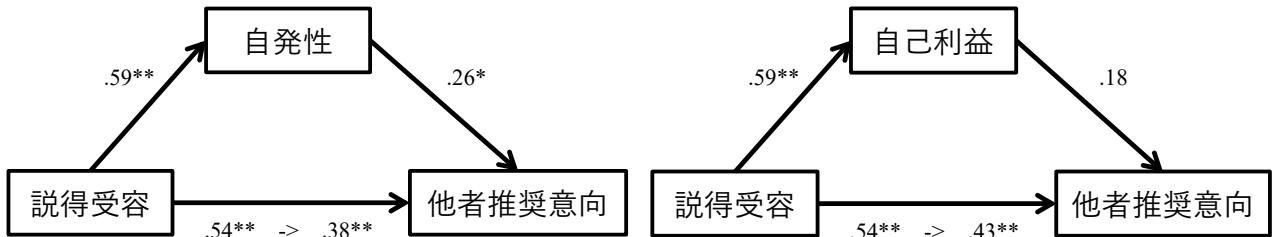


図 4 実験群における自発性の媒介効果

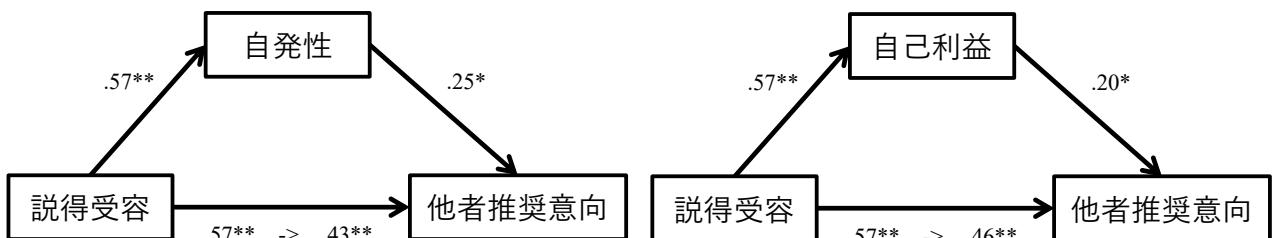


図 5 統制群における自発性の媒介効果

発性を媒介変数として投入することで $\beta = .54$ が $\beta = .38$ と減少しており、媒介変数による効果が見られた。

統制群についても実験群と同様の媒介分析を実施し、図 5 に分析結果を示した。図 5 は統制群についても説得受容から他者推奨意向 ($\beta = .57, p < .01$) と自発性 ($\beta = .57, p < .01$) に有意な正の効果があることを表している。また、向社会的行動の向社会性認知から他者推奨意向 ($\beta = .25, p < .05$) に有意な正の効果が示された。加えて媒介分析をおこなったところ、媒介効果がみられなかった ($z = 1.81, n.s.$)。

分析結果から説得記事による介入がある場合には説得受容と他者推奨意向の関係性に自発性が影響を与えることが示された。本結果より自発性の高い人に向けた向社会的行動の説得はより他者推奨意向を生み、向社会的行動を伝搬させると考えられる。また、自発性を高める介入がより向社会的行動の他者推奨を促す可能性がある。一方で統制群においては説得受容と他者推奨意向の関係に自発性が与える影響が小さいことが分かった。統制群では説得受容の項目で一般的な向社会的行動の意向を測定していることから、自発的に向社会的行動を実施する状況について想定出来ていない可能性が考えられる。

自己利益の媒介効果

同様にして自己利益が向社会的行動の説得受容から他者推奨意向を媒介するのかどうかを検証した。向社会的行動の他者推奨意向を従属変数、説得受容を独立変数、自己利益を媒介変数とおいて実験群の媒介分析を行った結果を図 6 に示した。分析結果は説得受容から他者推奨意向 ($\beta = .54, p < .01$)、自己利益 ($\beta = .59, p < .01$) にそれ

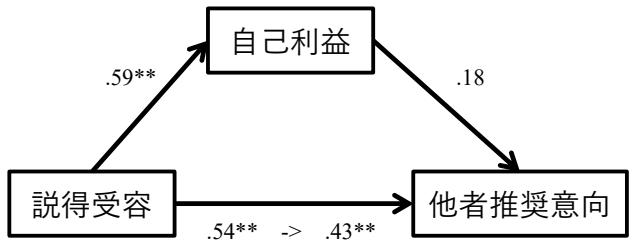


図 6 実験群における自己利益の媒介効果

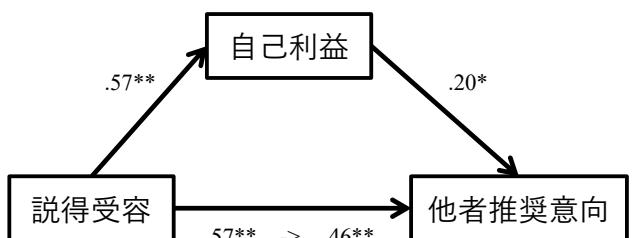


図 7 統制群における自己利益の媒介効果

表 3 高い他者推奨意向を示した回答者の理由回答

理由	実験群 (n=46)	統制群 (n=55)
運動が楽しいと思うから	76.1 %	65.5 %
健康のためになると思うから	97.8 %	98.2 %
一緒に運動する仲間が欲しいから	39.1 %	36.4 %
運動をすること自体に意義を感じるから	80.4 %	76.4 %
その他	8.7 %	7.3 %

それ有意な正の効果を示したが、自己利益から他者推奨意向 ($\beta = .18, n.s.$) に有意な正の効果が示されなかった。本要因についても自己利益を媒介変数とした媒介分析を行ったところ媒介効果が有意ではなかった ($z = 1.41, n.s.$)。統制群についても図 7 のように同様の分析を行ったが自己利益に関して媒介効果は見られなかった ($z = 1.65, n.s.$)。本結果より向社会的行動が自身にとって利益をもたらすと感じる度合が他者推奨意向に与える影響は小さいことが示された。

高い他者推奨意向を示した回答者の理由

他者推奨意向の質問調査項目について「非常にそう思う」「ややそう思う」と回答し、他者推奨意向が高いと判断できる実験群 46 人、統制群 55 人についてその理由に関する質問を行った結果を表 3 にまとめる。結果より他者推奨意向が高いユーザのうち「運動をすること自体に意義を感じるから」を選択したユーザが両群とも 70% 以上と高く、実験群は統制群と比べて 4% 高いことが示された。次にここで回答された運動の意義が社会的意義を指しているかを明らかにするために実験群、統制群のそれぞれで向社会性認知

を従属変数として他者推奨意向の高い群と低い群を 2 標本 t 検定で比較した。その結果、実験群と統制群の両方で他者推奨意向が高い群は低い群と比べて向社会性認知が有意に高いことが示された（実験群: $t(82) = -5.7, p < .01$, 実験群: $t(101) = -4.1, p < .01$ ）。よって向社会的行動の他者推奨意向が高い人はより行動の社会的意義を理解して行動の他者推奨意向に至っていることが示された。また、回答者の多くは「運動が楽しいと思うから」や「健康のためになると思うから」についてもその回答理由に挙げていた。従って、楽しさや健康といった理由が基盤にあった上で、社会的意義をあわせて運動をすることを他者に推奨していると考えられる。

以上の結果より本検討では向社会的行動の説得受容から他者推奨意向へ至る過程を向社会性認知と自発性が媒介することが示された。

4. おわりに

本研究の目的は向社会的行動の促進に向けた説得的介入において向社会的行動の他者推奨意向を導く要因を明らかにすることである。本目的のため、本研究では行動変容のステージが説得受容（介入によって説得を受け入れた状態）と他者推奨意向（人に対象行動を推奨する状態）の 2 段階で進行すると仮説を立ててその媒介変数を明らかにするために WEB アンケート調査を実施した。また、本調査においては向社会的行動の 1 つであるパトロールランを対象行動として選んだ。アンケート調査結果より各被験者の行動に対する社会的な意義の理解と行動に対する自発性が説得受容と行動の他者推奨意向を媒介することが分かった（向社会性認知: $z = 3.43, p < .01$, 自発性: $z = 2.08, p < .05$ ）以上のように向社会的行動の他者推奨を促すためには行動の自発性や社会的意義の理解を向上させる介入が有効であることが示唆された。本研究結果を情報システムによる行動変容に活用することで、一部の個人に向社会的行動を促進する介入の効果が他者にも波及し、結果的に人々が助け合う社会の実現に寄与できる可能性がある。今後の課題として説得的介入による態度変容の長期的な効果、および行動面での効果についても検証を行う必要がある。また、他者推奨行動は他者とのかかわりの中で生じるものであることから個人の特性のみならず集団やコミュニティの特性を活用することによって向社会的行動の他者推奨を促す技術について検討できる可能性が考えられる。

参考文献

- [1] Caprara, G. V., Kanacri, B. P. L., Gerbino, M., Zufianò, A., Alessandri, G., Vecchio, G., Caprara, E., Pastorelli, C. and Bridgall, B.: Positive effects of promoting prosocial behavior in early adolescence: Evidence from a school-based intervention, *International Journal of Behavioral Development*, Vol. 38, No. 4, pp. 386–396 (online), 10.1177/0165025414531464 (2014).
- [2] 曹蓮, 山崎悠大, 柏本幸俊, 上坂大輔. 向社会的行動の般化に及ぼす心理的要因～社会的意義づけされた軽い運動に焦点を当て～, 電子情報通信学会ヒューマンコミュニケーション基礎 (HCS) 2022 年 12 月研究会 (2022).
- [3] Harjumaa, M. and Oinas-Kukkonen, O.: Persuasion Theories and IT Design, pp. 311–314 (online), 10.1007/978-3-540-77006-0_37 (2007).
- [4] Aldenaini, N., Alqahtani, F., Orji, R. and Sampalli, S.: Trends in Persuasive Technologies for Physical Activity and Sedentary Behavior: A Systematic Review, *Frontiers in Artificial Intelligence*, Vol. 3 (online), 10.3389/frai.2020.00007 (2020).
- [5] Orji, F. A., Vassileva, J. and Greer, J.: Evaluating a Persuasive Intervention for Engagement in a Large University Class, *International Journal of Artificial Intelligence in Education*, Vol. 31, No. 4, pp. 700–725 (online), 10.1007/s40593-021-00260-4 (2021).
- [6] 柏本幸俊, 曹蓮, 山崎悠大, 上坂大輔. 向社会的行動促進に向けた心理プロセス調査と AI による行動誘発・習慣化に向けた展望, 情報処理学会 IoT 行動変容学研究グループ 第 2 回研究会, p. 33–39 (2022).
- [7] Slattery, P., Finnegan, P., Vidgen, R. and Land, L.: Persuasion in prosocial domains: Explaining the persuasive affordances of volunteering websites (2014).
- [8] Bontis, N., Booker, L. D. and Serenko, A.: The mediating effect of organizational reputation on customer loyalty and service recommendation in the banking industry, *Management Decision*, Vol. 45, pp. 1426–1445 (2007).
- [9] Al-Ansi, A., Olya, H. G. and Han, H.: Effect of general risk on trust, satisfaction, and recommendation intention for halal food, *International Journal of Hospitality Management*, Vol. 83, pp. 210–219 (online), <https://doi.org/10.1016/j.ijhm.2018.10.017> (2019).
- [10] 井上史子, 沖 裕貴, 林 徳治中学校における自主性尺度項目の開発, 教育情報研究, Vol. 21, No. 3, pp. 13–20 (2006).
- [11] 清水裕士. フリーの統計分析ソフト HAD : 機能の紹介と統計学習・教育, 研究実践における利用方法の提案, メディア・情報・コミュニケーション研究, No. 1, pp. 59–73 (2016).