

萌え要素の加減算に基づくキャラクタの類似度判定手法の提案

Determining similarities between comic personages based on addition and subtraction of Moe-factors

松井 俊樹^{1*} 朴 炳宣² 松下 光範¹
Toshiki Matsui¹ Byeongseon Park² Mitsunori Matsushita¹

¹ 関西大学総合情報学部

Faculty of Informatics, Kansai University

² 関西大学大学院総合情報学研究科

Graduate School of Informatics, Kansai University

Abstract: The aim of our study is to calculate similarities between Comic personages by focusing on Moe-factors. “Moe” means strong feelings of affection, often towards an ideal personage in comic and anime. In the existing similarity calculation method between comic’s personages, it is difficult to search a similar personage based on the operation of the constituent factors such as “a comic personage like Conan, but with passionate.” To meet the requirement, this paper proposes a similarity calculation method by adding and subtracting Moe-factors to a known comic personage by focusing on Moe-factor as a component of the user’s interest. Our user study showed that the proposed method enabled calculating similar comic personages with an accuracy of 70.0%.

1 はじめに

近年では、アニメやコミックに登場するキャラクタに対する注目が高まっている。また、東によると現代ではキャラクタの特定の見た目(e.g.,セーラー服)や内面特徴(e.g.,無口)といった断片的な構成要素が重要視されている[1]。キャラクタの構成要素の中でも特に、ユーザの関心を引きつける要素(e.g.,猫耳, ツンデレ)は、この萌えという言葉を用いて、「萌え要素」や「萌え属性」と呼ばれて親しまれており、ユーザのキャラクタの好みに大きく影響を与えている。本稿では、この萌え要素に着目し、ユーザが好むキャラクタを検索する手法の実現を目指す。萌え要素をクエリとしてキャラクタを探す場合、ユーザは自身が好む萌え要素を列挙して、挙げられた要素を持つキャラクタを探す方法が考えられる。しかし、ユーザが必ずしも自らが好む萌え要素を明確に認識しているとは限らず、外在化が難しい場合も存在する[2]。そのため、情報検索手法の研究においては、ユーザの情報要求の曖昧性に対応した検索が可能であることが望ましい[3]。クエリ入力の負担を減らし検索を効率的に行う手段として、ユーザの既知のキャラクタをもとに類似したキャラクタを探す方法がある。例えば「荒っぽく無鉄砲だが、正義感

の強いキャラクタ」が好きなユーザが、好みにあったキャラクタを探す際、すでにその好みに該当するキャラクタを知っているなら、そのキャラクタと似ているキャラクタを探すことで、「荒っぽい」、「無鉄砲」、「正義感が強い」などの要素を列挙する必要がなくなる。本稿では、このような既知キャラクタをもとに類似したキャラクタを探す検索方式に着目する。

類似キャラクタを判別する手法として、朴らは性格診断法であるエゴグラムを用いて類似する性格のキャラクタを判定する手法を提案している[4]。しかし、朴らの手法で扱うことができるのはキャラクタ構成要素のうち性格のみであり、キャラクタの見た目に基づく構成要素は考慮されていない。また、類似度判定の際に「コナンに熱血を足したキャラクタ」のような、既存キャラクタの構成要素に操作を加えた情報を検索の基準にすることができない。

そこで本研究では、キャラクタの内面だけでなく見た目も含めた、萌え要素の加減算に基づくキャラクタの類似度判定手法を提案する。

2 関連研究

河谷らは萌えを基準としたキャラクタの検索・分類の方法についてまとめている[5]。この研究では、萌えの対象となるキャラクタの特徴を画像から抽出し、そ

*連絡先：関西大学総合情報学部
〒569-1095 大阪府高槻市霊山寺町 2-1-1
E-mail: mat@res.kutc.kansai-u.ac.jp

の特徴をもとにキャラクタを検索する手法を提案している。キャラクタの顔画像から、視覚的に判別可能な萌え要素である目、顔、髪、の3つの領域を対象とし、減色処理と色クラスタリング処理によって各領域をレイヤーに分け、それぞれの特徴量を求めた。これにより目領域、顔領域、髪領域の特徴が抽出された。その後、抽出した特徴をキャラクタの画像に紐づけることにより、好みのキャラクタと似た特徴を持つキャラクタ画像の検索を実装した。しかし、河谷らの手法は画像をもとにしたものであるため、扱うことができる特徴が見た目で判別できるものに限定されている。

白石らは、絵本に関するレビューの文章から絵本が持つ感情的特徴を抽出し、感情を加味した絵本の分類手法を試みている [6]。まず、WordNet Affect というソーラスで単語間の関係を保持している概念辞書をもとに、喜び、悲しみ、恐れ、怒り、嫌悪、驚きの6つの感情を示す単語から感情単語表を作成した。次に、形態素解析により分割された絵本のレビューに対し感情単語表にある単語との一致数を感情ごとにカウントし、カウント結果を用いて絵本ごとの特徴ベクトルを作成し、クラスタリングによる分類を行なった。この手法による分類の分析の結果、絵本が有する感情の一端を捉えることができた。本稿ではこの手法をもとに、萌え要素の辞書を作成し、キャラクタごとの萌え要素の特徴ベクトルを作成することでキャラクタの類似度判定を行う。

キャラクタ情報に基づくコミック検索の実現を目的として、朴らはキャラクタの構成要素の中でも性格に着目した類似キャラクタ判定手法を検討している [4]。この手法では、性格診断法であるエゴグラムをもとに、性格を表す単語・短文にベクトルを付与した辞書を作成した。次に、ユーザによる編集が可能な自由参加型 Web 百科事典のキャラクタに関する文書から辞書に含まれる語や短文を抽出し、それらの語が持つベクトルの値を集計した。その後、その値をもとに類似度を測りキャラクタの性格タイプを導き出した。朴らの手法による分類の結果を、人間がキャラクタの立場でエゴグラムを用いた性格タイプの判別を行なった結果と比較し、55.0%の分類精度を示した。これにより、自由参加型 Web 百科事典から抽出した情報に基づき作成されたベクトルを用いることで、類似キャラクタの判別が可能であることが示唆された。本稿ではこの手法をもとに、自由参加型 Web 百科事典のキャラクタの説明に関する文書から特徴を抽出を行う。

3 データセットの作成

キャラクタ間の類似度判定を行うにあたり、キャラクタが持つ萌え要素をベクトルにより表現する。その

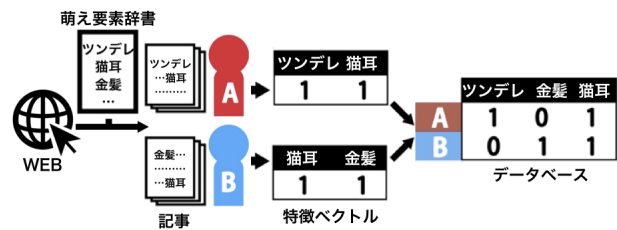


図 1: 特徴ベクトルの作成

ため、キャラクタの内面、見た目の特徴を表す萌え要素を収集し、萌え要素辞書を作成する。次に、キャラクタに関する説明が含まれた文書から、萌え要素辞書に含まれる表現を抽出し、抽出された表現を基底としてキャラクタごとの特徴ベクトルを作成した (図 1 参照)。本章では、上記手順の実装方法について示す。

3.1 対象データ

キャラクタが持つ萌え要素を抽出するため、萌え要素を示す表現を抽出し辞書を作成する。萌え要素を表現する語には、アニメやコミックの文化で生まれた独特の表現 (e.g., ツンデレ, ボクっ娘) が存在する。また、萌え要素は作品の制作者側ではなく、ユーザ側によって見出されていくものである。そのため本稿では、ユーザによる編集が可能であり、独特の表現に関する記述が豊富に存在する自由参加型 Web 百科事典の Wikipedia¹, ニコニコ大百科², ピクシブ百科辞典³, アニヲタ wiki (仮)⁴を対象とし、それらの Web サイトで萌え要素として扱われている表現を収集した。

自由参加型 Web 百科事典では、萌え要素・萌え属性は同義語として扱われているため、各自由参加型 Web 百科事典で「萌え要素」または「萌え属性」として扱われている語を収集する。各 Web サイトごとに、本稿における萌え要素の収集条件を以下に示す。

- Wikipedia: category 「萌え属性」に属する語
- ピクシブ百科事典: 「萌え要素」の記事でまとめられた語
- ニコニコ大百科: 「萌え要素・萌え属性の一覧」にまとめられた語
- アニヲタ wiki (仮): 「萌え要素」, 「萌え属性」のタグがついた語

¹Wikipedia (<https://ja.wikipedia.org/wiki/>: 2018 年 10 月 28 日時点確認)

²ニコニコ大百科 (<http://dic.nicovideo.jp/>: 2018 年 10 月 28 日時点確認)

³ピクシブ百科事典 (<https://dic.pixiv.net/>: 2018 年 10 月 28 日時点確認)

⁴アニヲタ Wiki (仮) (<https://www49.atwiki.jp/aniwotawiki/>: 2018 年 10 月 28 日時点確認)

表 1: 表記揺れ統一例

統一前	統一後
タレ目, たれ目, 垂目, 垂れ目	タレ目
ドジっ子, ドジっ娘, ドジっ娘	ドジっ子
地味子, 地味っ子	地味子
狐耳, きつね耳, キツネ耳	狐耳
ロングヘアー, ロングヘア, 長髪	ロングヘア

	減算する要素 黒髪	赤髪	加算する要素 金髪	...
操作前	1	1	0	...
操作後	0	1	1	...

1から0へ 0から1へ

図 2: ベクトルの操作

3.2 萌え要素辞書作成

収集した 1182 語の萌え要素のうち、重複する表現、公序良俗に反する表現、表現単体で特徴を表していないもの (e.g., うなじ, 髪の色), キャラクタの内面, 見た目の特徴を表していない表現 (e.g., 温泉回, 世界観) を取り除いた 854 語を萌え要素辞書に使用する表現とした。漢字・ひらがな・カタカナの差や、表記揺れ、別名称が存在する表現については、各表現ごとに weblio 辞書⁵と 3.1 節で使用した自由参加型 Web 百科事典の記事を確認し、表記揺れとして記載されている表現を一つのカテゴリに統一した (表 1 参照)。統一前の表現が抽出された際、統一後の表現が抽出されたものとして処理する。例えば、表 1 に示した「タレ目」の場合「タレ目」、「たれ目」、「垂目」、「垂れ目」のいずれの表現が抽出された場合においても、「タレ目」が抽出されたとする。

また、萌え要素として収集した表現の中には「純情」、「純真」、「純粋」のように、意味の類似した語が存在する。これらの語については、weblio 辞書で類似語検索を行い、類語として分類されている語を、同一の意味カテゴリの要素として統一した。辞書作成に用いた単語・短文が weblio 辞書の類語検索辞書に存在しなかった場合、語の意味を weblio 辞書と自由参加型 Web 百科事典で検索し、説明文の中に類語や別表記が存在しないかを確認した (表 2 参照)。

これらの処理により、萌え要素の辞書として 465 の意味カテゴリを持つ「萌え要素辞書」を作成した。

3.3 特徴ベクトル作成

「漫画全巻ドットコム⁶」の漫画発行部数ランキングに掲載された 102 作品に登場する 508 キャラクタを対象とし、キャラクタに関する説明が含まれた文書を萌え要素辞書作成と同一の Web サイトから収集した。まず、キャラクタごとに収集した文書から、萌え要素辞書に含まれる表現を抽出する。この際、語尾の表記揺

れに関しては MeCab による形態素解析を行い原形に統一した。

次に、各キャラクタごとに抽出された表現を基底として one-hot ベクトルを作成した。このとき、抽出対象となる辞書として使用した 465 次元から、抽出数が 0 の基底語を除去することで、362 次元に圧縮した。本稿ではこれをキャラクタの特徴ベクトルと呼ぶ。

4 提案手法

提案手法では、前章で作成したキャラクタの特徴ベクトルを用いて、ユーザが指定したキャラクタ (以下、基準キャラクタと記す) と類似するキャラクタを提示する。ユーザは検索条件を入力する場合、「ルフィ + 金髪」、「悟空 - 熱血」のように、基準キャラクタに対して加減算を行う萌え要素 (以下、変更要素と記す) を設定する。基準キャラクタの特徴ベクトルに対して、加算する変更要素の次元を 1 に、減算する変更要素の次元を 0 に操作する。例えば、特徴ベクトルの基底語として「黒髪」を持ち、「金髪」を持たない基準キャラクタに対し、「金髪」を加算する変更要素として、「黒髪」を減算する変更要素に指定した場合、図 2 のように操作する。その後、操作された特徴ベクトルと類似度が高い特徴ベクトルを持つキャラクタを判定し、ユーザに対して提示する。

4.1 関連要素の設定

ユーザが変更要素を指定する際、内面特徴については表現の多様性が高いため (e.g., 「根暗」と「陰気」)、ユーザが指定した変更要素がキャラクタの基底語ではない場合が想定される。しかし、ユーザが表現の多様性を考慮して変更要素の指定を行うのは困難である。そこで、変更要素と関連の強い基底語 (以下、関連要素と記す) についても同様の操作を行う。例えば、「無愛想」の関連要素として「無口」と「冷淡」が設定されていた場合、「無愛想」が加算する変更要素として指定された場合、「無愛想」、「無口」、「冷淡」の次元を 1 に変更する。関連要素を設定するにあたり、内面特徴を示す

⁵weblio 辞書 (<https://www.weblio.jp/>: 2018 年 11 月から 2018 年 12 月確認)

⁶<https://www.mangazentan.com/> (2018 年 12 月 20 日時点確認)

表 2: 辞書検索の例 (weblilo 辞書)

収集した萌え要素	意味	類似語
冷血漢 脳筋 おちゃっぴい	人情味にかけている人。冷酷な人間 思考が単純・バカ, といった性格の者を評して言う語 おしゃべりで, 茶目っ気のあるさま	冷酷 単純, バカ おしゃべり, 茶目っ気

表 3: パターンの例

基準キャラクタ	プラスする要素	マイナスする要素	正解キャラクタ
サンジ うずまきナルト 三千院ナギ 結城リト 高良みゆき	真面目・無愛想・常識人 短気・傲慢 冷淡 気弱・勇気・面倒見が良い あざとい・ツインテール	女好き 孤独・ツツンヘア 自己中心的・お調子者 鈍感 賢い・眼鏡	佐藤潤 エドワード・エルリック ソーニャ 沢田綱吉 吉川ちなつ

萌え要素間での関連性の強さを分析する。一般に、文書中で共起関係がある語は同じ文脈を共有していると考えられている。よって、共起ネットワーク上でエッジが結ばれたノード同士には意味的な関連性が認められる場合がある [7]。そこで本稿では、各キャラクタを説明する文書においての内面特徴の共起関係に着目し、共起ネットワークを作成することで関連の強い基底語の分析を行った。共起ネットワークとは、文書に出現する語をノードとし、共起関係があるノード同士をエッジで結びつけることにより、ノード同士の関係性を可視化するネットワークである。共起ネットワークの作成には、樋口が作成・公開しているテキストマイニングツール、KH Coder を使用した [8]。KH Coder はテキスト型のデータを統計的に分析するためのフリーソフトウェアである。共起関係は、Jaccard 係数によって算出した。

共起ネットワークの作成に使用するデータは、特徴ベクトルを作成する際に収集した 508 キャラクタの文書と、特徴ベクトルの 362 次元のうち、内面特徴を表す 153 の要素を対象とした。KH Coder により文書を形態素解析し、辞書に含まれる表現のうち内面特徴を示す表現の共起回数を抽出する。共起関係の集計単位はキャラクタごととした。不必要に多くのノードと結びつき、関連要素の数が膨大になることで、萌え要素の加減算の結果がユーザの意図と離れることを防ぐため、本稿では最小限に簡潔化されたネットワークの作成を目指した。そのため、KH Coder による共起ネットワークの作成条件は出現数 5 以上、Jaccard 係数 0.1 以上の最小スパニング・ツリーとした。KH Coder の機能であるサブグラフ検出により、共起ネットワーク上で相対的に強く結びついているグループが自動的に色分けされており、ネットワーク全体から見ると同じ色で分けられた部分は強い共起関係がある。本稿では、作成した共起ネットワーク図上で隣接するノード同士を関連要素とした (図 3 参照)。例えば「内向的」の場

合は「穏やか」と「恥ずかしがり」が、「残忍」の場合は「戦闘狂」と「意地悪」が関連要素となる。

4.2 類似度の判定

ユーザにとっては、変更要素として指定した要素が、検索において最も重視する基底語である。そのため基準キャラクタと比較キャラクタの類似度は、特徴ベクトル中の変更要素の次元の一致率で算出した。この値が同じ比較キャラクタ同士の順位づけは、変更要素の次元以外の一致率で決定した。一致率の算出には \cos 類似度を用いた。 \cos 類似度は、二つの n 次元のベクトル間の距離を測る際に用いられ、値が 1 に近いほど二つのデータが似通っていることを表す。

5 実験

5.1 実験目的

本稿の提案手法では、自由参加型 Web 百科事典から収集した萌え要素、キャラクタの説明に関する文書から抽出した情報に基づいて萌え要素の加減算を行なっている。人が判断する萌え要素の加減算の結果と、提案手法が提示する萌え要素の加減算の結果に一致性が見られるかを確認することで、提案手法の有効性を検証した。

5.2 正解データの作成

人が判断する萌え要素の加減算結果をアンケートで収集し、正解データとする。アンケートは 10 代から 20 代の男女 6 名 (男性 4 名, 女性 2 名) を対象として行った。アンケートでは「見た目と内面の特徴が近くなるよう、“基準キャラクタ + 眼鏡 - 恥ずかしがり屋 = 解

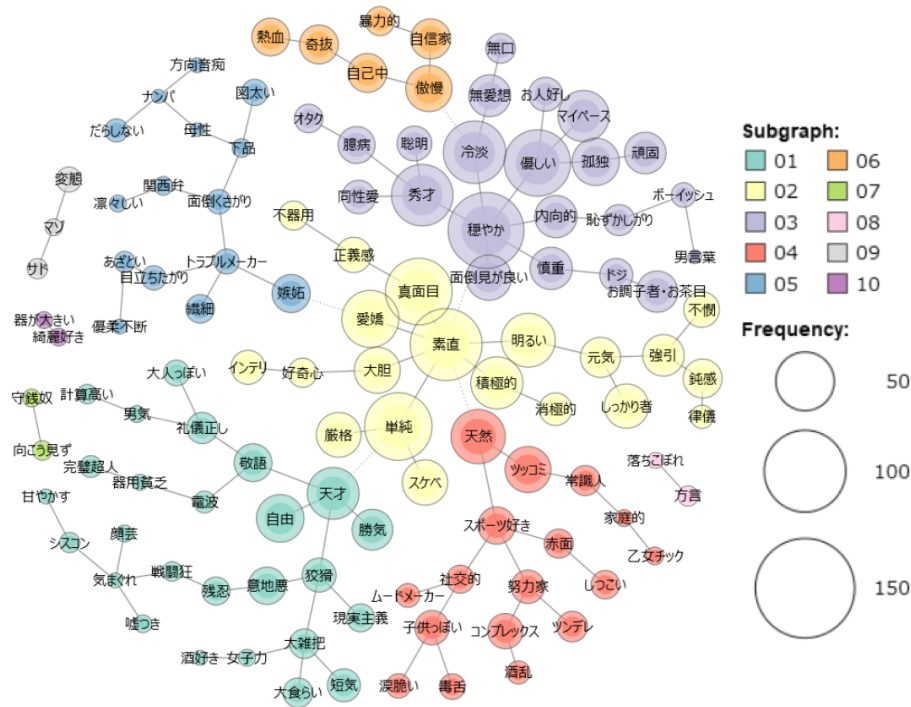


図 3: 共起ネットワーク

答キャラクター”という形式に当てはまる計算式を作ってください」という指示を与えた。このアンケートにより基準キャラクター、解答キャラクター、変更要素の組が20パターン収集された。得られたパターンの一例を表3に示す。

5.3 実験概要

実験は、3.3節で収集した508キャラクターに、アンケートで挙げられたキャラクターのうちでそれに含まれない13キャラクターを加えた521キャラクターを対象に、特徴ベクトルを作成した。アンケートによって収集された20の基準キャラクターと変更要素をもとに、4章で述べた手法を用いて基準キャラクターの特徴ベクトルを操作した。特徴ベクトルを操作した基準キャラクターと、特徴ベクトルを作成した521キャラクターの類似度判定を行う。類似度判定の結果、20の各パターンごとに520のキャラクターの中で上位3位までに解答キャラクターが提示されれば正解とした。

5.4 実験結果

提案手法と正解データを比較した結果、解答キャラクターが1位として提示されたのは45.0% (20キャラクター中9キャラクター)であった。2位として提示されたのは

20.0% (20キャラクター中4キャラクター)であった。3位として提示されたのは5.0% (20キャラクター中1キャラクター)であった。上位3位までに提示されたのは70.0% (20キャラクター中14キャラクター)であった。提案手法による結果を表4に示す。さらに、変更要素を加減算する前の基準キャラクターと解答キャラクターの類似度の順位平均と、変更要素に操作を加えた基準キャラクターと解答キャラクターの順位平均を比較した。変更要素の加減算する前の解答キャラクターの類似度の順位平均は126.3位 (最高4位, 最低432位, 分散10268.21)であり、変更要素の加減算を適用した後の順位平均は9.6位 (最高1位, 最低46位, 分散216.54)であった。

6 考察

実験により人手で作成された萌え要素の加減算結果を正解データとし、本手法の有効性を検証した結果、精度は70.0%であった。また、変更要素操作前から操作後で順位の分散は10268.21から216.54へと収束し、類似度順位が下がったのは20パターン中1パターンのみだった。これらの結果から、本稿で提案した萌え要素の加減算に基づくキャラクター類似度判定手法に有効性があることが示唆された。

一方で、解答キャラクターを上位3位以内に提示できなかった6パターンのうち5パターンでも解答キャラクターの類似度順位の上昇幅は限定的であった。これは、

表 4: 解答キャラクタ順位結果

パターン番号	変更要素操作前	変更要素操作後
1	130 位	1 位
2	149 位	38 位
3	10 位	1 位
4	165 位	2 位
5	79 位	1 位
6	4 位	43 位
7	54 位	1 位
8	322 位	2 位
9	180 位	46 位
10	41 位	1 位
11	59 位	3 位
12	72 位	21 位
13	113 位	2 位
14	33 位	2 位
15	179 位	11 位
16	82 位	1 位
17	171 位	1 位
18	160 位	13 位
19	91 位	1 位
20	432 位	1 位

アンケート回答者が想定した基準キャラクタ、解答キャラクタが持つ萌え要素と、提案手法が判断した基準キャラクタ、解答キャラクタの基底語に差異があったことで、変更要素の次元の一致率が下がったことが要因となっている。アンケート回答者と提案手法の判断に差異が生まれた原因として、キャラクタが所持する萌え要素を表現した特徴ベクトルを作成する際の、萌え要素の抽出手法の精度が影響していると考えられる。

萌え要素を抽出する際、提案手法では萌え要素辞書の表現と一致した表現のみを抽出している。そのため文脈によって表現された萌え要素が存在すると、不要な要素を抽出してしまう場合や、抽出すべき要素が抽出できない場合がある。そのため、萌え要素を抽出する際、萌え要素の抽出の際に行う形態素解析の後に構文情報を参考とした係り受け解析を行うことを検討する。

また、提案手法では特徴ベクトルを作成する過程において one-hot 表現を用いているため、特徴ベクトルの基底語間に重要度の差が存在しない。抽出された萌え要素のうちどの要素がユーザにとって重視されているのかを考慮した類似度判定を行うため、文書からキャラクタの萌え要素を抽出する際、Bag-of-words や TF-IDF を用いてキャラクタの各基底語に対して重み付けを行うことで、類似度判定において変更要素に指定された要素の一致率を算出する際、加算する変更要素として指定された要素の重みが重いものを上位に提示することで、よりユーザが重視した観点に基づいた類似判定が可能であるかについて検討したい。

7 おわりに

本研究では、既知キャラクタをもとに類似したキャラクタを探す検索方式に着目し、萌え要素の加減算に基づくキャラクタの類似度判定手法を提案した。人手で作成したパターンを正解データとした実験から、提案手法の有効性が示唆された。今後は、文脈を考慮した萌え要素の抽出や、ユーザが重要視するキャラクタの基底語に対して重み付けを行い、萌え要素の加減算結果の精度向上を目指す。

参考文献

- [1] 東 浩紀: 動物化するポストモダン—オタクから見た日本社会—, 講談社 (2001).
- [2] 土方 嘉徳: 推薦システムにおけるインタラクション研究へのいざない, ヒューマンインターフェース学会誌, Vol. 15, No. 2, pp. 131–134 (2013)
- [3] R. S. Taylor: Question-negotiation an information-seeking in libraries, *College and Research Libraries*, (1968).
- [4] 朴 柄宣, 居林 香奈枝, 松下 光範: エゴグラムに基づいたコミックキャラクタの性格分類, 第 32 回人工知能学会全国大会論文集, 1J3-02 (2018).
- [5] 河谷 大和, 柏崎 礼生, 高井 昌彰, 高井 那美: アニメにおける人物顔画像の萌え因子特徴評価と検索分類システムへの応用, 映像情報メディア学会技術報告, Vol. 34, pp. 113–118 (2010).
- [6] 白石 諒, 大竹 恒平, 生田目 崇: レビューに含まれる感情を用いた児童向け絵本分類の検討, 情報処理学会第 79 回全国大会 (分冊 2), pp. 533–534 (2017).
- [7] 佐藤 進也, 福田 健介, 菅原 俊治, 栗原 聡: 文書ストリームにおける語のバーストと共起ネットワークにおけるクラスタ構造の関係について, 情報処理学会論文誌データベース (TOD), Vol. 48, No. SIG14(TOD35), pp. 69–81 (2007).
- [8] 樋口 耕一: 言語研究の分野における KH Coder 活用の可能性, 計量国語学会, Vol. 31, No. 1, pp. 36–45 (2017).

知識獲得の教材としてのコミックの利用についての一検討

An Investigation of Using Comics as Teaching Materials for Knowledge Acquisition

松岡航平^{1*} 西原陽子² 山西良典²

¹ 立命館大学大学院情報理工学研究科

¹ Graduate School of Information Science and Engineering, Ritsumeikan University

² 立命館大学情報理工学部

² College of Information Science and Engineering, Ritsumeikan University

Abstract: Comics have many useful knowledges in their texts. The purpose of readers for reading comics is to enjoy the stories and the images. The readers might not recognize the knowledges in the comics. It is required that extraction and visualization of knowledges to support the readers to recognize them. In this paper, we propose a method to extract knowledges from the texts of comic. The method extracts the knowledges by classifying the texts of speech into two classes: classes for texts with/without knowledge. The method takes a set of texts of comic as input. Then, the method transforms each text into an averaged word vector using distributed representations of words. The averaged word vectors are assigned labels for classification and trained with support vector machine. After training, the method obtains a classification model. We experimented with the proposed method. The averaged accuracy was 0.83, the averaged precision was 0.81, and the averaged recall was 0.85.

1 はじめに

コミックのセリフやコマには日常生活で使える知識や、知っておくと為になる知識が多く含まれている。このような知識を学習可能な知識と呼ぶ。学習可能な知識が含まれるコミックのセリフとコマの例を図1に示す。図1では、一人のキャラクターが話している様子がセリフと共に描かれている。このセリフからは、大ぐま座と小ぐま座は1年中北極星の周りを回っているという知識が得られる。このことは、コミックを読むことでストーリーを楽しむだけでなく、学習可能な知識を獲得することも可能になる、つまりコミックを知識獲得の手段の一つとして利用できることを示している。

しかし、コミックを読む本来の目的はストーリーを楽しむことであると考えられる。ストーリーを楽しんでいると、知識が含まれていることに明確に気づかないこともあると予想される。このためコミックの読み手は知識をそれと認識しにくいという問題を抱えている。また、コミックの同じ話を何度も読み返さないことも多いと予想される。したがって、あとでその知識

を利用または応用するほど記憶に残らない可能性がある。コミックに含まれる学習可能な知識を明確に認識すること、そして記憶に残すためには、その知識の部分を抽出し、読み手に可視化することが必要と考えられる。

そこで、本研究ではコミックのセリフに含まれる常識的知識が書かれたセリフを抽出する方法を提案する。本研究では学習可能な知識が書かれたセリフを解説文と呼ぶ。知識の一つとして常識的知識があり、自動獲得やユーザによる獲得を支援する研究が進められている。なぞなぞを用いた常識的知識を獲得するゲームに関する研究 [4] や、Web 上のテキストデータからコモンセンス表現を収集する手法の研究 [1] などがある。他にもコンピュータとの会話において常識的知識を応用しようとした研究 [7] もある。既存研究においてはコミックのセリフや会話文から学習可能な知識の抽出を試みた例は少ない。本研究では、機械学習を用いてセリフの中に学習可能な知識が含まれるものとそうでないものを分類する手法を提案する。

既存研究において、コミック中のセリフから人物像を表す役割語という情報を抽出し分類する研究 [5] や、アニメの会話を用いて外国語のリスニングスキルを訓練支援をする研究がある [6]。コミックやアニメを応用

*連絡先：立命館大学大学院情報理工学研究科
〒525-8577 滋賀県草津市野路東 1-1-1
E-mail: is0280sx@ed.ritsumeikai.ac.jp

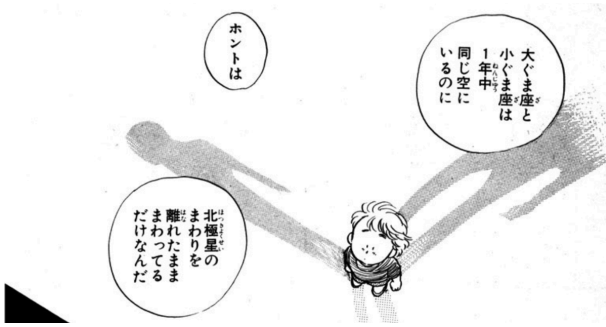


図 1: 学習可能な知識を含むコミックのセリフの例。Manga109[2] より引用。「愛さずにはいられない」©よしまさこ。

した学習教材に関する研究は広がりを見せている。これまでは外国人を対象とした研究が多かったが、本研究によって日本人にとっても学習教材として使える可能性が見出されることになる。

2 提案手法

提案手法の流れを説明する。提案手法ではコミックから抽出されたセリフテキストを入力としてとる。1つの吹き出しに含まれるセリフを一単位とする。セリフの中に含まれる単語に対して、fastText¹を用いて作成した単語分散表現モデルを用いて、各単語の分散表現を獲得する。続いて、セリフの中に含まれる単語の分散表現の平均をとり、セリフの平均ベクトルを作成する。最後に、セリフが解説文か非解説文かを分類する学習モデルを作成する。

2.1 入力: 吹き出しの中に含まれるセリフテキスト

提案手法に与える入力吹き出しの中に含まれるセリフテキストである。図1では吹き出しが3つあり、それぞれ「大ぐま座と小ぐま座は1年中同じ空にいるの」「ホントは」「北極星のまわりを離れたまままわってるだけなんだ」というテキストが得られる。この場合、提案手法にはこれら3つのテキストが与えられる。

2.2 単語分散表現モデルの作成

セリフテキストをベクトルとして表現するために、単語分散表現モデルを作成する。本研究では、あるセリ

¹<https://fasttext.cc/>

フテキストが学習可能な知識を含む解説文であるならば、テキストに含まれる単語は非解説文の単語とは異なっていると仮定する。つまり、解説文であれば専門的な単語であったり、会話では使われないような単語が含まれるという意味である。この特徴を単語分散表現を用いて表現し、解説文と非解説文への分類を試みる。

単語分散表現モデルを作成するために、Wikipediaのdumpデータを学習に使用する文書として用いる。Wikipediaのデータからテキストを抽出し、MeCab²、辞書はNEologd³を用いて分かち書きして単語ごとに区切る。さらに、単語ごとに区切られたデータに対してfastTextで学習させることにより、単語分散表現モデルを作成する。

2.3 セリフテキストの平均ベクトルへの変換

セリフテキストを単語分散表現を用いて平均ベクトルへ変換する。一つのセリフテキストの平均ベクトルを得る流れを図2に示す。

セリフテキストごとの平均ベクトルを求めるために、単語ごとのベクトル値を求める必要がある。セリフテキストを形態素解析し、単語に分割する。形態素解析には先ほどと同じく、MeCabとNEologdを用いる。得られた単語をベクトルで表すために、前節で作成した単語分散表現モデルを用いることで、単語ベクトルを獲得する。得られた単語ベクトルをセリフテキスト内に含まれる単語の分だけ足し合わせ、単語数で割ることで1つのセリフテキストの平均ベクトルを得る。

2.4 解説文か非解説文に分類するモデルの作成

セリフテキストを解説文か非解説文に分類するモデルを作成するために学習データを用意する。前節で用意した平均ベクトルに対し、解説文か非解説文かを表すラベルを付与したデータを学習データとして用いる。解説文である場合は1、非解説文である場合は0を付与する。

解説文か非解説文かを表すラベルを付与した平均ベクトルを用いて分類モデルを作成する。分類モデルの作成にはサポートベクタマシンを用いる。

3 提案手法の評価実験

前章で作成した分類モデルとテスト用のセリフテキストを用いて、分類精度を評価する実験を行う。実験

²<http://taku910.github.io/mecab/>

³<https://github.com/neologd/mecab-unidic-neologd/>

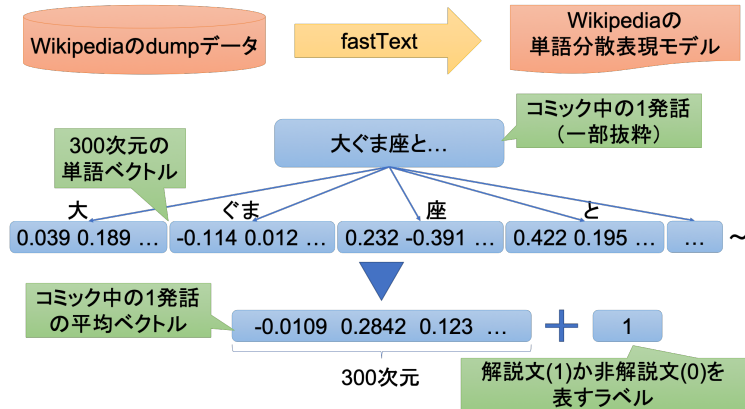


図 2: 単語分散表現を用い、一つのセリフテキストの平均ベクトルを得る流れ

の目的は、入力されたコミックのセリフテキストが解説文か非解説文のいずれかに正しく分類できているかを確認することである。

3.1 実験手順と設定

コミックのセリフテキストを用意し、提案手法を用いて分類し、分類精度を評価した。

本稿では「Dr. STONE 1」 [3] 内の発話を手作業で書き起こしたものをデータとして用いる。全 851 件の発話データは解説文が 44 件、非解説文が 807 件であった。

単語分散表現の次元数は 300 とした。解説か非解説であるかのラベルは人手により与えた。第一著者と第二著者の 2 名が各セリフテキストを読み、常識的知識が含まれるセリフテキストにラベルの 1、そうでないテキストにラベルの 0 を付与した。2 名の意見が一致したラベルのみを採用した。

サポートベクタマシンは python のライブラリである scikit-learn にあるものを用いた。サポートベクタマシンのカーネルは rbf、コストパラメータは 10.0、ガンマ値は 0.1 とした。

解説文と非解説文の数に偏りがあったため、本実験では非解説文を解説文と同じ 44 件ずつに分け、19 個の非解説文からなる文書を作成した。19 個の非解説文の文書にそれぞれ解説文の文書を結合し、19 件の文書を作成した。サポートベクタマシンでの学習と予測においては、1 件の文書を与えるごとに 7 割を学習データ、3 割がテストデータとなるように設定した。学習データ、テストデータの選定はランダムに行われた。

3.2 評価項目

評価する項目は以下の通りである。

- E1 予測したデータ数の内、実際に正解であるデータ数の割合 (正解率)
- E2 解説文であると予測したデータ数の内、実際に解説文であるデータ数の割合 (適合率)
- E3 解説文のデータ数に対し、解説文と予測したデータ数の割合 (再現率)

3.3 実験結果

解説文と非解説文を 44 件ずつ含む文書 data1 から data19 に対して行なった学習モデルによる分類精度結果を表 1 に示す。前節の評価項目について、全 19 回の分類精度の結果からそれぞれ平均を算出した。正解率の平均は 0.83、適合率の平均は 0.81、再現率の平均は 0.85 であった。

4 おわりに

本研究ではコミックのセリフに含まれる学習可能な知識が含まれる文を抽出する方法を提案した。提案手法には、入力として吹き出しの中に含まれるセリフテキストが与えられる。単語分散表現モデルを用いて、入力されたセリフテキストを平均ベクトルへ変換し、得られた平均ベクトルに対し、解説文か非解説文かを表すラベルを付与したデータを学習データを準備する。その後、サポートベクタマシンを用いて分類モデルを作成する。評価実験においてその分類精度を確認したところ常識的知識が含まれる文を分類するモデルの正解率の平均は 0.83、適合率の平均は 0.81、再現率の平均は 0.85 であった。今後は分類に失敗したセリフの考察を行い、学習教材としてのインタフェース開発などが課題としてあげられる。

表 1: 学習モデルによる解説文と非解説文の分類精度

	正解率	適合率	再現率
data1	0.85	0.85	0.85
data2	0.89	0.93	0.87
data3	0.93	0.92	0.92
data4	0.96	1.00	0.92
data5	0.85	0.76	1.00
data6	0.89	0.92	0.86
data7	0.67	0.50	0.89
data8	0.74	0.75	0.69
data9	0.81	0.80	0.73
data10	0.85	0.79	0.92
data11	0.74	0.69	0.75
data12	0.81	0.80	0.86
data13	0.85	0.91	0.77
data14	0.78	0.80	0.80
data15	0.74	0.67	0.73
data16	0.88	0.88	0.93
data17	0.85	0.92	0.80
data18	0.81	0.71	1.00
data19	0.78	0.87	0.87
平均	0.83	0.81	0.85

Words with Comic Scenes, International Symposium on Affective Science and Engineering, 2018.

- [6] Junjie Shan, Yoko Nishihara, and Ryosuke Yamanishi, A System for Japanese Listening Training Support with Watching Japanese Anime Scenes, 22nd International Conference on Knowledge-Based and Intelligent Information & Engineering Systems (KES2018), Procedia Computer Science, Vol.126, pp.947-956, 2018.
- [7] 篠原 宜道, 渡部 広一, 河岡 司, 常識判断に基づく会話意味理解方式, 言語処理学会第 8 回年次大会発表論文集, 2002.

参考文献

- [1] Svetoslav Dankov, Rafal Rzepka, Kenji Araki, Applying a contextual approach for collecting common sense statements to English and Bulgarian, 情報処理学会研究報告自然言語処理 (NL), Vol.113, No.188, pp.105-111, 2008.
- [2] Azuma Fujimoto, Toru Ogawa, Kazuyoshi Yamamoto, Yusuke Matsui, Toshihiko Yamasaki, Kiyoharu Aizawa, Manga109 dataset and creation of metadata, International Conference on Pattern Recognition workshop MANPU (The First International Workshop on coMics ANalysis, Processing and Understanding), 2016.
- [3] 稲垣 理一郎, Boichi, Dr.STONE 1, 集英社, 2017.
- [4] 中原 和洋, 山田 茂雄, 日本でのコモンセンス知識獲得を目的とした Web ゲームの開発と評価, Unisys 技報, Vol. 30, No.4, pp.295-305, 2011.
- [5] Yoko Nishihara, Kohei Matsuoka, Ryosuke Yamanishi, Effects Comparison between English and Chinese Speakers in Learning Japanese Role

絵本における登場キャラクターの行動に着目した感情極性評価 Evaluating Emotion polarity focusing on the character 's behavior in a picture book

伊良原 諒^{1*} 松下 光範¹
Ryo Irahara¹ Mitsunori Matsushita¹

¹ 関西大学総合情報学部
Faculty of Informatics, Kansai University

Abstract: The aim of the research is to develop a search system of picture books that have close transition patterns in terms of the emotional polarity received from the story. Our method calculates and visualizes the transition of the emotional polarity for each character appeared in the story to make the searcher grasp them. Currently, picture book search systems have been developed that use not only the bibliographic information of the book but the content information such as texts and images. However, the transition pattern of the emotional polarities which calculated automatically may differ from the impression the reader received from the story. This is because various characters' actions are mixed in the text of a picture book. To solve the problem, we focus on the action sequence of each character and calculates the transition of the emotional polarity based on the sequences character by character. To verify the validity of the method, we conducted an experiment to evaluate the accuracy of extracted verbs and calculated emotional polarity. Our experiment so far revealed the necessities of estimation of a speaker of each utterance and adjustment of the values of the used emotional polarity dictionary.

1 はじめに

現在、絵本は膨大な数が出版されており、全国出版協会・出版科学研究所の2016年版出版指標年報によると、1年あたりの新刊絵本の冊数は1500冊を超えている[8]。そのため、非常に多くの絵本の中から購入する絵本を選択することが難しい。

現在、絵本選択の手段として、絵本情報サイト「絵本ナビ¹」のように、タイトルや著者名、出版社などの書誌情報を基にした検索や全ページの試し読みなど、絵本のレビューの投稿や閲覧などが利用可能である。

絵本の購入目的には、情操教育のための絵本の読み聞かせ²や、絵本から癒しや心理的变化を受ける[7]ための購入などが挙げられる。そのため、絵本の選択の意図は多岐に渡る。これらの目的に応じた検索を実現するため、絵本の内容情報を基にした検索手法や子供の興味・関心を基にした検索手法など様々な検索の手法が必要であるが、現在利用される絵本検索サイトで

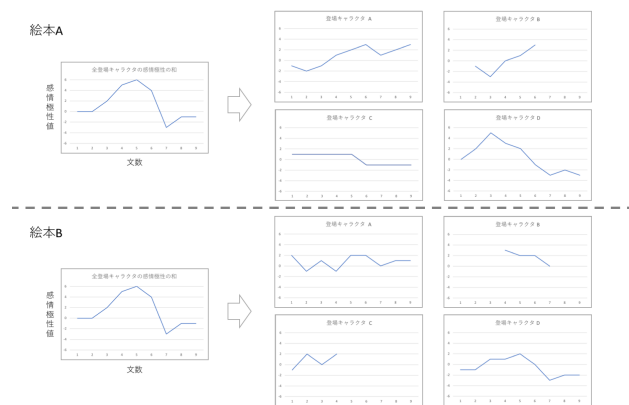


図 1: ストーリーとキャラクターごとの感情極性の例

はこのような検索手法は実現されていない。この問題に対して、意図に沿った絵本を発見するために様々な絵本検索手法が提案されている。

絵本に出現する単語の感情極性の推移を算出することで絵本同士の類似度をはかり、類似絵本を検索を行う手法を検討が行われている。しかしストーリーには複数のキャラクターの感情極性が混在するため、出現するすべての単語から感情極性の推移を算出すると、読み

*連絡先：関西大学総合情報学部
〒569-1095 大阪府高槻市豊山寺町 2-1-1
E-mail: mat@res.kutc.kansai-u.ac.jp

¹絵本ナビ <https://www.ehonnabi.net> (2019/2/14 確認)

²文部科学省 http://www.mext.go.jp/a_menu/shotou/youji/asobou/try/yomikikase.htm (2019/2/14 確認)

手がストーリーから受ける感情極性の推移と異なる結果になることがある。例えば、図1では絵本A,絵本Bのそれぞれにおいて、左側には絵本における全てのキャラクタの感情極性が混合したグラフを示し、右側には絵本における登場キャラクタごとに感情極性を分類したグラフを示している。全てのキャラクタの感情極性を混合したグラフを比較すると絵本Aと絵本Bのグラフは形状が類似していることが読み取れるが、登場キャラクタごとに感情極性を分類したグラフを比較すると、キャラクタごとの感情極性のグラフはそれぞれ異なる形状であることが読み取れる。このようにキャラクタの感情極性が混合している場合、各キャラクタの感情極性が類似していなくても、ストーリー展開が類似していると判定される問題が考えられる。

本研究では各キャラクタに着目したストーリー展開を基に絵本の検索を目指す。これを実現することで、より詳細な絵本の類似度を判定することが可能になる。さらに映画や小説などにおいて、人はストーリーに登場するキャラクタに感情移入することが明らかになっていること [1] から、登場キャラクタごとの感情の流れを把握することで、ユーザがストーリーから受け取る感情極性の推移に近い絵本検索も行えると考えられる。

そこで、本稿ではストーリーを各登場キャラクタの行動に関する記述の連続として表現されるものと仮定し、行動を表す動詞をキャラクタごとに分類して感情極性を付与することでストーリー展開の把握を試みる。

2 関連研究

藤田らは、子どもの興味に合う内容および子どもの年齢に応じた読みやすさの絵本を提案するため、絵本の内容情報を基にした絵本検索システム「ぴたりえ」を提案している [9]。子供の興味に合う内容の絵本を提案するため、グラフ構造を索引とした高速汎用検索アルゴリズムを実装し、特徴量による絵本の検索を可能にしている。これによって検索に厳密に一致する絵本を提案するだけでなく、内容語や書誌情報など様々な要因を用いた検索を実現している。今後の展望として、ハッピーエンドやバッドエンドのようなストーリー展開をもとにした検索や、読み聞かせにかかる時間をもとにした検索の必要性が述べられている。

安尾らは、絵本の中に出現する単語の感情極性に着目し、その推移を検索に用いることで類似した絵本を検索可能にする手法を提案している [11]。テキスト中に出現する感情極性が類似している絵本同士はストーリー展開が類似しているという仮定のもと、ページ単位で感情極性を計数している。しかし、全てのキャラクタの感情極性が混合していることから、立場が対立するキャラクタが存在する場合に両方の極性を取得する

ことが起こり、ストーリーの感情極性評価に影響を与えられられる。

絵本検索に関する研究では、テキストデータや書誌情報などを用いて、単語の難易度やストーリー展開に着目した絵本の検索方法が提案されている。しかし、登場キャラクタごとの感情極性を基にした絵本の検索手法は実現されていない。

ストーリー分析に関する研究として、Reaganらは、小説における感情極性の変化は、大きく分けて6種類に分類できることを明らかにした [3]。この研究では感情極性の変化を把握するため、感情極性の時系列変化を示す曲線であるエモーションナルアークに着目し、ストーリーの分析を行った。

吉田らは、小説全体の雰囲気を感じの観点から捉えた選別を行うため、感情を用いて小説の俯瞰分析を行っている [6]。文章中に現れる感情表現に対して、感情の極性や感情の種類を判断する辞書を使用して文書の感情の変遷を取得している。しかし、抽出された感情表現が誰の感情表現であるか判断できないといった問題や、感情表現辞典においてどの感情の種類に分類するか不明瞭な語句が存在するという問題がある。

以上のようにストーリー分析は文章に含まれる単語の感情極性や感情の種類を評価し、それらの時系列変化を把握することで行われている。これらの手法では、ストーリー全体を俯瞰的に把握することができる。しかし、登場キャラクタごとにストーリーを分析する研究は行われていない。

レビュー文やSNSの投稿、発話などからユーザの感情抽出を行う研究が多く存在する。その際、感情を評価する基準として Positive や Negative 等の極性や R.Plutchik の基本感情モデル [2] のように複数の感情に分類を行い、感情を抽出するものがある。

熊本らは、新聞記事を読んだ人々が感じる印象の記事から抽出するテキスト印象マイニング手法を提案した [10]。印象語辞書を構築し、記事の印象値を算出した。被験者実験を行い、それぞれの印象尺度における印象値を人手で付与した印象値と比較し、それらの平均誤差から提案手法の精度を評価した。

これらのように感情抽出に関する研究では、感情極性辞書や感情表現辞典など、様々な感情モデルを用い、テキストからの感情抽出が試みられている。本稿では、これらの手法から本研究に適した感情抽出手法の選定を行う。

3 提案手法

絵本テキストデータからの、登場キャラクタおよび動詞の抽出・感情極性付与を行う。その手順を図2に示す。

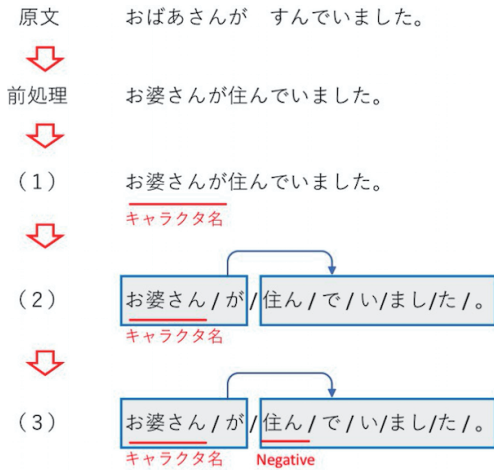


図 2: 提案手法における処理の例

テキストデータの前処理として、鉤括弧で囲まれた文(以下、台詞文と示す)を1つの文とし、その他の文は“。”を基準に1文とする区切りを定めた。例外として、鉤括弧の中に3文以上が含まれている場合、係り受け解析に失敗することがあるため、“|”の中身を“。”で区切り、それぞれを1文とした。また、“～と、赤頭巾は言いました”のように台詞文と結合している後ろに現れる表現は、1文とした。

登場キャラクタとする名詞の選定条件として、(1) 絵本テキストの文において主格となる、(2) 鉤括弧で括られた台詞の発話者となる、を共に満たす名詞とした。条件(2)を選定条件に設定した理由として、例えば、「雨が降る。」という文があった場合、条件(1)のみでは文中の主格となる語が“雨”であるため、登場キャラクタでないものを登場キャラクタとして選定してしまう場合が考えられるためである。

ここで、“一/匹/目/の/子豚”のように複数の形態素で表される登場キャラクタが存在する。これを防ぐため、ユーザ辞書に一意的な登場キャラクタ(以下、一意キャラクタ)を全て登録した。さらに、“三匹の子豚”のような、複数の登場キャラクタをまとめて指す単語(以下、複合キャラクタ)は、ユーザ辞書に登録し、それらが指す各々のキャラクタと動詞を人手で対応づけた。

また同じ表現でも、異なるキャラクタを指している場合がある。例えば、「子豚が食べられました。」、「次の子豚が現れる。」、「子豚が食べられる。」という表現の場合、“子豚”という表現が何匹目の子豚を指示しているのか機械的に判断することが困難であるため、人手で判断し対応付けを行い、係り受け解析を行った。

係り受けから判定した動詞の感情極性を評価するための辞書を、日本語極性辞書(用言編)[5]と、PNtable[4]の中から選定する。選定の条件として、登場キャラクタが係る動詞の感情極性を取得することが必要である

(例) そこへ、一匹目の子豚が逃げてきました。



図 3: 正解データの作成例

ため、登録されている動詞の数が多感情極性辞書が適切だと考え、日本語極性辞書(用言編)とPNtableの動詞登録数の比較を行った。PNtableに含まれる動詞の数は4140語であり、感情極性辞書(用言編)は638語であった。動詞の登録数の多いPNtableを使用する。

PNtableでは、Negativeとして登録されている語数がPositiveとして登録されている語数に対して多いため、PositiveとNegativeの語数が均等に近づくよう、全ての単語の感情極性値の平均が0、分散が1となるz-scoreによる正規化を行い感情極性値として再付与した辞書を作成した。作成した辞書と係り受けで抽出した動詞を用いて感情極性を判定する。

4 実験

提案手法での動詞抽出精度と感情極性付与精度を評価するため、提案手法から得られたデータと、人手で作成した正解データを比較する実験を行った。

登場キャラクタごとに感情極性を付与するためには、登場キャラクタとそれに係る動詞が紐づくことが必要である。そのため、登場キャラクタを主語とする動詞の抽出精度を評価する実験(i)を行う。また、PNtableは単語の感情極性を評価するものであるが、それが絵本の登場キャラクタの感情極性と一致するかを検証する必要がある。そのため、抽出された動詞の感情極性の精度を評価する実験(ii)を行う。実験に使用した絵本と、それぞれの作品に付与したタグ(識別名)を表??に示す。以下ではそれぞれの作品を示す場合、タグによって表記する。

4.1 正解データの作成

実験に用いる正解データの作成について記述する。提案手法での登場キャラクタを主語とする動詞の抽出と抽出した動詞の感情極性が正しく取れているかを判定するために人手でこれを評価し正解データの作成を行った。評価者は、20～23歳の大学生男女9名であっ

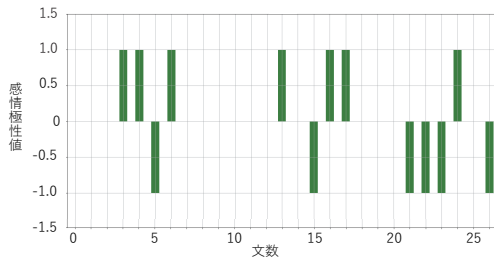


図 4: 感情極性のみの正解データのグラフ

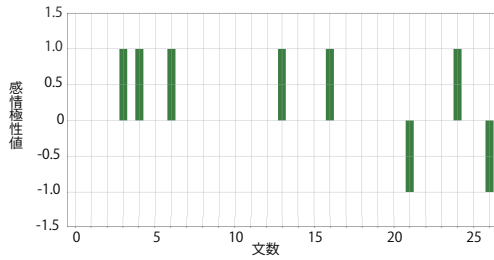


図 5: 本研究で用いる正解データのグラフ

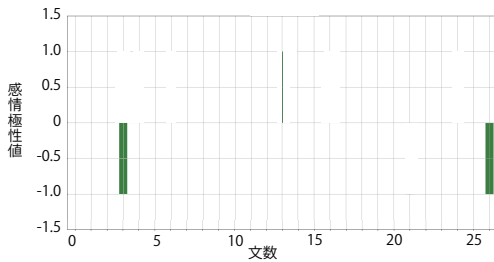


図 6: 提案手法データのグラフ

た．1冊の絵本に対して3人の評価者が評価し，そのうち，2人以上の評価が一致したものを正解データとして採用した．前処理後の絵本テキストデータの各文に対して，動詞部分の抽出と登場キャラクターごとの文の Positive か Negative の感情極性の判定したものを評価とする．動詞部分の抽出では，絵本テキストデータの各文の動詞部分に評価者が線を引き，その行動の主語となる登場キャラクターを記すタスクを課した．感情極性の判定では，評価者がキャラクターにとって動詞が Positive であるか Negative であるかを判定し記述するタスクを課した．正解データの作成例を図3に示す．感情極性を評価するタスクから得られた感情極性値をグラフ化し，その一部を図4に示す．

得られたデータを精査したところ，他のキャラクターの動作が感情極性値に影響を与えている場合があった．そこで本実験では，提案手法がキャラクターを主体とする動作による感情極性の変化を取得できるかを検証する．そのため，人手によって取得した感情極性値のうち，キャラクター自身の動作が起因する変化のみを正解

表 1: 実験 (i) における動詞の抽出率

	Precision	Recall
実験条件 a	0.84	0.31
実験条件 b	0.85	0.38

データとした．その正解データをグラフ化した一部を図5に示す．この処理を1冊分の絵本テキストデータに対して行う．

4.2 動詞抽出精度の比較実験

実験 (i) では，提案手法を用いて抽出した動詞と正解データに含まれる動詞の一致率を求めた．ここで，複合キャラクターのストーリー展開に対する影響を確認するため，一意キャラクターのみを係り受け解析に用いる手法を実験条件 a とし，一意キャラクターと複合キャラクターを用いる手法を実験条件 b として各々の条件下での抽出精度を比較した．実験 (i) における動詞の抽出率を表1に示す．

動詞抽出数を比較した結果，実験条件 b の方が多く抽出を行えた．これは，複合キャラクターの表現が数多くテキスト内に含まれるためであると考えられる．複合キャラクターが登場する作品では，実験条件 b 用いることでより正確に動詞を抽出できることがわかった．

また感情極性の推移を正確に把握するためには，感情極性の付与対象となる動詞を多く抽出することが望ましいが，動詞の再現率は実験条件 a では31%，実験条件 b では38%であり，感情極性の推移の把握に十分とは言い難い．これは，絵本に多く含まれている鉤括弧で括られた台詞文には主語が記述されない場合が多く，台詞文中の動詞と登場キャラクターとを紐づけることができなかつたためだと推察される．

4.3 感情極性付与精度の比較実験

実験 (ii) では，実験 (i) の実験条件 b で正解データと一致した動詞のみを対象に，動詞の感情極性と正解データを比較し，その一致率を評価した．提案手法から得られたデータをグラフ化した一部は図6のようになった．対象となる動詞の表記を PTable に登録されている表記と一致させる処理を行った．その結果，対象とした動詞の41%は PTable に登録されていない動詞であったため実験対象から除外した．実験 (ii) における感情極性の一致率を表2に示す．実験結果から，PTable に登録されている感情極性の多くは正解データの感情極性と異なることが明らかになった．特に正解データ

表 2: 実験 (ii) における感情極性の一致率

		提案手法	
		Positive	Negative
人 手	Positive	0.07	0.41
	Negative	0.01	0.13
	Even	0.04	0.32

では Positive であるが、提案手法では Negative である場合が多くみられた。

5 発話者の推定

本稿では、テキストから動詞を抽出する際、登場キャラクターがその動詞の主語であることを動詞の抽出条件とした。しかし、実験において鉤括弧で括られた台詞文中では、主語が省略されることが多いため、動詞の抽出制度が低くなることが明らかになった。そのため、台詞文において主語が省略される場合、それを補うことにより動詞の抽出精度を向上させることができると考え、これを実装した。

台詞文における主語を推定するにあたり、平常文の台詞で省略される主語は発話者であることが多いことに注目する。絵本において“(主語)が言いました。”という表現では、この文の「言う」の前後に現れる台詞文の発話者は“(主語)”であると推定できる。また、“(主語)は～と言いました。”という表現では、“～”の主体は“(主語)”であると推定できる。このことから、「言う」という表現に着目し、発話者の推定を試みる。これを実現するため、本節では、台詞文の発話者推定の手法を考察し、動詞の抽出制度の向上を図る。

実装の条件として、1. テキストの地の文において「言う」とその主語を抽出できたとき、その文の前後に台詞文があること、2. 「言う」の直前に「そう」が含まれること、3. 1. で抽出した主語が台詞文中に含まれないことの3点を定めた。

条件 1. は、上に述べた着目点に沿って定めた。条件 2. は“桃太郎はそう言いました。”という文において「そう」は前の台詞文を指すため、台詞中の“(主語)”から前の台詞文の発話者の推定が可能であると考え定めた。条件 3. は、発話者自身の名前が台詞中に含まれることは自己紹介をする際など限定的な場面であり、頻度が少ないと考え定めた。

そのため、条件 1. から 3. の全てに当てはまる場合は「言う」の主語を、その前の台詞文の発話者とし、条件 1. と 3. に当てはまる場合は「言う」の主語を、前後の台詞文のうち、発話者の名前を含まない台詞文の発話者とした。

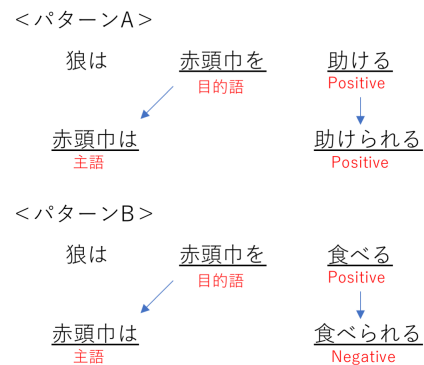


図 7: 目的語に対する感情極性付与のパターン

これらの条件に沿って、台詞文の発話者の推定を行い、人手による発話者の推定結果と比較した。いくつかの文では、人手による判断と一致する発話者を推定することができ、発話者推定が動詞の抽出精度の向上に繋がること示唆された。しかし、複数の台詞が連続した場合や一つの台詞文に複数の推定結果が重複した場合に正しく推定できないことがわかった。今後は発話が行われている場面に、どのキャラクターが存在するのかに着目することで本稿で推定できなかった発話者の推定が可能になると考えられる。これらの処理を行うことで、感情極性の精度が向上すると考えられる。

6 今後の展望

6.1 キャラクターの関係性

図 4 に示した感情極性のグラフは、ぶた 1 に登場する狼の感情極性を表したものである。このグラフは、読み手が感じるストーリー中のキャラクターの感情極性を最も反映していると考えられる。これは他の登場キャラクターの動作が、そのキャラクターの感情極性に影響を与えることを加味した評価を行ったためである。しかし、正解データと提案手法から得られたデータはそのキャラクターのみの動作から感情極性を評価したものであった。今後、キャラクターの関係性を考慮し組み合わせることで、図 4 のようなグラフに近づくと考える。そのため、提案手法においても、キャラクターごとの関係性を考慮する必要がある。

6.2 感情極性の評価

実験から、登録されている単語や感情極性値は人手の判断と異なることが示された。辞書に登録されていない単語があることや登録されている感情極性値が異

なることが原因であると考えられる。そのため、辞書に登録されていない単語の登録や、人手による判断と異なる感情極性の単語の特徴を調査し辞書の感情極性の変更を行う必要がある。それに加えて本節では(1) 目的語のキャラクタ、(2) 動詞以外の品詞に着目して、感情極性の評価の向上が期待される処理を述べる。

(1) に関して、提案手法では、登場キャラクタが主語である場合に紐づく動詞に対して感情極性の付与を行った。しかし、正解データを分析した結果、目的語にある登場キャラクタも感情極性付与の対象となっていた。そのため、登場キャラクタが目的語として出現する場合に対しても感情極性の評価を行うことで精度が向上すると期待される。図7は、動詞に紐づく目的語を主語とし、動詞の部分を受け身の表現にする処理を示している。パターンAは、主語の登場キャラクタと目的語の登場キャラクタに係る動詞の感情極性が同じである。パターンBは、主語の登場キャラクタと目的語の登場キャラクタに係る動詞の感情極性が異なる場合である。このように目的語のキャラクタを主語に置き換えた場合、動詞は受け身の表現になると考えることができる。パターンA、パターンBそれぞれに対応する受け身の表現の単語を辞書に登録する。それにより登場キャラクタが目的語として出現する場合に紐づく動詞に対して、感情極性の付与が行えるようになり、感情極性評価の精度の向上が見込める。

(2) に関して、提案手法では動詞から感情極性を評価していたが、動詞の前の記述によって動詞単独の感情極性評価とは異なる評価がされるものがある。例えば、“思う”という単語は、動詞単独の感情極性評価ではNegativeと判断されているが、“思う”の前に“嬉しく”が記述されていた場合、Positiveと評価されることが考えられる。このような語に考慮した感情極性評価を行うためには、動詞以外の品詞と組み合わせることで人手の判断に近い感情極性が取得できると考えられる。

7 おわりに

本稿では、登場キャラクタごとの行動に着目した感情極性の評価手法を検討した。この手法の有用性を検証するために、提案手法による動詞抽出の精度と、付与される感情極性の精度の評価を行った。今後、感情極性辞書にない単語の登録を行うこと、及び人手による判断と異なる感情極性の単語の特徴を調査し辞書の感情極性の変更を行うことにより、感情極性評価の精度向上に取り組む。また今回の手法では、登場キャラクタが主格である場合のみを感情極性の評価対象としたが、登場キャラクタが目的格である場合も考慮した手法を検討し、読み手の受ける感情極性により近いストーリー把握を目指す。

参考文献

- [1] Davis, M. H.: A Multidimensional Approach to Individual Difference in Empathy, *JSAS Catalog of Selected Documents in Psychology*, Vol. 10, p. 85 (1980).
- [2] Plutchik, R.: *The Emotions*, University Press of America (1991).
- [3] Reagan, A. J., Mitchell, L., Kiley, D., Danforth, C. M. and Dodds, P. S.: The emotional arcs of stories are dominated by six basic shapes, *EPJ Data Science*, Vol. 5, No. 1, p. 31 (2016).
- [4] Takamura, H., Inui, T. and Okumura, M.: Extracting Semantic Orientations of Words Using Spin Model, *Proceedings of the 43rd Annual Meeting on Association for Computational Linguistics*, ACL '05, Stroudsburg, PA, USA, Association for Computational Linguistics, pp. 133–140 (2005).
- [5] 小林的ぞみ, 乾健太郎, 松本裕治, 立石健二, 福島俊一: 意見抽出のための評価表現の収集, *自然言語処理*, Vol. 12, No. 3, pp. 203–222 (2005).
- [6] 吉田知世, 小林一郎: 感情極性に基づく文書の俯瞰分析への取り組み, 第73回全国大会講演論文集, Vol. 2011, No. 1, pp. 387–388 (2011).
- [7] 笹倉剛: 絵本を中心としたビブリオセラピーに関する研究, *神戸親和女子大学言語文化研究*, Vol. 9, pp. 21–39 (2015).
- [8] 全国出版協会・出版科学研究所: 2016年版出版指標年報 (2016).
- [9] 藤田早苗, 服部正嗣, 小林哲生, 奥村優子, 青山一生: 絵本検索システム「ぴたりえ」~子どもにぴったりの絵本を見つけます~, *自然言語処理*, Vol. 24, No. 1, pp. 49–73 (2017).
- [10] 熊本忠彦, 河合由起子, 田中克己: 新聞記事を対象とするテキスト印象マイニング手法の設計と評価, *電子情報通信学会論文誌. D, 情報・システム*, Vol. 94, No. 3, pp. 540–548 (2011).
- [11] 安尾萌, 服部正嗣, 藤田早苗, 松下光範: 物語の類型に着目した絵本の類似探索手法に関する一検討 (ヒューマンコミュニケーション基礎), *電子情報通信学会技術研究報告*, Vol. 116, No. 436, pp. 103–108 (2017).

選択肢により痛みの表現を支援する タッチパネル型インタフェース

Touch Panel Interface to Support Explanation of Pains by Choosing Options

岡部裕太^{1*} 西原陽子¹ 山西良典¹

¹ 立命館大学情報理工学部

¹ College of Information Science and Engineering, Ritsumeikan University

Abstract: This paper proposes a touch panel interface to support explanation of pain by choosing options. Medical staffs of rehabilitation must pain information of patients. Patients would explain their pains with words and gestures. The explanation of pain depends on expression of a patient. If a patient does not explain his/her pains accurately, a medical staff may not have enough treatment to the patient. The proposed interface provides patients as interface users with means to explain their pains. The users input four types of pain information: the occurring condition, the body part, the aspect, and the degree of pain. The four types of information can be input by choosing options. The interface users can explain their pains independently on their vocabulary. We had an experiment to verify the efficiency of the proposed interface. We found that the proposed interface could support users to explain their pains.

1 はじめに

リハビリテーションの現場において、医療者が患者の痛みの情報を正確に知ることは、適切な施術を行う上で重要である。リハビリテーションを行う医療者は、痛む時の条件、痛む場所やその内容、および程度を患者から聞き取り、患者の希望する動作が行えるようになることを目指して施術を行う。痛みの内容を正確に伝えることができるほど、適切な施術を行うことができる。

患者の痛みは患者にしかわからない主観的な情報である [3]。したがって、痛みの情報が伝えられるかどうかは患者の表現力や語彙力に依存すると考えられる。例えば、表現力がまだ乏しい幼児であると、自分の感じている痛みを大人が表現するほどには適切に表現することはできないと考えられる [8]。また、成人であっても方言を主に使用している場合、方言を解さない医療者にはその情報は伝わりにくいと考えられる [7]。表現力や語彙力によらず痛みを適切に表現可能な仕組みが求められている。

そこで本研究では自由記述ではなく選択肢を用い、表現力や語彙力に依存しない痛みを表現するための入力インタフェースを提案する。入力する項目はリハビリ

テーションの施術前に医療者が患者に質問する項目とし、痛みが生じる条件、痛みが生じる部位、痛みの内容、痛みの程度の4項目とする。各項目の入力に適切な選択肢を用意し、インタフェース上に示す。インタフェースのユーザは選択肢を選択することにより、自らが抱える痛みを表現できるようになる。

既存研究においてがん性疼痛を把握するためのシステムや [4]、押しボタンにより痛みの記録を簡便に行うことができるシステム [1] などが提案されている。リハビリテーションに関わる医療者は患者が日常生活を送る上での不具合、できるようになりたいことをサポートするため、痛いというだけでは適切な施術を行うことが困難になる。本研究ではリハビリテーションの場での利用を想定し、痛みの程度だけでなく、痛みが生じる条件、痛みが生じる部位、痛みの内容、痛みの程度の4項目を合わせて記録するものを提案する。

2 痛みの表現を支援するインタフェースの提案

痛みの表現を客観的にかつ簡単に行うことができるインタフェースを提案する。

*連絡先：立命館大学情報理工学部メディア情報学科
〒525-8577 滋賀県草津市野路東 1-1-1
E-mail: nisihara@fc.ritsumeikai.ac.jp



図 1: 痛みの記録を表示する画面

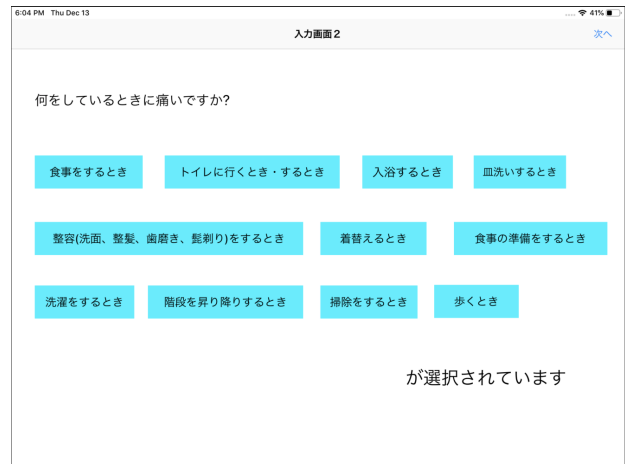


図 2: 痛みが生じる条件を入力する画面

2.1 提案するインタフェースの構成

提案するインタフェースは3つの画面から構成される。3つの画面を図1, 図2, 図3に示す。

インタフェースのユーザは初めに図1の画面が与えられる。図1の画面では、過去に記入した痛みの記録を時系列順に表示する。ユーザは右上の「追加」を選択することにより図2の画面に移動できる。図2の画面では、痛みが生じる条件を入力することができる。条件は日常生活をする中で痛みが生じる条件を選択する。条件を入力した後で右上の「次へ」を選択することにより図3の画面に移動できる。続いて、図3の画面では、痛みの生じている部位、痛みの内容、痛みの程度を入力できる。全ての項目を入力した後で右下の「データ追加」を選択することによりデータが保存される。

本研究では macOS 10.13.6, Xcode 10.1 上で Swift 4.2.1 を用いてインタフェースを実装した。データの保存には UserDefaults を利用した。図2と図3の画面で入力される項目について詳細を説明する。

2.2 痛みが生じる条件の入力

リハビリテーションを受ける患者の多くは日常的に行う動作が行えなかったり、動作を行おうとすると痛みが生じるため動作ができないなどの困難を抱えている。何もしていない状態では痛みを感じないが、ある動作を行おうとすると痛みが生じることがあるため、それを解消することを希望する方が多い。そこで、どのような動作を行おうとすると痛みが生じるかを回答してもらう。これにより患者の痛みの状態を把握することができる。加えてリハビリテーションを担当している施術者が異なった場合にも、施術者間で大きく異なる施術を実施することにもつながる。

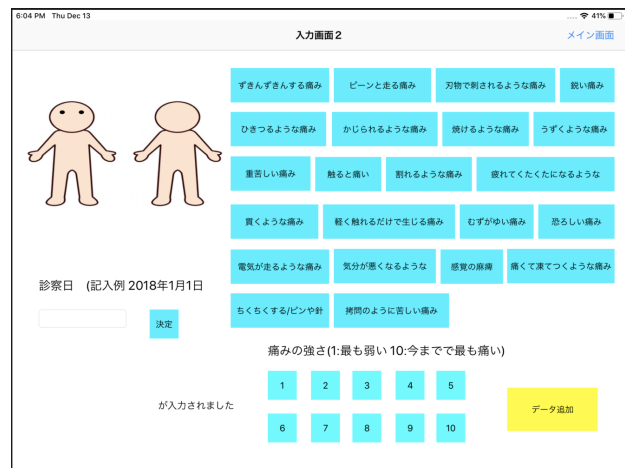


図 3: 痛みの生じている部位、痛みの内容、痛みの程度を入力する画面

インタフェースのユーザは、選択肢を選択することにより痛みが生じる条件を入力する。選択肢を図2に示す。選択肢は11個ある。これらの選択肢は Activity of Daily Living (ADL) および Instrumental Activity of Daily Living (IADL) にまとめられた中から用意された [6]。ADL は食事・更衣・移動・排泄・整容・入浴など生活を営む上で不可欠な基本的行動を指す。図2では食事をするとき、トイレに行くとき・するとき、入浴するとき、整容(洗顔、整髪、歯磨き髭剃り)をするとき、着替えるとき、階段を昇り降りするとき、歩くときが該当する。IADL は道具を使って行われる動作で、電話使用・調理・買い物などの行動を指す。図2では皿洗いする時、食事の準備をするとき、洗濯をするとき、掃除をするときが該当する。

2.3 痛みが生じる部位の入力

前の画面で選択した痛みの生じる条件に対応し、痛みが生じる部位を入力する。インタフェースのユーザは、人の全身を示した画像の前面と背面の2つを見て、痛みが生じる部位を1つ選択し、入力する。痛みが生じる部位を画像の1点を指し示すことにより、言語で表現するよりも簡単に表現することが可能となる。部位は42箇所に分かれている。1箇所ごとの区切りはタッチパネルで選択できるボタンの大きさにより定められている。

2.4 痛みの内容の入力

痛みが生じる部位の選択と同様に、前の画面で選択した痛みの生じる条件に対し、痛みの内容を入力する。痛みの内容を言葉で表現すると、その適切さや粒度は患者の語彙力に左右される。また、これまでに経験したことがない痛みであると、その内容を適切に表現することは難しい。ただ単に「痛い」としか表現できない可能性もある。

痛みの内容の表現を容易にすること、そしてリハビリテーションの施術者が痛みの内容の理解を容易にすることを目的とし、痛みの内容を客観的に評価する手法が複数提案されている [2]。本研究では、最も広く用いられている手法である MacGill Pain Questionnaire [5] を用いて痛みの内容の入力を支援する。

MacGill Pain Questionnaire には全部で 22 種類の痛みの表現が記載されている。22 種類の表現を図 3 に示す。

2.5 痛みの程度の入力

前の画面で選択した痛みの生じる条件に対し、痛みの程度を入力する。インタフェースのユーザは、1 から 10 までの数字の 1 つを選択することにより、痛みの程度を入力する。数字による表現はリハビリテーションにおける痛みの程度の確認において頻繁に用いられている方法である。10 を「今までの最も強い痛み」とし、1 を「今までの最も弱い痛み」とする。

3 提案インタフェースの評価実験

提案インタフェースを評価する実験を行なった。実験の目的は、提案インタフェースにより痛みの表現が支援できているかを確認することであった。

3.1 実験手順

実験は次の手順に従い行われた。

1. 被験者を 2 つの群に分け、それぞれを実験群と統制群とする
2. 実験者は被験者に過去に経験した痛みを 5 つ思い出すことを依頼する
3. 被験者はインタフェースを用い、経験した痛みについて入力する。
4. 実験者は入力が終了するまでの時間を計測する。

被験者は 21 歳から 22 歳までの男性 14 名であった。実験群を 7 名、統制群を 7 名とした。

実験群が用いたインタフェースは提案インタフェースであった。統制群が用いたインタフェースは、提案インタフェースで入力する項目を紙に自由記述させるものであった。以下、比較インタフェースと呼ぶ。

3.2 評価項目

本実験により得られるデータは、14 名の被験者がそれぞれ 5 件ずつ入力した合計 70 件の痛みの表現、および入力開始から終了までの時間の 2 種類であった。この 2 種類のデータを用いて、以下の項目を評価することにより提案インタフェースの評価を行った。

1. 入力開始から終了までにかかる平均時間
2. 提案インタフェースで利用された選択肢の割合
3. 比較インタフェースで記述された単語の種類数

3.3 実験結果

各被験者が入力開始から終了までに要した時間を表 1 に示す。提案インタフェースを用いた被験者は平均 471 秒、比較インタフェースを用いた被験者は平均 626 秒を要した。平均値の差の検定を行ったところ 2 つの平均値の差に有意な差があることがわかった ($t=-2.45$, $p=0.03 < 0.05$)。

提案インタフェースを用いた被験者が入力した例を表 2 と表 3 に示す。被験者により選択された条件の選択肢は 11 個中 9 個で、割合は 82% であった。痛みが生じる部位の選択肢は 42 個中 14 個が選択された。割合は 33% であった。痛みの内容の選択肢は 22 個中 11 個が選択され、割合は 50% であった。痛みの程度の選択肢は 10 個中 9 個が選択され、割合は 90% であった。

比較インタフェースを用いた被験者が入力した例を表 4 と表 5 に示す。痛みが生じる条件は 24 種類であっ

表 1: 実験時間

提案インタフェース	333 秒	897 秒	361 秒	468 秒	497 秒	356 秒	388 秒	平均 471 秒
比較インタフェース	751 秒	713 秒	727 秒	488 秒	520 秒	907 秒	801 秒	平均 626 秒

表 2: 提案インタフェースを用いた被験者の入力結果 (1)

条件	部位	内容	程度
食事をするとき	顔	鋭い痛み	8
トイレに行く時・するとき	下半身	拷問のように苦しい痛み	7
入浴するとき	左太もも	ずきんずきんする痛み	4
歩く時	左足	痛くて凍てつくような痛み	10
整容 (洗顔, 整髪, 歯磨き, 髭剃り) をするとき	顔	鋭い痛み	2

表 3: 提案インタフェースを用いた被験者の入力結果 (2)

条件	部位	内容	程度
皿洗いするとき	下半身 (腰, お尻)	ピーンと走る痛み	6
入浴する時	手 (右)	痛くて凍てつくような痛み	5
トイレに行く時・するとき	下半身	焼けるような痛み	6
食事をするとき	顔	触ると痛い	3
歩く時	左足	うずくような痛み	2

表 4: 比較インタフェースを用いた被験者の入力結果 (2)

条件	部位	内容	程度
野球練習中	右腕の手首	ポーンと痛い	手を上げられないくらい
日常生活ずっと	歯	キンキン	歯を叩きたくなるくらい, 歯を叩いて他の痛みでごまかしたくなるくらい
歩く時	太もも	ぶつけた時のいたさ	普通くらい
立っている時	股間	キーン	うずくまって何もできないくらい
ずっと	腹	ギョルギョル	トイレから出られない, すこしうずくまるくらい

表 5: 比較インタフェースを用いた被験者の入力結果 (2)

条件	部位	内容	程度
膝を曲げる時	右膝の靭帯	ズキズキくるような痛み	曲げるのが苦痛に感じる程度
右手で全ての作業をするとき	右手小指の第一関節	ジーンとくるような痛み	血の気が引くくらい
歩く時	右ふくらはぎ	ずっとつっているような痛み	動かす時に違和感を覚える程度
しばらく立っている時に急にかがんだとき	腰	ジーンととるような痛み	耐えられるが痛いと言に出してしまう程度
歩く時	すね	ズキズキくる痛み	すねの一部が腫れ上がり歩行する時に引きずる程度

た。痛みが生じる部位は 23 種類であった。痛みの内容についてはオノマトペを用いて表現していることが多く、35 個中 20 個でオノマトペが使用されていた。割合は 57%であった。痛みの程度については「できない」や「できないくらい」という表現を用いていることが多かった。35 個中 10 個が該当した。割合は 29%であった。

4 結果に対する考察

得られた実験結果を用いて 3.2 に示した評価項目について評価をし、考察を行う。

4.1 入力所要時間に対する考察

表 1 に示した入力所要時間は、提案インタフェースを用いた実験群が平均 471 秒であり、比較インタフェースを用いた統制群が平均 626 秒であった。提案インタフェースを用いた実験群の方が $626 - 471 = 155$ 秒だけ早かった。提案インタフェースを用いた実験群はすべての入力を選択肢を選択することにより、痛みの表現を考える必要があまりない。一方で比較インタフェースを用いた統制群は全ての入力を言葉で表現するため痛みの表現を考える必要がある。表現を考える時間が短縮されたため、入力所要時間が短縮されたと考えられる。この結果から、提案インタフェースを用いることで痛みの表現を短時間でできることがわかった。

4.2 提案インタフェースで用いられた選択肢の割合に対する考察

提案インタフェースでは痛みが生じる条件、痛みが生じる部位、痛みの内容、痛みの程度の 4 つの項目を選択

肢により入力する。それぞれの項目に対して用いられた選択肢の割合を調べたところ、痛みが生じる条件は 82% (= 11/9)、痛みが生じる部位は 33% (= 14/42)、痛みの内容は 50% (= 11/22)、痛みの程度は 90% (= 9/10) であった。各被験者は過去の経験を思い出すことにより痛みの表現を行なった。自分の経験が該当する選択肢がないと回答した被験者はいなかった。この結果から、患者が抱える可能性がある痛みに対して表現を支援することができることがわかった。

4.3 比較インタフェースで記述された単語の種類数に対する考察

比較インタフェースでは痛みが生じる条件、痛みが生じる部位、痛みの内容、痛みの程度の 4 つの項目を自由記述により表現する。それぞれの項目に対して種類数を調べたところ、痛みが生じる条件は 24 種類、痛みが生じる部位は 23 種類であった。痛みの内容については 57% (= 20/35) の回答でオノマトペを使用していた。痛みの程度については 29% (10/35) の回答で「できない」「できないくらい」という表現を用いていた。

痛みが生じる条件については第三者間で認識が大きく異なるような表現は見られなかった。「歩く時」「立っている時」などの表現があった。提案インタフェースで用意した選択肢とは表現が若干異なっていたが、条件の判定が人により異なることは少ないと考えられる。このことから、条件の項目については患者のボキャブラリーに左右されにくい項目であるということが示された。

痛みが生じる部位については、条件と同様に、第三者間で認識が大きく異なるような表現は見られなかった。「右腕の手首」「左足」などの表現があった。これらの表現の認識が人により異なることは少ないと考え

られる。このことから、部位の項目について患者のボキャブラリーに左右されにくい項目であるということが示された。

痛みの内容については、オノマトペを用いて表現する被験者が多かった。用いられたオノマトペは、ズキズキ(11回)、ジンジン(4回)、ジーン(3回)、ヒリヒリ(2回)であった。被験者が自由記述において痛みの内容をオノマトペで表現した理由としては、痛みを感じる時にどんな風に痛いと言われて回答する際、オノマトペを普段から用いることが多いためと考えられる。ある痛みに対して適切なオノマトペを当てはめられるかどうかは、患者さんのボキャブラリーに依存する。幼い子供や日本語を母語としない人にとってはオノマトペを用いた痛みの内容の伝達は困難である可能性がある。今回の被験者は成人男性であったためオノマトペの利用は適切に行われたと考えられる。今後、オノマトペが適切に利用できない可能性がある被験者に対しても実験を行い、痛みの内容の表現に対する考察を進めていく。

痛みの程度については、「できない」「できないくらい」を用いて表現する被験者が多かった。ある動作ができないという表現は数値表現ができず、順位づけをすることが困難である。また、リハビリテーションの施術により回復してきた場合でも、どの程度回復したのかを数値的に評価することが困難と考えられる。数値により表現を試みることで、それぞれの痛みの順位づけが行える可能性、回復の程度を評価できる可能性が明らかになった。

5 おわりに

本研究では選択肢により痛みの表現を支援するタッチパネル型インタフェースを提案した。提案したインタフェースでは、痛みが生じる条件、痛みが生じる部位、痛みの内容、痛みの程度を選択肢の選択により入力させる。リハビリテーションの施術を受ける患者さんに入力してもらうことで、施術者が施術前に患者さんの痛みの状態やどのような動作ができるようになりたいのかの情報を把握するために利用してもらうことを想定した。提案インタフェースを用いた評価実験を行ったところ、入力所要時間が短縮されること、痛みの表現が適切に行えることが明らかになった。今後は提案したインタフェースを実際のリハビリテーションの場で用いていただき、施術に対して得られる効果を検証していくことが課題である。

謝辞

本研究を進めるにあたり、岡橋さやか先生、鎌田真由美先生、和田拓弥先生に貴重なご意見をいただきました。羽鹿諒氏にはインタフェースの実装においてアドバイスをいただきました。記して謝意を申し上げます。

参考文献

- [1] 深谷 陽子, 安藤 詳子, 稲垣 聡美, 宮崎 雅之, 中村 みゆき, 澤井 美穂, 野田 幸裕, 神里 みどり, 簡便な操作で痛みの強さを記憶する痛み計の臨床試行, *Palliative Care Research*, Vol.1, No.1, pp.201-205, 2006.
- [2] 濱口 眞輔, 痛みの評価方法, *日本臨床麻酔学会誌*, Vol.31, No.4, pp.560-569, 2011.
- [3] 上田 祐也, 清水 祐一郎, 坂口 明, 坂本 真樹, オノマトペで表される痛みの可視化, *日本バーチャルリアリティ学会論文誌*, Vol.18, No.4, pp.455-463, 2013.
- [4] 平澤 宏卓, 安藤 詳子, 光行 多佳子, 山本 陽子, 黒田 宏美, 山田 和正, 大日方 五郎, がん性疼痛アセスメントにおけるタブレット PC 版「ペインメモリー」の有用性の検討, *Palliative Care Research*, Vol.10, No.1, pp.312-317, 2015.
- [5] Melzack R, The McGill Pain Questionnaire, Major properties and scoring methods. *Pain* 1, pp.277-299, 1975.
- [6] 角 徳文, ADL・IADL の評価尺度, 総合リハビリテーション, Vol.45, No.8, pp.853-855, 2017.
- [7] 山本 悠人, 北村 達也, 岩城 裕之, 今村 かほる, 身体部位を表す方言語彙の表示システム, *日本語教育方法研究会誌*, Vol.22, No.3, pp.56-57, 2016.
- [8] 山下 利之, 近藤 真悟, 岡崎 章, 小児看護におけるプレパレーションのための痛み評価ツールの開発, *人間工学*, Vol.51, No.6, pp.435-440, 2015.

外食・惣菜がメインでも栄養バランスを整える 食品推薦システムの開発

Development of a Food Recommendation System for Supporting Nutrient Balance

大畑 拓己^{1†} 西原 陽子¹ 山西 良典¹
Takumi Ohata, Yoko Nishihara, Ryosuke Yamanishi

¹ 立命館大学 情報理工学部

¹ College of Information Science and Engineering, Ritsumeikan University

Abstract: Having foods considered nutritional balance is important for both physical and mental health. Due to the change of lifestyle, people often have lunch and dinner at restaurants and buy foods at markets. The nutritional balances of those tend to be not good. This paper proposes a food recommendation system for supporting the nutrient balance. Users of the proposed system input logs of foods that they have ate. The proposed system calculates intake nutrients and energies, then recommends foods that support to adjust the nutrient balance. We had evaluation experiments with the proposed system. The proposed system could support users to adjust the nutrient balance.

1 はじめに

外食産業やファストフード、コンビニエンスストアの発展に伴い、その手軽さゆえに多くの人々が外食や持ち帰り惣菜(以下、中食という)を大いに活用している。食の選択肢が広がっていくとともに、多様な食品の中から栄養素のバランスのとれた食品を選択することは心身の健康を保つために大切である。しかし、食品の購入の際にエネルギー以外の栄養素を意識しているものは少ない [1]。

栄養所要量 [2] という、厚生労働省による性別・年齢・運動レベルによって決まる栄養素の摂取目安がある。その中でも最も大切と言われている三つの栄養素はタンパク質・脂質・炭水化物の三大栄養素である。三大栄養素の偏りは脳の働きが鈍りや、体力の低下などの健康に悪い影響が及ぼす。現代の日本では、炭水化物が減少し、脂質が増加する栄養摂取パターンに接近している傾向がある [3]。

外食・中食を食べるのは男性の一人暮らしの人の割合が高くなっている [1]。そして、男性で栄養成分を気にしている割合は、女性の気にしている割合の半分である [4]。さらに、栄養成分を気にしている人の多くは「エネルギー」を参考にしている人がほとんどで、三大栄養素はあまり見られていない。その理由は、摂取す

べき三大栄養素の摂取目標を把握していないということであると考えられる。また、自分の摂取した三大栄養素を把握するためには食べたものから栄養素の計算をする必要があり、その作業は手間がかかり面倒なため、継続が難しいと考えられる。

レコーディングダイエットというログを記録するダイエット手法があるように、食事を記録することは健康に影響を与える [5]。1日単位で栄養素を計算して把握している人は少ないため、1日単位で栄養素の過不足の可視化をすることは栄養素のバランスの改善を支援するのではないだろうか。そして、計算した栄養素の値をもとに、栄養素のバランスを整える食品を推薦することは栄養素のバランスの改善を支援するのではないだろうか。

食事を推薦するシステムが数多く提案されているが [6][7]、その多くは自炊をして生活している人を対象としたシステムである。しかし、実際には外食・中食を活用して生活している人が多いため、外食・中食をメインとして生活する人を対象とした食品推薦システムは有用であるのではないかと考える。

外食産業では近年、各企業のウェブサイト上で、販売している食品の栄養素情報を公開する企業が増えていく。しかし、栄養素の情報を見ようとすると各企業のウェブサイトを訪問しないと見れないため、面倒である。そこで、各店舗の栄養素を統合したデータベースで管理できると栄養素情報参照の手間が省けるので

†連絡先: 立命館大学情報理工学部
〒525-8577 滋賀県草津市野路東 1-1-1
E-mail: is0298pv@ed.ritsumeik.ac.jp

はないかと考える。本研究では、外食食品データベースを利用し、栄養素の過不足を可視化することや、栄養バランスを整える食品を推薦することにより、栄養所要量の摂取目標を満たす食生活を送ることを支援することを目的とする。

2 提案するシステムのデザイン指針

提案するシステムの目的は栄養素の過不足の指摘や、栄養バランスの補正を行える食品の推薦により、栄養所要量の摂取目標を満たす食生活を送ることを支援することである。

これを達成するためには、以下の機能が必要と考えられる。1つ目は、食事のログ(以下、**食事ログ**という)の入力できる機能である。2つ目は、食べたものの栄養素の可視化する機能である。3つ目は、栄養バランスを整える食品の推薦する機能である。

2.1 食事ログを入力する機能

栄養素の過不足の指摘や栄養バランスの補正をする食品を推薦するためには、ユーザの食べたものの記録を入手する必要がある。そこで、ユーザに食事ログを入力してもらうための機能を提供する。

2.2 栄養素を可視化する機能

栄養素の過不足を指摘するためには、ユーザが食べたものの栄養素を評価し可視化する必要がある。評価する栄養素は三大栄養素とし、可視化する方法はレーダーチャートとする。正しい栄養バランスであれば、レーダーチャートの形が正三角形となる。このため、一目で栄養素の過不足がわかる。

2.3 食品を推薦する機能

栄養バランスの補正を行うために食品を推薦する。推薦する食品としては朝食、昼食、夕食のいずれかが考えられる。どの食事を推薦するかを決定するために、著者が所属する大学の大学生の中で一人暮らしをしている者にアンケートを行った。アンケートでは、1週間の中でどの食事で外食または中食を摂取するかを問うた。回答者は15名であった。アンケートの結果、最も外食・中食をとる回数は「昼食」が最大となった。この結果は既存研究の結果とも一致する[8]。このことから、提案システムでは昼食に摂取する食事を推薦することを決定した。

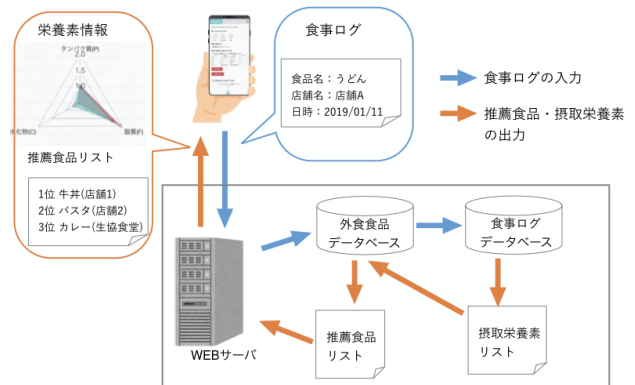


図1: 提案システムの概要

提案システムでは前日の夕食と当日の朝食の栄養素を算出し、それを元にし昼食で摂取する食品を5つ推薦する。前日の夕食と当日の朝食の栄養素のバランスが摂取目標を満たしている場合は、三大栄養素の摂取目標の中央値に近い食品を5つ推薦する。前日の夕食と当日の朝食の栄養素のバランスが摂取目標を満たしていない場合は、三大栄養素のうち最も摂取目標から外れている栄養素を補正する食品を推薦する。

3 提案システムの実装

提案システムはブラウザ上で動作するウェブアプリケーションとして実装する。提案システムの概要を図1に示す。ユーザはブラウザからWebページを開き、Webページ上で食事ログを入力することにより、食品の推薦を受けることができる。

3.1 食事ログの入力

食事ログは、ユーザID、店舗名、食品名、サイズ、食事日時、データ作成日時、エネルギー(kcal)、タンパク質(g)、脂質(g)、炭水化物(g)、の情報からなるとする。食事ログを入力する方法として3種類の方法を用意する。それぞれレシート入力、食品検索入力、手動入力の3種類である。詳細について説明する。

3.1.1 レシート入力

購入した外食・中食のレシートをOCRを用いて入力する。OCRにより認識された文字列を整形し、店舗名、食品名、サイズを抽出する。外食食品データベース内の店舗名、食品名、サイズとマッチングをとる。合致したものを食事ログとして保存する。

表 1: 三大栄養素の摂取割合目標と中央値

栄養素	目標下限値	目標上限値	中央値
たんぱく質	13	20	16.5
脂質	20	30	25.0
炭水化物	50	65	57.5

外食食品データベースに登録されている外食食品データは、ID、店舗名、食品名、サイズ、メニュータグ、エネルギー (kcal)、タンパク質 (g)、脂質 (g)、炭水化物 (g)、食塩相当量 (g)、三大栄養素のエネルギー産生割合である P_r , F_r , C_r 、三大栄養素のエネルギー産生割合の摂取目標からの差分である P_{diff} , F_{diff} , C_{diff} 、三大栄養素の摂取目標からの差分の和である PFC_{diff} の情報からなるとする。外食食品データは 2658 件登録されている。

3.1.2 食品検索入力

レシートがない食品の場合は検索により入力を行う。ユーザが店舗名、食品名、サイズのいずれかの項目またはその全ての項目を入力すると、入力された文字列が各項目に含まれる食品のみを絞り込んで表示する。検索結果の一覧を表示し、食べたものを選択してもらい、食事ログとして保存する

3.1.3 手動入力

検索によっても摂取した食品が得られない場合は手入力により入力を行う。入力を必須とする項目は、食品名、店舗名、三大栄養素 (3種類)、エネルギー量の合計 6 項目になる。

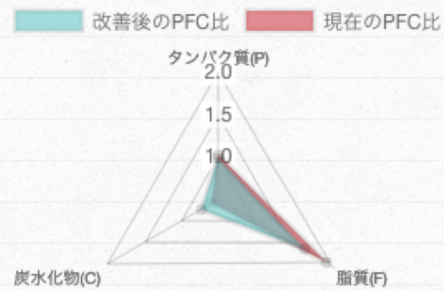
3.2 栄養素情報の算出と可視化

入力された食事ログを用いて、摂取した栄養素を算出する。本研究では、栄養素をたんぱく質、脂質、炭水化物からなる三大栄養素とエネルギーとする。

1 日で摂取すべき三大栄養素の割合はエネルギーとの関連により定められている。18 歳以上 29 歳以下の男性の三大栄養素の摂取割合目標と中央値を表 1 に示す。はじめに栄養素の割合を算出し、割合を満たしているかを可視化する。

タンパク質、脂質、炭水化物の重さをそれぞれ P_w , F_w , C_w とする。文献 [9] に記載されている数式を参考にし、三大栄養素がエネルギーを産生する割合を算出する。ある食品のエネルギーを E とした時の、タンパク質、脂質、炭水化物がエネルギーを産生する割合を

オススメ食品を食べた時のPFC変化



過去 1 週間の摂取カロリー

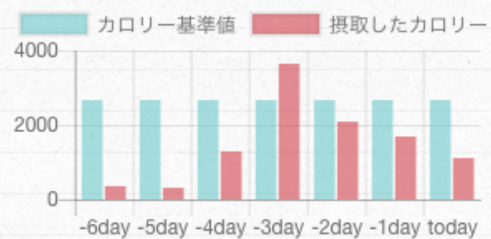


図 2: 栄養素表示の例

P_r , F_r , C_r とし、算出には式 (1)、式 (2)、式 (3) を用いる。 P_r , F_r , C_r の範囲はそれぞれ 0 以上 100 以下となる。

$$P_r = \frac{P_w * 4 * 100}{E} \quad (1)$$

$$F_r = \frac{F_w * 9 * 100}{E} \quad (2)$$

$$C_r = 100 - P_r - F_r \quad (3)$$

算出された栄養素情報を可視化する。図 2 に栄養素表示の例を示す。栄養素表示にはレーダーチャートを用いる。栄養素表示は三大栄養素のそれぞれが、栄養所要量の中央値の値になった時に正三角形となるようなレーダーチャートを提示する。赤いレーダーチャートが現在の P_r , F_r , C_r で、青いレーダーチャートが推薦食品を食べた時の P_r , F_r , C_r を示している。

エネルギーは過去一週間の各日の摂取量を棒グラフを用いて表示する。赤い棒グラフはユーザが摂取したエネルギー、青い棒グラフは摂取基準を示している。

3.3 推薦食品の出力

算出した栄養素の割合の情報を基にし、推薦する食品を決定し出力する。

3.3.1 推薦する食品候補のリストの作成

1日を次のように3分割し、それぞれ朝食・昼食・夕食とする。

- 朝食…6:00 - 11:00 にとった食事
- 昼食…11:00 - 16:00 にとった食事
- 夕食…16:00 - 翌朝 6:00 にとった食事

昼食を推薦する日の「前日の夕食」と「当日の朝食」のエネルギー量を E_{before} とする。外食食品データベースの各食品のエネルギー量を $E_{suggest}$ とする。外食食品データベースから、 $E_{before} + E_{suggest} < E_{target}$ を満たす食品を取得し、リスト S_{list} に追加する。

3.3.2 栄養所要量の差分の算出

栄養素の摂取割合が表 1 に従うように食品を推薦する。そのために現在の栄養所要量の目標との差分を算出する。

たんぱく質、脂質、炭水化物の摂取割合の目標値からの差分を、それぞれ $P_{diff}, F_{diff}, C_{diff}$ とする。 P_r, F_r, C_r の目標上限と下限は表 1 の通りである。 P_r, F_r, C_r が目標上限と下限を超える場合、目標の範囲内の場合と、3つの場合分けを行い、 $P_{diff}, F_{diff}, C_{diff}$ をそれぞれ次の式 (4)、式 (5)、式 (6) を用いて算出する。

$$P_{diff} = \begin{cases} P_r - 13 & (P_r < 13) \\ 0 & (13 \leq P_r \leq 20) \\ 20 - P_r & (20 < P_r) \end{cases} \quad (4)$$

$$F_{diff} = \begin{cases} F_r - 20 & (F_r < 20) \\ 0 & (20 \leq F_r \leq 30) \\ 30 - F_r & (30 < F_r) \end{cases} \quad (5)$$

$$C_{diff} = \begin{cases} C_r - 50 & (C_r < 50) \\ 0 & (50 \leq C_r \leq 65) \\ 65 - C_r & (65 < C_r) \end{cases} \quad (6)$$

$P_{diff}, F_{diff}, C_{diff}$ の絶対値の差を PFC_{diff} とし、次の式 (7) を用いて算出する。

$$PFC_{diff} = |P_{diff}| + |F_{diff}| + |C_{diff}| \quad (7)$$

3.3.3 食品の推薦

昼食を推薦する日の「前日の夕食」と「当日の朝食」の PFC_{diff} を、 PFC_{diff_before} とする。同様にして、外食食品データベースの各食品の PFC_{diff} を算出し、 $PFC_{diff_suggest}$ とする。昼食を推薦する日の「前日の

分類	食品名	店舗	食べたい
主食	牛黒カレー (大盛)	吉野家	<input type="checkbox"/>
定食	おしんこ鮭朝食 (大盛)	すき家	<input type="checkbox"/>
定食	鯖塩焼きセット	立命館大学生協食堂	<input type="checkbox"/>
主食	塩豚おろしぶっかけ (大)	はなまるうどん	<input type="checkbox"/>
主食	ウェンディーズパーガー (トリプル)	Wendys First Kitchen	<input type="checkbox"/>

図 3: 食品推薦の例

夕食」と「当日の朝食」と「推薦食品」の PFC_{diff} を算出し、 PFC_{diff_after} とする。

$PFC_{diff_before} = 0$ の時は三大栄養素の摂取目標を満たしている。摂取目標を満たしている場合には、 $PFC_{diff_suggest} = 0$ である食品を推薦する。推薦手順は以下の通りである。

1. S_{list} から $PFC_{diff_suggest} > 0$ である食品を除く。
2. S_{list} から店舗が被らないようにして 5 つの食品を選択する。

反対に、 $PFC_{diff_before} > 0$ の時は三大栄養素の摂取目標を満たしていない。三大栄養素のバランスを補正する食品を推薦する。そのために、 $P_{diff}, F_{diff}, C_{diff}$ の基準からの差が最も大きいものを N_{adjust} とし、優先して補正する。手順は以下の通りである。

1. S_{list} から $PFC_{diff_suggest} = 0$ である食品を除く。
2. N_{adjust} が目標上限値を上回る時、 S_{list} から $PFC_{diff_suggest}$ が小さい順に N_{adjust} の目標値からの差分が負の値になる食品を最大 5 品選択する。ただし、 $PFC_{diff_after} \leq PFC_{diff_before}$ を満たす食品とする。
3. 同様に、 N_{adjust} が目標下限値を下回る時、 S_{list} から $PFC_{diff_suggest}$ が小さい順に N_{adjust} の目標値からの差分が正の値になる食品を最大 5 品選択する。ただし、 $PFC_{diff_after} \leq PFC_{diff_before}$ を満たす食品とする。

図 3 に食品推薦の例の図を示す。これは、夕食にすき家の牛丼 (並盛)、朝食にミスタードーナツのエンゼルフレンチ、オールドファッションを食べた時に推薦された 5 品である。最大 5 つの食品それぞれに対して、主食・主菜・副菜・汁物・定食の分類、食品名、店舗を表示する。

表 2: グループごとの三大栄養素・エネルギー量の結果

実験グループ	PFC_{diff}	エネルギー量
A(提案システム)	42.67	1692.2
B(比較システム 1)	49.13	1408.1
C(比較システム 2)	44.98	1450.5

4 評価実験

提案システムにより栄養素の摂取バランスの改善が支援できるかを評価する実験を行なった。

4.1 実験手順

本実験は以下の手順で行った。

1. 被験者を募り 3つのグループに分けた。グループ A, グループ B, グループ C とした。
2. 2019/01/11 の夕食 (16 時) から, 2019/01/16 の昼食 (16 時) までの 5 日間システムを利用してもらった。グループ A は提案システム, グループ B は比較システム 1, グループ C は比較システム 2 を使用して食事ログを入力した。
3. 5 日間の食事ログを入力した後に, アンケートに回答してもらった。

被験者は立命館大学情報理工学部 に所属する男性 21 名であった。年齢は 18 歳から 29 歳までであった。21 名を 7 名ずつの 3 つのグループに分けた。

用いたシステムは 3 つあった。提案システムは本論文で提案したシステムである。比較システム 1 は提案システムから食品推薦の機能を除いたシステムであった。比較システム 2 は提案システムから食品推薦の機能と、栄養素の可視化の機能を除いたシステムであった。

提案システムを用いたグループ A の被験者には、推薦食品を食べたいと思ったかを昼食前に送信してもらった。各グループには食事ログを入力する際、自炊をしたり、外食食品データベースにない食品を食べた場合は、レシピサイトから類似する料理を探してもらい、そのレシピから算出した栄養素とエネルギーを入力してもらった。

得られた食事ログおよび栄養素の情報から以下の項目をグループ間で比較することによりシステムの評価を行なった。

1. 平均 PFC_{diff}
2. 平均エネルギー

4.2 実験結果

表 2 にグループごとの評価項目を示す。三大栄養素改善日数が最も多いのは提案システムを用いたグループ A であった。平均 PFC_{diff} が最も小さいのは提案システムを用いたグループ A であった。エネルギー量においては、比較システム 1 を用いたグループ B が最も少なくなった。

5 考察

三大栄養素の改善日数は提案システムを用いたグループ A が最も多くなった。比較システム 1 と提案システムの差は食品推薦機能の有無であった。提案システムは 7 名の被験者にそれぞれ 5 回ずつ昼食の推薦をした。7 名の被験者は平均 1.38 回、推薦された食品を食べたいと思ったと回答したが、実際に推薦された食品を食べた回数は 0 回であった。しかし、三大栄養素のバランスが類似した食品を選択することがあった。類似した食品を選択した例を表 4 に示す。同じおにぎりのカテゴリから食品を選んだことが示されている。食品を推薦することにより、推薦食品を意識して昼食の選択を行うようになり、その結果として三大栄養素の改善日数が最も多くなったと考えられる。したがって、食事推薦は三大栄養素のバランス改善を支援することがわかった。

エネルギー量については最も少なかったのはグループ B であった。グループ B の被験者のエネルギー量が少なかった理由としては、表 3 に示すように朝食を抜き、夕食と昼食のみ摂取する人が多かったことが原因として考えられる。各被験者の朝食、昼食、夕食の接種回数を調べたところ、全てのグループで多くの被験者が朝食を摂取していなかった。グループ A では朝食をとっていない分を昼食で補おうとするため、摂取エネルギー量が多くなったと考えられる。この結果から各人の食事生活に応じた推薦を検討すべきであることがわかった。

6 おわりに

本研究では、外食・惣菜がメインでも栄養バランスを整える食品推薦システムを提案した。提案システムは、外食産業の栄養素データベースを元に、栄養素の過不足を指摘することや、栄養バランスの補正を行える食品を推薦することにより、栄養所要量の摂取目標を満たす食生活を送ることを支援するシステムである。評価実験の結果、食品推薦を行った被験者に対しては摂取する三大栄養素バランスの改善が確認された。こ

表 3: グループ B のあるユーザの 1 日分の食事ログ

店舗名	食品名	サイズ	ログ日時	区分	E_{nergie}	P_w	F_w	C_w
はなまるうどん	カレーライス	大	2019-01-12 22:53	夕	518	5.8	27.3	62.4
ミスタードーナツ	オールドファッション		2019-01-12 22:58	夕	372	12.9	10.5	57.9
コンビニ	フィットチーネグミ		2019-01-12 22:58	夕	267	9.3	13.6	26.9
コンビニ	ポディメンテ		2019-01-12 23:00	夕	608	25.3	19.5	82.3
ローソン	悪魔のおにぎり		2019-01-13 10:18	昼	608	25.3	19.5	82.3

表 4: 推薦した食品と摂取した食品の比較

	食品名	P_{diff}	F_{diff}	C_{diff}
推薦	昆布おむすび	-5.57	-16.77	24.34
摂取	おかかおにぎり	-4.32	-18.13	24.45

の結果から、提案システムにより食品推薦を行うことは健康的な食生活を支援することを確認できた。

今後は食品を推薦する際に個人の嗜好を考慮していきたい。推薦を行なった際の満足度 [10] を考慮することにより、より質の高い推薦が行える可能性がある。加えて 1 品だけを推薦するのではなく献立を推薦するように改善していきたい。Kuo ら [11] は食材を入力することで、よく一緒に食べられている料理の組み合わせの献立を出力するシステムを提案している。提案システムにおいても献立を推薦することにより、より栄養バランスを整えることを支援していきたい。

謝辞

本研究の一部はすかいらーく財団の助成を受けた。記して謝意を申し上げる。

参考文献

- [1] 平成 27 年国民健康・栄養調査報告. <https://www.mhlw.go.jp/stf/houdou/0000142359.html> 2019 年 1 月 23 日に閲覧.
- [2] 報道発表資料「日本人の食事摂取基準(2015 年版)策定検討会」の報告書. <http://www.mhlw.go.jp/stf/houdou/0000041733.html> 2019 年 1 月 23 日に閲覧.
- [3] 平成 25 年国民健康・栄養調査報告. <https://www.mhlw.go.jp/bunya/kenkou/eiyou/h25-houkoku.html> 2019 年 1 月 23 日に閲覧.
- [4] 調査から見える日本の食卓. https://www.nhk.or.jp/bunken/research/yoron/20161001_7.html 2019 年 1 月 23 日に閲覧.
- [5] 岡田斗司夫. いつまでもデブと思うなよ. 新潮社, 2007.
- [6] 伊原啓晃, 玉井森彦, 安本慶一. 複数要因を総合的に考慮した夕食レシピ推薦システムとその評価. *CSEC-64*, Vol. 32, No. 2014-02-27, pp. 1-8, 2014.
- [7] 高橋 淳, 伊藤 孝行, 植田 嗣也. 栄養の食べ合わせを考慮した料理推薦機構の試作. 研究報告知能システム (ICS), Vol. ICS-164-2, No. 2011-10-21, pp. 1-6, 2011.
- [8] 金塚永華, 川村公子, 戸塚優衣, 栗原久. 大学生の食生活と総合的健康状態との関連について. 東京福祉大学・大学院紀要, Vol. 8(2), No. 2018/3/25, pp. 221-229, 2018.
- [9] エネルギー産生栄養素バランスとは. <https://www.glico.co.jp/navi/e04.html> 2019 年 1 月 23 日に閲覧.
- [10] Kenta Oku, Shinsuke Nakajima, and Jun Miyazaki. 状況依存型ユーザ嗜好モデリングに基づく Context-Aware 情報推薦システム. *Sci. Technol.*, Vol. 48, No. June, pp. 162-176, 2007.
- [11] Man-Kwan Shan Suh-Yin Lee Fang-Fei Kuo, Cheng-Te Li. Intelligent menu planning: recommending set of recipes by ingredients. In *Proceedings of the ACM multimedia 2012 workshop on Multimedia for cooking and eating activities*, pp. 1-6, 2012.

味表現を修飾する副詞出現頻度に関する地域差の分析

Analysis of Regional Differences on Frequency of Adverbs for Taste of Foods

横川舜^{1*} 西原陽子¹ 山西良典¹ 吉田光男²
Shun Yokogawa, Yoko Nishihara, Ryosuke Yamanishi, Mitsuo Yoshida

¹ 立命館大学 情報理工学部

¹ College of Information Science and Engineering, Ritsumeikan University

² 豊橋技術科学大学 情報・知能工学系

² Department of Computer Science and Engineering, Toyohashi University of Technology

Abstract: Restaurant staffs often use social networking services like Twitter and Instagrams as tools for spreading their advertisements. They appeal their restaurants by using photos and phrases to attract the potential customers. We believe that the attractive phrases should be depending on the regions where the restaurants exist. We focus on especially adverbs for tastes of foods are different depending on the regions. In this paper, we propose a method to analyze the regional differences on frequency of adverbs for taste of foods, and applied the method to tweet texts. We found that some adverbs were used for the specific taste of foods and in the specific regions.

1 はじめに

飲食店の広告や口コミ情報の投稿に Twitter や Instagram などのソーシャルネットワーキングサービスが利用されることが増えている [7]. 飲食店の店員が自店舗やおすすめ商品を宣伝するために写真をアップロードし、コメントを添えて投稿することや、飲食店のお客さんが実際に店を訪れ、食べたものについて同様のことをすることが一般的になりつつある。飲食業界の競争は激しく、利益向上や集客力向上のためには広告により自店舗の個性やメニューをアピールすることが重要になってきている。

自店舗のメニューやその個性をアピールする際に、言語的な表現にも配慮をすることが重要と考えられる。関西では「とてもおいしい」は「めっちゃおいしい」のように標準的な表現から言語表現が変化する。「めっちゃ」は方言の一種である。方言は地域ごとに異なる特徴を持つ言葉である [5]. 関西の人にとっては広告文の中に「とてもおいしいです」と書かれているよりは「めっちゃおいしいで」と書かれている方が馴染みがあり、広告や広告を出している店舗に親しみを覚える可能性がある。すなわち、その地域で頻繁に使われている表現を用いるように広告文を微修正することが望ましいと考えられる。

本研究では飲食店が広告に用いると考えられる味表現に着目し、味表現を修飾する単語である副詞の地域差を分析するための手法を提案する。味表現とは味覚を表現する際に使われる単語やフレーズを指す。地域ごとに好まれる味覚や特徴的な食事は異なっており、レストランのメニューを分析することで地域ごとの食の特徴を明らかにした研究がある [8]. 一方で、味表現を修飾する際に使われる副詞について地域差を分析するための手法や、分析した結果については数が少なく、明らかにされていないところが多い。地域差を明らかにすることにより、共通した味表現に対して地域ごとに受け入れやすい副詞を活用することに繋がる。そのため、本研究では味表現そのものではなく、それを修飾する副詞の地域差を明らかにすることに主眼を置く。

単語の地域性を分析した研究や、地域に特徴的な単語を抽出する手法に関する研究は多く存在する。例えば、特定の地域での出現頻度が相対的に高い語句を抽出する手法 [6] や、地名に対する市民の関心動向を分析した研究 [1] がある。他にも疾患を表す単語と地域との関連を分析し、ある疾患の発症地域を推定、予測する研究もある。例えば、熱中症と考えられるツイートと実際に搬送された地域の相関を分析した研究や [2], 検索ワードの履歴からインフルエンザの流行を公的機関と同等の正確さでより早期に予測する研究がある [3]. 特に Twitter や Instagram などのソーシャルネットワーキングサービスでは、ユーザの日々の行動が投稿される

*連絡先: 立命館大学情報理工学部
滋賀県草津市野路東 1 丁目 1 - 1
E-mail:nishihara@ed.ritsumeikan.ac.jp

ため地域性の高い情報が得やすい特徴がある。ツイートデータの中で用いられる単語を用いて方言の地理的な分析を行った研究 [4] もある。これらの研究はある単語とある地域の関連に着目している。本研究でも味表現と地域の関連に着目するが、この関連に加えて副詞の使われ方の特徴を分析する点が異なっている。

2 提案する分析手法

本章では味表現を修飾する副詞を分析する手法を説明する。その概要は以下の手順の 1 から 6 である。

1. 分析するテキストを形態素解析し、分かち書きを行う。単語の原型と品詞の情報を獲得する。
2. 単語の中から味表現に一致するものを抽出する。
3. 抽出された味表現を修飾する副詞を抽出する。
4. 味表現が抽出されたテキストが投稿された地域情報を獲得する。
5. 地域ごとに味表現と副詞のペアの頻度を算出する。
6. 地域ごとの味表現を修飾する副詞の特徴を獲得する。

2.1 テキストからの単語情報の獲得

テキストを形態素解析し分かち書きを行う。その結果から単語の原型と品詞の情報を獲得する。

本研究では分析するテキストとして Twitter のツイートを用いる。ツイートはユーザの個人の意見が位置情報付きで表現されることが多く、本研究での分析に適している。形態素解析には MeCab を用い、辞書は NEologd を用いる。

2.2 味表現に一致する単語の抽出

獲得した単語と味表現の単語リストとの対応をとり、味表現に一致する単語を抽出する。本研究で用いた味表現のリストを表 1 に示す。これらの味表現は「おいしいを感じる言葉 2018」[9] を参考にして用意した。味表現の数は全部で 97 個であった。

2.3 味表現を修飾する副詞の抽出

抽出された味表現を修飾する副詞を抽出する。本研究では MeCab と NEologd により副詞と判定され、かつ抽出された味表現の直前に出現する単語を、味表現を修飾する副詞として抽出する。

表 1: 味表現のリスト

飽きのこない, 味わい深い, あっさり, 後味すっきり, 後味のよい, 後を引く, 脂っこい, 脂の乗った, 甘い, 甘辛い, 甘口, 甘さ控えめ, 甘じょっぱい, 甘ずっぱい, 甘味のある, 薄味, うす塩, うすすら, 旨辛, うま味のある, 塩味のある, 香りのよい, 香る, 薫る, 辛い, 辛口, 芳しい, 甘露, 凝縮した, キレのある, クセになる, クセのない, 口に広がる, クリーミー, 激辛, 濃い味, 香ばしい, 香味, コクうま, 極甘, 極旨, 極辛, コクのある, コク深い, こっくり, こってり, 酸味のある, さっぱり, サワー, さわやかな, 塩辛い, 渋い, しみうま, しみ込んだ, 渋味のある, しょっぱい, シンプルな, スイート, すっきり, すっぱい, すっぱ辛い, スパイシー, スモーキー, 繊細な, ソルティ, ダシの効いた, 淡白, 複雑な, ふくよかな, フルーティ, 芳醇な, ほのかな, ほろ苦い, ほの甘い, ほんのりした, マイルド, まったり, まろやかな, みずみずしい, やさしい, 超すっぱい, デリシヤス, 特濃, 苦みのある, 濃厚な, 濃密な, 鼻に抜けるような, やみつきになる, 華やかな, 余韻のある, ビター, よくしみた, 美味, リッチな, ピリ辛, 風味豊かな, 深みのある

2.4 味表現が投稿された地域情報の獲得

味表現が投稿された地域情報を獲得する。本研究では味表現が投稿された位置にもっとも近い県庁所在地を、その味表現が投稿された地域情報として獲得する。投稿された場所の都道府県と獲得された県庁所在地の都道府県は必ずしも一致しないが、異なる県であっても隣接する二つの地域は表現が類似する可能性があると考えたため、本研究ではもっとも近い県庁所在地を地域情報とする。

本研究では位置情報付きのツイートデータを用いる。はじめにツイートデータに付随している緯度経度の情報を抽出する。続いて、日本の 47 都道府県の各県庁所在地の緯度経度を用いて、各県庁所在地との距離を算出する。距離が最も短い県庁所在地を一つ選択する。選択された県庁所在地が所属している都道府県を、味表現が投稿された地域情報として獲得する。

2.5 地域ごとの味表現を修飾する副詞の特徴の獲得

地域ごとに味表現と修飾する副詞のペアの頻度を算出する。頻度が閾値以上であったペアに絞り込み、地

域ごとの副詞の特徴を獲得する。

3 提案手法を用いた分析および結果

本章では提案手法を用いて実際に分析をした結果を示す。

3.1 分析の詳細

提案手法を用いてツイートデータを分析し、地域ごとに味表現と副詞のペアを抽出する。

本論文で用いたツイートデータは2017年1月1日から2017年11月29日のおよそ11か月の間に得られたツイートデータである。ツイートの合計は24,862,718件である。

本論文では全ての味表現ではなく「あっさり」「こってり」「さっぱり」の3種類の味表現に対して得られた結果を分析する。この3種類の味表現はツイートデータの中で出現頻度が高く、それぞれの出現頻度は「あっさり」は300回、「こってり」は275回、「さっぱり」は275回である。本論文において、出現頻度の単位は回数とする。ペアの出現頻度の閾値は2回とする。

本論文では全ての地域ではなく、「東京」「愛知」「大阪」「京都」「福岡」「北海道」「沖縄」の7つの県庁所在地に対して得られた結果を分析する。これらの県庁所在地を選択した理由は次の通りである。東京は他地域からの人材流入が激しい地域であることから、特徴が見られないと考える。愛知・大阪・京都・福岡は、味付けに特徴のある地域である。愛知であれば、味噌を使った味噌カツや味噌煮込みうどん、大阪・京都であればダシのきいた薄味の関西風が存在する。このことから味表現を修飾する副詞に対し、特徴が見られる可能性があると考え。北海道・沖縄に関しては都道府県内で最北と最南の位置する地域であり、言葉の使い方が本州とは異なっている部分がある。このことから味表現を修飾する副詞に対し、特徴が見られる可能性があると考え。

3.2 分析の結果

表2に7つの地域で3つの味表現に対して、修飾する副詞が2回以上出現した例の一部を示す。2回位以上出現した副詞は98種類であった。出現回数の合計は989回であった。

表3に、味表現「こってり」を修飾した副詞を地域ごとに示す。表4に、味表現「あっさり」を修飾した副詞を地域ごとに示す。表5に、味表現「さっぱり」を修飾した副詞を地域ごとに示す。表6には実際のツイート例を示す。

表 2: 出現した副詞とその出現頻度 (回数)

副詞	出現頻度
いつも	96
めっちゃ	50
意外と	46
初めて	44
あっさり	41
少し	41
ちょっと	41
とても	35
こってり	33
色々	24
これから	24
本当に	24
さっぱり	20
そう	19

表2に示した中で、最も頻度が高い副詞は「いつも」である。そのあとは「めっちゃ」「意外と」が続いた。表2には「あっさり」「こってり」「さっぱり」も掲載されている。これはMeCabとNEologdにより副詞と判定され、3種類の味表現のいずれかの直前に出現したため含まれている。

4 分析結果に対する考察

前章で得られた分析結果に対して考察を行う。

4.1 各味表現を修飾する副詞の地域差

表3に各地域で「こってり」を修飾した副詞とその頻度を示す。7つの地域のうち「沖縄」は条件を満たす副詞が存在しなかったため掲載していない。東京、京都、北海道の3つの地域では、「いつも」が最も使われていた。ツイートの例としては、「天下一品高円寺天一は裏切らない。たまにはあっさり頼んでみようかな、と直前まで思ってるのに、いつもこってり頼んじゃう」があった。京都、北海道でも同様のツイート例が見られた。これらの3つの地域では修飾する副詞の地域差は見られなかった。

一方で、大阪は「こってり」が副詞として用いられることが多く、愛知は「めっちゃ」、福岡は「あっさり」であり、それぞれ異なっていた。各ツイートの例は以下の通りであった。「新メニューの背脂カレー。更に辛味オイルをかけてこってりこってり。」(大阪)、「またラーメン巡りめっちゃこってり系だったけど思ったよりアッサリしてて美味しい」(愛知)、「今日のメは博多

表 3: 「こってり」に対する地域ごとの副詞と出現頻度(回数)

地域	副詞	頻度
東京	いつも	14
	初めて	13
	こってり	12
	めっちゃ	10
	少し	8
	やはり, ちょっと, これから	7
	そう, ほんとに	6
大阪	こってり	8
	あっさり	7
	とても	4
	本当に, もちろん, とりあえず, 少し, めっちゃ	3
	もう, 初めて	2
愛知	めっちゃ, なかなか	2
京都	いつも	4
	本当に, りん	2
福岡	あっさり	5
	ちょっと, もう	2
北海道	いつも, どう, ホロホロ, あっさり	2

の方に教えてもらった shinshin さんへ. あっさりこってりが美味しゅうございました。」(福岡). これらの3つの地域では, 味表現「こってり」を修飾する副詞に地域差がある可能性が分かった.

表4に各地域で「さっぱり」を修飾した副詞とその頻度を示す. 7つの地域のうち, 沖縄を除く6つの地域で「意外と」が頻度の1位または2位に現れた. ツイートの例としては, 「見た目とは裏腹に意外とあっさりした味付け」(東京)があった. 他の5つの地域でも同様のツイートの例が見られた. これらの6つの地域では「あっさり」を修飾する副詞の地域差は見られなかった.

沖縄は「いつも」が「こってり」を修飾する副詞として用いられることが多かった. ツイートの例としては, 「いつもあっさり塩味しか GET してないような …」(沖縄)があった. 「いつも」はそのほかの地域でも頻度の上位に見られている(東京, 大阪, 福岡でも2位). したがって, 沖縄だけで特徴的に使われている副詞とは言いがたい. この結果から味表現「あっさり」については, 修飾する副詞に地域差がない可能性がわかった.

表5に各地域で「さっぱり」を修飾した副詞とその頻度を示す. 7つの地域のうち, 東京, 大阪, 愛知, 京都の4つの地域で「いつも」が頻度の1位に現れた. ツイートの例としては, 「こちらで奥ちゃんが肉類を食べたいと. いつもさっぱり系だからね」(東京)があった.

表 4: 「あっさり」に対する地域ごとの副詞と出現頻度(回数)

地域	副詞	頻度
東京	意外と	18
	いつも	17
	とても	10
	あっさり, こってり, めっちゃ	9
	ちょっと, 少し	8
	かなり, 初めて	7
大阪	意外と	8
	いつも	6
	どう	5
	少し, しっかり, とても	4
	こってり, たっぷり, さっぱり, ちょっと	3
愛知	意外と	6
	かなり, ちゃんと, これから, こっている, とても, 初めて, ちら	2
京都	こってり	5
	意外と, 少し	4
	あっさり	3
	初めて, ドロドロ	2
福岡	意外と	6
	いつも, 結構, ふわふわ, とても, 比較的	3
北海道	意外と	4
	そう, 初めて	3
	これから, まだ, ぜひ, いつも, こってり	2
沖縄	いつも	3

他の4つの地域でも同様のツイートの例が見られた. これら4つの地域では「さっぱり」を修飾する副詞の地域差は見られなかった.

一方で, 福岡は「すぐ」が副詞として用いられることが多く, 北海道は「めっちゃ」, 沖縄は「ジメジメ」であり, それぞれ異なっていた. 各ツイートの例は以下の通りであった. 「飲んですぐさっぱりした感じがあって飲みやすい」(福岡), 「いわゆる焦がし醤油かな? 見た目はすごいけど, めっちゃさっぱりしてて, シメには最適でした!」(北海道), 「グレープフルーツ+ライム+ユーカリでジメジメすっきりリフレッシュ!」(沖縄). 2つ目の例については, 抽出された副詞が味表現を修飾しているが, 1つ目と3つ目の例については修飾がなかった. 1つ目の例については抽出された副詞は味表現ではなく副詞の前の動詞を修飾していると考えられる. 3つ目の例については抽出された副詞は味表現を修飾しているわけではなく, 反対に味表現の方

表 5: 「さっぱり」に対する地域ごとの副詞と出現頻度(回数)

地域	副詞	頻度
東京	いつも	28
	ちょっと	13
	すっきり	12
	さすが	11
	めっちゃ, さっぱり	9
	これから, スッキリ, 少し	8
	とても	7
大阪	いつも	7
	ちょっと, 初めて	4
	まだ, あっさり, めっちゃ	3
	本当に, ぜひ, すぐ, 少し	2
愛知	いつも	5
	ちょっと	4
	本当に, 相変わらず, いっぱい, めっちゃ	2
京都	いつも	4
	そう	2
福岡	すぐ, すっきり	3
	いやいや, 少し, いつも, やっぱり, チョット	2
北海道	めっちゃ	4
	そう, いつも, すっきり	2
沖縄	ジメジメ	4
	さっぱり	2

が副詞を修飾していると考えられる。この2つの例から係り受け解析を丁寧にすることで味表現を修飾する副詞を絞り込むことができたことがわかった。

4.2 考察全体のまとめ

分析した3種類の味表現の中で、地域差が見られたのは「こってり」を修飾する副詞であった。味表現「こってり」については、大阪、愛知、福岡では修飾する副詞がそれぞれ異なっていた。広告文を作る際に地域ごとに馴染む表現を考える上では、この知見を活かすことができると思われる。

一方で「あっさり」「さっぱり」について修飾する副詞の地域差は見られなかった。また係り受け解析を行うことにより、味表現を修飾する副詞を絞り込む必要があることがわかった。

表 6: 実際のツイート例。味表現と修飾する副詞、および地域情報を併せて示す。

天下一品高円寺天一は裏切らない。たまにはあっさり頼んでみようかな、と直前まで思ってるのに、いつもこってり頼んじゃう(東京)
新メニューの背脂カレー。更に辛味オイルをかけてこってりこってり。(大阪)
またラーメン巡りめっちゃこってり系だったけど思ったよりアッサリしてて美味しい(愛知)
今日のメは博多の方に教えてもらった shinshin さんへ。あっさりこってりが美味しゅうございました。(福岡)
見た目とは裏腹に意外とあっさりした味付け(東京)
いつもあっさり塩味しかGETしてないような...(沖縄)
こちらで奥ちゃんが肉類を食べたいと。いつもさっぱり系だからね
飲んですぐさっぱりした感じがあって飲みやすい(福岡)
いわゆる焦がし醤油かな?見た目はすごいけど、めっちゃさっぱりしてて、シメには最適でした!(北海道)
グレープフルーツ+ライム+ユーカリでジメジメすっきりリフレッシュ!(沖縄)

5 おわりに

本論文では味表現を修飾する副詞の地域差を分析する手法を提案した。実際に Twitter のツイート文を用いて分析し、結果に対して考察した。その結果、地域差が見られる味表現とそうでない味表現があることがわかった。分析で用いた3種類の味表現については「こってり」のみ地域差が見られ、そのほか2つの味表現については修飾する副詞の地域差が見られなかった。

今後は、そのほかの味表現に対してもデータ量を増やし同様の分析を行い、地域差が見られる味表現とそうでない味表現に対して考察を加えていく。実際に地域ごとに味表現を修飾する副詞を、地図上に可視化することで、さらに特徴を発見していきたい。最終的には、入力に地域と味表現を選択することにより、味表現に対して修飾しているその地域に馴染みのある副詞を推薦するというインターフェースを作成することにより、広告・宣伝文句の作成に直接的な支援を行っていきたい。

参考文献

- [1] 安藤 有生, 関 洋平, 市民のツイートを用いた分散表現に基づく地名に対する市民の関心の傾向の可視化, 知能と情報, Vol.30, No. 6, pp. 804-814, 2018
- [2] 布施 明, 坂 慎弥, 立澤 裕樹, 吉野 雄大, 萩原 純, 布施 理美, 宮内 雅人, 横田 裕行, SNS における熱中症の共起ワードと熱中症救急搬送者数との関係性についての検討, 日本臨床救急医学会雑誌, Vol.19, No. 6, pp. 703-710, 2016 .
- [3] Jeremy Ginsberg, Matthew H. Mohebbi, Rajan S. Patel, Lynnette Brammer, Mark S. Smolinski & Larry Brilliant, Detecting Influenza Epidemics using Search Engine Query Data, Nature, No. 457, pp.1012-1014, 2009.
- [4] 桐村 喬, 峪口 有香子, 岸江 信介, 方言の分析資料としてのマイクロブログデータ アンケート調査との整合性, 2014 年度日本地理学会春季学術大会, セッション ID: 703, 2014.
- [5] 小林隆, 篠崎晃一 (編), ガイドブック方言研究, ひつじ書房, 2003.
- [6] 奥健太, 西崎剛司, 服部文夫, 地域限定性スコアに基づく位置情報付きコンテンツからの地域限定語句の抽出, 情報処理学会論文誌データベース (TOD), Vol.5, No.3, pp.97-116, 2012.
- [7] 小野寺翼, ソーシャルメディア時代の広報, 情報の科学と技術, Vol.68, No.4, pp. 156-159, 2018.
- [8] 田中 優里, 山西 良典, 西原 陽子, レストラン検索サイトのメニュー名の統計的分析による食の地域性の考察, ARG 第13回 Web インテリジェンスとインタラクション研究会, WI2-2018-23, pp.61-64, 2018.
- [9] おいしいを感じる言葉 2018, <http://www.bmft.jp/pdf/services/kotoba.pdf> (2018年2月21日アクセス確認).

調理の失敗低減を目的とした 動作中心型料理レシピ提示手法の提案

Action-Centered Recipe Presentation for Reducing Cooking Failure

大杉 隆文^{1*} 松下 光範²
Takafumi Ohsugi¹ Mitsunori Matsushita²

¹ 関西大学大学院総合情報学研究科
Graduate School of Informatics, Kansai University

² 関西大学総合情報学部
Faculty of Informatics, Kansai University

Abstract: The aim of this research is to reduce cooking failures of a user who only has little experience of cooking. User-contributed recipe sites are convenient for such the user, however, they may face difficulty to understand the recipe because recipe contributors often post their recipes without explaining some cooking processes assumed that they are common. To solve the problem, this paper proposes a system that presents recipes as a form of the flow graph. With this system, a user who only has little experience of cooking will understand recipes easily. To confirm the usefulness of the proposed system, we conduct a comparative experiment between cooking with a user-contributed recipe site and that with the proposed system.

1 はじめに

近年、特定の誰かが料理レシピを投稿するのではなく、誰でも自身で作成した料理レシピを投稿できるユーザ投稿型料理レシピサイト（以下、料理レシピサイトと記す）が普及している。料理レシピサイトの一つである cookpad¹ の料理レシピ数は年々増加しており、2018 年 12 月現在、料理レシピは約 300 万品が投稿されている²。料理レシピサイトに料理レシピを投稿するユーザは、自身の料理知識や技術レベルを基準に料理レシピを作成している。そのため、同じ料理の料理レシピであるとしても投稿するユーザごとに記述の形式が異なる。これにより同じ料理でも料理レシピによって調理の難易度が変化し、調理者は料理レシピサイトから自身のレベルに合わせて料理レシピを選択する必要がある。しかし、投稿されている料理レシピ数は膨大であるため、難易度が異なる料理レシピが複数存在する。このことから調理者が選択した料理レシピが自身に必ず適しているとは限らないため調理者が自身のレベル

より高い料理レシピを選択する必要が生じ、味や見た目が悪いといった失敗を引き起こす。こうした失敗を避ける方法として料理本のアドバイスや料理レシピサイトのアドバイスが存在している。一般的に料理は、材料とそれを加工する異なる動作の連続（調理工程）で構成されている。料理を行う際に、調理工程全体の流れを把握することで調理者自身が行う行動や必要な材料を把握することができる。しかし、料理歴が短い調理者は、調理に慣れていないため調理工程を把握することが困難である。一つの動作がどのように次の動作に繋がるかを連想できないため、料理歴が短い調理者はしばしば失敗を引き起こす。例えば、料理における前提知識の不足が失敗の原因として挙げられる。これによって引き起こされる失敗は調理者自身に起因する。料理レシピサイトでは調理者に一定水準の知識があることを前提として、調理の詳細が省略されている場合がある。これは用語説明の詳述によって料理レシピが冗長になることを避けるための措置である。料理レシピに記述はないが工程としては内在しており、調理者はその工程を補って調理する必要がある。例えば、食材を炒める際にはフライパンを用い、フライパンで焼く前には油を引くといったことが一般的である。しかし、料理レシピサイトの料理レシピには“フライパン

*連絡先：関西大学大学院総合情報学研究科
〒569-1095 大阪府高槻市霊山寺町 2-1-1
E-mail: k715328@kansai-u.ac.jp

¹クックパッド株式会社: cookpad <http://cookpad.com/>

²<http://pdf.irpocket.com/C2193/BX1b/JHMN/zzTh.pdf>
(2018 年 12 月 13 日確認)

を使用する”, “フライパンに油を引く”といった記述が省略されている場合がある。調理に関わる専門用語が記述されることで料理歴が短い調理者が料理レシピを正しく理解できない場合がある。例えば, “飴色になるまで”, “桂むき”といった料理に関する用語が使用された際, それらの表現を自身の経験や知識で補って解釈する必要がある。このような場合でも, 料理歴が長い調理者は自身の経験から省略されていた場合でも動作を把握し調理を行うことができる。しかし, 料理歴が短い調理者は前提知識を身につけていない場合が多く, 調理用語を理解できなかったり省略されている調理工程を把握できなかったりするため失敗を引き起こす。このことから料理歴が短い調理者は料理歴が長い調理者に比べて調理の失敗を引き起こすことが多い。上述した背景の下, 本研究は料理歴が短い調理者を対象とし, 調理者自身に起因する問題の解決を目指す。課題解決の方法として, 料理歴の短い調理者が料理レシピを参考にして調理を行った場合に引き起こす失敗を低減可能な料理レシピの提示を行う。本稿では, 調理者の失敗を低減するためにフローグラフ形式で料理レシピを表示するインタフェースを提案する。

2 関連研究

調理者の調理技術の習得に関する研究が行われている [4, 7]。平島らは, 調理動作の中でも使用頻度の高い「切る」が調理を行う際に重要であり, 「切る」を習得することができない場合, 調理を行うことが難しいとした [4]。そこで, 「切り方」の難易度, その要因について分析を行うために, 大学生, 専門学校生を対象に「切り方」の知識とそれを行う自信度について調査した。分析の結果, 大学や専門学校の調理実習の授業においても基礎的な包丁技術を学ぶ必要があるとした。

また, 自動調理に関する研究も行われている [3, 1] Bollini らはロボットに自動で調理させることを目指している [1]。この研究では物体の知覚, 物体の操作, ロボットの言語理解のための技術を組み合わせ, 調理環境を用意し, 性能の実験を行った。ユーザが料理レシピ文と材料をロボットに入力として与えた場合, 調理を自動で行うロボットを開発した。実験の結果, システムは用意した 15 の料理レシピのうち 2 種類を完成させることができた。手順の読み込みに失敗した理由の例としては材料の省略をあげている。

料理レシピの構造の研究として Ohsugi らは調理手順テキストに存在する表現の曖昧性に着目し, それらが発生する原因を調査した [9]。この研究では, 調理手順テキストの中から料理に関係のある用語を手動で抽出し, タグを付与している。その中でも動作に着目し, 動作とそれに付随する要素 (e.g., 食材, 道具) について

分析を行った。その結果, 表現の曖昧性は同じ動作でも付随している要素が異なることが原因であるとした。

河内らは料理レシピをフローで表現することのメリットを述べている [5]。例えば (1) 動作の概要を直感的に把握できる, (2) 材料や器具が必要な箇所を把握できる, (3) 手順の分岐の流れと合流箇所を明確に把握できるなどである。

3 デザイン指針

料理歴が短い調理者は前提知識を身につけていない場合が多く, その場合は理解できない用語を Web で調べ, それを参考にする。例えば, cookpad では下ごしらえや料理用語など料理に関わる基本の情報のページが用意されている³。しかし, 調理者は調理中に時間に関わる工程が発生する場合があるため, 調理に集中する環境が必要である。料理歴の短い人は調理を行いながら複数の Web ページやコンテンツを見る必要があり, 調理だけに集中できない。例えば, 火を使った調理をする場合の「3分炒める」や「沸騰するまで煮込む」といったものがあげられる。このときに調理者は調理に集中していることが重要であり, 用語の検索や用語の説明を読むなど他の作業を行うことで指定された通りに調理できない場合がある。これに対して料理歴が長い調理者は料理に慣れているため, 他の作業を行いながら時間に関わる工程を行うことができる。一方, 料理歴が短い調理者は料理に不慣れなため, 他の作業を行いながら時間に関わる工程を上手く行うことができず, 調理に失敗する。

また, 既存の料理レシピサイトは従来型の料理レシピの形式 (テキストと画像) を使用していることが多い。しかし, 河内らは調理手順をフローで表示することで流れを理解しやすくなると主張している [5]。他にも過去の研究では料理レシピはフロー形式で提示されており [2, 10, 11, 6], 本研究でもそれに倣う。加えて, 既存のテキストと画像を軸に構成されている料理レシピでは一つの動作がどのように次の動作に繋がるか明確でないため, 調理者は自身がどこまで調理を行ったかを直感的に把握できない。そのため, 調理者は直感的に調理工程を把握することができない。

これらのことから本研究で提案システムのデザイン指針を (1) 知識や経験の不足を補えること, (2) 調理の手順と同一ページでアドバイスやコツなどの補足情報が確認できること, (3) 一つの動作を一つの手順として扱うこと, (4) 直感的に調理工程を把握できること, とし, これを満たすために手順をフローグラフ形式で表現する方法を採用した。

³https://cookpad.com/cooking_basics (2019年1月21日確認)

4 実装

4.1 システムの構成

3章で述べたデザイン指針を基に料理レシピの提示システムを構築した。提案システムの構成を図1に示す。システムは料理レシピDB, 料理レシピごとの料理に関する用語, データ成形モジュール, 提示画面, 入力判断部, 使用データ, 画面生成部から構成される。以下にシステムが実行する処理の流れを説明する。まず, ユーザがセレクトボックスで選択した料理レシピをシステムの入力判断部で処理する。次に入力された料理名と合致するものが使用データに格納されている場合, システムは画面生成部に処理を移行させる。最後に使用データに基づき画面生成部で画面の生成を行い, 提示画面をユーザに提示する。

また, 料理レシピのデータとして使用する料理レシピDBにはクックパッド株式会社が国立情報学研究所を通じてデータ提供を行っているもの(クックパッドデータセット)を使用した⁴。これは2014年9月30日までにcookpadに投稿された料理レシピ約172万品のデータが使用可能になっているものであり, MySQL形式のデータベースである。これらのデータから前田らは料理に関する用語だけを自動で抽出することを行っている(図1破線部参照)[8]⁵。この研究では, 手順文書を手続きの流れを考慮してフローグラフで表現し, それぞれの関係性を明確にすることを目的としている。そこで, 料理レシピから固有表現を抽出し, 料理レシピをフローグラフで表現している。固有表現とは, 人参やじゃがいもといった食材や包丁やまな板といった道具, 切るや炒めるといった動作など, 料理に関する用語のことを指す。これらの固有表現に食材や道具などを識別できるタグをつけたデータを作成している。本研究では, この固有表現にタグづけされたデータ(図1中料理に関する用語参照)を使用する。このデータをさらに成形し(図1中データ成形モジュール参照), 提案システムで使用する。

4.2 システムに必要なデータ

提案システムでは3章で述べたデザイン指針(4)を満たすためには, 調理手順の中でも動作や食材などが識別可能である必要がある。そこで, 料理レシピDBから料理の固有表現を抽出しタグづけされたデータを用いる。タグづけされたデータは料理レシピDBに格納されている各料理レシピに対してJSON形式で保存

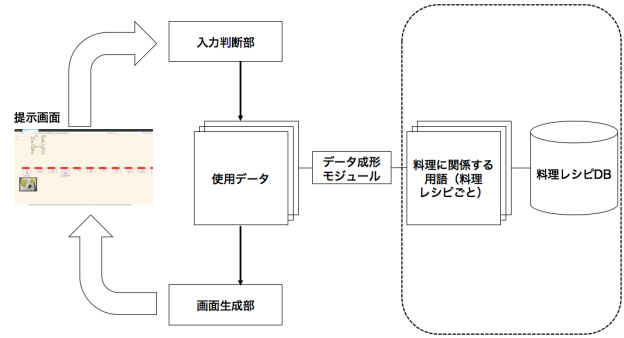


図1: 提案システムの構成図

されている。加えて, タグづけされたデータの中には手順だけではなく, cookpadに記載されている食材名や調理に関するアドバイスの部分なども格納されている。タグづけされたデータの中から食材名, 量, 手順のデータを提案システムに用いた。

以下に提案システムで用いるデータ成形の流れを示す。まず, タグづけされた各料理レシピのデータから提案システムで使用する料理レシピを選定した。本実験で使用する料理レシピは豆腐ハンバーグとサラダとした。しかし, タグづけされたデータは動作ごとに手順がまとまっているわけではない。そこでデータ成形モジュールを用いて提案システムで使用する料理レシピに対して, 動作とその動作に必要な要素(e.g., 食材, 道具)を一つのグループとする処理を行った。最後に, データ成形モジュールで取得したデータをJSON形式で保存し, 動作を一つのグループとしたデータを料理レシピごとに用意することでユーザに提示する画面を生成する。

4.3 システムの実装

4.2節で述べたデータ成形モジュールのプログラムにはAnaconda3-2.4.1をプラットフォームとし, Pythonのversion 3.5.1を用いて開発を行った。加えて, Pythonのライブラリとしてjsonとcollectionsを使用した。入力判断部と画面生成部と提示画面はWebアプリケーションとして開発を行った。開発にはHTMLとCSS, JavaScriptを用いた。JavaScriptのライブラリとして, D3.js⁶のversion 3.5.17, jQuery⁷のversion 3.2.1, Tippy.js⁸のversion 2.5.4, anime.js⁹のversion 2.2.0を用いた。

作成した提案システムのプロトタイプを図2に示す。提案システムは料理名を選択するセレクトボックス(図

⁴クックパッド株式会社(2015):クックパッドデータ。国立情報学研究所情報学研究所データリポジトリ。(データセット)。<https://doi.org/10.32130/idr.5.1>(2019年1月4日確認)

⁵<http://www.ar.media.kyoto-u.ac.jp/data/recipe/>(2019年2月27日確認)

⁶<https://d3js.org/>(2019年1月4日確認)

⁷<https://jquery.com/>(2019年1月4日確認)

⁸<https://atomiks.github.io/tippyjs/>(2019年1月4日確認)

⁹<http://animejs.com/>(2019年1月4日確認)

2中①参照), 調理に必要な材料を示す材料部(図2中②参照), 選択した料理名の料理レシピをフローグラフで示すフローグラフ部(図2中③参照)で構成されている。セレクトボックスには図1中の使用データに格納されている料理レシピ名が表示されている。まず, ユーザはそれらの中から自身が調理する料理レシピを選択する。ユーザが料理レシピを選択すると, それに合わせた材料と手順がそれぞれ表示される。フローは調理に関係のある用語がそれぞれブロックになっている。動作のブロックは赤色もしくは橙色でその他の要素のブロックは白色で表示している。一つの動作を一つのグループとし, 動作ブロックをグループの一番上に, その動作に必要な要素ブロックは動作ブロックの下に表示する。例えば, 図2の一番左のグループでは動作「水切りする」が一番上の橙色のブロックになっており, その下に材料「豆腐」が白色のブロックになっている。このように配置することで3章で述べたデザイン指針(3)を満たすことが可能になる。

また, 動作を並べたフローにすることで, ユーザは調理を行う前にどのような流れで調理を行う必要があるかを把握することができる。フローが分岐している部分は, 動作を行っている間に他の動作を行うことができることを示している。例えば, 図2の動作「水切りする」を行っている間に, 動作「冷ます」まで行うことを示している。調理の流れを把握する上でテキストだけが記述されている場合, 理解や想像ができない表現や用語があるため, 各動作のグループに対して動作の終了時の画像と注意喚起やコツをテキストで付与した(以下, 画像表示機能と記す)。ユーザは動作ブロックにマウスオーバーすることでそれを確認し, 用語や表現が理解できる(図2中④参照)。画像表示機能によって, 3章で述べたデザイン指針(1), (2)を満たすことができる。

また, ユーザに対して次に行うべき手順を把握させるために, 提案システムでは終了した手順をクリックすることでその手順が半透明になり, 次の手順のグループが上下に動くようにした(図3参照)。こうすることによってユーザは自身がどの手順まで行ったか, 次に何をすべきかを把握することができる。加えて, 直感的に調理工程を把握させるために, 提案システムではユーザの動作中にできた中間生成物(e.g., 炒め終わった玉ねぎ, 油を敷いたフライパン)が次に使われる場所を提示するようにした。中間生成物が次の手順以降で使用されている場合, 中間生成物の対応箇所ブロックは桃色に変化する(図3参照)。こうすることで, ユーザは食材や道具の状態の変化を把握でき, 次に使うべきものを理解することができる。中間生成物の表記はデータ元では省略されていることが多いため, 著者が手動で中間生成物のデータを作成した。これらの機能によってデザイン指針(4)を満たすことができる。

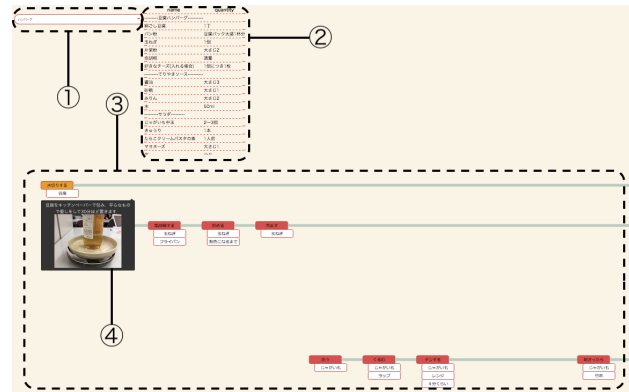


図2: 提案システム

5 評価実験

5.1 目的と概要

提案システムがユーザの調理の失敗を低減することが可能かを明らかにするため, 既存のテキストベースの料理レシピサイトである cookpad と提案システムとの比較を行った。この実験では, 実験参加者はテキストと画像ベースの料理レシピサイトである cookpad と提案システムのいずれかを使用し, 調理を行う。実験は大学生を対象に cookpad を参照して調理を行う cookpad 群, 提案システムを参照して調理を行う提案システム群の2群に振り分けた。実験は本研究の対象ユーザである料理歴の短い調理者を対象に, cookpad 群4人, 提案システム群4人の計8人に実施した。

5.2 実験準備

実験の準備として調理してもらった料理の品目を決定した。この実験では, 料理に慣れていない実験参加者が行ったことのない工程(e.g., 豆腐の水抜きを行う, ハンバーグのタネをこねる)が含まれる「豆腐ハンバーグとサラダの作成」を課題として実験参加者に課した。加えて, 手際よく行うためには時間を有効に活用する必要があるため, 本実験では調理が完成するまでの時間に余裕があれば洗い物を行うように指示した。今回は豆腐ハンバーグとサラダの料理レシピを使用した。調理環境として関西大学総合情報学部D棟3階のキッチンスペースを使用し, 材料は料理レシピに記述されている分量を用意した。実際の調理環境に近づけるために調味料は計量されていない状態で用意し, 実験参加者が自身で計量する形式にした。準備した調理道具を以下に示す。菜箸, ゴミ箱, 包丁, まな板, ボウル(大), ボウル(小2個), 計量カップ(2個), 計量スプーン(大きじ, 小さじそれぞれ1本ずつ) タオル(2枚), 薄手袋, 皿(4枚), スプーン(2本), フォー

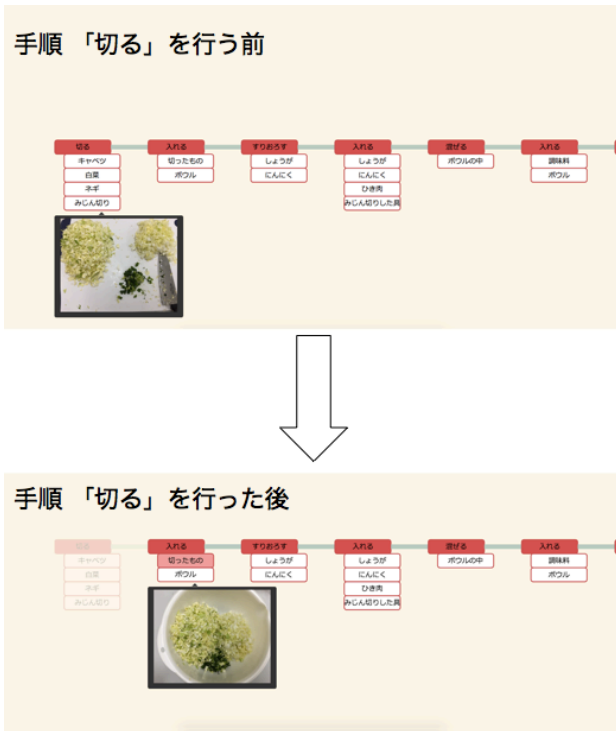


図 3: 一つの手順を行う前と行った後

ク, 電子レンジ, 重し, 直径 24 cm フライパン, 電熱コンロ, キッチンペーパー, サランラップ, 高さ調節用の作業台を用意した. 上記の道具以外を実験参加者に要求された場合, 観察者が渡す形式にした. 加えて, 実験参加者の許可を得て, 調理時の様子を録画した.

5.3 実験手続き

まず参加者を募集するにあたり, 調理経験の有無, 料理歴などを確認するために, Google アンケートを用いて事前に調査を行った. 質問項目を以下に記す.

1. 名前を記述してください
2. 年齢を回答してください
3. 料理歴はどの程度ありますか (ex. 一年, 学校の家庭科のみ, 親の手伝いをしていた)
4. どの程度の頻度で料理を行いますか (ex. 週 3 回, 月 2 回)
5. 料理をする際に参考にする媒体はありますか (Yes / No)

(5-1) 質問 (5) で Yes と回答した方は何を参考にするか記述してください (ex. cookpad, 料理本)

6. 食材に対するアレルギーを持っていますか (Yes / No)

(6-1) 質問 (6) で Yes と回答した方はアレルギーに関して具体的に記述してください (ex. 小麦, 卵)

7. ハンバーグを作ったことがありますか (Yes / No)

(7-1) 質問 (7) で Yes と回答した方はどの程度作ったことがあるか教えてください (ex. 全部, 混ぜるところだけ手伝った)

(7-2) 質問 (7) で Yes と回答した方は頻度はどのくらいですか (ex. 月に一度くらい, 今までに一回)

8. 連絡先を回答してください (実験参加していただく際以外に使用いたしません)

これらの中から料理歴と料理頻度, ハンバーグの調理経験の有無が同程度の 2 人を 1 組とし, 4 組 (計 8 人) を実験参加者に選出した.

次に実験の手順について述べる. 実験参加者に対して実験開始前に課題を説明し, cookpad 群にはサービスの使用方法, 提案システム群にはシステムの使用方法の説明を行った. その後, 実験参加者に調理を行ってもらった. 調理中に理解できないことがある場合, cookpad 群は cookpad で参照できる情報, 提案システム群は提案システムで参照できる情報のみを参照することを許可した. ハンバーグとサラダを調理し, 器への盛り付けが完了した段階で終了とした. 最後に, 参加者の料理中の思考を確認するために半構造型の調理後インタビューと成果物 (ハンバーグ) のアンケート評価を行った. インタビュー項目を以下に記す.

1. 料理してみてどうだったか
2. 自身の思った通りに調理ができたか
3. どのようにレシピを見ていたか
4. 調理しやすかったか
5. 手際よくできたか
6. 改善点はどこか

成果物の評価に関しては官能評価を用いたアンケートを行った. アンケートの評価項目を以下に記す.

- 水分の量 (1. 少ない, 2. やや少ない, 3. 適切, 4. やや多い, 5. 多い)
- 固さ (1. 柔らかい, 2. やや柔らかい, 3. 適切, 4. やや固い, 5. 固い)

- 焼き加減 (1. 生, 2. やや生, 3. 適切, 4. やや焼きすぎ, 5. 焼きすぎ)
- 見た目 (1. 悪い, 2. やや悪い, 3. どちらとも言えない, 4. やや良い, 5. 良い)

インタビューは参加者の許可を得て録音した。インタビューとアンケート評価の終了後実験を終了した。また、成果物の妥当的な評価を行うために、実験者、実験参加者とは異なる2名に成果物の評価を行ってもらった。この2名は全ての実験参加者の成果物を評価した。評価項目は実験参加者と同じ評価項目を用いたが、「見た目」の項目については、成果物を上から俯瞰した形で撮影した写真を評価者に提示し、評価してもらった。

5.4 実験結果

評価基準を(1)調理を手際よくできたか、(2)調理者がどのように調理を行ってどう感じたか、(3)完成した調理物がどのようなものか、に決定し、これらに対して定量的評価と定性的評価を行った。(1)に対して、定量的な評価として料理の完成時までの調理時間の平均値をそれぞれ比較した(図4参照)。それぞれの群に有意な差は見られなかった($t(6) = 1.097, p = 0.3148, n.s.$)。定量的な評価では両群の間に差はなかった。定性的な評価として、インタビューで調理者の主観評価を確認した。調理後インタビューの「手際よくできたか」の質問に対して、cookpad群は4人全員が「手際がよくなかった」と回答した。「玉ねぎを切る動作の効率が悪かった」、「テレビ番組とかとのスピードの違いを感じそれと比較してしまった」などの意見が得られた。一方で、提案システム群は4人中2人が「手際がよくなかった」と回答した。「思った以上にテンポよくいかなかった」、「どの程度で焼けるかの感覚がわからなかった」などの意見が得られた。残りの2人は明確に「手際がよかった」とは回答しなかったが、実験参加者Gは「手際がよかった部分もあれば、悪かった部分もある」、「玉ねぎを焦がさずに炒めながら他の作業を進められたが、ピーラーの使い方がわからなかった点は手際が悪かった」という意見が得られた。実験参加者Hは「手際について難しかったが、シンプルな工程だったので調理しやすかった」という回答が得られた。以上の手際に関する評価から、提案システムはcookpadと比較して、手際よく調理できる傾向があることが示唆された。

(2)に対して、定性的な評価として、実験参加者の調理姿の観察とインタビューで調理者の主観評価を確認した。2品を同時に作る並行作業についてcookpad群は4人中3人が並行作業を行っていた。実験参加者

Aは並行作業を行っていなかった。インタビューから実験参加者Aは「調理する前は並行作業を行おうとしていたが、一つの料理に手がかかったため集中せざるを得なくなった」という回答が得られた。提案システム群については、4人全員が提案システムが提示した順番で調理を行った。さらに、その中の2人は料理レシピを見て、調理中の待ち時間に先の工程で調理が可能な部分の動作を行う様子が見られた。

また、「調理しやすかったか」という質問に対して、cookpad群の中で明確に「調理しやすかった」と回答した人は4人中2人で「思ったほど難しくなかった」といった意見が得られた。「調理しやすかった」と回答しなかった実験参加者からは「料理経験がないのでわからない」、「携帯だと変わるかもしれないがパソコンだったので特に不自由はなかった」などの意見が得られた。提案システム群については4人中3人が「調理しやすかった」と回答した。「調理の腕は置いといて調理はしやすかった」、「部分的には難しかったが調理しやすかった」、「並行で調理することはわかりやすかった」などの意見が得られた。「調理しやすかった」と回答しなかった実験参加者からも「(画像と説明があるため)最初は全然迷わなかった」という意見が得られた。提案システム群4人中2人は調理しやすかった理由として「やるべき動作が最初に書かれていること」、「やるのがシンプルだった」を挙げた。以上の評価から、提案システムは調理者自身が思った通りに調理を行うことに対して有効であることが示唆された。

(3)に対して、定量的な評価として、実験参加者と評定者2名による成果物の評価の比較を行った。表1, 2, 3に調理者と評定者2名の各評価項目に関するアンケート結果を示す。評価項目の水分の量、固さ、焼き加減については“3. 適切”を一番良い評価とし、そこから離れている分だけマイナスとして得点化を行った。例えば水分の量が“2. やや少ない”だった場合、マイナス1得点とした。

実験参加者が評価した評価値(水分の量、固さ、焼き加減の合計値)の得点平均に対して、Wilcoxonの順位和検定を行った。その結果、それぞれの群の実験参加者の得点平均の場合 $W(4, 4) = 13.5, p = 0.1143, n.s.$ 、それぞれの群の評価者1のみの比較の場合 $W(4, 4) = 9, p = 1, n.s.$ 、それぞれの群の評価者2のみの比較の場合 $W(4, 4) = 13.5, p = 0.1143, n.s.$ であった。

また、評価項目の見た目については5段階のリッカート尺度を用いて評価した。実験参加者と評価者がそれぞれ評価した得点平均に対して、Wilcoxonの順位和検定を行った。その結果、それぞれの群の実験参加者の得点平均の場合 $W(4, 4) = 10, p = 0.7429, n.s.$ 、それぞれの群の評価者1のみの比較の場合 $W(4, 4) = 13, p = 0.1714, n.s.$ 、それぞれの群の評価者2のみの比較の場合 $W(4, 4) = 12.5, p = 0.3143, n.s.$ であった。以上の

表 1: 実験参加者のアンケート結果

実験参加者	水分の量	固さ	焼き加減	見た目
A	4	1	3	1
B	3	2	3	4
C	4	1	3	2
D	3	3	3	3
E	5	2	2	1
F	3	3	2	2
G	3	1	1	1
H	1	2	2	4

表 2: 評価者 1 のアンケート結果

実験参加者	水分の量	固さ	焼き加減	見た目
A	3	2	2	4
B	2	5	3	5
C	3	3	3	4
D	2	5	3	5
E	5	1	3	2
F	2	4	3	2
G	3	3	3	3
H	2	5	3	5

成果物に対する評価から、成果物に対しては cookpad 群と提案システム群の間に有意な差は見られなかった。一方で、今回の実験では cookpad 群が平均的に成果物の評価が良い傾向が見られた。

6 議論

提案システムで複数品の調理を同時に行う際にはフローの分岐は有効であることが示唆された。調理姿の観察から提案システム群の調理者全員が手順通りに料理を行った。実験参加者 G からは「列で分岐していることによってどの料理のどの食材に対して行う作業かを判断できる」、実験参加者 H からは「並行で調理を

表 3: 評価者 2 のアンケート結果

実験参加者	水分の量	固さ	焼き加減	見た目
A	3	2	3	2
B	3	3	3	5
C	3	2	2	4
D	3	4	3	5
E	5	1	2	1
F	3	3	4	2
G	4	2	2	1
H	1	5	3	5

調理時間平均

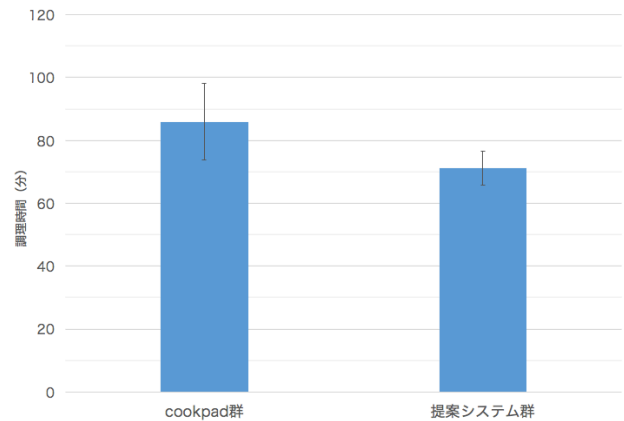


図 4: それぞれの群の調理時間の比較

行うことがわかりやすかった」という意見が得られた。しかし、この分岐は著者自身が作成したため、調理の効率に対して最適化されているとは言えない。そのため、今後は複数品を調理する際に最適化できる方法を検討する必要がある。

動作ごとのブロックわけについてインタビューからブロックにわけることによって調理手順を把握しやすくなることが示唆された。実験参加者 F からは「テキストベースで表示されている料理レシピには冗長に記述されているため何をすべきかわからないことがある」という意見が得られた。実験参加者 G からは「行う動作が記述されていることで自身が何をすべきかを判断できた」という意見が得られた。

また、「どのようにレシピを見ていたか」という質問項目に対して、実験参加者 E は「一度インタフェース上で見える範囲を確認した上で調理を行った」と回答した。実験参加者 F は「最初に全体を見てどの料理をどういう時間で作るかを確認した上で調理を行った」と回答した。F は自身が何も行うことがない状態 (e.g., ハンバーグを焼いている時間) に次の工程を確認していた。実験参加者 G は「自身の行っている動作とその次の動作を確認した」、実験参加者 H は「自身の行っている動作のみを確認した」と回答した。これらのことから、レシピの見方については現在行うべき手順を全員が把握できており、次の動作を意識させることができたこと示唆される。

動作終了時の画像が出る機能についてインタビューから調理者がイメージを把握するために効果的であることが示唆された。提案システム群の 4 人全員がこの機能を使用していた。実験参加者 G からは「終了時の画像が出ることによって、焼く際に洗い物ができた」という意見が得られた。G は終了時の画像の焼き色になるまで時間があると判断し、他の動作を行っていた。

また実験参加者 F は「自身がイメージできない工程については画像を参考にした」、実験参加者 E は「食材の見た目について画像を参考に動作を行った」と回答した。一方で、実験参加者 E は画像を見て調理を行ったが、キュウリをみじん切りする際に皮を剥くといった本来必要のない動作を行う様子が見られた。このことから調理者に誤解を与えない動作終了時の画像を提示する必要がある。

7 結論

本研究の目的は料理歴が短い人を対象に料理の失敗を低減することである。本稿では、料理レシピの提示方法として調理動作に着目し、調理動作をグループとし、それらを料理の手順通りに繋げることでフローグラフで表現するシステムの開発を行った。提案システムの有効性を確認するために提案システムと既存の料理レシピサイトである cookpad の比較実験を行った。比較実験の結果、提案システムは cookpad と比較して、手際よく調理できる傾向があることが示唆された。また、調理者自身が並行で調理を行う際に有効であることが示唆された。

現在の提案システムでは調理者が行う動作については支援できなかった。例えば、提案システム群の実験参加者からは「火の加減がよくわからなかった」、「じゃがいもを潰す際に上手いかなかった」などの意見が得られた。そこで、調理終了の画像の 1 枚だけではなく連続写真や gif 動画、もしくは失敗のリカバリ方法などを提示することで支援できる可能性がある。

また、調理後インタビューから「調味料や材料、道具が全て揃えられていることから調理を行うことが容易であった」という意見が得られた。本来、調理者は自身の調理環境で調理を行う。その際、調理に必要な調理器具や材料を選択する。そのため、実際の調理環境に近づけた実験を行うためにはダミーとなる調味料や材料、調理器具を用意する必要がある。

謝辞

本研究では、国立情報学研究所の IDR データセット提供サービスによりクックパッド株式会社から提供を受けた「クックパッドデータセット」を利用した。記して謝意を示す。

参考文献

- [1] Bollini, M., Tellex, S., Roy, N. and Rus, D.: Interpreting and Executing Recipes with a Cooking Robot, *In the 13th International Symposium on Experimental Robotics*, pp. 481–495 (2013).
- [2] Hamada, R., Ide, I., Sakai, S. and Tanaka, H.: Structural analysis of cooking preparation steps in Japanese, *Workshop on Information Retrieval with Asian Languages*, pp. 157–164 (2000).
- [3] 濱園侑美, 小林一郎, 麻生英樹, 長井隆行, 中村友昭, 持橋大地: ロボットの調理動作学習へ向けた取り組み, *情報処理学会第 77 回全国大会講演論文集*, Vol. 2015, No. 1, pp. 501–502 (2015).
- [4] 平島円, 磯部由香, 堀光代: 大学および専門学校生における「切り方」の難易度, *日本調理学会誌*, Vol. 50, No. 3, pp. 104–113 (2017).
- [5] 河内一行, 川端昌子, 鈴野弘子, 永島伸浩: 応用自在な調理の基礎フローチャートによる系統的実習書 *日本料理編 (改訂版)*, 家政教育社 (2015).
- [6] 吉川祐輔, 宮下芳明: グラフィカルデータフローによる調理レシピプログラミング言語の提案, *情報処理学会研究報告*, Vol. 2010-HCI-139, No. 4, pp. 1–7 (2010).
- [7] 北野直子, 我如古菜月, 川上育代, 池上由美, 沼田貴美子, 中嶋名菜, 江藤ひろみ: 大学生における調理に対する意識の現状と料理教室参加後の調理に対する意識および調理技術の変化, *日本食生活学会誌*, Vol. 22, No. 4, pp. 308–314 (2012).
- [8] 前田浩邦, 山肩洋子, 森信介: 手順文書からの意味構造抽出, *人工知能学会論文誌*, Vol. 32, No. 1E, pp. 1–8 (2017).
- [9] Ohsugi, T. and Matsushita, M.: Investigation of recipe components to resolve ambiguity in cooking procedure, *In Proc. 6th Asian Conference on Information Systems*, pp. 68–73 (2017).
- [10] Walter, K., Minor, M. and Bergman, R.: Workflow Extraction from Cooking Recipes, *In Proc. Computer Cooking Contest*, pp. 207–216 (2011).
- [11] Yamakata, Y., Imahori, S., Maeta, H. and Mori, S.: A method for extracting major workflow composed of ingredients, tools and actions from cooking procedural text, *Multimedia & Expo Workshops* (2016).

深層学習を用いた客観指標との比較による個人特有の嗜好の抽出 Extraction of Individual-Specific Preferences by Comparing with Objective Indicators using Deep Learning

渡邊 有輝^{1*} 砂山 渡² 畑中 裕司² 小郷原 一智²
Yuki Watanabe¹ Wataru Sunayama² Yuji Hatanaka² Kazunori Ogohara²

¹ 滋賀県立大学工学部電子システム工学科

¹ Department of Electronic Systems Engineering, The University of Shiga Prefecture

² 滋賀県立大学工学部

² School of Engineering, The University of Shiga Prefecture

Abstract: In recent years, studies to capture the trends of the world and the personal characteristics of SNS users have been promoted from SNS data. However, simply trying to capture the characteristics of individuals is mixed with the general characteristics of people around the world, and it is difficult to extract individual characteristic features. Therefore, in this research, we propose a method to extract personal preferences from the difference with this index, using the taste of many people in the world as an objective index. In other words, we construct a network that classifies comment sets in SNS into items within the field by deep learning, targeting one field. Then, personal preference is evaluated based on which item the individual comment is classified into many items.

1 はじめに

近年, SNS (Social Networking Service) のデータから世の中の傾向を捉える研究 [1][2][3] やユーザの個人的特徴を捉える研究 [4][5][6] が進められるようになってきた. しかし, 単純に個人の特徴を捉えようとすると, 世の中の人々がもつ一般的な特徴と混ざってしまう, 個人特有の特徴を抽出することができない. そこで本研究では, 世の中の多くの人の嗜好を客観指標として, この指標との差分として個人の嗜好を抽出する方法を提案する. すなわち, ある嗜好を表す分野を対象として, SNS におけるコメント集合をその分野内の項目 (構成要素) に分類するネットワークを深層学習により構築した上で, 個人のコメントがどの項目に多く分類されるかをもとに, 個人の嗜好を評価する.

2 深層学習による個人特有の嗜好抽出システム

この章では, 構築したシステムについて, システムの構成と個人特有の嗜好抽出までの流れについて述べる.

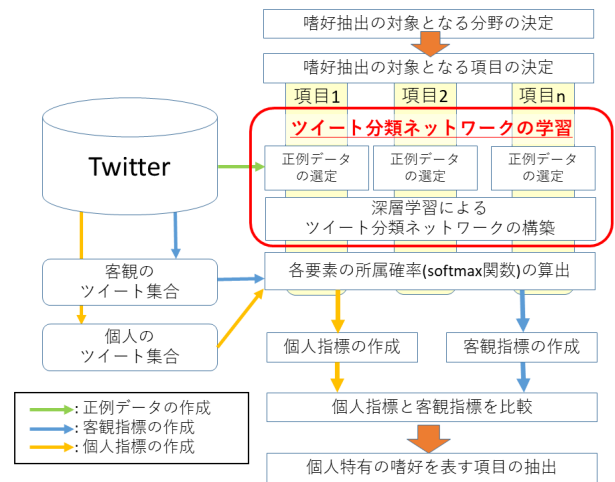


図 1: 提案する個人特有の嗜好抽出システムの構成

2.1 システムの構成

本研究で提案するシステムの構成を図 1 に示す. 分野内の項目 (構成要素) の正例データを深層学習の 1 つである DNN (Deep Neural Network) を用いて学習し, 分類ネットワークの重み付けを行う. そして, 構築した分類ネットワークに, 全体のテキスト集合を入力データとした場合の分類結果から, 一般の嗜好度

*連絡先: 滋賀県立大学工学部電子システム工学科
〒 522-8533 滋賀県彦根市八坂町 2500
E-mail: oi23ywatanabe@ec.usp.ac.jp

を表す客観指標の作成を行う。また、個人のテキスト集合を入力データとした場合の分類結果から、個人の嗜好度を表す個人指標の作成を行う。最後に、客観指標と個人指標の比較を行うことで、個人特有の嗜好抽出を行う。

2.2 抽出項目となる分野、項目の決定

本研究では、特定の分野における嗜好の抽出を行っている。そのため、嗜好抽出の対象となる分野と項目をあらかじめ決める必要がある。今回は、料理に関する嗜好抽出システム、アイドルグループ「乃木坂 46」のメンバーに関する嗜好抽出システムの構築を行った。

2.3 各項目の正例データの選定

本研究では、Twitter で投稿されたメッセージであるツイートを用いる。スマートフォンの利用と旅行消費に関する調査(2017)では、各 SNS での使用目的に関する調査をおこなっており、Twitter は「本音で語る」、「ちょっとした思い付き」、「趣味の活動」などが多く、等身大の自分自身を表現する目的の利用者が多いという結果が出ている。また、日本での利用者が 4500 万人を超えているという点からツイートを入力データとして用いている。

深層学習に用いる正例データを、以下の条件で収集した。料理の正例データとして各料理 1000 件、合計 3.5 万件を収集した。また、乃木坂 46 の正例データとして各メンバー 500 件、合計 2 万件を収集した。今回は、他の項目の正例データを負例データとして取り扱うため、負例データの収集は行わない。

- 項目名を含むツイート
- 他の項目名を含まないツイート
- 重複したツイートは 1 ツイートのみ使用
- 料理の場合は「おいしい」「うまい」を含むツイートを条件に追加

2.4 文章分類ネットワークの構築

本研究では、文章分類を深層学習を用いて行う。深層学習とは、中間層を 2 層以上に多様化したニューラルネットワークである。深層学習による文章分類の構成を図 2 に示す。学習ネットワークの構造は、基本的に入力層と 2 つ以上の中間層、出力層で構成されており、各層は情報を格納するノードで構成されている。また、ノード同士は情報を伝える線であるエッジで連結しており、情報伝達時にはエッジに伝わる情報を強化したり、減衰させたりする値である重みが与えられる。

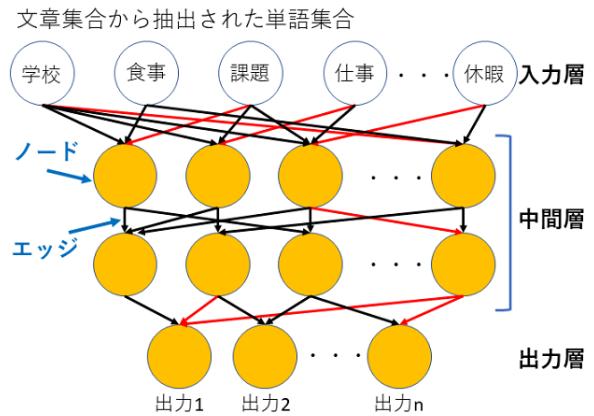


図 2: 深層学習による文章分類の構成

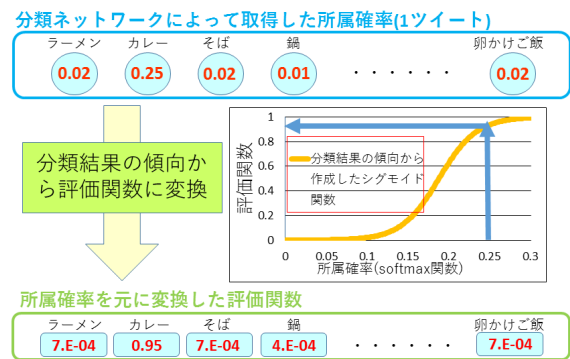


図 3: 分類結果を用いた指標の作成

本研究では、深層学習のライブラリのひとつである Deep Learning 4j(DL4j)[7] を利用して深層学習の実装を行う。

深層学習の入力データや抽出対象の文章は、データを数値化する必要がある。そのため、文章を形態素解析し、BoW(Bag of Words) への変換を行う。BoW とは、各文章での単語の出現頻度をベクトルで表現したものである。また、形態素解析には Igo[8] を用い、形態素解析用の辞書として MeCab[9] で使用される ipadic を使用した。

2.5 分類結果による指標の作成

分類結果を用いた指標の作成の流れを図 3 に示す。

本研究では、一般的な嗜好度を表す客観指標と、個人の嗜好度を表す個人指標の作成を行った。客観指標には、料理の客観指標には「おいしい」「美味しい」「うまい」「美味い」のいずれかを含むツイートを 150 万件、

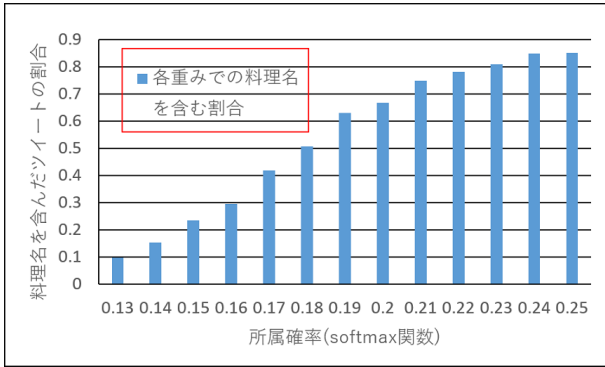


図 4: 料理名を含んだツイートの割合

表 1: ツイートの文章が料理名のみだった場合の softmax 関数値

料理名	所属確率	料理名	所属確率	料理名	所属確率
ラーメン	0.2067	ピザ	0.2551	ハンバーガー	0.2556
カレー	0.2397	おでん	0.2351	肉まん	0.2542
そば	0.1724	ハンバーグ	0.2021	牛丼	0.2121
鍋	0.2212	みそ汁	0.2110	コロケ	0.2214
寿司	0.2195	刺身	0.2173	とんかつ	0.1949
うどん	0.2164	たこ焼き	0.2501	サンドイッチ	0.2581
からあげ	0.1926	ステーキ	0.2160	豚汁	0.2115
焼肉	0.2070	チャーハン	0.2400	卵焼き	0.1759
パスタ	0.1945	オムライス	0.2238	肉じゃが	0.2380
おにぎり	0.1864	天ぷら	0.2094	親子丼	0.1855
ポテト	0.2427	お好み焼き	0.2175	卵かけご飯	0.2754
餃子	0.1981	シチュー	0.2455	所属確率の平均値	0.2201

乃木坂 46 の客観指標には、各メンバー名を含むツイート 90 万件の分類結果を用いて指標の作成を行う。また、個人指標には、個人のツイートの分類結果を用いて指標の作成を行う。

文章分類時の各出力層での所属確率を表す softmax 関数の値と、ツイート内容が各項目名のみだった場合の softmax 関数の値から指標の作成を行う。客観指標に用いたツイートの分類結果を softmax 関数の値ごとに、各項目名につき 100 件確認し、ツイート内に各項目名が含まれている割合を調べた。料理名を含んだツイートの割合を図 4 に示す。また、ツイートが各料理名のみだった場合、乃木坂 46 の各メンバーのみだった場合の softmax 関数の値を表 1、表 2 に示す。

以上の結果から、料理に関する分類ネットワークでは、softmax 関数値が 0.2201 の時に、乃木坂 46 に関する分類ネットワークでは、softmax 関数値が 0.1433 の時に評価関数が 0.8 となるようなシグモイド関数を作成する。

softmax 関数の値から評価関数への変換をツイートごとに行い、評価関数の合計値をそれぞれ客観指標、個人指標として用いる。作成した客観指標の上位 10 項目を表 3、表 4 に示す。

表 2: ツイートの文章が乃木坂 46 のメンバー名のみだった場合の softmax 関数値

メンバー名	所属確率	メンバー名	所属確率	メンバー名	所属確率
秋元真夏	0.1227	斉藤優里	0.1273	樋口日奈	0.1556
生田絵梨花	0.1350	阪口珠美	0.1528	星野みなみ	0.1186
伊藤かりん	0.1104	桜井玲香	0.1380	堀未央奈	0.1297
伊藤純奈	0.1218	佐々木琴子	0.1589	松村沙友理	0.1479
伊藤理々杏	0.1396	佐藤楓	0.1360	向井葉月	0.1385
井上小百合	0.1406	白石麻衣	0.1533	山崎怜奈	0.1777
岩本蓮加	0.1500	新内真衣	0.1498	山下美月	0.1420
梅澤美波	0.1415	鈴木絢音	0.1494	吉田綾乃クリスティー	0.1634
衛藤美彩	0.1530	高山一実	0.1625	与田祐希	0.1288
大園桃子	0.1486	寺田蘭世	0.1395	若月佑美	0.1467
川後陽菜	0.1407	中田花奈	0.1653	渡辺みり愛	0.1447
北野日奈子	0.1418	中村麗乃	0.1400	和田まあや	0.1299
久保史緒里	0.1584	西野七瀬	0.1391	所属確率の平均値	0.1433
齋藤飛鳥	0.1448	能條愛未	0.1474		

表 3: 料理に関する客観指標

順位	料理名	客観指標	順位	料理名	客観指標
1	ラーメン	29568.07	6	寿司	10332.29
2	カレー	19388.25	7	からあげ	8646.55
3	鍋	14959.47	8	パスタ	8585.22
4	そば	12834.28	9	焼肉	7340.67
5	うどん	11158.08	10	ポテト	7293.95

2.6 個人特有の嗜好抽出

前節の方法を用いて客観指標と同様に個人指標の作成を行い、客観指標と比較することで個人特有の嗜好抽出を行う。

今回は、個人のツイート 100 件を用いて個人指標の作成を行った。また、客観指標と個人指標をそれぞれ割合に置き換え、客観指標と個人指標での各項目の割合差から嗜好抽出を行う。本研究では、割合差の偏差値が 55 以上の項目を個人特有の嗜好として抽出している。料理と乃木坂 46 に関する嗜好抽出の一例を表 5、表 6 に示す。表 5 から、この利用者の個人特有の嗜好として「ハンバーグ」「たこ焼き」「ピザ」が抽出された。また、表 6 から「阪口珠美」「渡辺みり愛」「能條愛未」が抽出された。

個人特有の嗜好を抽出することにより、自分のツイートを入れた際は、主観視した場合と客観視した場合の結果を比較することが可能となり、主観だけでは分からなかった新たな嗜好の発見につながることを期待できる。また、自分以外の利用者のツイートを入れた場合は、共通の嗜好を持つ人を探すツールとして利用できるのではないかと考えている。

表 4: 乃木坂 46 のメンバーに関する客観指標

順位	メンバー名	客観指標	順位	メンバー名	客観指標
1	西野七瀬	115306.07	6	松村沙友理	30514.93
2	齋藤飛鳥	107105.05	7	与田祐希	27300.86
3	白石麻衣	77860.25	8	梅澤美波	24729.78
4	山下美月	34051.79	9	星野みなみ	22543.10
5	生田絵梨花	33150.39	10	山崎怜奈	21395.45

表 5: 料理に関する嗜好抽出

料理名	客観指標の割合	個人指標の割合	偏差値	抽出結果
ハンバーグ	2.19%	17.07%	81.05	
たこ焼き	1.97%	15.87%	79.01	
ピザ	2.31%	15.42%	77.34	
親子丼	1.45%	2.30%	51.77	
豚汁	1.18%	1.75%	51.19	
肉じゃが	1.06%	1.41%	50.73	
卵焼き	1.49%	1.83%	50.69	
コロッケ	1.44%	1.76%	50.66	
ハンバーガー	1.20%	1.45%	50.53	

3 個人特有の嗜好抽出結果の妥当性の検証実験

本章では、個人の嗜好抽出結果の妥当性の検証実験について述べる。本研究では、深層学習を用いて Twitter 利用者個人のツイート集合を分類し、分類結果から客観指標と個人指標を作成し、比較することで個人特有の嗜好抽出を行うことを目的としている。そのため、指標を用いた抽出結果が客観的に評価されている必要がある。今回は、本研究で構築した料理の嗜好抽出システムと乃木坂 46 の嗜好抽出システムを用いて検証を行った。

3.1 抽出結果が収束するために必要なツイート件数の検証

本実験で構築したシステムによる、個人特有の嗜好度を示す偏差値が収束するために必要なツイート件数の検証を行った。5 人の Twitter 利用者のデータを用意し、偏差値が何件で収束するのか検証を行った。今回は、ツイート数を 20, 50, 100, 200, 500, 1000, 2000 件にした場合に分けて検証を行った。

料理と乃木坂 46 に関するシステムでのツイート件数と嗜好度を示す偏差値の関係をそれぞれ図 5、図 6 に示す。

図 5、図 6 の結果から、1000 件が妥当なツイート件数であると判断した。また Twitter 利用者一人当たりの平均ツイート数が約 1400 件であるという結果も理由の一つである。個人の特徴を捉えることが必要なため、ツイート件数が多いに越したことはないが、必要なツ

表 6: 乃木坂 46 に関する嗜好抽出

メンバー名	客観指標の割合	個人指標の割合	偏差値	抽出結果
阪口珠美	1.17%	21.99%	99.40	
渡辺みり愛	1.35%	5.14%	58.99	
能條愛未	1.20%	3.56%	55.62	
斉藤優里	1.30%	2.87%	53.71	
向井葉月	0.71%	2.06%	53.20	
伊藤かりん	0.86%	1.97%	52.64	
星野みなみ	2.46%	3.32%	52.04	
中村麗乃	0.98%	1.77%	51.88	
吉田綾乃クリスティー	0.60%	1.34%	51.75	
和田まあや	0.88%	1.59%	51.70	

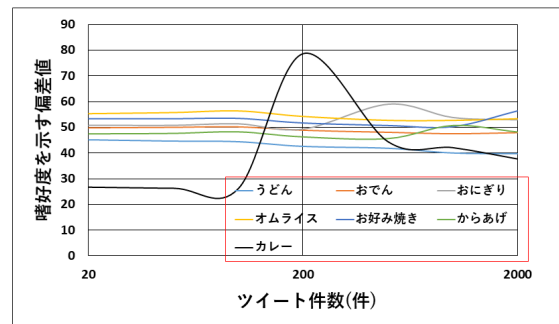


図 5: 料理に関するツイート件数と嗜好度を表す偏差値の関係

weet件数を増やしてしまうと一部の利用者には利用できないシステムになってしまう。よって、嗜好抽出に必要なツイート件数は 1000 件と結論付けた。

現在の手法における問題点として、大きく二つ挙げられる。一つ目は、ツイート件数が 1000 件に満たない場合の対処法である。解決策としては、利用者がリツイートといいねを行った他の利用者のツイートを入力データに追加することが挙げられる。リツイートは他の人のツイートに対して、良い情報や気に入った情報などを引用する場合に用いられており、また、いいね機能はそのツイートをお気に入り登録できる機能である。よって、個人特有の嗜好との関連性は下がる可能性はあるが、双方ともに利用者が興味を持っているツイートであるといえるので、入力データを増やすという点では利用できるのではないかと考えている。

二つ目は、ツイートの投稿時間を考慮していないという点である。今回のシステムでは、最新のツイートと 1000 件前のツイートを同等に扱っている。しかし、嗜好は常に同じではないため、最新のツイートの方が個人の嗜好抽出としては重視すべきである。そのため、分類結果の所属確率だけでなく、ツイートの投稿時間を評価関数へ変換する際の要素に追加することで、システムの精度向上と、抽出結果の収束に必要な件数の減少が期待できると考えている。

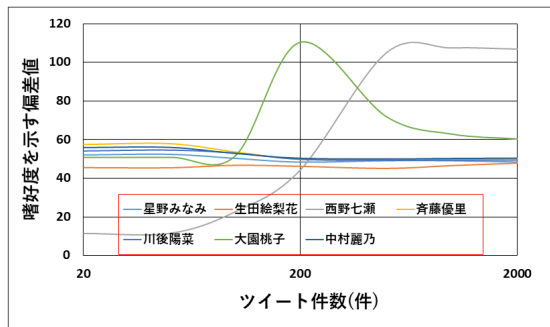


図 6: 乃木坂 46 に関するツイート件数と嗜好度を表す偏差値の関係

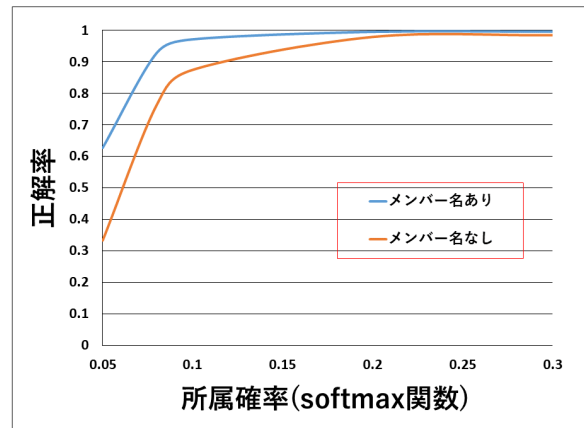


図 8: 乃木坂 46 に関する所属確率別での分類率

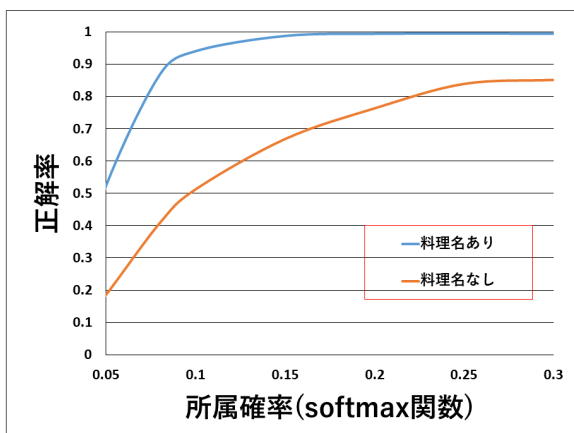


図 7: 料理に関する所属確率別での分類率

表 7: 深層学習で学習された各料理と関連が高い単語 (上位 10 単語)

ラーメン	カレー	寿司	餃子	ハンバーガー
カップ	インド	回転	王将	サンドイッチ
天ぷら	ナン	ネタ	宇都宮	バーガー
麺	キーマ	はま	皮	マクドナルド
チャーシュー	無印	シロー	浜松	マック
煮干	スパイス	回る	包む	ピエロ
揚げたて	レトルト	魚	水	ラッキー
蕎麦	グリーン	サーモン	モチ	アメリカ
タンメン	ステーキ	すし	焼ける	モス
こってり	チキン	北海道	祭り	パンズ
コシ	ルー	マグロ	クリームチーズ	コーラ

3.2 各項目の関連語に関する検証

分類ネットワークにおいて、項目名以外の単語から好きな項目を判定できるか検証を行った。各項目の関連語が学習されていると確認できれば、深層学習を用いた効果があると判断できるためである。

テストデータとして、各料理の正例データ 300 件 (合計 10500 件)、各乃木坂 46 メンバーの正例データ 250 件 (合計 1 万件) を用いて「醤油ラーメンおいしい」のような項目名を含む場合と、「醤油おいしい」のような項目名を削除した場合での分類率を比較した。分類結果をそれぞれ図 7、図 8 に示す。また、「ラーメンおいしい」などの項目名以外に関連語が含まれていないツイートを除くため、分類時の所属確率 (softmax 関数値) 別での分類率から比較を行っている。

図 7 から、料理名を含まない場合でも 6~8 割の分類は可能であるという結果が得られた。また図 8 から、乃木坂 46 のメンバー名を含まない場合でも 9 割の分類は可能であるという結果が得られた。

次に各料理の関連語を抽出し、関係のありそうな単

語が学習されているのか検証を行った。ツイート内容が、入力層に登録している単語 1 単語のみであった場合の分類結果から関連語の抽出を行った。料理の分類ネットワークでは学習データ 3.5 万件の内、20 回以上出現している 2438 単語を入力層に登録している。また、乃木坂 46 の分類ネットワークでは学習データ 2 万件の内、10 回以上出現している 2436 単語を登録している。深層学習で学習された各料理と関連が高い単語、乃木坂 46 各メンバーと関連が高い単語をそれぞれ表 7、表 8 に示す。

表 7、表 8 を見ると、結果から、関連している単語が学習されていることが確認できた。しかし、上位に他の項目が入っている場合も見られる。関連語として他の項目があるということは、文章分類によって選別する項目が独立しておらず、他の項目と強い関係があるということである。そのため、深層学習を用いる理由の一つである各項目の特徴語の学習が難しくなっている。

この問題を解決する手法としては、深層学習を用いて関連の深い項目をまとめるという手法が挙げられる。各項目の関連語に共通の単語が含まれている場合、そ

表 8: 深層学習で学習された乃木坂 46 各メンバーと関連が高い単語 (上位 10 単語)

秋元真夏	生田絵梨花	伊藤かりん	西野七瀬	佐々木琴子
真夏さん	いくちゃん	将棋	卒業	ダム
佐藤	ピアノ	甥	年間	アニメ
まなったん	アンコール	純奈	年内	松村
諦める	ロミオ	寺田	なっちゃん	塩
さんま	グルメ	有能	発表	小百合
技術	葵	フォーカス	エース	ルックス
真	ミュージカル	手塚	みなみ	真夏
松村	ジュリエット	蘭世	久保	物語
楓	陽菜	理々杏	中村	ガン
若月	川後	話	居る	芋

それぞれの項目は関連度が高いと予測できる。そこで、これらの項目を 1 つにまとめて学習を行うことで、ネットワークの簡潔化と分類の精度向上が期待できる。この手法では、現在の手法では抽出できなかった特徴語を取得できるのではないかと考えている。

4 結論

本研究では、Twitter 利用者を対象として深層学習による文章分類を用いた個人特有の嗜好抽出を行うシステムを構築した。また、システムによる抽出結果の妥当性の実験として各項目の関連語からも嗜好が抽出できるのか検証を行い、6~8 割のツイートでは分類が可能であったことから、抽出結果の妥当性が確認された。これにより、目視では困難であった個人特有の嗜好抽出が可能となり、個人の嗜好について客観的に見ることが可能となった。

今回のシステムは文章分類時の softmax 関数値を用いている。そのため、抽出する項目が独立しているほど有用なシステムであるため、個人特有の趣味を抽出するシステムの構築などに応用できると考えられる。

これを受けて今後の課題として、項目に関連した単語をより強く抽出する方法の検討や、時系列を考慮した抽出方法の検討を行い、より汎用性の高いシステムの構築を目標としていきたい。

参考文献

- [1] 谷 季恵, 松村 嘉之: Twitter 上の情報拡散がもたらす商品販売効果推定モデルの提案, 精密工学会 学術講演会講演論文集, pp. 3-4(2016)
- [2] 福井 一喜: 東京大都市圏に居住する若者の観光・レジャーにおける SNS 利用 「SNS 映え」を超越する若者たち, E-journal GEO, 14 巻 1 号 pp. 1-13(2019)

- [3] 澤山 郁夫, 三宅 幹子: 大学生の独り言的ツイートは独り言なのか 発話傾向との関連から, パーソナリティ研究, 27 巻 1 号 pp. 31-41
- [4] 山根 宏彰, 萩原 将文: SNS における統計情報による文章の嗜好推定, 日本知能情報ファジィ学会 ファジィシステムシンポジウム講演論文集, pp. 780-783(2013)
- [5] 原 侑平, 原 元司: トピックモデルを用いた Twitter ユーザの性別判定, 第 80 回全国大会講演論文集, pp. 537-538(2018)
- [6] 浅妻 佑弥, 山下 晃弘, 松林 勝: SNS 上への発言の特徴分析に基づくユーザの属性推定, 第 80 回全国大会講演論文集, pp. 541-542(2018)
- [7] DeepLearning4j, (URL)<https://deeplearning4j.org/index.html>
- [8] Igo, (URL)<https://github.com/sile/igo>
- [9] MeCab, (URL)<http://taku910.github.io/mecab/>

深層学習における学習ネットワークからの分類パターンの抽出

Extraction of Classification Patterns from Deep Learning Networks

安藤 雅行^{1*} 河原 吉伸^{2,3} 砂山 渡⁴ 畑中 裕司⁴
Masayuki ANDO¹ Yoshinobu KAWAHARA^{2,3} Wataru SUNAYAMA⁴ Yuji HATANAKA⁴

¹ 滋賀県立大学大学院工学研究科

¹ Graduate School of Engineering, The University of Shiga Prefecture

² 大阪大学産業科学研究所

² The Institute of Scientific and Industrial Research

³ 理化学研究所革新知能統合研究センター

³ RIKEN Center for Advanced Intelligence Project

⁴ 滋賀県立大学工学部

⁴ School of Engineering, The University of Shiga Prefecture

Abstract: In deep learning, there is a problem that concrete classification patterns for deriving reasons for classification are often incomprehensible. In this paper, we propose a classification patterns extraction system from deep learning networks and verified the effectiveness of the system. The proposed system takes out learning networks from the learning result of deep learning and extracts classification patterns from the learning networks. Then the system displays the extracted classification patterns so that users of the system can interpret the learning networks. In verification experiments, the significance of the extracted classification patterns was estimated by chi-square test. The results showed that users of the system can extract classification patterns effective for interpretations of the learning networks by using the proposed system.

1 はじめに

インターネットの普及に伴い、また、SNS (Social Networking Service) の出現によって、画像、テキスト、数値データが大規模になり、その処理や情報の抽出に機械学習が使用されるようになってきた。しかし、従来の機械学習は大量のデータから規則などを学習し、分類・予測を行う際、データのどの特徴(画像なら色や形など)に注目するかは人間が指定する必要があった。そこで注目されるようになってきた技術が、深層学習である。深層学習は近年流行りだした機械学習であり、学習を行う層(入力データの規則などを学習する部分)を多層化している。これにより、より人間の脳の学習に近い段階的な学習ができ、従来の機械学習と比べて学習の精度が高いという利点がある。

一方で、その深層学習による予測・分類基準が人間には不明な点が問題になってきている。特に、医療分野や自動運転では、その分類基準の理解は安全性において

重要視されている。仮にテキスト分野においても深層学習の判断基準をより深く理解できれば、医療分野において新人とベテランの書いた電子カルテの違いから、良い電子カルテを書く方法を容易に理解でき、企業においても良い報告書や企画書を書く方法を短時間で習得できるなど、深層学習の新しい活用が期待される。

本研究では、構造が複雑になる代わりに、単語の出現の時系列や順序も考慮した学習が可能な、再帰的深層学習 (Recurrent Neural Network や Long short-term memory など) を使用し、テキスト集合の学習によってネットワークの層に付けられた重みの値を取り出し、各層を構成するノードと呼ばれる要素を持つ情報を、単語として表現する。そして層間の単語の結びつきに時間の流れ(単語のテキスト中での出現順序)を当てはめることで、単語の順序を考慮した組み合わせとしての分類のパターンの抽出を行うシステムを提案する。

以下本論文では、2章で関連研究について述べる。3.2章で深層学習による分類パターンの解釈支援システムの構成と詳細について述べる。4章で提案システムの評価実験について述べ、5章で本論文を締めくくる。

*連絡先: 滋賀県立大学大学院工学研究科 電子システム工学専攻
安藤雅行

〒522-8533 滋賀県彦根市八坂町 2500
E-mail: oh23mandou@ec.usp.ac.jp

2 関連研究

インターネットの普及などにより、急速に大規模化しつつあるテキストへの対策として活用され始めているのが、深層学習を用いたテキストマイニングシステムである [1, 2]. 深層学習とは、一般に多層から構成されるニューラルネットワークを用いた学習を指し、例えば、深層学習の応用モデルである畳み込みニューラルネットワーク [3] の出現により、画像を用いた場合に限らず多くの場面で高い分類性能を実現できることが報告されている。

その一方で、深層学習は、その出力を導いた根拠についての解釈が困難であることも知られている。画像認識においては、この問題に対する研究も最近進められており、例えば、入力画像に対応する畳み込みニューラルネットワークにおける層間のスコアの勾配を計算することでネットワークの可視化を行う方法 [4] や、学習済みのネットワーク中間層のノード情報を用いて、対応する画像中の画素への寄与度を計算することにより画像の分類に重要な部位を表示する方法などが提案されている [5].

しかし自然言語への深層学習の適用においては、上記のような画像認識における方法を直接適用できない。そこで、アテンションと呼ばれる手法を用いた研究 [6, 7] が注目されている。アテンションとは、深層学習において分類・予測を行う際、出力に直接結びつく入力を探る手法で、このアテンションにより、出力に貢献する特徴は何かを視覚的にわかりやすくなっている。最新の研究では、アテンション計算を層ごとに行い、より分類・予測精度を高めた研究 [8] や、アテンションのみで構築された深層学習 [9] なども登場している。しかし、アテンションはあくまで入力と出力の関係のみに注目し、内部でどのような学習が行われているかは考慮していない。

そこで、自身の学士の研究 [10] では、テキストベースの深層学習について、層ごとの学習の流れを単語情報として表し、人間が理解できる形に直すことで、分類基準の理解のための、学習ネットワークの解釈を支援するシステムの開発を目的とし、一定の成果を得ることができた。一方で、この時使用した深層学習が、構造は単純だがテキストの単語の有無だけを特徴とし、単語の出現の時系列や順序を一切考慮しないものだったため、学習ネットワークの解釈が一定までしか得られなかった。

したがって、本研究ではこのような問題意識の下、文章(テキスト)の分類問題を例として、時系列関係を含めた分類に寄与する出力ごとの特徴を抽出できるように、再帰的ニューラルネットワークを用いた学習によってネットワークの各層に付けられた重みの値を抽出し、タイムステップごとの中間層が学習した情報を

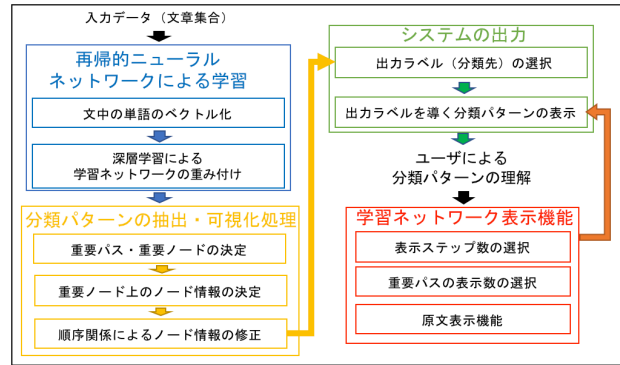


図 1: 分類パターン抽出システムの構成

文章中の単語として表現する。そこから、単語の順序関係を考慮した組み合わせとしての分類のパターン、つまり出力を導くルール抽出を行うシステムの構築を目指す。

3 深層学習の重みを用いたテキストの分類パターン抽出システム

本章では、本研究で開発した深層学習の重みを用いたテキストの分類パターン抽出システムについて、システムの構成とその詳細について述べる。

3.1 分類パターン抽出システムの構成

分類パターン抽出システムでは、まず、図 1 に示すように、各分類先ごとにラベル付けしたテキスト集合を RNN にて分類し、その分類先を導いた学習ネットワークと、学習ネットワーク上の重みから、提案システムの分類パターンの抽出・可視化処理によって各出力(分類先)を導くネットワーク上のパスと、パス上の各ノードの情報(そのノードで学習された単語)の決定、表示を行う。最後に、システムの利用者は、システムによって得られた学習ネットワークの表示を自分が見やすいように調整し、分類パターンを抽出する。そして分類パターンの意味を理解しやすくするための機能を利用できる。

3.2 深層学習による学習ネットワークの形成

3.2.1 文中の単語のベクトル化

深層学習で学習を行う前に、テキストデータは文中の単語を抽出したあと、単語は One hot 法 [11] と呼ばれる手法に従い単語ベクトルの羅列に直される。そして、文中の各単語を単語ベクトルに置き換え、深層学

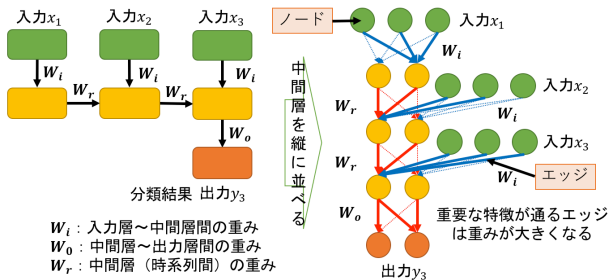


図 2: RNN の学習ネットワークと学習の様子

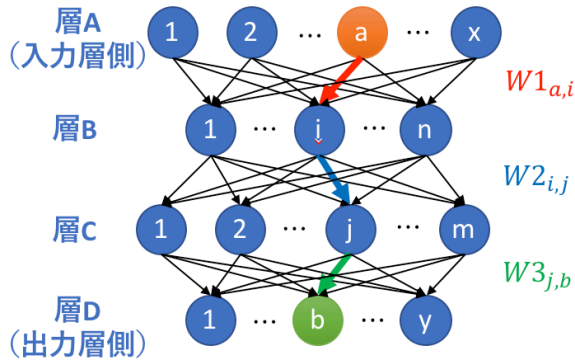


図 3: 4 層全結合型ネットワーク

習への入力データとする。なお、抽出される単語は名詞する。これは、文章の特徴とその順序関係をより学習・抽出しやすくするためである。

3.2.2 学習によるネットワークの重み付け

One hot 法によって単語ベクトルの羅列に変換され、さらにラベル付けされたテキストデータは、RNN の学習にて、それぞれの出力ラベル (分類先) を導くネットワークへの重み付けがされていく。その様子を図 2 に示す。入力文章は各単語がベクトル化され、タイムステップごとに単語ベクトルが順番に入力されていく。また、RNN での分類時は、最後の単語が入力されたタイミングで、出力層から出力される。

3.3 学習ネットワークからの分類パターンの抽出・可視化処理

本提案システムの分類パターンの抽出・可視化処理では、RNN によって得られた学習ネットワークから、各出力を導く最も関係が強いパス (ネットワーク上のエッジの繋がりの線) を決定する処理を行う。この出力と繋がりが強いパス (重要パスと呼ぶ) の決定について、図 3 に示した 4 層全結合型ネットワークモデルを例として、具体的な手順を述べる。

まず、ある分類先 (図 3 の例では層 D のノード b) に到達するパスについて、パス上のエッジについた重みの積で定義される、重要度と呼ばれる値を算出する。図 3 の入力層のノード i から太矢印のパス上のエッジに付いた重みを $W1_{a,i}$, $W2_{i,j}$, $W3_{j,b}$ とすると、このパスの重要度 $P_{a,i,j,b}$ は以下の式 (1) で導かれる。

$$P_{a,i,j,b} = W1_{a,i} \times W2_{i,j} \times W3_{j,b} \quad (1)$$

そして、式 1 で出力 b に到達する全てのパスの重要度を計算して比較し、最も値の大きいパスを、重要パスと決定する。出力 b の重要パスの重要度を S_b とし、その計算式 (2) を以下に示す。図 2 の右図のように、出力層と中間層間の重み W_0 のあと、中間層間の重み W_r を繰り返し辿ることで、過去の間層を遡ることになる。

$$S_b = \max_{i,j,b} B_{a,i,j,b} \quad (2)$$

次に、重要パス上のノードについて、ノードの情報を決定の処理を行う。ノード情報の決定方法は、出力ごとの重要パスを決定した時と同様に、入力層からノード情報を決定したい中間層ノード間の重み W_i の大きさから、ノード間の結びつきの強さを求める。ただし、図 2 に示す様に、各中間層にはそれぞれ個別の入力層が対応している。ここで、RNN の入力には One hot 法による単語ベクトルを用いているため、入力層ノードにはノードごとに 1 種類の単語が対応していることになる。そこで、重要パス上の入力層ノードの単語を参照し、その単語をノードの情報と決める。最後に、ノード情報の単語について、前後の中間層ノード情報の単語と、原文中でその順番で表示されているかどうかを確認し、されていないなら単語の重要度を下げる。

こうして各出力を導く重要パスとパス上のノード情報を表示した学習ネットワークから、過去の間層の重要ノードから現在の中間層の重要ノードの単語より、時系列を考慮した特徴の並びとしての分類パターンを抽出することができる。

3.3.1 出力ラベルを導く分類パターンの表示

本研究で開発した分類パターンの抽出システムでは、分類先に強く結びつく、重要パスの集合としての学習ネットワーク上に重要ノードの情報が表示される。例として、5 種類のお菓子の作り方に関するテキストの分類を行った場合の、システムのメイン画面を図 4 に示す。表示分類先は「マカロン」とする。このネットワークは、RNN が学習した情報を、過去から順番に中間層上に表示したものである。過去から出力層直前ま

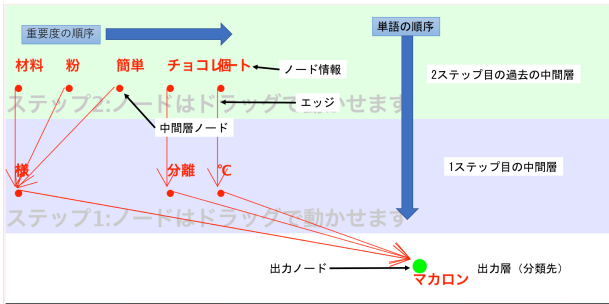


図 4: システムの画面

表 1: 学習ネットワーク表示機能

機能名	効果
表示するステップ数の選択	RNN において、何ステップ過去までの中間層を表示するか決定する
重要パスの表示数の選択	分類先ごとに、何本の重要パスを表示するか決定する
原文表示機能	分類パターン（順序を考慮した単語の組み合わせ）が、原文中でどのように出現しているかを表示する

での中間層の重要ノード情報（単語）の並びは、分類パターンを示し、重要パス 1 本につき 1 つの分類パターンが表示されていることになる

図 4 では、入力層を除いて選択したステップ数（図 4 では 2）だけ中間層と、出力層が表示され、出力層には分類された分類先を示す出力ノードが表示されている。また、分類先を示す出力ノードには、そこから選択したステップ数の長さで、選択した重要パス数だけ重要パスが表示されている。そして、重要パス上の各中間層ノードには、ノード情報として単語が表示されている。

3.4 システム上での学習ネットワーク表示機能

システムには、利用者が分類パターンの抽出・理解を行いやすいように、その表示内容を変更できる機能がある。その主なものを表 2 に示す。

4 分類パターン抽出システムの有効性の検証実験

本章では、深層学習の重みを用いた分類パターンの抽出システムにより、学習ネットワークから抽出でき

表 2: テキストデータの詳細

データ名	内容
童話 5 種類	日本の童話「かぐや姫」「鶴の恩返し」「さるかに合戦」「桃太郎」「浦島太郎」の、それぞれの概要やあらすじなどについて書かれたテキストをネット上 ¹ から 1 種類あたり 50 テキストずつ用意した。

表 3: 有意性のある分類パターンの総数

分類先名	カイ二乗検定での分類パターン総数	分類パターン総数
かぐや姫	93,211	15,527,540
鶴の恩返し	9,208	15,527,540
さるかに合戦	24,839	15,527,540
桃太郎	46,749	15,527,540
浦島太郎	53,334	15,527,540
平均	45,468	15,527,540

る分類パターンが、テキストデータを理解する（学習ネットワークの解釈を行う）を目的とした上で、各分類先に特徴的な分類パターンが抽出できているかを検証した実験について述べる。なお、本実験では、分類パターンを「2 種類の異なる単語の、順序を考慮した組み合わせ」と定義する。単語を 2 種類にした理由は、最も基本的な分類パターンであるためである。

4.1 実験準備

4.1.1 使用テキストデータと学習モデル

分類パターン抽出の対象とするテキスト「童話 5 種類」について、詳細を表 2 に示す。また、計算上の分類パターンの総数（抽出単語数 × 抽出単語数 - 抽出単語数）とその中で「ある分類パターンが、ある分類先に特有のものである」という仮説を有意水準 5% で検定したカイ二乗検定で、有意性があると推定された分類パターン数を表 3 に示す。

続いて、深層学習として使用した RNN モデルの概要を表 4 に示す。システム上での表示設定は、学習ネットワーク表示機能により、表示ステップ数 2、重要パス数 20 とする。また、テキスト「童話 5 種類」に対する RNN の分類精度は 100% であった。

4.2 実験手順

実験の対象者は著者 1 名とした。対象者はテキスト「童話 5 種類」について、本研究で構築した分類パター

¹Google: <https://www.google.co.jp> で「童話名 あらすじ」と検索し、表示された上位 50 位のサイトの本文をテキストとして利用

表 4: 学習モデルの詳細

データ名	抽出した 単語集合	入力層 ノード 数	中間層 ノード 数	出力層 ノード 数
童話 5 種 類	名詞 3,941 種の単語	3,941	50	5

表 5: 抽出された分類パターン

順位	かぐや 姫	鶴の恩 返し	さるか に合戦	桃太郎	浦島太 郎
1 位	人→命	家→鶴	自分→ 一緒	腰→日 本一	心→者
2 位	前→一 つ	日→子 供	柿→甲 羅	中→男	家→者
3 位	竹→一 つ	粗末→ 金	蟹→成 長	人→日 本一	声→事
4 位	年→一 つ	昔→彼	白→外	退治→ 日本一	魚→事
5 位	男→地	日→者	家→外	中→日 本一	二→誰 か
6 位	人 → 人々	家→不 思議	芽→甲 羅	成長→ 達	乙姫→ 浜
7 位	者→一 つ	前→子 供	木→仇	桃→日 本一	村→心
8 位	姫→幸 せ	心→帰 り	蜂→仇	日→日 本一	的→奥
9 位	頃→一 つ	矢→金	的→物 語	桃→幸 せ	玉手箱 →煙
10 位	始まり →駄目	三→百	自分→ 上	山→人	浦島→ 心

表 6: カイ二乗検定による分類パターン

順位	かぐや 姫	鶴の恩 返し	さるか に合戦	桃太郎	浦島太 郎
1 位	竹→姫	鶴→娘	種→柿	桃太郎 →退治	浦島→ 太郎
2 位	竹→月	鶴→家	柿→種	桃太郎 →鬼	浦島→ 玉手箱
3 位	姫→月	鶴→一	種→木	鬼→桃 太郎	太郎→ 玉手箱
4 位	竹→中	羽→娘	種→白	鬼→退 治	浦島→ 中
5 位	月→姫	一→晩	種→実	退治→ 鬼	浦島→ 乙姫
6 位	中→姫	晩→娘	柿→栗	桃→鬼	浦島→ 人
7 位	竹→人	鶴→晩	木→柿	退治→ 桃太郎	太郎→ 中
8 位	姫→人	一→羽	種→栗	桃→桃 太郎	太郎→ 乙姫
9 位	中→月	娘→一	柿→白	桃→退 治	乙姫→ 玉手箱
10 位	人→姫	日→羽	種→家	川→鬼	浦島→ 日

ン抽出システムによって、各分類先ごとに 20 個分類パターンを抽出した。そして抽出された分類パターンとカイ二乗検定で有意性があると推定された分類パターン上位 20 個（カイ二乗値順）とを比較し、抽出された分類パターンが他の童話にはない各童話に特有の時系列パターンとなっているかを検証した。

4.3 結果と考察

抽出された分類パターンのうち、重要度順に上位 10 個を表 5 に示す。また、比較対象として、カイ二乗検定で有意性があると推定された分類パターンのうち上位 10 個を表 6 に示す。

まず、表 5 をみると、「かぐや姫」では人→命（4 人のうち一人が命を失った）や竹→一つ（竹の一つが光っていた）など、「鶴の恩返し」では家→鶴（家に人に化けた鶴が来た）や粗末→金（機を売って粗末な暮らしに金が入ってきた）など、「さるかに合戦」では柿→甲羅（投げた柿が蟹の甲羅に当たった）や蟹→成長

(子蟹が成長した)など、「桃太郎」では腰→日本一(腰に日本一のきびだんごを付けていた)や中→男(桃の中に男の子がいた)など、そして「浦島太郎」では心→者(心の優しい若者がいた)や乙姫→浜(乙姫に別れを告げて元の浜に帰ってきた)など、それぞれの物語特有の特徴を持つ分類パターン(赤字で表記)が平均7個ほどと、多く抽出できたことがわかる。よって、本研究の提案システムでは、利用者が分類先に特有の分類パターンを抽出することができるようになると言える。

一方で、表6をみると、カイ二乗検定での分類パターンは、「かぐや姫」では竹→姫、「鶴の恩返し」では鶴→娘など、物語の大筋に関係がありそうな単語を含むものが多く、こちらの方が学習ネットワークの解釈に有効そうに見える。しかし、各分類先ごとの分類パターンには、意味が重複している余計な分類パターン(青字で表記)が上位に来ており、それらは学習ネットワークの解釈においては余分と考えられる。そして、カイ二乗値だけではどれが余分でどれが重要な分類パターンかを判別することは難しい。よって、学習ネットワークの解釈という点では、提案システムによる抽出された分類パターンの方が有効的と言える。

以上より、実験結果から、本研究で開発したシステムでは、カイ二乗検定での分類パターンと比較することで、深層学習の学習ネットワークから、学習ネットワークの解釈に役立つと思われる、分類先に対して特有の特徴を表す、有効性のある分類パターンの抽出が可能であると言える。

5 おわりに

本研究では、複数のテキストデータの分類を、単語の順序関係を学習できる深層学習であるRNNで行い、学習ネットワークの解釈を行うための、分類パターンの抽出システムの構築を目的とした。本研究の特徴として、重みを辿ることで、RNNの学習ネットワーク内の情報の伝達を過去に向かって探索している点が挙げられる。提案システムの有効性を確かめる検証実験では、提案システムで抽出された分類パターンとカイ二乗検定で有意性があると推定された分類パターンを比較し、提案システムによって抽出された分類パターンの方が、テキストの特有の特徴を理解するのに有効的であると結論づけた。今後の研究では、抽出された分類パターンからテキスト自体への解釈を行えるよう、分類パターン同士の組み合わせを用いた学習ネットワークの解釈支援を目標とする。

参考文献

- [1] ボレガラ ダヌシカ, “自然言語処理のための深層学習”, 人工知能学会誌, Vol.29, No.2, pp.195-201, 2014
- [2] Ebru Arisoy, Tare N. Sainath, Brian Kingsbury, Bhuvaba Ramabhadran, “Deep Neural Network Language Models”, In Proceedings of the NAA-CLHLT Workshop, Will We Ever Really Replace the N-gram Model?, pp.20-28, 2012
- [3] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, “Gradient-based learning applied to document recognition”, In Proceedings of the IEEE, 1998
- [4] Matthew D. Zeiler and Rob Fergus, “Visualizing and understanding convolutional networks”, In Proceedings of ECCV ’14, pp.818-833, 2014
- [5] 西銘 大喜, “ディープニューラルネットワークによる画像からの表情表現の学習”, 第29回人工知能学会全国大会, 3L4-3, 2015
- [6] M Daniluk, T Rocktaschel, J Welbl, S Riedel, “Frustratingly Short Attention Spans in Neural Language”, ICLR, 2017
- [7] A.Vaswani, N.Shazeer, N.Parmar, J.Uszkoreit, L.Jones, A.N. Gomez, L. Kaiser, and I. Polosukhin, “Attention Is All You Need”, CoRR, vol. abs/1706.03762, 2017
- [8] Jonas Gehring, Michael Auli, David Grangier, Denis Yarats, and Yann N. Dauphin. “Convolutional sequence to sequence learning”, arXiv preprint arXiv:1705.03122v2, 2017
- [9] A.Vaswani, N.Shazeer, N.Parmar, J.Uszkoreit, L.Jones, A.N.Gomez, L.Kaiser, I.Polosukhin, “Attention Is All You Need”, In the Annual Conference on Neural Information Processing Systems (NIPS), 2017
- [10] 安藤雅行, 河原吉伸, 砂山渡, 畑中裕司, “深層学習における分類パターンの解釈支援”, 2018年人工知能学会合同研究会 SIG-AM-20-02, pp.1-6, 2018
- [11] M. Feurer, A. Klein, K. Eggenberger, J. Springenberg, M. Blum, and F. Hutter. “Efficient and robust automated machine learning”, In Neural Information Processing Systems (NIPS), 2015

辺ベクトルを用いた経路推薦問題の解法における 目的関数の改善に関する提案

Proposal to Improve Objective Function for Route Recommendation Problem Using Edge Vector-based Formulation

柴田祐樹¹ 高間康史^{1*}
Hiroki Shibata¹ Yasufumi Takama¹

¹ 首都大学東京大学院システムデザイン研究科

¹ Graduate School of System Design, Tokyo Metropolitan University

Abstract: This paper proposes to improve the objective function used for the route recommendation problem using the Edge Vector-based formulation. In this formulation, there are real and virtual edges. Real edges represent the routes of a given problem, and virtual edges are used to transform a non-complete graph like a real world road map into a complete graph. However, the effects of real and virtual edges to the objective function are not symmetric, which decreases the performance of the optimization when a state variable includes virtual edges. This paper shows an alternative formulation of the objective function that satisfies the symmetry on the effect of real and virtual edges. The effectiveness of the proposed method is shown by the experiments on a real world data set and an artificial data set.

1 はじめに

観光旅行をする際、ガイドブックや旅行会社が提供する観光ツアーを利用するのではなく、Social Networking Service (SNS) などの Web から入手した情報を用いて旅行者自身で観光計画を立案することも増えてきている。観光情報サイトなどでは、エリアなどの指定条件を満たす観光スポットをランキング形式で提示するサービスを提供しているが、知名度の高いスポットや宣伝に力を入れているスポットが上位に表示される傾向がある。ユーザの検索行動として、ランキング上位の数件のみしか確認しないことが一般的であるため、結果として個人の趣味嗜好に合った観光地を訪れる機会の損失を招き、一部の観光地へ旅行客が集中するなど、旅行客、観光地側双方に不利益を生じる。こういった問題を防ぐためにも、ユーザの好みに応じたスポットを選び、限られた時間で効率よく巡回する経路を提案する観光経路推薦手法は重要であり、研究が進められている [1][2][3]。

与えられた節点全てを巡回する最短経路を求める問題は Traveling Salesman Problem (TSP) [4] として古くから研究されている。NP-Hard [5] に属する TSP を解くためには近似解法が用いられ、代表的なものには

2-Optimization (2-Opt) 法 [6]、Simulated Annealing (SA) [7] がある。さらに、スポットに価値を付加し、与えられた総負荷制限の中で価値の総和が最大となる経路を選ぶ問題として、Selective Traveling Salesman Problem (STSP) [8] が定式化されている。文献 [1]、[3] では STSP に基づく問題の定式化において、スポットの価値を Web から取得したデータを用いて推定する方法を提案している。また、文献 [2] は STSP を拡張し、訪れた時間に依存して価値が変化する問題を最適観光経路問題として定式化することを提案している。

以上のように、TSP および STSP を拡張した定式化は広く用いられているが、STSP では節点にスポットを割当て、節点間の辺の重みに、スポット間を結ぶ最短経路等の固定的な負荷 (コスト) を割り当てるため、特定の交通機関や道路しか用いられないことに等しい。これにより、途中経路選択に対するユーザの価値観を考慮できず、また、ユーザの集中による混雑の発生等の問題も考えられ、Web から得られるデータが潤沢となっている近年の状況において適切なモデル化であるとは言えない。これに対し文献 [9] ではスポットのみでなく、途中の経路における街並みや景観等もユーザにとって重要な因子であるとし、これらを考慮可能な観光経路推薦手法 (以降 Edge Vector: EV と呼ぶ) を提案している。スポットおよび経路に関する価値や負荷すべてを辺に対する重みとして定式化することで、ス

*連絡先: 首都大学東京大学院システムデザイン研究科
〒191-0065 東京都日野市旭が丘 6-6
E-mail: ytakama@tmu.ac.jp

ポットと経路の情報を統一的に扱い、最適化の過程において両者を同時に調整しながら最適経路の発見を可能としている。

EV において、最適化手法として SA が用いられる。最適化の過程では問題を構成する経路に対応しない辺 (仮想辺) が用いられ、仮想辺が実際の経路に対応する辺 (実辺) に置換されるように目的関数を定式化することで非完全グラフに対応しつつ、実辺上のスポットを探索可能としている。しかしながら、目的関数に対する実辺・仮想辺の影響が対称でないことが置換の過程に影響し、最適化効率の低下が予測されるため、対称性が最適化に及ぼす影響の調査、およびそれに基づいた定式化の改良を行う必要があると考える。

以上の背景のもと、本稿では、対称性が最適化に与える影響を数学的に考察し、対称性の低い定式化では最適化効率が低下する場合があることを示す。この問題に対し、実辺・仮想辺について対称性の高い目的関数の定式化を提案し、最適化効率を改善できることを示す。人工データセットおよび文献 [1] で用いられているテーマパークにおける観光客の行動履歴から作成されたデータセットを用いた評価実験により、提案する定式化の有効性を示す。

2 関連研究

2.1 STSP に基づく経路推薦

TSP には全ての節点を一度ずつ訪れるという制約があるため、スポットに対して取捨選択性のある観光経路推薦問題にこの定式化をそのまま適用することはできない。既存研究では、訪れる節点に関する制約の代わりに、通過する経路の負荷の合計 (総所要時間等) に対して制約を設け、スポットに設定された価値の総和を最大化する問題である STSP [8] の定式化を利用、あるいは拡張したのことが多い [1][2][3]。

文献 [1] では、待ち時間が頻繁に発生するテーマパークにおける経路推薦手法を提案している。アトラクションを節点 (その集合を \mathcal{N} とする)、巡回する節点の総数を D として、推薦する経路を節点の列 $n = (n_i \in \mathcal{N} | i = 1, \dots, D)$ 、各節点を訪れた時刻を $t = (t_i \in \mathbb{R} | i = 1, \dots, D)$ と表現している。 n_1, n_D はそれぞれ固定的な始点、終点であり、最適化の過程で変化することはないが、巡回する節点の総数 D は最適化の過程で変化する。 U をユーザの集合とし、各アトラクションにはユーザごとに異なる価値 (好み) $f_{\text{int}} : U \times \mathcal{N} \rightarrow [0, \infty)$ 、全ユーザで共通の人気度 $f_{\text{pop}} : \mathcal{N} \rightarrow [0, \infty)$ 、ある時刻における待ち時間 $t_{\text{queue}} : \mathcal{N} \times \mathbb{R} \rightarrow [0, \infty)$ 、アトラクションの所要時間 (待ち時間を除く) $t_{\text{dur}} : \mathcal{N} \rightarrow [0, \infty)$ が定義され

る。また、全てのアトラクション間について、移動時間 $f_c : \mathcal{N}^2 \rightarrow [0, \infty)$ が定義されている。以上の設定において、ユーザ $u \in U$ に対してテーマパーク内のアトラクションを巡る最適な経路を推薦する問題を、式 (1) で与えられる所要時間 $C_c \in [0, \infty)$ の制約のもと、式 (2) に定義される目的関数 F を最大化する n を求める問題と定式化している¹。

$$\sum_{i=1}^D \{f_c(n_i, n_{i+1}) + t_{\text{dur}}(n_i) + t_{\text{queue}}(n_i, t_i)\} < C_c \quad (1)$$

$$F = \sum_{i=1}^D \left(\frac{f_{\text{int}}(u, n_i) + f_{\text{pop}}(n_i)}{t_{\text{queue}}(n_i, t_i)} \right) \quad (2)$$

解法としては Monte Carlo Tree Search [10] を拡張した PersQ と呼ばれる手法が提案されている。STSP では式 (1) の制約の他に、各スポットを訪れる回数が一度以下である制約、始点と終点を必ず訪れる制約、経路が繋がっている制約を満たす必要があるが、PersQ はこれらを満たす解のみを探索するように実装されている。

文献 [2] では、STSP をスポットにおける価値の時間的変化も考慮するように拡張した定式化を提案している。文献 [3] は、ユーザと対話的に経路を決定していく支援システムである CTPlanner を提案している。問題の定式化には STSP を用い、辺の重みには Google Map 等の経路案内サービスを用いてあらかじめ算出した値を移動負荷として設定している。解法には遺伝的アルゴリズム [11] を用いている。

2.2 辺ベクトルを用いた観光経路推薦手法の定式化

本節では EV の定式化と、同手法で用いられる近似解法について述べる。また、各用語の命名が文献 [9] において曖昧であると考えられるものについては、対応関係を示しながら、本稿で新たに定義したものをを用いて説明する。

\mathcal{N} を節点の集合、 \mathcal{E}_p を辺の全体集合とする。 \mathcal{E}_p は問題に与えられた辺 (定義辺) と後述する自己回帰辺、仮想辺²を含む。辺 $e \in \mathcal{E}_p$ に対し、 $n_{\text{st}}(e)$ を e の始点、 $n_{\text{ed}}(e)$ を e の終点としたとき、これらは $\bigcup_{e \in \mathcal{E}_p} (n_{\text{st}}(e), n_{\text{ed}}(e)) = \mathcal{N}^2$ の関係を満たすものとする。

$|\mathcal{N}|$ 個の辺からなる経路を含む辺ベクトルを $\mathbf{e} := (e_i | i = 1, 2, \dots, |\mathcal{N}|)$ と定義する。 $n_{\text{st}}(e_1)$ は経路の始

¹式 (1) と式 (2) による問題の定義は文献 [1] に書かれているものと多少異なるが、説明を簡潔にするため最低限必要な条件のみを満たすように記述している。

²文献 [9] では実在辺、自己ループ辺、非存在辺と呼ばれる

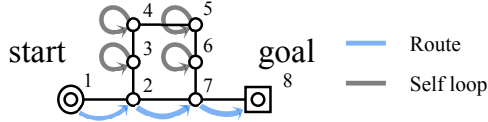


図 1: グラフの例.

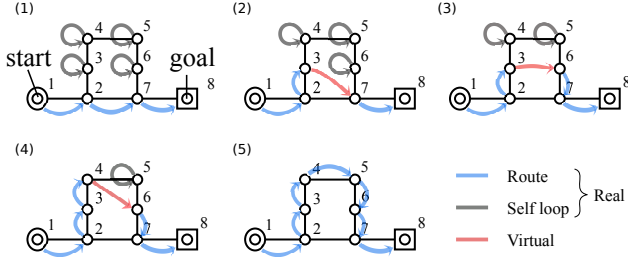


図 2: 仮想辺を用いた経路拡張の例. (1) から (5) にかけて経路が拡張されていく様子を描いている.

点, $n_{ed}(e_{|\mathcal{N}|})$ は経路の終点であり, 最適化中変化しない. 辺 $e \in \mathcal{E}_p$ のうち, $n_{st}(e) = n_{ed}(e)$ であるものを自己回帰辺と呼ぶ. 辺ベクトルは自己回帰辺を含むことで, 要素数を固定としながら様々な経路を表現することができる. 要素数が固定であるため, 後述するように Gibbs Sampling [13] を適用可能となる.

図 1 では, 定義辺を黒線で表し, 輪を描いているものが自己回帰辺, 青で示すものは辺ベクトルに含まれる定義辺である. この図は, 節点 1, 2, 7, 8 をたどる経路の例であり, 対応する辺ベクトルは $((n_{st}(e_i), n_{ed}(e_i)) | i = 1, 2, \dots, 8) = ((1, 2), (2, 7), (3, 3), (4, 4), (5, 5), (6, 6), (7, 8))$ となる.

また, この手法では, 仮想辺を用いることで, 2 辺に対する操作のみにより経路の変更をしながら最適化を行う. 図 2 に示す経路変更の例では, 赤線により仮想辺を示している. 定義辺上に乗らない仮想辺を経由することで, 本来は 4 つの自己回帰辺を定義辺に変換する必要がある変化を, 2 辺への置換を 4 回繰り返すことで表現している. 本稿では, 定義辺と自己回帰辺を合わせて実辺と呼び, その全体集合を \mathcal{E} ($\subset \mathcal{E}_p$) とする. 仮想辺の集合は \mathcal{E} の補集合となる.

経路は Boltzmann 分布 [12] から確率的に生成されるとモデル化し, 最適解を SA により近似的に求める. e の確率変数を e_A , その分布 $p(e_A)$ を式 (3) により定義する. $A = \{1, 2, \dots, |\mathcal{N}|\}$ は辺ベクトルの各要素に対する添字の集合である. 本稿では, 確率モデルの表記に文献 [14] で提案されている添字記法を用いる. この記法において, 同じ添字集合を添字に持つ確率変数は独立同分布に従う.

$$p(e_A) = \frac{\exp(-\frac{1}{kT}\phi(e_A))}{\sum_{y_A} \exp(-\frac{1}{kT}\phi(y_A))} \quad (3)$$

ここで, \sum_{y_A} は y_A の状態空間についての総和を示

す. T は温度である. また, k は Boltzmann 定数であり, 問題ごとに適切な値を設定する. $\phi(e_A)$ は目的関数であり, この値が小さくなる程, e_A の出現率が高くなり, T が小さくなるほどこの傾向が強くなる. SA ではこの性質を用いて最適解の近似解を生成するため, 目的関数の値が小さな状態が望ましい経路を表すように $\phi(e_A)$ の定式化を行う.

$\phi(e_A)$ を式 (4) により定義する. 文献 [9] の定式化と多少異なるが, 本質的には同じであり, 3 節における考察と提案手法の説明に同一の式を利用可能とするためにこの式を用いる. $\alpha_s = 0$ としたものが, EV の定式化に対応する³.

$$\phi(e_A) = F_{res} \left(\sum_{i \in A} f_c(e_i) + \alpha_s \sum_{i \in A} f_s(e_i) - C_c \right) - \sum_{i \in A} f_d(e_i) + \alpha_{as} \sum_{i \in A} f_s(e_i) \quad (4)$$

$f_c(e)$ は e の負荷, $f_d(e)$ は e の価値であり, 両関数の値は辺の向きによらず, 以下の条件を満たすものとする. 2 つ目の条件は自己回帰辺に対するもの, 3 つ目の条件は後に説明する STSP の制約を無制限に超えた経路が生成されることを防ぐものである.

$$\forall e \notin \mathcal{E}, f_c(e) = f_d(e) = 0,$$

$$\forall e \in \mathcal{E}, n_{st}(e) = n_{ed}(e) \Rightarrow f_d(e) = f_c(e) = 0,$$

$$\forall e \in \mathcal{E}, n_{st}(e) \neq n_{ed}(e) \Rightarrow 0 \leq f_d(e) < f_c(e)$$

$f_s(e)$ は仮想辺に対する負荷であり, 式 (5) により定義される.

$$f_s(e) = -f_c(e) + \sum_{s \in S(e)} f_c(s) \quad (5)$$

ここで, $S(e)$ ($\subset \mathcal{E}$) は辺 e の終始点を結ぶ最短経路を構成する辺の集合である. また, $S(e) = \emptyset \Rightarrow f_s(e) = \infty$ を満たすものとする. これにより, 仮想辺には実辺からなる最短経路負荷以上の負荷が割り当てられるため, 仮想辺を含む状態に対し, 常に実辺のみからなる目的関数の値がより低い状態が存在することになり, 最適化の途中で仮想辺は実辺に置換される傾向となる. また, 前述した SA の性質により, $T \rightarrow 0$ の極限において仮想辺はすべて実辺に置換される. 最短経路負荷は Floyd-Warshall 法 [15] により求めることができる. この手法は最大で $O(|\mathcal{N}|^3)$ の計算量, $O(|\mathcal{N}|^2)$ の記憶量を要するが, 事前に求めて置くことが可能であるため, 仮想辺の導入によるこの計算負荷の増加は実用上問題とならない. EV の定式化において, α_{as} は仮想辺

³文献 [9] ではさらに f_s に対し微小な定数項を加算している等の違いがあるが, それらの影響がほぼ無いことを事前実験において確認しているため, 詳細については省略する.

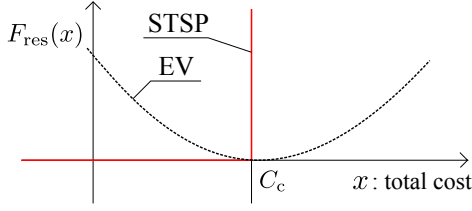


図 3: 制約関数の概形. STSP: STSP, EV: 文献 [9].

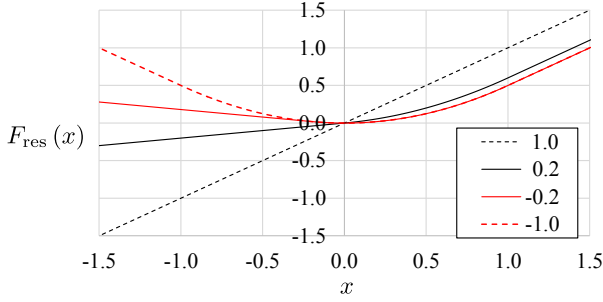


図 4: $\sigma^2 = 1, \alpha_g \in \{1.0, 0.2, -0.2, -1.0\}$ における $F_{\text{res}}(x)$ のグラフ.

を消滅させるための罰則係数に対応するものであり、1 以上で、かつ仮想辺が最終的に問題なく消滅する程度に 1 に近い値を用いる。

$F_{\text{res}}: \mathbb{R} \rightarrow \mathbb{R}$ は経路負荷に対する STSP の制約を表現する関数であり、 C_c は制約値である。観光経路推薦においてこの制約値は、経路長や旅行時間に対応する。STSP, EV における制約関数の概形を図 3 に示す。STSP では経路長が C_c を超えることを認めないため、目的関数は C_c を堺に無限大の値を取る。このような目的関数を用いた最適化では、経路長が C_c 付近となる状態の近傍探索における最適化効率が低下する。また、興味のあるスポットがあれば、多少の制約違反が許容される観光旅行においてこのモデル化は妥当であるとは言えない。これに対し、EV ではある程度の誤差を認め、最適化効率を向上させたいうで、より現実的なモデルとしている。

EV と 3 節で説明する提案手法で共通して利用する制約関数の定式化を式 (6) に示す。

$$F_{\text{res}}(x) = \begin{cases} \alpha_g x + c_1 & , x < x_m \\ c_2 x^2 + c_3 x & , x_m \leq x \leq \sigma^2 \\ x + c_4 & , \sigma^2 < x \end{cases} \quad (6)$$

各設定値を、 $c_1 = c_2 x_m^2 + c_3 x_m - \alpha_g x_m$, $c_2 = (1 - c_3)/(2\sigma^2)$, $c_3 = \max(\alpha_g, 0)$, $x_m = (\alpha_g - c_3)/(2c_2)$, $c_4 = c_2(\sigma^2)^2 + c_3\sigma^2 - \sigma^2$ と定めたとき、 $\alpha_g = -1$ において、この関数は EV のものに対応する。 σ^2 は制約違反の許容量を決定するパラメータであり、これにより違反の許容度合いを調節することが可能である。図 4

に α_g の値による F_{res} の形状の違いを示す。

SA において、Boltzmann 分布を近似するために Gibbs Sampling (GS) [13] を用いる。経路変更を行うための最小要素は二つの辺であるため、GS において利用される部分空間として辺ベクトル中の 2 辺を選ぶ。この 2 辺に対する添字を $i, j (\in A)$ としたとき、これら以外の添字 $A \setminus \{i, j\}$ で定義される変数により条件付けられた分布は式 (7) と書ける。

$$p(\mathbf{e}_{\{i,j\}} | \mathbf{e}_{A \setminus \{i,j\}}) = \frac{\exp(-\frac{1}{kT}\phi(\mathbf{e}_A))}{\sum_{\mathbf{y}_{\{i,j\}}} \exp(-\frac{1}{kT}\phi(\mathbf{y}_{\{i,j\}}, \mathbf{e}_{A \setminus \{i,j\}}))} \quad (7)$$

添字 A の部分集合 M を添え字に持つ確率変数は対応する部分空間を構成し、式 (8) の性質が満たされる [14].

$$p(\mathbf{e}_{A \setminus M}) = \sum_{\mathbf{y}_M} p(\mathbf{y}_M, \mathbf{e}_{A \setminus M}), M \subset A \quad (8)$$

ここで、 $\sum_{\mathbf{y}_M}$ は \mathbf{y}_M の状態空間についての総和を示す。この状態空間には 2-Opt 法に対応する、経路としての制約を満たすものを与える。

操作対象となる 2 辺の添字集合 $\{i, j\}$ は状態変化が起こらないような無駄な GS の試行を削減するように選ばれる。EV ではこの選択法による GS の適用法を近似的な棄却サンプリングと呼んでいる。この選択法では目的関数の変化量を近似的に見積もり、予測される変化量の小さい方から K 番目以内までの状態への変化を与えるものの中から $\{i, j\}$ を無作為に選択する。SA の基礎となる Markov Chain Monte Carlo 法の枠組みにおいては、提案分布 [16] による探索範囲を近傍の状態に制限する方法と見る方が適切であるため、本稿ではこの手法を近傍探索法と呼ぶ。

3 対称性に対する考察と提案手法

式 (4) 中のパラメータについて、EV では $\alpha_s = 0, \alpha_{as} \approx 1 \wedge \alpha_{as} > 1$ としており、仮想辺の負荷 f_s は制約関数 F_{res} に含まれないため、仮想辺と対応する実辺の入れ替えに対して目的関数の値が大きく変化する。仮想辺と実辺に関して非対称な定式化となっている。

非対称性が最適化に及ぼす影響について、以下により定義される $F_c(\mathbf{e}_A), F_s(\mathbf{e}_A)$ を用い、次のように考察を行う。

$$F_s(\mathbf{e}_A) = \sum_{i \in A} f_s(\mathbf{e}_i)$$

$$F_c(\mathbf{e}_A) = \sum_{i \in A} f_c(\mathbf{e}_i)$$

ある状態 e_A からの $\phi(e_A)$ の変化量を $\Delta\phi$, $F_s(e_A)$ の変化量を ΔF_s , $F_c(e_A)$ の変化量を ΔF_c と表し、経路の様々な遷移過程について、仮想辺を対応する実辺集合に置換したとき、およびその逆の置換を行ったときの $\Delta\phi$ について考える。式 (4) から、 $\Delta\phi$ を式 (9) で近似する。

$$\Delta\phi = \xi\Delta F_c + \alpha_s\xi\Delta F_s + \alpha_{as}\Delta F_s \quad (9)$$

$\xi = F'_{res}(F_c(e_A) + \alpha_s F_s(e_A) - C_c)$ としている。なお、 f_d は STSP の制約に関係なく、 F_{res} に含まれないため、負荷の対称性に関する本議論においては考慮に入れない。ここで、 e_A に、仮想辺 $e^v \notin \mathcal{E}$ を足したことによる ϕ の変化量を $\Delta\phi^{+v}$ 、対応する実辺上の最短経路である $S(e^v)$ ($\subset \mathcal{E}$) を足したことによる変化量を $\Delta\phi^{+s}$ とする。 $\Delta\phi^{+v}, \Delta\phi^{+s}$ は式 (5) による f_s の定義と式 (9) を用いれば、それぞれ式 (10)、式 (11) と書ける。

$$\Delta\phi^{+v} = \alpha_s \xi f_s(e^v) + \alpha_{as} f_s(e^v) \quad (10)$$

$$\Delta\phi^{+s} = \xi f_s(e^v) \quad (11)$$

以上の結果より、 $\alpha_s \approx 0, \alpha_{as} \approx 1$ と設定する EV の定式化では、 $F_c(e_A) \geq \sigma^2$ において $\xi = 1$ となるため、 $\Delta\phi^{+v}$ と $\Delta\phi^{+s}$ はほぼ等しいが、 $F_c(e_A) < \sigma^2$ においては $F_c(e_A)$ の値が小さくなるほど差異が大きくなることがわかる。つまり、このような定式化を用いた探索においては、仮想辺と対応する実辺の入れ替えに関して目的関数の値が大きく変化するため、GS における状態変化の過程は、2 辺以上の同時操作により実辺のみを用いた探索と、仮想辺を経由した場合で大きく異なることとなり、最適化効率の低下が予測される。

この問題に対し、本稿では仮想辺を実辺と対称に扱う、 $\alpha_s = 1, \alpha_{as} \approx +0$ の設定を用いることで最適化効率を改善する目的関数の定式化を提案する。この設定においては、式 (10)、式 (11) より、 ξ の値に関係なく $\Delta\phi^{+v}$ と $\Delta\phi^{+s}$ がほぼ等しくなることがわかる。この場合、仮想辺と対応する実辺の入れ替えに関して目的関数の値がほぼ変化しないため、仮想辺を経由した場合における最適化効率の低下が起きないと考える。

完全に対称性を満たすとき、置換に関して目的関数の値が変化しなくなり、仮想辺の実辺への置換が起きづらくなる。つまり、定義辺上のスポットの探索が行われにくくなるため、僅かに対称性を崩すために、 α_{as} に微小な値を設定する。この値の設定方法に、ユーザ毎に値が異なる f_d は関与しないため、問題のグラフに対して事前実験を行うことにより、全てのユーザに対して利用可能な共通の値を定めることができる。

対称性が満たされることが最適化の結果に及ぼす具体的な影響を考察する。辺ベクトルの状態は、Random Walk [16] と同様の性質により温度が高いときには経

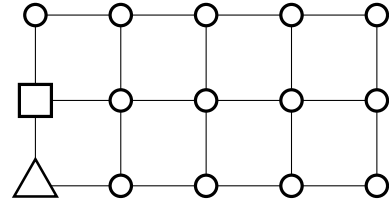


図 5: 5 × 3 の節点からなる升目構造グラフの人工データセット。四角は始点、三角は終点をそれぞれ表す。

路長が長い状態を取り、そこから温度が下がるにつれ、仮想辺を含む状態と、対応する実辺に置換された状態を交互に取りながら短い経路に遷移していくが、上述の通り、EV の定式化に対応する $\alpha_s \approx 0, \alpha_{as} \approx 1$ とするとき、仮想辺を含む状態への変化が起りづらくなる場合があるため、経路は長い状態のまま硬直する可能性がある。その結果、最終的な経路長が長くなりやすいと考えられる。この問題が実際に起こり、対称性の高い定式化において解決できることは 4 節の評価実験により示す。

対称性の議論からは外れるが、 α_g の設定次第で最適化効率の向上を望めるため、このパラメータの設定方法についても述べる。 α_g を大きくするほど、 $F_{res}(x)$ は単調な直線に近くなり STSP の制約が機能しづらくなるため、 α_g には小さな値を用いることが望ましい。また、 $\alpha_g \approx 0$ とすることで、対称性が満たされる条件、つまり、 $\alpha_s = 1$ ならば $F_c(e_A) + F_s(e_A) < C_c$ である状態 e_A について $\phi(e_A)$ は $F_c(e_A), F_s(e_A)$ によらず一定となり、そのような経路間を滑らかに探索できるため、計算精度の向上が期待できる。また、 $\alpha_g > 0$ の設定では、経路長に対して目的関数は常に正の傾きを持つため、より短い経路長を持つ状態ほど目的関数の値が小さくなる。このことから、 $\alpha_g \approx 0 \wedge \alpha_g > 0$ と設定することで、無駄な経路を削減可能な、かつ精度の高い探索となることが期待できる。

4 評価実験

4.1 人工データセットによる実験

3 節の考察で予想したとおり、既存手法では生成される辺ベクトルを e_A^* としたとき、その経路の長さ $F_c(e_A^*)$ が C_c よりも長くなりやすいことと、提案する定式化を用いた場合にこれを解決できることを検証するため、図 5 に示すような升目構造を持つ人工データセットを用い実験を行った。図において、四角、三角はそれぞれ始点、終点を表す。実験では 20×20 のサイズを用いたが、スペースの関係上図では 5×3 としている。また、 $\forall e \in \mathcal{E}, f_c(e) = 1 \wedge f_d(e) = 0$ としている。

表 1: 様々な (α_g, α_s) において提案する定式化を用い生成された辺ベクトル e_A^* の経路長 $F_c(e_A^*)$ と計算時間の平均値.

(α_g, α_s)	平均経路長	平均計算時間 [s]
$(-1, 0)$	106.9	10.2
$(-1, 1)$	100.0	5.42
$(0.01, 0)$	106.7	9.96
$(0.01, 1)$	1.000	2.80

$(\alpha_g, \alpha_s) \in \{(-1, 0), (-1, 1), (0.01, 0), (0.01, 1)\}$ における, $C_c = 100, \sigma^2 = 50$ とした場合の, 128 回の試行における $F_c(e_A^*)$ の平均値と, 計算時間⁴ の平均値を表 1 に示す. $t \in [0, 1]$ を, $t = 0$ が計算開始, $t = 1$ が計算終了を意味する計算の進行度としたとき, 温度変化は $T(t) = 10^{-3t}$ により与え, t が 5×10^{-5} 進む毎に標本抽出を行い, $\alpha_{as} = 0.005$, Boltzmann 定数には文献 [9] の指針に従い $k = 1.2$ を選び, 近傍探索の範囲には $K = 16$ を用いた. 擬似乱数の生成には Mersenne Twister [17] を用いた.

表 1 より, 文献 [9] の定式化に対応する $(\alpha_g, \alpha_s) = (-1, 0)$ では 3 節での議論の通り, C_c の値より長い経路が生成される傾向にあることがわかる. $(\alpha_g, \alpha_s) = (0.01, 0)$ の場合でもこの傾向は変わらないことから, 目的関数を経路長に対して単調増加とただけではこの傾向は改善されないことがわかる. 一方, 3 節での議論に従い $\alpha_s = 1$ とした $(\alpha_g, \alpha_s) = (-1, 1)$ のときはほぼ期待される経路長が得られていることがわかる. また, $(\alpha_g, \alpha_s) = (0.01, 1)$ の実験結果では始点・終点を直接結ぶ長さ 1 の経路が生成される傾向にあることがわかる. 本実験では $\forall e \in \mathcal{E}, f_d(e) = 0$ としており, 経路長を増加させるメリットがないため妥当な結果といえる. これらの結果から, $\alpha_s = 1$ とする対称性の高い設定において, 目的関数の形状から期待される解を得られることがわかる. また, 生成される経路長が短いほど計算時間が減少しており, 無駄な経路を削減する α_g の設定と合わせれば, 提案する定式化は高速化にも貢献することがわかる.

4.2 実データセットによる実験

文献 [9] における評価実験と同様に, 文献 [1] で用いられているデータセットを用いて実験を行った. 評価指標には文献 [9] と同様, Recall を用いる. 既存手法および

⁴C++ 14 で実装し, Ubuntu 16.04.5 上において, g++ version 5.4.0 により最適化オプション `-O3` を付加してコンパイルし, 同 OS 上で倍精度浮動小数点数を用いて計算した結果に基づく. 個々の実験はすべて単一スレッドにより実行され, CPU には Ryzen Thread Ripper 1920X (<https://www.amd.com>) を用い, 主記憶は DDR4-2400 をバス幅 256 bit で接続し動作させた.

表 2: PersQ[1] と提案する定式化 ($EV(\alpha_g, \alpha_s)$) を用いた結果の Recall. $EV(-1, 0)$ が文献 [9] の定式化におけるものに対応する. C: Cali. Adv, H: Hollywood, D: Disneyland M: Magic King., E: Epcot.

	C	H	D	M	E
PersQ	0.483	0.482	0.332	0.440	0.472
$EV(-1, 0)$	0.291	0.421	0.231	0.388	0.401
$EV(0.01, 0)$	0.406	0.461	0.405	0.450	0.530
$EV(0.01, 1)$	0.441	0.472	0.423	0.475	0.555

提案手法により生成された経路の Recall を表 2 に示す. PersQ は文献 [1] で提案された手法, $EV(\alpha_g, \alpha_s)$ が提案する定式化である. $EV(-1, 0)$ が文献 [9] における定式化に対応する. $EV(-1, 0)$ においては, $\alpha_{as} = 1.005$ を, それ以外の設定においては $\alpha_{as} = 0.005$ を用いた. 温度変化は $T(t) = 10^{-4t}$ に従うとし, t が 5×10^{-4} 進む毎に標本抽出を行った. 擬似乱数の生成には 4.1 節における実験と同様に Mersenne Twister を用いた. その他の設定値については文献 [9] と同様である.

表 2 から, 提案する定式化を用いることで性能が向上することを確認できる. $\alpha_s = 1$ とする対称性の高い設定において精度が向上していることがわかるが, 特にこのデータセットでは, α_g の調整が性能改善に貢献していることがわかる.

5 おわりに

本稿では, 辺ベクトルによる定式化を用いた経路推薦手法において, 実辺と仮想辺の対称性が最適化効率に与える影響を数学的に考察し, 高い対称性を持つ設定が可能な目的関数の定式化を提案した. 評価実験により, 考察から予測される最適化効率の低下が再現されることを示し, 提案する定式化において対称性の高い設定とした場合にこの問題を解決できることを示した. また, 提案する定式化を用いることで, 実データセットを用いた評価実験では既存の定式化を用いたものを上回る性能を達成可能なことを示した. 今後, 実際の道路ネットワークにおける検証を行うことで, より現実に即した解法とすることが期待できる. また, 本手法は確率的生成モデルとして定式化されているため, 同モデルに対する学習則を利用しスポットの価値を推定できる可能性がある. 確率的生成モデルにおいて推論と学習に利用する手法は本質的に同じであり, 価値の推定と推論を同一のアルゴリズムにより実装できれば理論的解析も行いやすくなるため, 検討する予定である.

謝辞

本研究の一部は JSPS 科研費 16K12535, 15H02780 の助成および首都大学東京傾斜的研究費(全学分) 学長裁量枠戦略的研究プロジェクト戦略的研究支援枠「ソーシャルビッグデータの分析・応用のための学術基盤の研究」によるものです。

参考文献

- [1] Lim, K., Chan, J., Karunasekera, S.: Personalized Itinerary Recommendation with Queuing Time Awareness, *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 325-334 (2017)
- [2] 松田善臣, 名嘉村盛和, 姜 東植, 宮城 隼夫: 最適観光経路問題とその解法, 電気学会論文誌 C, Vol. 124, No. 7, pp. 1507-1514 (2004)
- [3] 倉田陽平, 有馬貴之: 対話的旅行計画作成支援システムの実装と評価, 第 25 回日本観光研究学会全国大会, pp. 173-176 (2010)
- [4] Gutin, G., Punnen, A.: The Traveling Salesman Problem and Its Variations, *Springer Science & Business Media* (2006)
- [5] Yannakakis, M.: Expressing Combinatorial Optimization Problems by Linear Programs, *Journal of Computer and System Sciences*, Vol. 43, No. 3, pp. 441-446 (1991)
- [6] Helsgaun, K.: General k-Opt Submoves for the LinKernighan TSP Heuristic, *Mathematical Programming Computation*, Vol. 1, No. 2-3, pp. 119-163 (2009)
- [7] Kirkpatrick, S., Gelatt Jr., D. C., Vecchi, M.: Optimization by Simulated Annealing, *Science*, Vol. 220, No. 598, pp. 671-680 (1983)
- [8] Laporte, G., Martello, S.: The Selective Traveling Salesman Problem, *Discrete Applied Mathematics*, Vol. 26, No. 2-3, pp. 193-207 (1990)
- [9] 柴田祐樹, 高間康史: 辺ベクトルを用いた観光経路推薦問題の定式化と焼きなまし法による解法の提案, 知能と情報, Vol. 31, No. 1 (2019)
- [10] Browne, C., Powley, E., Whitehouse, D., Lucas, S., Cowling, I. P., Rohlfshagen, P., Tavener, S., Perez, D., Samothrakis, S., Colton, S.: A Survey of Monte Carlo Tree Search Methods, *IEEE Transactions on Computational Intelligence and AI in Games*, Vol. 4, No. 1, pp. 1-43 (2012)
- [11] Vose, D. M.: Modeling Simple Genetic Algorithms, *Journal of Evolutionary Computation*, Vol. 3, No. 4, pp. 453-472 (1995)
- [12] Metropolis, N., Rosenbluth, A., Rosenbluth, M., Teller, A.: Equation of State Calculations by Fast Computing Machines, *The Journal of Chemical Physics*, Vol. 21, No. 6, pp. 1087-1092 (1953)
- [13] Geman, S., Geman, D.: Stochastic Relaxation, Gibbs distributions, and the Bayesian Restoration of Images, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. PAMI-6, No. 6, pp. 721-741 (1984)
- [14] Shibata, H., Takama, Y.: Index-based Notation for Random Variable and Probability Space, *Journal of Advanced Computational Intelligence and Intelligent Informatics*, Vol. 23, No. 4, 4 pages (2019)
- [15] Chan, M. T.: More Algorithms for All-Pairs Shortest Paths in Weighted Graphs, *Proceedings of the 39th Annual ACM Symposium on Theory of Computing*, pp. 590-598 (2007)
- [16] C. M. Bishop: Pattern Recognition and Machine Learning, Chapter 11, *Springer*, 8th printing (2006)
- [17] Matsumoto, M., Nishimura, T.: Mersenne Twister: A 623-Dimensionally Equidistributed Uniform Pseudorandom Number Generator, *ACM Transactions on Modeling and Computer Simulation*, Vol. 8, No. 1, pp. 3-31 (1998)

空間座標情報を用いた単語の分散表現獲得

Learning Word Embeddings Using Spatial Information

城光英彰 岡隆之介 内出隼人 伍井啓恭 大塚貴弘

Hideaki Joko, Ryunosuke Oka, Hayato Uchihe, Hiroyasu Itsui and Takahiro Ohtsuka

三菱電機株式会社 情報技術総合研究所

Information Technology R&D Center, Mitsubishi Electric Corporation

Abstract: This study proposes a word embedding learning method for performing both facility and region name search. In the previous method, the word embeddings were learned based on the context in which the words emerged. However, because the previous method missed the spatial information, the word embeddings that share similar contexts (e.g., universities) become similar even though the objects of the words were different (e.g., “東京大学” (The University of Tokyo) and “京都大学” (Kyoto University)). In the proposed method, the word embeddings are learned based on the spatial information data, which comprise both the object names and coordinates. Therefore, even if the words share similar contexts, the proposed method can learn different word embeddings for different words if those objects are different. The proposed method is evaluated using the synonym search task. As a result, it was observed that the MRR for the evaluation data improved using the proposed method in comparison with the previous method. Furthermore, the proposed method improved the Mean Reciprocal Rank by 177% maximally, i.e., from 0.151 to 0.418, in comparison with the previous method.

1. はじめに

施設名や地名の検索などにおいて、単語の文字面だけでなく「意味」を考慮した検索が求められている。「意味」を考慮した検索には、分散表現の有効性が知られている[1][2]。一般に、この分散表現は、「類似した意味の単語は、その文脈の単語（文脈単語）の分布も類似する」という分布仮説[3]に基づき獲得される。しかし、既存手法では、空間の情報を考慮しないため、異なる対象を表す分散表現が類似する問題がある[4]。例えば、「東京大学」と「京都大学」は対象としては異なるが、どちらも「大学」という共通の特徴を持つために、その単語の周辺文脈には似たような単語が出現するため、二つの単語は類似した分散表現となる。その結果、例えば、東京大学への行き方を検索しようと「東大」（東京大学の略称）をクエリ単語として入力したときに、京都大学への行き方が検索されてしまう問題が生じる。この問題を解決するために、本研究では、空間座標情報を用いた分散表現獲得手法を提案する。空間座標の情報は個別の対象ごとに異なると想定されるため、この情報を活用すれば異なる対象については、相違した分散表現が得られると考えられる。提案手

法は、自然言語で記述された名称とその空間座標情報からなる大規模データ（空間座標 DB）から分散表現を学習する。本研究では、空間座標 DB として、ウィキペディア日本語版から取得した地名・施設名とその空間座標からなるデータを用いた。

評価は同義語検索タスクにより行った。具体的には、クエリ単語（例えば「東大」と各検索対象単語（例えば、「東京大学」や「京都大学」）について、分散表現のコサイン類似度を算出し、これに基づき検索対象単語をランキングし、そのランキング精度を平均逆順位（MRR）により評価した。

2. 既存手法（Skip-gram モデル）

分布仮説のもとで分散表現を獲得する研究には、Mikolov et al.[1] の Skip-gram モデルがある。本節では、Skip-gram モデルについて、先行研究[2]を参考に概説する。Skip-gram モデルは分布仮説に基づき単語の意味を表す数値ベクトル（分散表現）を獲得する手法である。ある単語 w_t が文章内の位置 t に存在した場合の、その文脈単語 w_{t+j} ($j \neq 0$) の発生確率 $p(w_{t+j}|w_t)$ を以下の式で与える。

$$p(w_{t+j}|w_t) = \frac{e^{v'(w_{t+j})^T v(w_t)}}{\sum_w e^{v'(w)^T v(w_t)}}$$

ここで、 $v(w_t)$ は単語 w_t の分散表現、 $v'(w_{t+j})$ は文脈単語の出現確率計算用のベクトルである。学習はテキストデータ内の全単語に対し行われる。そのため、尤度目的関数 l_{SG} は、以下の式で定義される。

$$l_{SG} = \sum_{t=1}^T \sum_{-c \leq j \leq c, j \neq 0} \log p(w_{t+j}|w_t)$$

ここで T はテキストデータのサイズ、 c は文脈窓のサイズである。学習時は、尤度目的関数 l を最大化する分散表現 $v(w)$ を求める。図1に分布仮説を用いた分散表現獲得手法のイメージを示す。中心単語「東大」の分散表現は、中心単語の文脈の単語分布により計算される。しかし、本手法では、原理的に、文脈単語の分布が類似する単語同士の分散表現は類似する。そのため、例えば「東京大学、京都大学」などの対象としては異なるが、文脈単語の分布が類似するものは、類似した分散表現となり問題である。

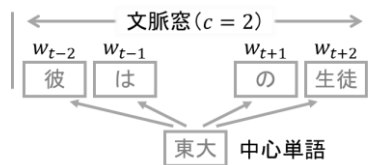


図1: 分布仮説を用いた分散表現獲得手法のイメージ図

3. 提案手法

3.1 Skip-space モデル

既存手法の問題を解決するために、本研究では、空間座標情報を用いた分散表現獲得手法 (Skip-space モデル) を提案する。提案手法は、自然言語で記述された名称とその空間座標情報からなる大規模データである空間座標 DB から分散表現を学習する。空間座標 DB のイメージを図2に示す。単語「東大」を含む対象の近傍に出現する対象の名称に含まれている単語 (近傍単語) と、単語「東京大学」の近傍単語はともに「文京」「本郷」などを含んでおり、類似している。一方で、「東京大学」と「京都大学」などの対象として異なるものは、空間座標 DB での近傍単語の分布が相違するため、相違した分散表現になる。その結果、第1節に例示した、東京大学への行き方を検索しようと「東大」をクエリ単語として入力した場合に、京都大学への行き方が検索されてしまう問題を解決できる。

以下では、提案手法の近傍単語の発生確率および尤度目的関数について述べる。対象 $x_c \in X$ に対し、 x_c からユークリッド距離が近い順に対象を k 個取得することを考える。取得する対象を近傍対象 $x_s \in S_{x_c}$ と呼ぶ。対象 x_c および近傍対象 x_s の名称の先頭から t 番目および u 番目の単語を $w_{x_c,t}$ および $w_{x_s,u}$ とする (図3を参照)。このとき、提案手法 Skip-space モデルでは、近傍単語の発生確率および尤度目的関数 l_{SS} を次のように変更する。

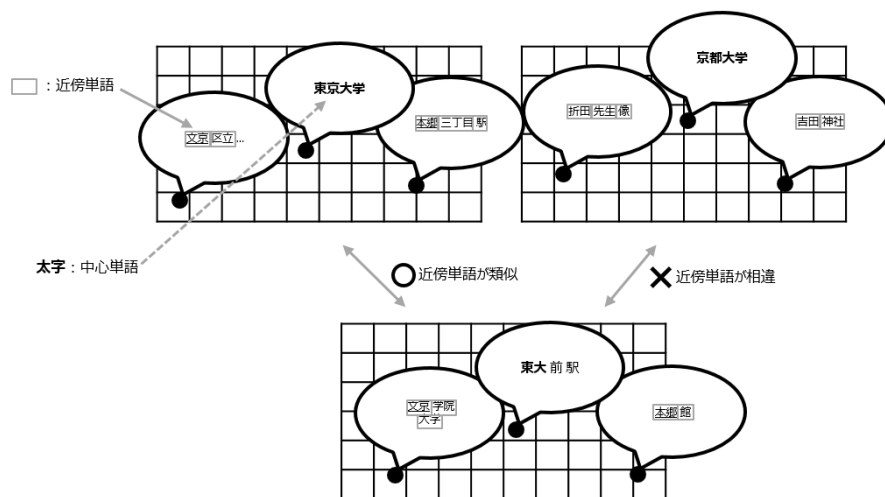


図2: 空間座標 DB のイメージ図。単語「東大」を含む対象の近傍に出現する対象の単語 (近傍単語) と、単語「東京大学」の近傍単語は類似する。

$$p(w_{x_s,u}|w_{x_c,t}) = \frac{e^{v'(w_{x_s,u})^T v(w_{x_c,t})}}{\sum_w e^{v'(w)^T v(w_{x_c,t})}}$$

$$l_{SS} = \sum_{x_c \in X} \sum_{x_s \in S_{x_c}} \sum_{1 \leq t \leq N_{x_c}} \sum_{1 \leq u \leq N_{x_s}} \frac{1}{N_{x_s}} \log p(w_{x_s,u}|w_{x_c,t})$$

ここで、 N_{x_c} および N_{x_s} は対象 x_c および近傍対象 x_s の名称を構成する単語の数である。また、 $\frac{1}{N_{x_s}}$ は各単語の近傍単語の出現数の偏り¹をなくすための正規化用の係数である。

3.2 Multi-task Skip-space モデル

Multi-task Skip-space モデルでは、空間座標データとテキストデータからマルチタスク学習[5]により分散表現を学習する。尤度目的関数 l_{SS+SG} が次のように変更される。

$$l_{SS+SG} = l_{SS} + l_{SG}$$

ここで、 l_{SS} および l_{SG} は、Skip-space および Skip-gram の尤度目的関数である。なお、文脈単語のベクトル v' は Skip-space による空間座標データ学習時と、Skip-gram によるテキストデータ学習時で、個別に与えられる。すなわち、各単語 w は二つの文脈単語のベクトル $v_{sg}'(w)$ 、 $v_{ss}'(w)$ を保持し、文脈単語および近傍単語の発生確率は、Skip-gram によるテキストデータ学習時においては、

$$p(w_{t+j}|w_t) = \frac{e^{v_{sg}'(w_{t+j})^T v(w_t)}}{\sum_w e^{v_{sg}'(w)^T v(w_t)}}$$

Skip-space による空間座標データ学習時においては、

$$p(w_{x_s,u}|w_{x_c,t}) = \frac{e^{v_{ss}'(w_{x_s,u})^T v(w_{x_c,t})}}{\sum_w e^{v_{ss}'(w)^T v(w_{x_c,t})}}$$

のように、変更される。

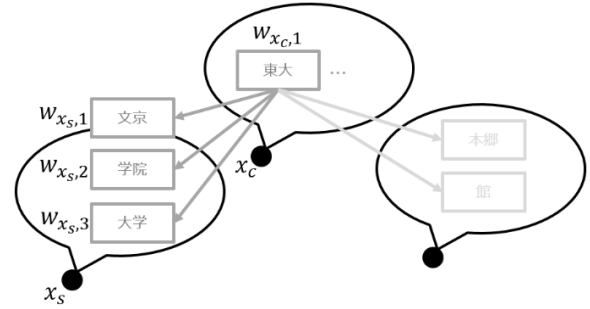


図 3: 空間座標 DB のイメージ図。単語「東大」を含む対象の近傍に出現する対象の単語（近傍単語）と、単語「東京大学」の近傍単語は類似する。

4. 関連研究

単語の分散表現を獲得する研究には様々なものがある。城光ら[2]は Skip-gram の拡張として、文脈限定 Skip-gram を提案した。文脈限定 Skip-gram は、文脈単語を特定の品詞を持つものや特定の位置に存在するものに限定し、各限定条件に対して分散表現を学習することで、単純な Skip-gram モデルに比べて高い精度で同義語獲得ができる。この手法は、品詞が同義語獲得に与える影響の分析が可能のため、解釈可能性が高く、さらに、限定条件の変更も容易であり、拡張可能性も高いという利点がある。しかし、この手法は単純な Skip-gram と同様に、この手法は空間の情報を考慮しないため、異なる対象について、相違した分散表現を得ることは難しい。

Skip-gram とは異なるアプローチで単語の分散表現を獲得する研究もある。Peters ら[6]は、Deep Bi-directional Language Model の隠れ層を加重平均することで、文脈を考慮した分散表現を獲得できるモデルである Embeddings from Language Models (ELMo) を提案した。Devlin [7] らは、Bi-directional Transformer [8] を用いて分散表現を獲得する Deep Bidirectional Transformers (BERT) を提案した。これらのモデルは、質問応答を含む様々な自然言語処理タスクにおいて state-of-the-art な性能を達成しており、用いるモデルも Skip-gram モデルより複雑なことから、異なる対象について、相違した分散表現を得ることができる可能性はある。しかし、これらのモデルで獲得した単語の分散表現を利用するには、都度その周辺文脈から分散表現を再計算する必要がある。再計算に必要な計算機の性能を考慮すると、機器への組み込みなどによる実用化は難しい。

これらの研究と異なり、提案手法は空間座標の情

¹ 対象ごとにその名称を構成する単語の数は異なるため、近傍単語の出現数の偏りが生じる。

報を活用できるため、異なる対象については相違した分散表現が得られるという優位点がある。また、単語の分散表現をその周辺文脈から再計算する必要がないため低算量であるという優位点もある。

5. 実験

実験では、まず、提案手法および既存手法により分散表現の学習を行う。用いた分散表現学習用データおよび手法については4.1節に示す。

評価は同義語検索タスクにより行う。具体的には、入力クエリと各検索対象の地名・施設名のそれぞれの分散表現からコサイン類似度を算出し、コサイン類似度の高い順に検索対象単語をランキングし、そのランキング精度を MRR により評価する。この評価実験のイメージを図4に示す。なお、地名・施設名が複数の形態素から構成される場合（例えば「都留文化大」など）は、各形態素の分散表現の平均値を名称のベクトルとする[9][6]。形態素解析は MeCab²を用いて行い、形態素辞書には IPAdic[10]を用いた。

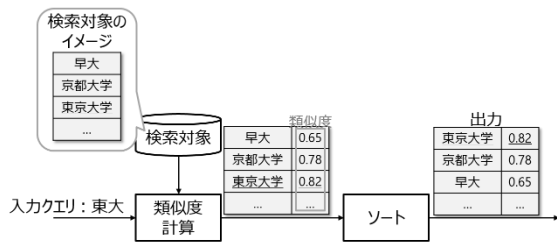


図4: 評価実験のイメージ図。入力クエリと各検索対象の地名・施設名について、分散表現のコサイン類似度を算出し、これに基づき検索対象単語をランキングし、そのランキング精度を MRR により評価する

4.1 分散表現学習用データ

用いた分散表現学習用データについて、表1に示す。「地名・施設名」を緯度・経度情報が付与されている記事の名称と定義する。提案手法 Skip-space モデルにおける学習用データは、ウィキペディア日本語版から取得した地名・施設名とその空間座標データからなる「地名・施設名 空間座標データ」である。既存手法における学習用データは、DBpedia Japanese から取得した地名・施設名と、その “dbpedia-

owl:abstract” プロパティから取得した概要テキストからなる「地名・施設名 概要テキストデータ」である。これらのデータに対し、第5節に示したのと同様の方法で形態素解析した。なお、地名・施設名概要テキストデータの取得は、このデータの述ベ形態素数が、地名・施設名 空間座標データの延ベ形態素数を越えた時点で終了し、提案手法と既存手法で用いたデータの延ベ形態素数が揃うようにした。

表1: 使用した分散表現学習用データ

分散表現学習用データの名称	種別	入手元	施設数	延ベ形態素数
地名・施設名 空間座標データ	空間座標データ	ウィキペディア 日本語版	122K	456K
地名・施設名 概要テキストデータ	テキストデータ	DBpedia Japanese	—	456K

4.2 分散表現学習手法

まず、各分散表現学習手法と、各々の手法に対し使用した学習データを表2に示す。SSとは提案手法である Skip-space モデルを「地名・施設名 空間座標データ」に対し適用したものである。SS+SGとは、提案手法である Multi-task Skip-space モデルを「地名・施設名 空間座標データ」と「地名・施設名 概要テキストデータ」に対し適用したものである。SGとは、既存手法である Skip-gram モデルを「地名・施設名 概要テキストデータ」に対し適用したものである。なお、学習データの詳細については4.1節を参照されたい。

次に、実験で使用したハイパーパラメータの設定を述べる。分散表現の次元数は、既存手法、提案手法ともに200とした。また、既存手法の文脈窓のサイズ $c = 5$ 、提案手法において取得する近傍の対象数 $k = 5$ とした。なお、学習の際には高速化のために Hierarchical softmax [11] による近似を行った。

² <http://taku910.github.io/mecab/>

表 2: 各分散表現学習手法と、各々の手法に対し使用した学習データ

分散表現 学習手法 の名称	使用した分散表現獲得モ デル	使用した分散表現学習用デ ータ	
		地名・施設 名 空間座 標データ	地名・施設名 概要テキス トデータ
提案手法: SS	Skip-space モデル	○使用	×未使用
提案手法: SS+SG	Multi-task Skip-space モデ ル	○使用	○使用
既存手法: SG	Skip-gram モデル	×未使用	○使用

4.3 評価用データ

用いた評価用データは、ウィキペディア日本語版から取得した「大学の略称データ」、DBpedia Japaneseから取得した「DBpedia 地名・施設名 略称データ」および「DBpedia 地名・施設名 別名データ」の三種類である。SS, SS+SG, SG の少なくとも一つの手法で分散表現が獲得されていない単語を含む地名・施設名をもつ評価データは、評価用データから除外した。評価用データの詳細を表 3 に示す。また、評価用データの例を表 4 に示す。各地名・施設名は意味 ID を付与されており、異なる地名・施設名が同一の意味 ID を持つ場合、その地名・施設名は同一の意味を持つ語（同義語）であることを示す。例えば、表 4 では、「東京大学」と「東大」がどちらも“1”という同一の意味 ID を持つ。そのため、「東京大学」と「東大」は同義語であることがわかる。

評価の際には、評価用データの中から一つを入力クエリとして、残りを検索対象として使用する。そのため、「大学の略称データ」においては、各クエリについて検索対象は 259 件、「DBpedia 地名・施設名 略称データ」では 62 件、「DBpedia 地名・施設名 別名データ」では、186 件となる。

表 3: 評価用データの詳細

名称	入手元	データ数
大学の略称データ	ウィキペディア 日本語版	260
DBpedia 地名・施設名 略称データ	DBpedia Japanese	63
DBpedia 地名・施設名 別名データ	DBpedia Japanese	187

表 4: 評価用データの例

地名・施設名	意味 ID
東京大学	1
東大	1
都留文化大	2
都留文大	2
都留文	2
...	...

6. 実験結果

既存手法と提案手法の評価結果を図 5 に示す。提案手法である SS と SS+SG では、既存手法 SG と比較し、全ての評価データにおいて MRR が向上していることがわかる。MRR の向上幅は「DBpedia 地名・施設名 別名データ」においても最も高く、既存手法 SG の 0.151 に対し、提案手法 SS+SG では 0.418 と、177%向上している。なお、各実験結果に対し、ウィルコクソンの符号付順位検定を適用したところ、「大学の略称データ」

「DBpedia 地名・施設名 別名データ」において、既存手法 SG と提案手法 SS, SS+SG の間において有意水準 1% で逆順位の母平均に有意差が認められた。これらの結果から、提案手法では既存手法と比べて、地名・施設名の「意味」の獲得精度が高まっていることがわかる。

入力クエリを「都留文大」とした場合の例を表 5 に示す。「都留文大」の同義語である「都留文」（共に正式名称は「都留文科大学」）が、既存手法では、全 259 件中第 98 位に現れるのに対し、提案手法では第 1 位に現れており、既存手法と比べ、提案手法では「都留文大」の意味の獲得精度が高いことがわかる。また、既存手法では、第 1 位が「大 大 大」であり（正式名称は「大阪大谷大学」）、「大」を多く含む地名・施設名であること、検索結果の上位 5 件全てが「大」を含むことがわかる。これは、既存手法では「都留」の意味を学習できておらず、そのため、「大」の意味に強く引っ張られた検索結果になったためと考えられる。これに対し、提案手法では、同義語である「都留文」が第 1 位に現れているだけでなく、「都留文科大学」の所在である山梨を所在地とする大学である「山梨学院大」が第 2 位に来ている。これは、提案手法では空間情報を考慮した分散表現が獲得できているためと考えられる。

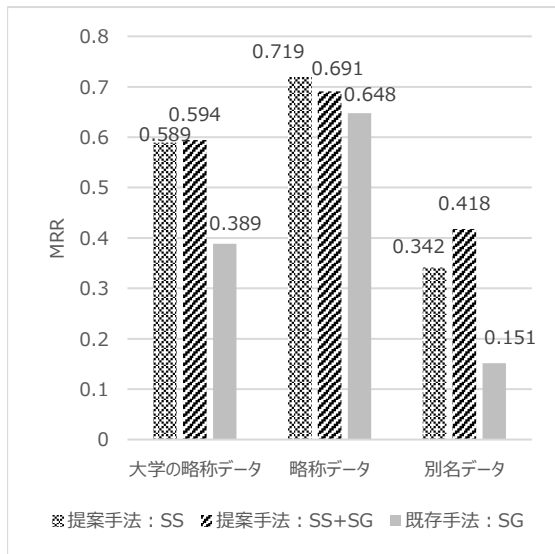


図 5: 既存手法と提案手法の評価結果. 提案手法である SS と SS+SG では, 既存手法 SG と比較し, 全ての評価データにおいて MRR が向上している

表 5: 入力クエリ「都留文大」に対する地名・施設名の検索結果. 括弧内は正式名称. 半角スペースは単語の区切りを表す. 「都留文大」の同義語である「都留文」(共に正式名称は「都留文科大学」)が, 既存手法では, 全 259 件中第 98 位に現れるのに対し, 提案手法では第 1 位に現れており, 既存手法と比べ, 提案手法では「都留文大」の意味の獲得精度が高いことがわかる.

順位	提案手法 : SS+SG	既存手法 : SG
1	都留文 (都留文科大学)	大 大 大 (大阪大谷大学)
2	山梨学大 (山梨学院大学)	学芸大 (東京学芸大学)
3	旭教大 (北海道教育大学旭川校)	弘大 (弘前大学)
4	旭教 (北海道教育大学旭川校)	大歯大 (大阪歯科大学)
5	東工大 (東京工科大学)	大薬大 (大阪薬科大学)

7. まとめ

本研究では, 施設名や地名の検索への応用を目的に, 空間座標情報を用いた単語の分散表現獲得手法を提案した. 既存の分散表現獲得手法では, 単語の出現文脈に基づき分散表現を獲得する. しかし, 既

存手法では, 空間の情報を考慮しないため, 異なる対象を表す分散表現が類似する問題がある. これに対して提案手法では, 自然言語で記述された名称とその空間座標情報からなる大規模データから分散表現を学習する. これにより, 異なる対象については, 相違した分散表現が得られる. 評価は同義語検索タスクにより行った. その結果, 提案手法は, 既存手法と比較し, 全ての評価データにおいて MRR が向上することがわかった. また, MRR の向上幅は「DBpedia 地名・施設名 別名データ」においても最も高く, 既存手法の 0.151 に対し, 提案手法では 0.418 と, 177% 向上した.

今後の課題を述べる. 本研究では, 空間座標情報としてウィキペディア日本語版から取得した地名・施設名とその緯度・経度情報を用いた. 今後は, 緯度・経度情報だけでなく, 三次元の座標情報や, 時間を考慮した座標情報を利用することで, 工場における設備名や設計図面における部品名の検索などに応用したい.

参考文献

- [1] Mikolov, T., Chen, K., Corrado, G., and Dean, J.: Efficient Estimation of Word Representations in Vector Space, In ICLR (2013)
- [2] 城光英彰, 松田源立, 山口和紀.: 文脈限定 Skip-gram による同義語獲得, 自然言語処理, Vol. 24, No. 2, pp. 187-204 (2017)
- [3] Zellig, H.: Distributional structure, Word, Vol. 10, No. 23, pp. 146-162 (1954)
- [4] 城光英彰, 松田源立, 山口和紀.: 同義語判定問題を用いた語義ベクトルの評価の検討, 第 10 回 インタラクティブ情報アクセスと可視化マイニング研究会 (2015)
- [5] Caruana, R.: Multitask Learning, Machine Learning, vol.28 (1997)
- [6] M. Peters, M. Neumann, M. Iyyer, M. Gardner, C. Clark, K. Lee, and L. Zettlemoyer, "Deep contextualized word representations," NAACL. 2018.
- [7] J. Devlin, M. Chang, K. Lee, and K. Toutanova, "Bert: Pre-training of deep bidirectional transformers for language understanding," CoRR, abs/1810.04805, 2018.
- [8] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. Gomez, L. Kaiser, and I. Polosukhin, "Attention is all you need," NIPS. 2017.
- [9] Tian, R., Okazaki, N. and Inui, K.: The mechanism of additive composition, CoRR, Vol. abs/1511.08407 (2015)
- [10] Asahara, M., and Matsumoto, Y.: ipadic version 2.7.0 User's Manual, Computational Linguistics Laboratory, Graduate School of Information Science, Nara Institute of Science and Technology (2003)
- [11] Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., and Dean, J.: Distributed representations of words and phrases and their compositionality, in Advances in Neural Information Processing Systems (2013)

人狼ゲームをプレイする自律エージェントの戦略に関する検討

Study of strategies for autonomous agents that play "werewolf game"

永山 翔滋^{1*} 阿部 穰太郎¹ 大矢 康介¹
Shoji Nagayama¹ Jotaro Abe¹ Kosuke Oya¹
阪本 浩太郎^{1,3} 渋谷 英潔² 森 辰則² 神門 典子^{3,4}
Kotaro Sakamoto^{1,3} Hideyuki Sibuki² Taturori Mori² Noriko Kando^{3,4}

¹ 横浜国立大学 大学院 環境情報学府 ² 同 環境情報研究院
¹Yokohama National University ²ditto
³ 国立情報学研究所 ⁴ 総合研究大学院大学
³National Institute of Informatics ⁴SOKENDAI

Abstract: "Werewolf game" is a popular multiplayer game in which "villagers" are trying to figure out who is a "werewolf" through conversation. Werewolves are usually pretending to be villagers. In this paper, we studied conversation in game logs in order to investigate how werewolves' cooperation contributed to winning percentage of werewolves' team. Since the number of "whispers" that are utterances via werewolves' private chat may be regarded as a measure of the werewolves' cooperation, we investigated the relation between the number of whispers and the winning percentage. As the result, we observed that the winning percentage of werewolves' team increased by 63 percentage point at most when the number of whispers of at least two werewolves was more than 106.

1 はじめに

人狼ゲームとは、対話を通して「村人」の中に潜伏した「人狼」を見つけ出す対戦型の多人数ゲームであり、近年、人狼知能プロジェクト¹など研究テーマとしても注目されている。人狼知能プロジェクトでは、「人間と自然なコミュニケーションを取りながら人狼をプレイするエージェント(AI)の構築」を究極の目標としており、エージェント同士の対戦成績を一つの指標として人工知能の研究開発に取り組んでいる。

人狼ゲームの詳細は3節に後述するが、人狼ゲームは村人陣営と人狼陣営に分かれて戦うチーム戦であり、対話というインタラクティブなコミュニケーションを通してチームの勝利条件を達成しようとする。しかしながら、ゲーム中の対話は基本的に全てのプレイヤーに伝わるオープンなものであり、特定のプレイヤーとだけ対話することはできない。そのため、味方陣営だけで相談するといったチーム全体での連携は困難であり、個人ごとの思惑が錯綜する点が人狼ゲームの特徴である。一方で、「人狼」となったプレイヤーには、「囁き」と呼ばれる人狼同士だけで対話ができる特別な対話チャンネルが用意されており、人狼同士の連携を可能にできる。これは、人狼以外のプレイヤーにはない強力なアドバンテージであり、人狼陣営が勝利するためには重要な要素であると考えられる。人狼ゲームの対話を分析した研究として、文献[1, 2]などが挙げられる

が、人狼同士の連携に焦点をあてたものは存在しない。我々は、文献[3]において、最低一人の人狼の発話数が101を超えることで人狼陣営が勝利する可能性が最大65ポイントに上がるという知見を得ているが、発話数を調整するためにどのように人狼同士が連携するかについては考察できていなかった。以上の背景から、本稿では「人狼」エージェントに焦点をあて、人狼同士の囁きが勝敗にどのような影響を与えるかを分析する。

本稿の構成は以下の通りである。2節で関連研究を述べ、本稿の位置付けを明確にする。3節で人狼ゲームの役職とルールを説明した後、人狼プレイヤーの基本戦術であるステルス人狼の観点から発話と注目が勝敗に影響を与える要因であることを述べる。4節では、分析対象とした人狼BBSについて説明をする。5節で文献[3]の分析結果を紹介した後、6節で、人狼同士の囁きが勝敗とどのような関係があるかを考察する。7節は結論である。

2 関連研究

人狼ゲームの対話を分析した研究として、稲葉らの研究[1, 2]がある。稲葉らの研究では、分析対象として、オンラインで行う掲示板型人狼ゲームである人狼BBS²のログを用いており、発話内容にその種類を分類するタグを付与し集計することで、タグの頻度により、ゲームの勝敗、襲撃や処刑される人の特徴を分析している。稲葉らは発話の内容に着目しており、一方で対象にはほとんど着目していない。対象とは、発話した人

*連絡先：横浜国立大学環境情報学府
〒240-8501 横浜市保土ヶ谷区常盤台 79-7 森 辰則研究室
E-mail: nagayama@forest.eis.ynu.ac.jp

¹<http://aiwolf.org/>

²<http://www.wolfg.x0.com/>

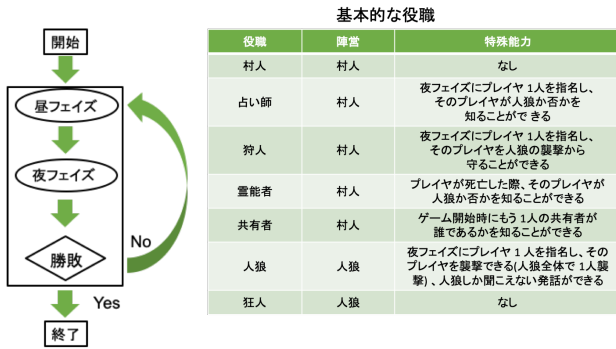


図 1: ゲームの流れと主な役職

が誰であるか、また誰に向けて発話しているかということである。例えば、発話内容タグには占い師のことに語るとして語るタグ (fortune telling) があるが、占い師が本物か偽物かを区別しない。占い師を信用する発話があった場合、本物の占い師であれば村人陣営の勝利に傾いていることを意味するが、偽物であれば人狼陣営の勝利に傾いている。そのため、我々は文献 [3] において、対象を明確にした上で勝敗に基づく分析を行った。その結果、発話数においては人狼の発話数が 101 を超え、また注目数においては占い師の注目数が 168 を下回る場合、人狼の勝率が高いことがわかった。本稿では、文献 [3] での分析をさらに発展させ、人狼の連携度についての分析を行う。

3 人狼ゲーム

3.1 役職とルール

図 1 にゲームの流れと基本的な役職³を示す。役職は、ゲーム開始時に各プレイヤーにランダムに割り当てられ、他のプレイヤーからは分からないようになっている。人狼ゲームはチーム戦で村人陣営と人狼陣営に分かれて戦う。人狼以外のプレイヤー数が人狼プレイヤーの数以下になると人狼陣営の勝利となり、それまでに人狼プレイヤーが全て処刑されると村人陣営の勝利となる。幾つかの役職には図に示した特殊能力があり、これを手がかりとして人狼プレイヤーを探す (村人陣営)、または見つからないようにする (人狼陣営)、というのがゲームの基本である。しかしながら、殆どのプレイヤーは特殊能力を持たない村人であるため、どちらの陣営にとっても彼らを如何にして説得するかが勝敗の鍵となる。

ゲームは「昼」と「夜」のフェイズに分かれて進行する。昼フェイズでは、誰が人狼であるかを議論し、人狼と思われるプレイヤー1人を多数決で処刑する (ゲームから脱落させる)。夜フェイズでは、自分の陣営を優位にするために特殊能力を行使する。例えば、占い師であれば任意のプレイヤー1人の役職が人狼であるか否かを知ることができ、人狼プレイヤーであれば任意のプレイヤー1人を襲撃する (ゲームから脱落させる) ことができる。以上で「1日」が終わり、勝敗がつ

³人狼ゲームには多くの特殊ルールがあり、ここで挙げた役職以外にも存在する。

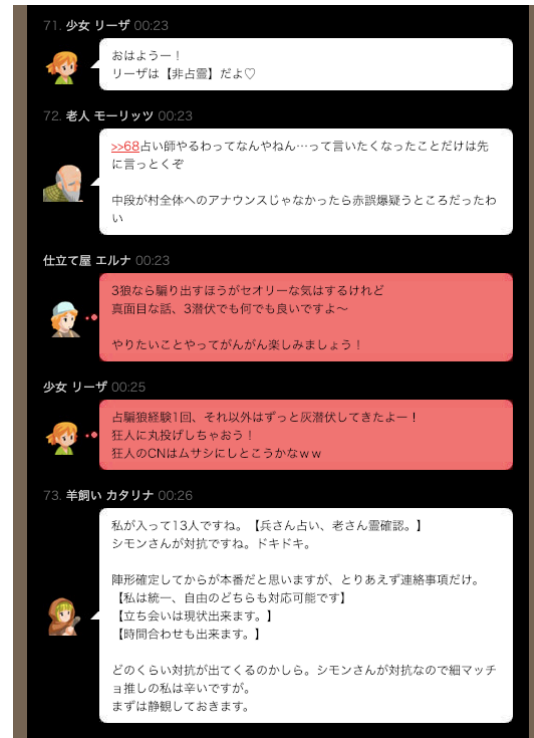


図 2: 人狼 BBS の例

かなければ「翌日」となる。つまり、1日に2人 (昼と夜で1人ずつ) プレイヤーが減っていくことになる⁴。

3.2 人狼の基本戦術

3.1 節のルールを踏まえて、人狼プレイヤーの視点からは、以下の二種類の基本戦術が考えられる。占い師や霊媒師などの役職を騙ることで特殊能力による (と思わせた) 誤情報を流す戦術 (騙り人狼) と、特殊能力がない村人として振る舞い、目立つ行動をしないことで処刑や占いなどを逃れる戦術 (ステルス人狼) である。騙り人狼は、村人を誤った結論に能動的に誘導できる利点があるが、騙られた相手からは人狼陣営⁵であることが分かってしまい、処刑や占いの対象になりやすいという欠点もある。一方で、ステルス人狼は、議論の誘導という点では受動的であるが、処刑や占いの対象に直接働きかけない⁶ため、相手に人狼であることが疑われにくい。ゲーム中に複数の人狼がいる場合、上記の戦術は分担されることが多い。全員が騙り人狼を行うと、本来1人しかいないはずの占い師などの役職が複数人存在することになり、役職を持っている全員を処刑することで人狼陣営が敗北する。また、全員がステルス人狼を行うと、占い師や霊媒師が本物でほぼ確定⁷し、村人陣営が有利となる。したがって、複数の人狼がいる場合、両方の戦術を分担する戦略がとられやすい。

⁴狩人が襲撃から守るなどの例外は存在するが基本的には2人ずつ減っていく。

⁵狂人の可能性もあるため人狼であることまでは分からない。

⁶人狼が占い師として騙りに出ている場合、処刑や占いの対象を直接左右する発言を行うことができる。

⁷狂人が騙る可能性がある。

表 1: 各村のデータ数と役職の人数

キャラクター数	ファイル数	役職ごとのキャラクター数						
		村人	占い師	狩人	霊能者	人狼	狂人	NPC
14人	89	6人	1人	1人	1人	3人	1人	1人
15人	33	7人	1人	1人	1人	3人	1人	1人
16人	309	8人	1人	1人	1人	3人	1人	1人
全体	431							

本稿では、プレイヤー個人の戦術に着目し、ステルス人狼に焦点をあてる。ステルス人狼の「村人として振る舞い、目立つ行動をしない」という基本行動は、実際のゲームにおいて簡単ではない。村人は特殊能力をもたないが、これは昼フェイズの議論において黙っていることを意味しない。ゲームの序盤では、これといった根拠がない状況で誰かを処刑しなければならない場面がしばしばあり、そういった状況では村人陣営にとってあまり有益な情報をもたらさないプレイヤーから処刑することが多い。結果、議論を黙って聞いているだけだったり、無益な発言をするだけのプレイヤーは処刑の対象となりやすい。したがって、ステルス人狼は、ある程度議論に参加して、村人陣営に有益である（と思わせる）発言をしなくてはならない。

ステルス人狼が、さらに難しいのは、有益すぎる発言もできないということである。これは本当の人狼プレイヤーを特定させる発言ができないというだけではない。村人陣営にとって非常に有益なプレイヤーであると思われた場合、本当に村人陣営であれば人狼から襲撃されるはずである。それにも関わらず、襲撃されていないという事実は、本当は人狼陣営であるという疑いを抱かせることになり、処刑や占いの対象となりやすくなる。したがって、ステルス人狼を実現するためには、「村人陣営に有益な発言をしつつも目立たない」というバランスを保った発言をする必要がある。

4 人狼 BBS

インターネットの掲示板を利用した人狼ゲームとして人狼 BBS がある。図 2 に人狼 BBS の画面を示す。1つの村⁸にはノンプレイヤーキャラクター (NPC) 1人を含めた 10人から 16人⁹が参加し、ゲーム内の 1日を現実の 1日で行う。各プレイヤーが 1日にできる最大発言数は 20回であり、またテキストベースで行うためノンバーバルコミュニケーション情報が一切考慮されないのも特徴である。人狼 BBS において、人狼は専用のチャット (囁き) で内通ができる。図 2 中の赤い吹き出しが囁きにあたり、人狼であるエルザとリーザが戦略について相談している。ゲームの早い段階で、人狼たちは囁きを使って、採用する戦略や、騙り人狼とステルス人狼の分担などを相談することが一般的である。

本稿では、分析対象として人狼 BBS:G 村から python の beautifulsoup¹⁰ を使ってログを収集した。人狼が 3人存在する村¹¹、かつ、ゲーム中に処刑や襲撃以外でプレイヤーが脱落していない¹²村を対象として、1つの

⁸人狼 BBS では 1回のゲームを 1つの村と呼称する。

⁹実際のプレイヤー数は 9人から 15人となる。

¹⁰<https://github.com/waylan/beautifulsoup>

¹¹キャラクター数が 13人から 16人の村が該当する。

¹²1日に 1回も発言しなかったプレイヤーは強制的に脱落させられる。また、自らゲームを止める場合もある。

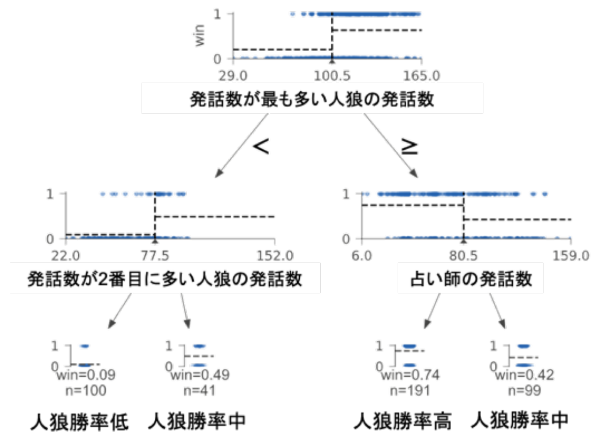


図 3: 発言数に基づく決定木 (ゲーム全体)

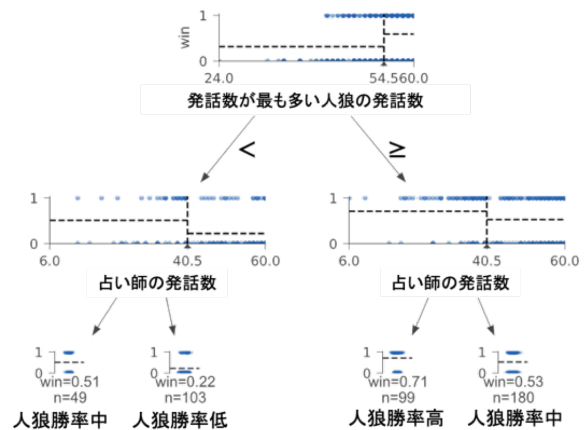


図 4: 発言数に基づく決定木 (ゲーム前半)

村を 1ファイルとして収集した。収集条件を満たしたのは、表 1の通り、14人から 16人の村について、全体で 431ファイル (243MB) となった。各ファイルにはゲーム開始前のプロローグ部分が含まれているが、役職が割り当てられる前の雑談であるため、本稿では分析対象からプロローグ部分を除外した。除外後の 1ファイルあたりの平均発言数は 70.7 発言となった。

5 プレイヤの発言数/注目数と勝率との関係の分析

本節では、文献 [3] の結果を紹介することで、発言数や注目数と勝率との関係を述べる。3.2 節で述べたように、ステルス人狼は「村人陣営に有益な発言をしつつ

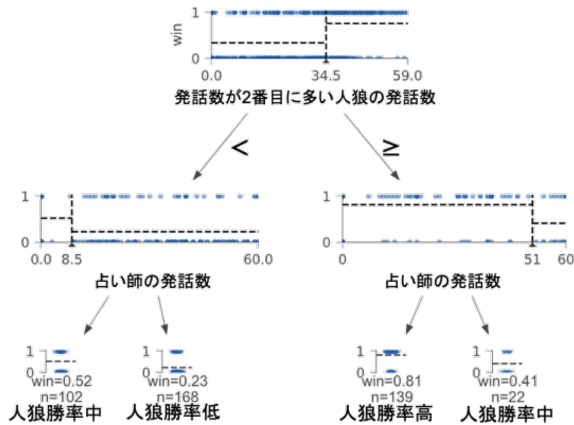


図 5: 発言数に基づく決定木 (ゲーム後半)

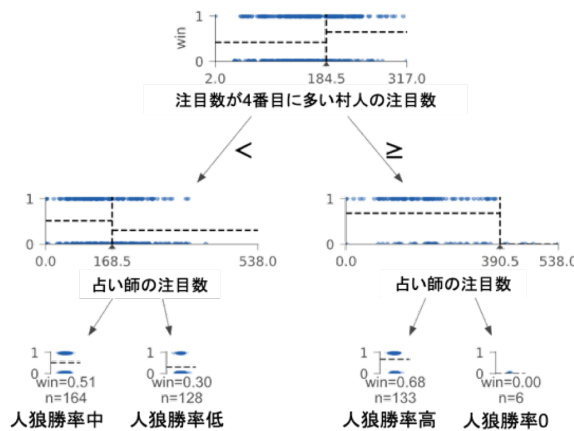


図 6: 注目数に基づく決定木 (ゲーム全体)

も目立たない」必要がある。発言の内容が有益かどうかを定量的に判断するのは難しい問題である。そのため文献 [3] では、目立っているかどうかに着目して分析を行った。目立っている程度を示す指標として、まず、発言が多いプレイヤーは目立っていると考え、あるプレイヤーの発言回数を発言数として用いた。もう一つの指標として、実際に他者の注目を浴びているかどうかを考慮し、あるプレイヤーが他のプレイヤーの発言で言及された回数を注目数として用いた。プレイヤーごとの発言数と注目数が勝敗にどう影響を与えるかを決定木を用いて調査した。16人のキャラクターを属性、各キャラクターがゲーム全体を通して何回発言したかを属性値、ゲームの勝敗をクラスとして、発言数に関する学習データを作成した。注目数についても同様に作成した。また、収集した431件のデータにおいてゲーム最終日の中央値が7日目であったため、3日目までをゲーム前半、4日目以降をゲーム後半として分割し、それぞれの期間における発言数と注目数の学習データを作成した。この時、人狼や村人など複数存在する役職は、回数が多い順にソートした。また、キャラクターが16人に満たない村では、不足分をダミーの村人がいるものとして学習データを作成した。決定木の作成にはpythonのライブラリであるscikit-learn¹³を用い、決定

¹³<https://github.com/scikit-learn/scikit-learn>

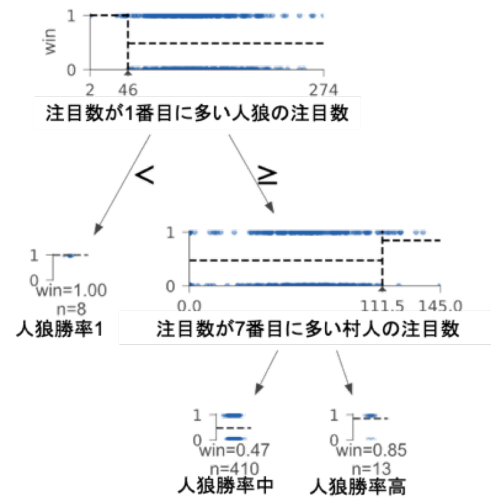


図 7: 注目数に基づく決定木 (ゲーム前半)

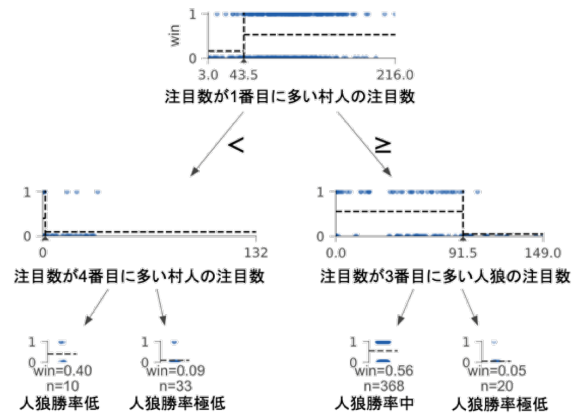


図 8: 注目数に基づく決定木 (ゲーム後半)

木の表示には dtreeviz¹⁴を用いた。

ゲームの全体、前半、後半の発言数と注目数による決定木を図3から図8にそれぞれ示す。図中の win は人狼陣営の勝率を表し、n は該当する村の数を表している。クラスである勝率を縦のパラメータとした。人狼陣営が勝った場合が1、村人陣営が勝った場合が0に対応している。属性である各役職の発言数、および注目数を横のパラメータに対応させている。発言数、注目数について、ある閾値によって、分岐が生じる。右側の枝が閾値以上の場合に対応し、左側の枝が閾値未満の場合に対応する。最終段に表示されているのは、クラスの平均と、該当するクラスの個数であり、平均が高いほど、人狼の勝率が高いと言える。図3であれば、発言数が最も多い人狼の発言数が54.5以上であり、占い師の発言数が40.5未満である場合、人狼陣営の勝率が71ポイントと高く、該当する村が9件あったことを示している。leave-one-out 交差検証による分類精度を表2に示す。ゲーム前半の着目数についての分類精度が低く信頼性に乏しいが、他は5割以上の精度である

¹⁴<https://github.com/parrt/dtreeviz>

表 2: 発話数と注目数に基づく分類精度

	全体	前半	後半
発話数	0.74	0.58	0.67
注目数	0.60	0.27	0.59

表 3: 囁き数に基づく分類精度

	全体	前半	後半
囁き数	0.62	0.58	0.57

ため、このデータに基づいて考察を行った。

発話数に基づく決定木を図3から図5に表示する。最初の分岐はどの決定木も人狼の発話数によってなされているため、人狼の発話数により勝敗が分岐することが確認できる。図3より、二段目の分岐にも人狼の発話数があり、人狼の発話数がさらに一定数以下であると人狼の勝率低下が見られる。また、このときの人狼の勝率は最低で9ポイントであり、発話数が最も多い人狼の発話数が閾値以上で占い師の発話数が閾値未満であれば、人狼の勝率は74ポイントとなり、65ポイント増加する。

ゲームの前半、後半における発話数に基づく決定木から、二段目の分岐には占い師の発話数があり、占い師の発話数が少ない方が人狼陣営の勝利に繋がる。つまり人狼が勝利するには初回から出来るだけ発話をし、占い師の発話数を制限する行動が必要である。

注目数に基づく決定木を図6から図8に表示する。注目数は特定の村人が人狼であると疑われているかどうか、占い師が本物であるかどうかの方に会話の主体があると思われる。図7より、ゲーム前半の注目数に基づく決定木をみると人狼の注目数は勝敗に関係しており、かつ一定数以下である場合、確定で勝利できる。しかし表2の精度を考慮すると、前半の人狼の注目数においてはあまり良い分類ではない。ゲームの後半に関しては、二段目の人狼の注目数による分岐がきているが、どちらに分岐しても勝率はそこまで高くない。つまり、ゲームの後半では人狼の注目数が低くても人狼の勝率陣営に有利であるとは言えない。以上から、発言数においては人狼の発話数が勝敗に関与しており、なおかつゲームの前半でもある程度の精度を持つので、この分析を基に特定のプレイヤーの発話数を操作する方法を考えていく。特に人狼が発話数を操作する際、他の人狼と作戦が被らぬよう、うまく連携する必要があり、またうまく連携が成されることで効率的に発話数を操作することが可能となる。人狼の連携について詳しくは6節で説明を行う。注目数においては思ったような成果は得られず、人狼への注目数が勝敗に与える影響があまり大きくない。特に注目数を減らす場合、発話数とは違い、処刑や襲撃などでゲームから除外しても制限することができない。そのため、注目数を上げないような発話内容や、役職の推定方法を今後考察していく。

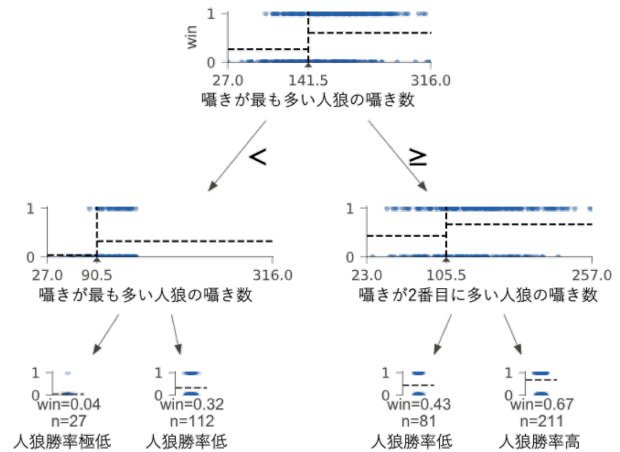


図 9: 囁き数に基づく決定木 (ゲーム全体)

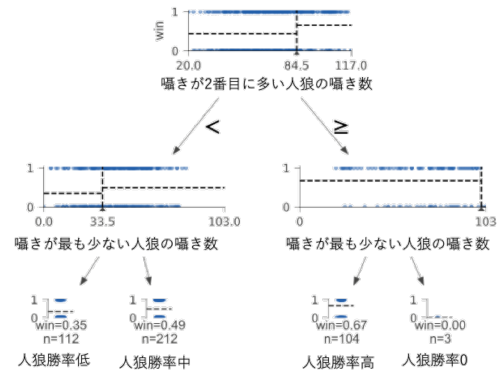


図 10: 囁き数に基づく決定木 (ゲーム前半)

6 人狼の連携度と勝率との関係の分析

6.1 囁き数

他の役職にはない人狼固有の能力として、4節で述べた「囁き」がある。囁きを利用することで秘密裏に人狼同士の連携が可能となり、勝敗に大きな影響を与えることができる。例えば、勝敗に影響を与える方法として、5節で述べたように、特定のプレイヤーの発話数を操作することが考えられる。特定のプレイヤーの発話数を操作する簡単な方法として、「問いかけ」がある。「問いかけ」とは、特定のプレイヤーに話しかけることで応答を募り、意図的に発話数を増加させることである。ただし一日にプレイヤーが発話できる数は有限であるため、一人で全てのプレイヤーの発話数を調整することは困難である。しかし自らの役職が人狼である場合には、囁きを介することで、人狼以外のプレイヤーに知られることなしに、発話数に余裕がある人狼に「問いかけ」をお願いし、人狼側が有利になるよう特定のプレイヤーの発話数を操作できる。各々の人狼が囁いた回数を囁き数として用い、人狼の囁き数が勝敗にどう影響を与えるかを決定木を用いて調査し

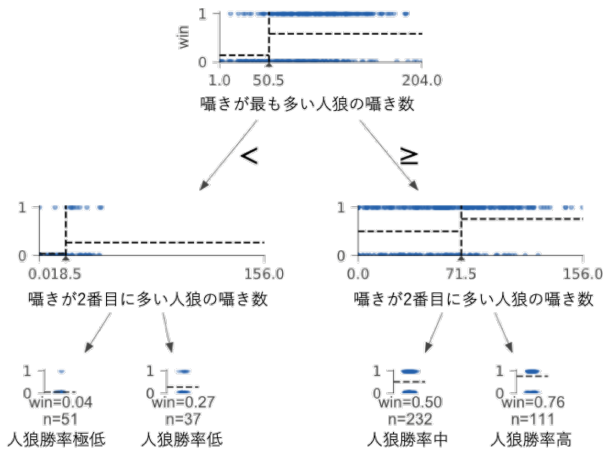


図 11: 囁き数に基づく決定木 (ゲーム後半)

た. 学習データの作成法は5と同様であり, 3人の人狼を属性, 各人狼がゲーム全体を通して何回囁いたかを属性値とした. 決定木の作成, 表示についても5と同様である. 囁き数による決定木を図9から図11にそれぞれ示し, leave-one-out 交差検証による分類精度を表3に示す. 属性である各人狼の囁き数を横のパラメータに対応させており, この決定木を基に連携度について考察する.

6.2 考察

囁き数が多いほど人狼同士の連携がとれていると単純にみなすことはできない. 例えば, 「問いかけ」や作戦の提案を行う際, 他の人狼が必ずしも自らの提案に乗ってくれるとは限らない. 自らの提案を受け入れてもらうには, 人狼ゲームの本質でもある対話を通して説得する必要がある. 対話をするということは, ある人狼の発言に対し他の人狼の応答がなされるので, 各人狼の囁きの数にはあまり差がない場合, 対話が成り立っていると考える. したがって, 人狼全体での囁きの数が多く, 各人狼の囁き数で差がないほど, 作戦や状況についてしっかり話し合い, 連携が取れていると仮定した. 図9より, ゲーム全体を通して, 囁き数が多い方が人狼の勝率は高い. 特に最初の分岐, 二段目の分岐ともに閾値以上である場合の人狼の勝率は67ポイントと高く, 最初の分岐, 二段目の分岐ともに閾値未満である場合に比べ, 勝率は63ポイントも上昇する. ゲームの前半では図10より, 最初の分岐は囁きが2番目に多い人狼であり, 閾値は84.5となっている. 二段目の分岐はどちらも囁きが最も少ない人狼によって成されているが, 最初の分岐で閾値未満の場合, 人狼の勝率は最大でも50ポイントである. 逆に最初の分岐が閾値以上の場合, 特異な三件を除いて人狼の勝率は高い. 最初の分岐において閾値以上の場合, 最も囁きが多い人狼もまた閾値以上の囁きを行なっているということであり, 少なくとも人狼二人は一定の対話を行なっていることが伺える. ゲームの後半では図11より, 最も囁きが多い人狼の囁き数で分岐が成されている. 二段目の分岐はどちらも二番目に囁きが多い人狼によって成されているが, 最初の分岐で閾値未満の場合, 人狼

の勝率は最大でも27ポイントであり, 高くはない. 逆に, 人狼の勝率が最も高くなるのは, 囁きが2番目に多い人狼の囁き数も多いときである. つまり, ゲームの後半においても対話を行い, 作戦についてよく話し合いをすることが勝率を上げる要因となる. まとめるに, ゲームの前半でも後半でも最低人狼二人の対話が一定数が行われている場合, 勝率は高くなる. ただし, 今回の解析は人狼の連携度という観点では, 囁き数の差についての評価が成されておらず, 別の処理が必要となる. また, 今回収集した人狼が三人存在する村では, 囁きのデータ数が40MBと少ないため, データとしては不十分であり, 解析に使用する村の数を増やさなければならない. しかし人狼が三人ではない村も解析の対象に含める場合, 役職に変動があり, それにより採用する戦略も変わってくるため, 正規化を行う必要がある. つまり人狼の数, および役職に寄らない正規化の条件を決定することが当面の課題である.

7 まとめ

本稿では, 「人狼」エージェントに焦点をあて, 文献[3]で得られた, 一人の人狼の発言数が101を超えることで人狼陣営が勝利する可能性が最大65ポイント上がるという知見に基づき, 人狼同士の囁き数がゲームの勝敗とどのように関係するかを分析した. 囁きを解析することで最低でも人狼二人の囁き数が106以上の時, 人狼陣営が勝利する可能性が最大で63ポイント上がることを確認した. 今後は発言内容および囁きの内容も考慮し, 研究を進めていく. また, 阪本らが開発している人狼ゲームプラットフォームLiCOS[4]ではプレイヤーの本心と欺瞞を対話とは別に収集可能なため, 人狼の連携度を考慮する際に活用していきたい.

謝辞

この研究は平成30年度国立情報学研究所公募型共同研究の助成を受けています.

参考文献

- [1] 稲葉通将, 鳥海不二夫, 高橋健一: 人狼ゲームデータの統計的分析. ゲームプログラミングワークショップ2012論文集, pp. 144-147, (2012)
- [2] 稲葉通将, 大畠菜央実, 高橋健一, 鳥海不二夫: 雑談ばかりしていると殺される? 人狼ゲームにおける発言行為タグセットの提案とプレイヤーの行動・勝敗の分析. 情報処理学会論文誌, vol. 57, No. 11, pp. 2392-2402, (2016)
- [3] 永山翔滋, 阿部稔太郎, 大矢康介, 阪本浩太郎, 渋谷英潔, 森辰則, 神門典子: ステルス人狼エージェント作成に向けた発言数と注目数の分析. 言語処理学会第25回年次大会, P3-30, (2019)
- [4] 阪本浩太郎, 渋谷英潔, 森辰則: プレイヤによる他者の役職推定過程を記録する人狼ゲームプラットフォームLiCOSの開発と欺瞞コーパスの収集. インタラクティブ情報アクセスと可視化マイニング第21回研究会, (2019)

プレイヤーによる他者の役職推定過程を記録する人狼ゲーム プラットフォーム LiCOS の開発と欺瞞コーパスの収集

Development of a “werewolf game” platform LiCOS to record players’ inference processes of other player’s role and making a deception corpus

阪本 浩太郎^{1,3*} 永山 翔滋¹ 飯塚 章裕¹
Kotaro Sakamoto¹ Shoji Nagayama¹ Akihiro Iizuka¹
渋谷 英潔² 森 辰則² 神門 典子^{3,4}
Hideyuki Sibuki² Tatunori Mori² Noriko Kando^{3,4}

¹ 横浜国立大学 大学院 環境情報学府 ² 同 環境情報研究院
¹Yokohama National University ²*ditto*
³ 国立情報学研究所 ⁴ 総合研究大学院大学
³National Institute of Informatics ⁴SOKENDAI

Abstract: “Werewolf game” is a popular multiplayer game in which “villagers” are trying to figure out who is a “werewolf” through conversation. LiCOS is a platform for multiple users to play an online real-time “werewolf game” in their web browsers. It needs to interactively exchange data among the users. Therefore, in this paper, we discussed the mechanism of LiCOS which performs interactive data communication among users, and reported the dialog log collected through an actual game.

1 はじめに

人狼ゲームとは、対話を通して「村人」の中に潜伏した「人狼」を見つけ出す対戦型の多人数ゲームであり、近年、人狼知能プロジェクト¹など研究テーマとしても注目されている。人狼ゲームにおける対話は、必ずしも協調的なものではなく、勝利のために他のプレイヤーを説得したり誘導したりすることが要求される。この説得や誘導には、虚偽を述べて騙すことも含まれている。例えば、「人狼」となったプレイヤーは、他のプレイヤーに正体を悟られないよう、普通の「村人」のふりをして「偽の推理」を披露したり、「占い師（人狼の正体を知ることができる）」を騙って他のプレイヤーを扇動したりすることで勝利を目指す。また、「村人」であっても「人狼」を焙り出すために、敢えて「村人」と認識しているプレイヤーを「人狼」と疑うような発言をする場合もある。

こういった、認識している事柄（以降、本心と呼ぶ）と異なる認識を他者に与えようとする発言を本研究では欺瞞と定義する。我々は、これからの対話システムには、ユーザーの命令に唯々諾々と従うのではなく、自発的に対話を進めていくことも必要であると考えており、システムが欺瞞的な発言をすることはその一つに該当すると考えている。こういった研究を進める上で、

欺瞞と本心を対応付けた対話コーパス（欺瞞対話コーパス）が必要である。それゆえ、我々は、人狼ゲームにおける対話を収集することで欺瞞対話コーパスを構築しようとしており、そのための人狼ゲームプラットフォーム LiCOS を開発した [1, 2]。

LiCOS は、Web ブラウザを介して、複数のユーザがオンラインでリアルタイムにゲームを行うためのプラットフォームであり、ユーザ間に生じるインタラクティブなデータのやり取りを処理する必要がある。そのため、本稿では、ユーザ間のインタラクティブなデータ通信を行う LiCOS の仕組みについて説明し、実際のゲームを通して収集された対話ログを考察する。

2 関連研究

従来研究で構築された対話コーパス [3, 4] では、発話者は基本的に誠実であり、本研究の焦点である、相手を騙すことを目的とした発話は収録されていない。人狼ゲームに関する従来研究 [5, 6] では、人狼 BBS² のログを利用したものが多い。人狼 BBS はオンラインで行う掲示板型人狼ゲームであるが、図 1 に示したような対応表は存在しないため、プレイヤーの本心は役割や投票などの行動から推測するしかない。一方で、

*連絡先：横浜国立大学環境情報学府
〒240-8501 横浜市保土ヶ谷区常盤台 79-7 森 辰則研究室
E-mail: nagayama@forest.eis.ynu.ac.jp

¹<http://aiwolf.org/>

²<http://ninjinix.com/>

異老 | %五宿 | M修旅虎兵宿妙 † 羊孫娘面
真偽 | 霊白白 | 灰白灰灰灰灰 † 白白狼狼
偽真 | 霊白白 | 白灰灰灰灰灰 † 白白狼狼

のように、人狼 BBS での発言中に表形式で推理や確定事実を述べるプレイヤーは一定数存在し、対応表に相当する形式で自分の推理などを記述している。対応表を導入した人狼プラットフォームは LiCOS の大きな特徴であり、プレイヤーの本心に関するデータを自然に収集できると考えている。

以上の考えから、我々は、文献 [1] において対応表を実装した LiCOS を提案し、文献 [2] において通信プロトコルと実際のゲームを通して収集された対話ログを報告した。本稿では、LiCOS がゲーム中のデータの流れをどのように管理しているかに焦点をあてて説明する。

3 LiCOS の仕様

ユーザーはウェブブラウザを通して人間のプレイヤーとしてゲームに直接参加するか、あるいはロボットをゲームに参加させることができる。

図 1 はウェブブラウザ上で人間のプレイヤーが参加したゲームのプレイ画面の例である [2]。左上に、ゲーム内の日にちとフェーズ、そのフェーズ終了までの残り時間を表示する。右上に、ゲーム中にプレイヤーが演じるエージェントの画像、名前（例では「Yihan」）、正体となる役職（例では「Villager (村人)」）と陣営（例では「Villagers (村人陣営)」）を表示する。左中央にプレイヤーの発言のタイムラインを表示する。下部に発言のための発言入力フォームを表示する。下部左側の公開用発言入力フォームから投稿した発言は、ゲーム参加者全体に共有される。下部右側の非公開用発言入力フォームから投稿した発言は、プレイヤー本人のみが閲覧できる。右中央に対応表を表示する。

プレイヤーは、タイムラインを通して他のプレイヤーと対話しつつ、他プレイヤーの正体の予想を対応表で容易にまとめることが可能である。対応表の入力を本心と仮定することで、プレイヤーの発言と本心の対応づけが得られる。次に、処刑先投票などで誰に投票したかによっても本心が現れると考えられる。さらに、非公開用発言入力フォームに入力された発言にも、本心が現れる可能性がある。発言、対応表入力、投票は全てタイムスタンプ付きでログとして保存されるため、これらの情報からタイムスタンプを基に発言と本心の対応づけを得る。発言と本心の対応の情報を記録すると、その中に欺瞞の状況も記録されている。

4 LiCOS のシステムの構成

図 2 は、LiCOS のシステムの構成 [1] を示している。ユーザはウェブブラウザを介してゲームサーバと WSS 通信 (WebSocket over SSL/TSL) を行うことで、ゲームをプレイする。ウェブブラウザとゲームサーバ間は

JSON-LD 1.1³形式のデータを投げ合う。また、ユーザは自身の代わりにロボットに戦わせることも可能であり、ロボットとゲームサーバ間のやり取りは基本的にウェブブラウザと同様である。開発状況としては、前半の人間同士でオンラインでゲームをプレイすることができる状態である。後半のロボットに戦わせられるようにすることについては今後行う。

5 ゲームの通信プロトコル

ゲームの通信プロトコルは、WSS を使用し、通信データは JSON-LD 1.1 形式で記述する。スキーマは、JSON Schema Draft-07⁴で記述し、<https://werewolf.world> にスキーマと例を公開している。データの構造は次の 7 種類である。

- systemMessage
- playerMessage
- boardMessage
- voteMessage
- flavorTextMessage
- scrollMessage
- errorMessage

systemMessage は、フェーズを切り替える際に伝える情報を記述する。playerMessage は、タイムラインに表示される 1 つの発言の情報を記述する。boardMessage は、プレイヤーが対応表に入力した情報を記述する。voteMessage は、プレイヤーの投票情報を記述する。flavorTextMessage は、フレーバーテキストを記述する。フレーバーテキストはゲームに不参加である特殊キャラクターのセリフのみで構成され、flavorTextMessage はそのようなセリフを playerMessage のリストとして内部に記述する。scrollMessage は、プレイヤーの画面スクロールの位置を記述する。errorMessage は、エラー情報を記述する。

6 ゲームプレイ時のデータの流れ

複数のクライアントとサーバの間は WSS を用いて非同期でデータを送受信する。一方で、ゲーム進行の都合により、朝フェーズから昼フェーズへ移行する、あるいはゲームの状況が勝敗条件を満たしたので結果を公表するといったように、一定時間が経過したり、何かの条件を満たした際に、同じゲームに参加している全てのクライアントにサーバは同時にデータを配らなければならない。そのような複雑なデータの流れを実現するために、Akka Streams⁵の MergeHub や BroadcastHub などを用いて図 3 が示すようにシステムを設計した。図の中央にある

³<https://json-ld.org/spec/latest/json-ld/>

⁴<https://json-schema.org/draft-07/json-schema-release-notes.htm>

⁵<https://doc.akka.io/docs/akka/current/stream/index.html>



図 1: ゲームのプレイ画面

WSS と書かれたシリンダーがサーバとクライアントの接続を示しており、そこから左側がサーバ、右側がクライアントである。細い矢印はデータの流れを示しており、サーバ側では WSS から届いたデータは MergeHub でマージされ、処理されてから、BroadcastHub を経由し WSS に届いている。Actor から出力されたデータは Queue を通り、処理されて、BroadcastHub を経由し WSS に届いている。MergeHub や Queue からデータを受け取って BroadcastHub に出力するデータの処理については、8 章で説明する。図では簡略化のために、ユーザを一人だけ示したが、実際には、ゲームをプレイ中の人数分だけ、MergeHub と BroadcastHub に別の WSS が接続している。クライアント側では、WSS から届いたデータがウェブブラウザ上で表示される。ウェブブラウザから出たデータは WSS に届く。ユーザとウェブブラウザの間ではインタラクションが起こる。このような設計にすることで、次の 4 つのパターンに分類される、ゲーム中のデータの流れを網羅的に扱うことができる。

1 つ目のパターンは、プレイヤーが発話をした場合のデータの流れである。プレイヤーが発話した場合、発話情報を持つ playerMessage は一度サーバに送られ、同一ゲームに参加する全クライアントに対しその playerMessage が同時に送られる。このようなデータの流れは、図 4 の経路を通ることで実現する。2 つ目のパターンは、他人が送信した playerMessage、もしくはサーバで発生したエラー情報を伝える errorMessage をクライアントが受け取る場合のデータの流れである。このようなデータの流れは、図 5 の経路を通ることで実現する。3 つ目のパターンは、クライアントからサーバに一方通行で伝える場合のデータの流れである。対応表の更新情報を持つ boardMessage、投票情報を持つ voteMes-

sage、スクロールの位置情報を持つ scrollMessage、クライアント上でのエラー情報を持つ errorMessage が、この場合のデータに該当する。このようなデータの流れは、図 6 の経路を通ることで実現する。4 つ目のパターンは、一定時間の経過などでサーバからクライアントに伝える場合のデータの流れである。時間経過や勝敗判定結果によって送られる、フェーズを切り替える際に伝える情報を持つ systemMessage と、フレーバーテキストを持つ flavorTextMessage が、この場合のデータに該当する。このようなデータの流れは、図 7 の経路を通ることで実現する。

7 同一ゲームに参加する全クライアントへのサーバからのデータ送信

ゲーム進行の都合上、一定時間が経過するごとに、あるいは勝敗が決定した際に、同じゲームに参加する全てのクライアントに同時にデータを配る必要がある。これを Akka Actors⁶ の Actor や Timers などを用いて図 8 のように設計した。図中のノードは状態を表しており、直線の矢印は状態遷移を表している。基本的には、Morning, Noon, Night を循環するが、勝敗が決まると、その時の状態に応じてそれぞれ MorningToEpilogue, NoonToEpilogue, NightToEpilogue に移動する。図の右側に示している、遷移後に送信するデータを図 3 の Queue に追加する。

⁶<https://doc.akka.io/docs/akka/current/index-actors.html>

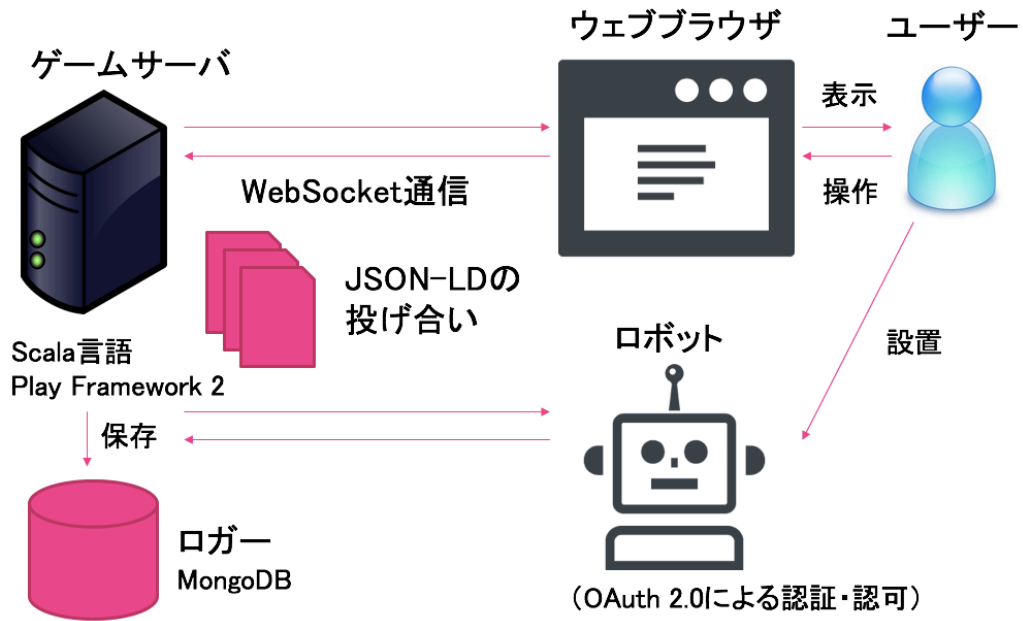


図 2: システム構成

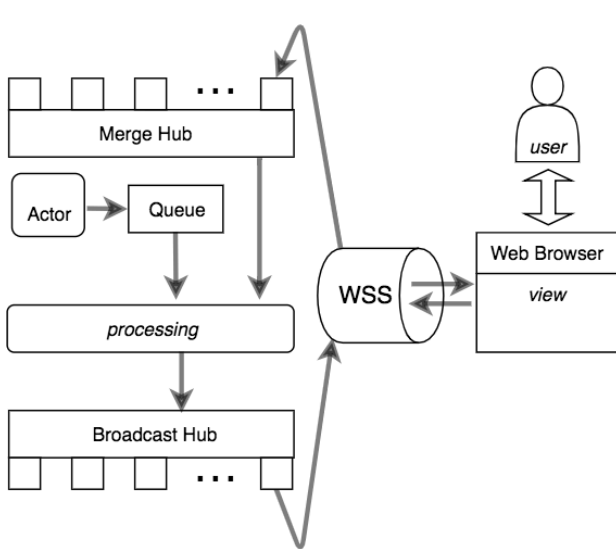


図 3: データの流れ

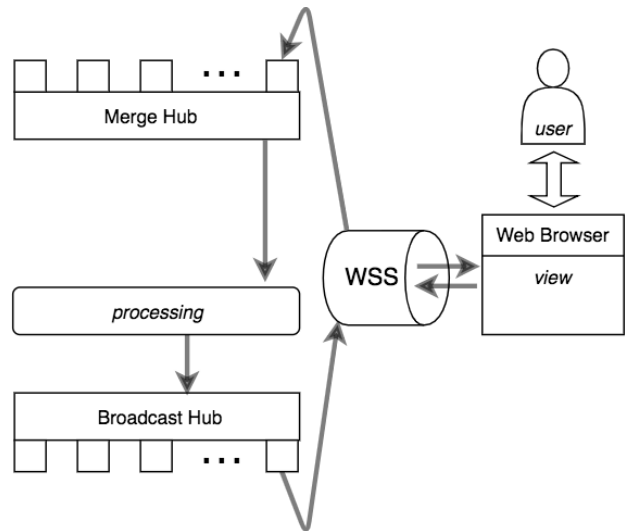


図 4: 自身の発話についての playerMessage データの流れ

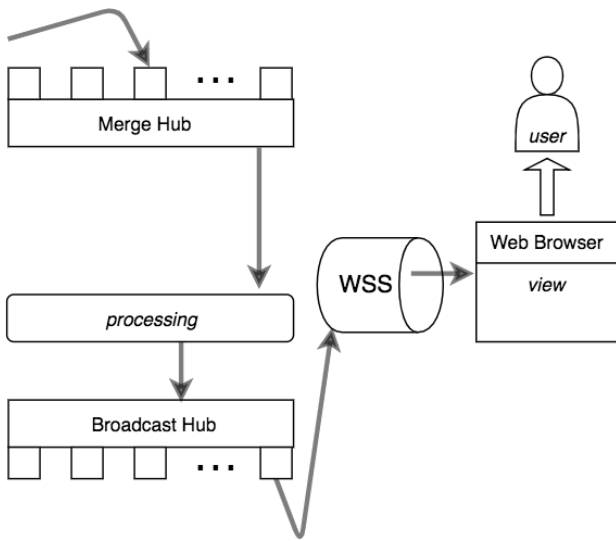


図 5: 他人の発話についての playerMessage データの流れ

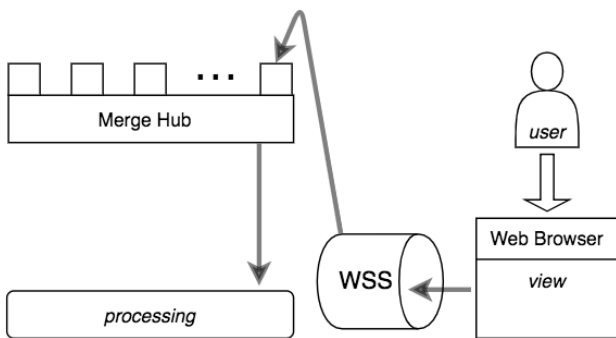


図 6: クライアントからサーバへ方通行のデータの流れ

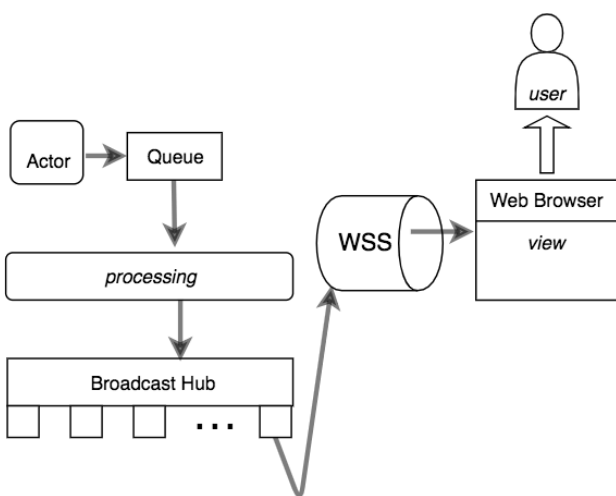


図 7: 時間経過などでサーバからクライアントに伝えるデータの流れ

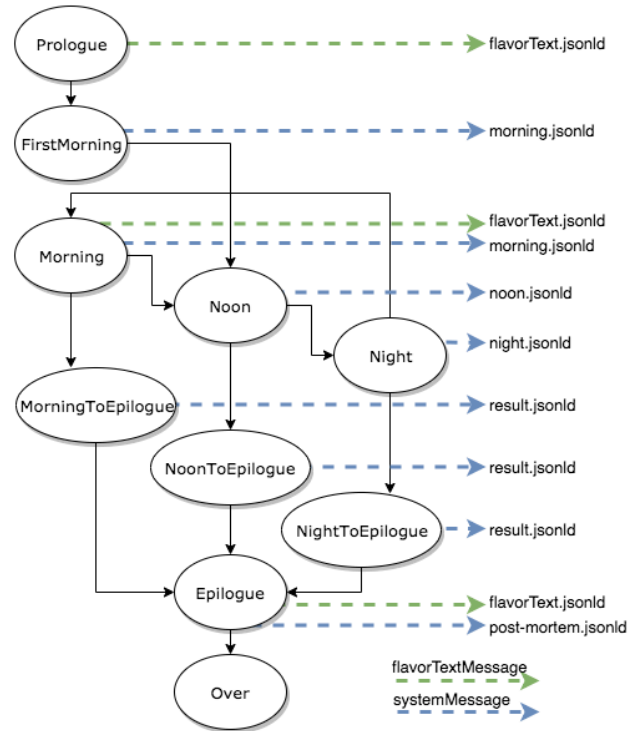


図 8: 同一ゲームに参加する全クライアントへのサーバからのデータ送信

8 処理の流れ

図 3 の MergeHub や Queue からのデータを受けて BroadcastHub に出力するまでの処理の流れについて説明する。まず、データをログに書き出す。次に、データをその構造に基づいて playerMessage, voteMessage などに分類する。データに含まれる値により、必要であればプレイヤーへのアクセスを制限する。例えば、他人のプライベートな発話は読まずに捨てる。分類されたデータごとに、必要があれば、データベースに保存する。例えば、voteMessage はあとで集計するため一度保存する。クライアントにデータを送る必要があれば、BroadcastHub に渡す。

9 ゲームログ

本節では、LiCOS によって収集された実際の対話ログを紹介する⁷。LiCOS を用いて人狼ゲームをオンラインで 8 人（内訳：人狼 1 人、占い師 1 人、その他 6 人が村人）でゲームを行った。ゲーム時間は 21 分 38 秒かかり、372 個 (2.3MB) のログファイルが出力された。発話は 64 回、対応表の更新は 59 回、投票は 22 回行われていた。このようなログファイルを 1 つ 1 つ開きながら 1 人 1 人の視点を追って欺瞞の出現に注目しながらゲームログを観察した。

⁷本稿で紹介する対話ログは、文献 [2] で述べたものと同一である。

ソースコード 1 は発話情報を保持する playerMessage からの抜粋であるが、人狼役の Fernando の初日の朝の発言であり、翌朝に占い師が誰であるか占い師自身に名乗り出させようとしている。

ソースコード 1: 人狼役の発話

```
1 {
2   "extensionalDisclosureRange" : [],
3   "phase" : "morning",
4   "date" : 1,
5   ...
6   "text" : {
7     "@value" : "明日になったら占えるから、結果を教えてください",
8     "@language" : "en"
9   },
10  "myAgent" : {
11    "@context" : "https://werewolf.world/context/0.2/agent.jsonld",
12    "@id" : "https://licos.online/state/0.2/village#1/agent#3",
13    "id" : 3,
14    "name" : {
15      "en" : "Fernando"
16    },
17    "image" : "https://werewolf.world/image/0.2/fernando.jpg",
18    "role" : {
19      "@context" : "https://werewolf.world/context/0.2/role.jsonld",
20      "@id" : "https://licos.online/state/0.2/village#1/role#werewolf",
21      "name" : {
22        "en" : "Werewolf"
23      },
24      "image" : "https://werewolf.world/image/0.2/werewolf.jpg"
25    }
26  }
27 }
```

ソースコード 2 は、別の playerMessage からの抜粋であるが、翌朝、このゲームの中でただ一人の占い師役である Nanyamka は自身が生存しているにも関わらず「占い師が殺されている説」を唱えている。

ソースコード 2: 占い師による欺瞞の例

```
1 {
2   "extensionalDisclosureRange" : [],
3   "phase" : "morning",
4   "date" : 2,
5   ...
6   "text" : {
7     "@value" : "占い師殺されてる説",
8     "@language" : "en"
9   },
10  "myAgent" : {
11    "@context" : "https://werewolf.world/context/0.2/agent.jsonld",
12    "@id" : "https://licos.online/state/0.2/village#1/agent#6",
13    "id" : 6,
14    "name" : {
15      "en" : "Nanyamka"
16    },
17    "image" : "https://werewolf.world/image/0.2/nanyamka.jpg",
18    "role" : {
19      "@context" : "https://werewolf.world/context/0.2/role.jsonld",
20      "@id" : "https://licos.online/state/0.2/village#1/role#seer",
21      "name" : {
22        "en" : "Seer"
23      },
24      "image" : "https://werewolf.world/image/0.2/seer.jpg"
25    }
26  }
27 }
```

```
26 }
27 }
```

以上のように、今回観察したログからは、占い師が自身の役職を隠すために、占い師が初夜に死んだことにする際に、欺瞞が現れた。

10 まとめと今後の予定

人狼ゲームプラットフォーム LiCOS におけるオンライン対戦の複雑なデータの流れに対応するシステムの開発について説明した。そして、8人で行ったオンライン対戦のゲームログに現れた欺瞞を報告した。

人狼は一般的にプレイヤー人数が7人程度以上いるとゲームとしての面白さが高まると思われるが、そのような対戦のログを観察するためには、1つ1つログファイルを開くとファイル数も視点数も多くとても大変であることがわかった。ログファイルを入力すると実際のゲーム画面で再現でき、各プレイヤーの視点を追えるようなシステムがあると欺瞞を観察がかなりしやすくなると考えられる。

今後は、欺瞞対話コーパスの作成と公開に向けた参加者を募集、ゲームログの収集、ゲームログからの各プレイヤー視点の再現システムの開発を行いたい。そして、ロボットのゲーム参加も可能にするインターフェースも開発する予定である。

謝辞

この研究は平成 30 年度国立情報学研究所公募型共同研究の助成を受けています。また、本研究を進めるにあたり、サーバマシンを提供してくださった rakumo 株式会社様、開発に協力してくださった横浜国立大学のサークル YNU WAIWAI の宇田川悠大氏に感謝します。

参考文献

- [1] 阪本浩太郎, 渋木英潔, 森辰則. 欺瞞対話コーパスの構築に向けた人狼ゲームプラットフォーム LiCOS の開発. 言語処理学会第 24 回年次大会発表論文集, pp. 885–888, 2018.
- [2] 阪本浩太郎, 永山翔滋, 石下円香, 渋木英潔, 森辰則, 神門典子. 人狼ゲームプラットフォーム LiCOS を用いた欺瞞対話コーパスのためのゲームログの収集. 言語処理学会第 25 回年次大会発表論文集, P8-6, 2019.
- [3] 東中竜一郎, 船越孝太郎, 荒木雅弘, 塚原裕史, 小林優佳, 水上雅博. テキストチャットを用いた雑談対話コーパスの構築と対話破綻の分析. 自然言語処理, 23(1), pp. 59–86, 2015.
- [4] 塚原裕史, 内海慶. オープンプラットフォームとクラウドソーシングを活用した対話コーパス構築方法. 言語処理学会第 21 回年次大会発表論文集, pp. 147–150, 2015.
- [5] 林友超, 馬場瑞穂, 宇津呂武仁. 役職確定情報に着目した人狼ログ・ダイジェストの作成. 第 30 回人工知能学会全国大会論文集, 2F4-2in2, 2016.
- [6] 稲葉通将, 鳥海不二夫, 高橋健一. 人狼ゲームデータの統計的分析. ゲームプログラミングワークショップ 2012 論文集 2012(6), pp. 144–147, 2012.

部屋配置とその出現数に着目した二段階賃料推定

Two-stage rent estimation using frequent subgraphs in floor plans

長谷川 優也 尾崎 知伸*
Yuya Hasegawa Tomonobu Ozaki

日本大学 文理学部
College of Humanities and Sciences, Nihon University

Abstract: The rent for rental property is determined by various factors such as area and age. In this paper, a two step rent estimation model is proposed in which the estimated rent value from the basic attributes will be corrected based on the layout information. In the model, each rental property is represented as a graph of room layout, and the number of occurrences of partial layouts (or subgraphs) is employed as attributes. In the evaluation experiments, in addition to the verification of the estimation accuracy, we extracted critical layouts to improve the interpretability.

1 はじめに

マンション等の賃貸物件の賃料は、専有面積や築年数に加え、駅までの距離や部屋の向き、周辺環境、間取り（部屋配置）など、様々な要因によって決定される。一般に、専有面積が大きいほど賃料は高くなり、築年数が古いほど賃料は安くなるなど、いくつかの要因に関しては賃料への影響は明らかである。その一方で、賃料増減の明示的な根拠を示すことが難しい要因として、間取り（部屋配置）があげられる。例えば、専有面積や築年数、駅までの距離等の条件がまったく同じ場合でも、部屋配置が異なれば賃料も異なることは一般的であるが、どの様な部屋配置が（いくつ）あれば、どの程度賃料が高く（または安く）なるのかを明示的に説明することは容易ではない。仮に、賃料に大きな影響を与える部屋配置を特定することができれば、物件の貸し手に対しては賃料設定に関する新たな根拠を、借り手に対しては設定された賃料の妥当性判断基準を、また物件の設計者に対しては新たな検討要素をそれぞれ提供することが可能となり、その影響は大きいと考えられる。

これらのことを背景に、本論文では、部屋配置と賃料との関係进行分析し、賃料に大きな影響を与える部屋配置を特定することを試みる。具体的には、専有面積や築年数などの基本的な属性から得られる賃料推定結果を間取り情報に基づき補正する、二段階賃料推定手法を提案する。その際、各間取りをグラフ化すると共に、新たな部分グラフ出現数の定義を提案し、物件（グラフ）中の部屋配置（部分グラフ）出現数を属性とし

て利用する。詳しくは後述するが、提案手法では、家賃推定モデルを2回の教師付き学習（回帰分析）で構築することとなり、より精度の高い家賃推定が期待できる。加えて、部屋配置のみを利用して構築される補正モデルの近似ルールを抽出することで、賃料に影響を与える部屋配置の明示的な特定を実現する。

本論文の構成は以下の通りである。2章で関連研究について言及する。3章では、本研究に関する形式的な定義を与え、二段階賃料推定手法を提案する。その後4章で、間取り図のグラフ化を通じた属性の準備について説明する。5章では、賃料の推定および賃料への影響が大きい間取りの特定に関し、それぞれ評価実験を行う。最後に6章でまとめを行い、今後の課題を述べる。

2 関連研究

これまでに種々の要因に着目した賃料の分析・推定に関する研究が報告されている [1, 2, 3, 4]。例えば文献 [1] では、重回帰分析を用い、物件の各詳細情報が賃料に与える影響を路線別に分析している。文献 [2] では、物件に対する詳細情報を専有面積や部屋数など物件そのものが持つ固有情報と、所在地や駅までの距離などの外的要因に依存する情報の2つに分類し、それぞれに対するモデルを構築・統合することで賃料を推定する手法を提案している。また外部データの利用に関しては、文献 [3] において、地震に関する地域危険度調査結果を用い、物件所在地の危険度が賃料に与える影響を建物の構造別、新旧の耐震基準別に分析している。文献 [4] では、物件詳細の客観的な情報に加え、実際に

*連絡先：日本大学文理学部情報科学科
〒156-8550 東京都世田谷区桜上水 3-25-40
E-mail: tozaki@chs.nihon-u.ac.jp

物件を見たときの感性などの主観情報も利用し、重回帰分析やサポートベクトル回帰 [5] を用いて分析を行っている。

一方、間取りに着目した分析もいくつか報告されている [6, 7, 8]。これらの研究では、部屋を頂点、そのつながりを辺とするグラフを用いて部屋配置（間取り）を表現し、構造的な側面からの分析を試みている。例えば文献 [6] では、大型マンションを対象にグラフに基づく間取りを類型化した上で順序付けを行い、間取りが賃料に与える影響を分析している。また文献 [7] では、頻出部分グラフマイニング [9, 10] の枠組みを利用した分析を展開している。複数の間取り図（グラフ）に共通して現れる部屋配置（頻出部分グラフ）を抽出し、相関ルール分析を用いてそれらの関係を分析するとともに、各頻出部分グラフを含むか否かを属性とした回帰モデルを構築し、部分グラフが賃料に与える影響を考察している。文献 [8] でも同様に、頻出部分グラフにおけるフリーパターン [11] を属性とした回帰木 [12] やモデル木 [13, 14] を構築することで、賃料に影響の強い部分グラフの抽出を試みている。

本論文では文献 [7, 8] と同様、頻出部分グラフマイニングの枠組みを利用した賃料推定を展開する。その際、属性値として、頻出部分グラフの出現の有無だけでなく、出現数を採用することを提案する。また部分グラフのみを利用して賃料の補正モデルを構築することで、より明示的な形で、部屋配置が賃料に与える影響を分析する。

3 賃料推定の定式化

3.1 準備

物件賃料の予測モデル構築に利用する属性の全体集合を $X = X_{basic} \cup X_{graph} \cup \{x_{mp}\}$ と表記する。ここで属性 x_{mp} は最寄り駅単位で算出する物件に対する相場を表す。また X_{basic} は、専有面積や駅からの距離など物件に関する基本的な属性（以降、基本属性と呼ぶ）の集合を、 X_{graph} は 4 章で導入する部屋配置に関する属性（以降、部屋配置属性と呼ぶ）の集合をそれぞれ表す。一方、予測対象である物件賃料属性を p と表記する。また、物件 t に対する属性 $a \in X \cup \{p\}$ の属性値を $a(t)$ と表記する。以上の準備を基に、本研究では、予測値（目的変数）の設定方法と利用属性（説明変数）が異なる複数の回帰モデルを考える。

最も基本的なモデルは、基本属性、部屋配置属性、相場のすべてを利用し、賃料を推定するモデル、すなわち

$$\hat{p} = f_{bgm}(X_{basic}, X_{graph}, x_{mp}) \quad (1)$$

である。

第二の基本的モデルとして、賃料 p を直接推定するのではなく、相場との比 $\frac{p}{x_{mp}}$ を推定するモデル $f_{bg/m}$

$$\frac{\hat{p}}{x_{mp}} = f_{bg/m}(X_{basic}, X_{graph}) \quad (2)$$

を考える。なおこの場合、物件 t に対する賃料は

$$\hat{p}(t) = f_{bg/m}(X_{basic}(t), X_{graph}(t)) \times x_{mp}(t)$$

と推定することができる。

3.2 二段階賃料推定

賃料推定に関する回帰モデル f_{bgm} と $f_{bg/m}$ はそれぞれ、一つのモデルを用いて賃料を推定する。これに対し本論文では、二つのモデル、すなわち基本属性を用いた賃料推定モデルと間取り情報を用いた賃料補正モデルを組み合わせたことを提案する。

提案する第一の推定モデル f_{bm+g} は、基本属性と相場を説明変数とするモデル f_{bm} の推定結果を、部屋配置属性を説明変数とするモデル f_g を用いて補正する構成をしており、形式的には下式で表現される。

$$\begin{aligned} \hat{p} &= f_{bm+g}(X_{basic}, x_{mp}, X_{graph}) \\ &= f_{bm}(X_{basic}, x_{mp}) + f_g(X_{graph}) \end{aligned} \quad (3)$$

本論文では、下記に示す方法により、 f_{bm+g} を 2 段階に分けて推定することを提案する。

1. まず、目的変数を p とし、基本属性 X_{basic} と相場 x_{mp} のみを利用して f_{bm} を推定する。
2. その後、第二の推定として、賃料と推定値の残差 $p - f_{bm}(X_{basic}, x_{mp})$ を目的変数とし、部屋配置属性 X_{graph} を利用してモデル f_g を構築する。

この様に推定を 2 段階に分けることは、2 回の教師付き学習を行うことに相当し、推定精度の向上が期待できる。また補正モデルである f_g は、部屋配置属性 X_{graph} のみを利用して構築されるため、部屋配置が賃料に与える影響をより直接的に反映していることが期待できる。

本論文では f_{bm+g} に加え、相場との比を推定する二段階推定モデルとして、 $f_{b/m+g}$ を提案する。

$$\begin{aligned} \frac{\hat{p}}{x_{mp}} &= f_{b/m+g}(X_{basic}, X_{graph}) \\ &= f_{b/m}(X_{basic}) + \frac{1}{x_{mp}} f_g(X_{graph}) \end{aligned} \quad (4)$$

モデル $f_{b/m+g}$ の推定も、 f_{bm+g} と同様の方法で実現する。すなわち、基本属性 X_{basic} を用いて相場との比

$\frac{p}{x_{mp}}$ を推定するモデル $f_{b/m}(X_{basic})$ を構築し、その後、残差を目的変数、部屋配置属性を説明変数としたモデル f_g を推定する。

4 部屋配置属性の構築

本章では、グラフとして表現される間取りの集合から、部屋配置属性 X_{graph} を抽出する方法、および各物件 t に対する属性値 $X_{graph}(t)$ の算出方法について形式的に説明する。

4.1 頻出部分グラフマイニングを用いた部屋配置属性の抽出

本論文では、関連研究 [6, 7, 8] と同様、部屋を頂点、その繋がりを辺とするグラフを用いて各部屋の間取りを定式化する。具体的には、12種の頂点ラベル（玄関、廊下、居室、水回り、収納、ウォークインクローゼット、台所、ダイニングキッチン、リビングダイニング、リビングダイニングキッチン、ベランダ、窓）及び5種の辺ラベル（ドア、引き戸、ガラス、無し、収納）を用い、ラベル付き無向グラフとして間取りを表現する。なお「水回り」には洗面室、風呂、トイレを含める。また「窓」と「収納」は部屋ではないが、部屋配置に関する主要な要素として採用している。一方、辺に対する「無し」ラベルは、空間を分ける仕切りとしてドアや引き戸が存在しない場合に相当する。

頂点・辺ラベルの全体集合を \mathcal{L} と表記し、ラベル付きグラフ $g = (V_g, E_g, \lambda_g)$ を、頂点集合 V_g と辺集合 $E_g \subseteq V_g \times V_g$ 、ラベル関数 $\lambda_g : V_g \cup E_g \rightarrow \mathcal{L}$ の3項組で表現する。2つのグラフ $g = (V_g, E_g, \lambda_g)$ と $p = (V_p, E_p, \lambda_p)$ に対し、条件

1. $\forall u \in V_p [\lambda_p(u) = \lambda_g(f(u))]$
2. $\forall (u, v) \in E_p$
 $\exists (f(u), f(v)) \in E_g \text{ s.t. } \lambda_p(u, v) = \lambda_g(f(u), f(v))$

を満たす単射関数 $f : V_p \rightarrow V_g$ が存在するとき、 p を g の部分グラフと呼び $p \sqsubseteq g$ と表記する。 n 個のグラフから構成されるデータベース $D = \{g_1, \dots, g_n\}$ に対し、グラフ p の支持度を D 中で p を含むグラフの割合、すなわち

$$S(p, D) = |\{g \in D \mid p \sqsubseteq g\}| / |D|$$

と定義する。利用者による頻度に関する閾値 σ ($0 < \sigma \leq 1$) に対し、条件 $S(p, D) \geq \sigma$ を満たすグラフ p を頻出部分グラフと呼ぶ。また D における頻出部分グラフの集合を

$$\mathcal{F}_D^\sigma = \{p \sqsubseteq g \mid g \in D, S(p, D) \geq \sigma\}$$

と表記する。

賃貸物件 t に対し、その間取りを表すラベル付き無向グラフを $g(t)$ と表記する。また、賃貸物件集合 T に対し、それらの間取り図グラフの集合を $D_g = \{g(t) \mid t \in T\}$ と表記する。本論文では、 D_g から得られる各頻出部分グラフ $p \in \mathcal{F}_{D_g}^\sigma$ を属性として利用する。すなわち $X_{graph} = \mathcal{F}_{D_g}^\sigma$ とする。

4.2 各物件に対する部屋配置属性の属性値

各頻出部分グラフ $p \in X_{graph}$ を属性とする場合、賃貸物件 t に対する p の属性値 $p(g(t))$ が必要となる。属性値 $p(g(t))$ として、グラフ $g(t)$ に対する p の支持度（出現数）を採用することが自然であると考えられるが、一般に、単一グラフ $g(t)$ における部分グラフ p の支持度は自明ではない。加えて、部屋配置を用いた賃料推定という応用を前提とした場合、これまでに提案されている支持度が必ずしも適しているとは限らない。以上のことを背景に、本論文では、単一グラフにおける部分グラフの新たな支持度を提案し、属性値計算に利用することを考える。以下、形式的な準備の後、既存の支持度 [15, 16] を導入し、次いで新たな支持度を提案する。

グラフ g と p に対し、 p と同型である g の部分グラフ、すなわち条件 $o \sqsubseteq g \wedge p \sqsubseteq o \wedge o \sqsubseteq p$ を満たす部分グラフ $o = (V_o, E_o, \lambda_o)$ を g における p の出現と呼ぶ。また g における p の出現の全体集合を

$$O(p, g) = \{o \mid o \sqsubseteq g, p \sqsubseteq o, o \sqsubseteq p\}$$

と表記する。以下 $O(p, g)$ を用い、グラフ g に対する p の支持度を複数定義する。

(1) 出現の有無に基づく支持度

第1の支持度 $S_{BIN}(p, g)$ は、グラフ g 中に p の出現が存在するか否かに基づくものであり、以下の様に定義される。

$$S_{BIN}(p, g) = \begin{cases} 0 & (O(p, g) = \emptyset) \\ 1 & (\text{otherwise}) \end{cases}$$

支持度 $S_{BIN}(p, g)$ は、グラフを用いた賃料推定に関する既存研究 [7, 8] を含め、幅広く用いられている。しかし、同じ部屋配置が1回しか現れない場合と2回以上現れる場合とを区別することができず、頂点・辺の数が小さい頻出部分グラフを属性とした場合、識別能力に関して問題があると考えられる。

(2) 極大独立集合に基づく支持度 [15]

第2の支持度 $S_{MIS}(p, g)$ は、頂点を共有せずに g 中に配置することのできる出現 o の最大数であり、以下の様に定義される。

$$S_{MIS}(p, g) = |MIS(p, g)|$$

ここで $MIS(p, g)$ は、 g における p の各出現 $o \in O(p, g)$ を頂点、(g における) 頂点を共有する出現同士を結んだものを辺とするグラフ、すなわち

$(O(p, g), \{(o_j, o_k) \mid o_j, o_k \in O(p, g), V_{o_j} \cap V_{o_k} \neq \emptyset\}, \lambda)$ の極大独立集合を表す。

(3) 最制限頂点に基づく支持度 [16]

第3の支持度 $S_{MRN}(p, g)$ は、 p 中の頂点 $v \in V_p$ の写像先集合 ($\varphi_o(v)$ の適用結果) に対する最小要素数であり、以下の様に定義される。

$$S_{MRN}(p, g) = \min_{v \in V_p} |\{\varphi_o(v) \mid o \in O(p, g)\}|$$

ここで $\varphi_o : V_p \rightarrow V_g$ は、出現 $o = (V_o, E_o, \lambda_o)$ に対し、条件

1. $\forall v \in V_p \Rightarrow \lambda_p(v) = \lambda_g(\varphi(v))$
2. $\forall (u, v) \in E_p \Rightarrow (\varphi(u), \varphi(v)) \in E_g$

を満たす頂点間の写像関数である。

$S_{MIS}(p, g)$ 同様、 $S_{MRN}(p, g)$ も出現の重複を考慮した支持度である。 $S_{MIS}(p, g)$ は、出現 o の単位での重複を許さないのに対し、 $S_{MRN}(p, g)$ は、部分グラフ p の頂点の単位で、重複を無視する形でカウントを行っている。両支持度は、出現数を考慮するという点で、出現数の違いを区別できない $S_{BIN}(p, g)$ の弱点を克服している。またそれぞれ支持度に関する逆単調性

$$\begin{aligned} \forall p \sqsubseteq q \rightarrow S_{MIS}(p, g) &\geq S_{MIS}(q, g) \\ \forall p \sqsubseteq q \rightarrow S_{MRN}(p, g) &\geq S_{MRN}(q, g) \end{aligned}$$

を満たすことが知られている。

その一方で、部屋配置という対象の性質を考えた場合、重複を許容しないことが必ずしも良いとは限らない。例えば、キッチンに1つの部屋だけが繋がっているグラフ g^1 と、キッチンに3つの部屋が繋がっているグラフ g^3 に対し、“キッチン-居室” という部分グラフ p の支持度はすべて $S_{MIS}(p, g^1) = S_{MIS}(p, g^3) = S_{MRN}(p, g^1) = S_{MRN}(p, g^3) = 1$ となり、支持度の差はない。しかし、 g^3 においてはキッチンへの導線が複数あり、また隣接した各部屋の利用方法にも複数の組み合わせが考えられることから、最も制約の強いキッチンだけを基準に支持度を考えることは必ずしも適切ではない。上記の議論を基に、本論文では、単一グラフにおける部分グラフの支持度として出現数に基づく支持度 $S_{NOO}(p, g)$ および頂点数の比に基づく支持度 $S_{ROV}(p, g)$ を提案する。

(4) 出現数に基づく支持度

本論文では、出現 $o \in O(p, g)$ の辺集合の種類数を支持度とすることを提案する。以下に提案する支持度 $S_{NOO}(p, g)$ の形式的な定義を示す。

$$S_{NOO}(p, g) = |\{E_o \mid (V_o, E_o, \lambda_o) \in O(p, g)\}|$$

$S_{MIS}(p, g)$ や $S_{MRN}(p, g)$ が、出現間で一部分でも重複を許容しないことに対し、 $S_{NOO}(p, g)$ では完全な重複以外を許容している。これにより、逆単調性 ($\forall p \sqsubseteq q \rightarrow S_{NOO}(p, g) \geq S_{NOO}(q, g)$) は成立しないが、出現の違いをより明確に表すことが可能となると考えられる。

(5) 頂点数の比に基づく支持度

支持度 $S_{NOO}(p, g)$ は出現の重複を考慮しないため、一ヶ所だけが異なるような出現を多数持つような大きな部分グラフに対して不当に大きな値を与えてしまう可能性がある。この問題を軽減するために、出現に含まれる総頂点数を部分グラフの頂点数で正規化した支持度 $S_{ROV}(p, g)$ を提案する。

$$S_{ROV}(p, g) = \frac{|\{v \in V_o \mid (V_o, E_o, \lambda_o) \in O(p, g)\}|}{|V_p|}$$

出現の重複に対して正規化の考え方を適用することで、重複を許容しない $S_{MIS}(p, g)$ や $S_{MRN}(p, g)$ と、重複を許す $S_{NOO}(p, g)$ との中間的な性質を持つ支持度が実現されることが期待できる。

5 評価実験

5.1 実験データと実験設定

実験には、株式会社 LIFULL が国立情報学研究所の協力により研究目的で提供している「LIFULL HOME'S データセット¹」を利用した。具体的には、間取り図を確認しにくい物件や二階建て構造を持つ物件等を除いた東京23区内の2部屋マンション400件および3部屋マンション561件の合計961件を選定している。

基本属性 X_{basic} として、データセット中で提供されている“徒歩距離”、“部屋面積”、“部屋階数”、“部屋の向き”、“部屋数”の5属性を採用した。各属性の統計量を表1に示す。なお“部屋の向き”は離散属性であり、その内訳は、東110件、西88件、南448件、北4件、南東111件、南西120件、北東15件、北西8件、不明57件である。

一方、物件 t の間取りに関するラベル付き無向グラフ $g(t)$ は、文献[17]で準備・使用されたデータを援用した。また、部屋配置属性 X_{graph} の抽出には、頻出部分グラフマイナー $gSpan[18]$ ² を利用した。具体的に

¹<https://www.nii.ac.jp/dsc/idr/lifull/homes.html>

²<https://www.cs.ucsb.edu/~xyan/software/gSpan.htm>

表 1: 基本属性の統計量

変数	賃料	徒歩距離	部屋面積	部屋階数
平均	232,936	691	87	4.8
標準偏差	269,021	404	42	4.7
最小値	55,000	80	31	0
中央値	158,000	640	77	3
最大値	2,050,000	2,720	3,56	43

表 2: 頻出部分グラフ数

サイズ	1	2	3	4	5
グラフ数	9	13	13	12	2

は、支持度パラメタを $\sigma = 450/961$ とし、49 の頻出部分グラフからなる集合 $X_{graph} = \mathcal{F}_D^{450/961}$ を導出した。サイズ (辺数) 別の頻出部分グラフ数を表 2 に示す。

5.2 推定精度

提案する二段階推定手法を評価するために、賃料推定モデルの構築実験を行った。具体的には、対象となる 961 件のデータ D を 861 件からなる訓練データ集合 D_{train} と 100 件からなるテストデータ集合 D_{test} に分割し、 D_{train} に線形重回帰モデルと XGBoost[19], Random Forest[20], LightGBM[21] をそれぞれ適用することで、賃料推定のための各モデル (f_{bgm} , $f_{bg/m}$, f_{bm} , $f_{b/m}$, f_g) を構築した。また推定精度の評価には、 D_{test} に対する平均絶対誤差 (Mean Average Error, MAE)

$$\frac{1}{|D_{test}|} \sum_{t \in D_{test}} |p(t) - \widehat{p}(t)| \quad (5)$$

を用いた。実験結果を表 3 に示す。

実験結果より、属性値として $S_{MIS}(p, g)$ を採用したモデル $f_{b/m+g}$ を XGBoost を用いて推定した場合に MAE = 3,768 となり、最も精度が高いことが分かる。また、推定手法別モデル毎の MAE 平均値の上位 3 件は、XGBoost を用いた $f_{b/m+g}$ (MAE 平均値 4,080), XGBoost を用いた f_{bm+g} (MAE 平均値 10,978), Random Forest を用いた f_{bm+g} (MAE 平均値 11,696) であった。

推定手法別にモデル f_{bgm} と f_{bm+g} , また $f_{bg/m}$ と $f_{b/m+g}$ の MAE 平均値をそれぞれ比較すると、線形重回帰モデルにおける $f_{bg/m}$ と $f_{b/m+g}$ 以外はすべて f_{bgm} より f_{bm+g} , また $f_{bg/m}$ より $f_{b/m+g}$ の方が精度が高いことが確認できる。モデル毎の MAE 平均値を算出すると f_{bgm} は 37,948, $f_{bg/m}$ は 40,166, f_{bm+g} は 26,162

, $f_{b/m+g}$ は 23,202 となり、 f_{bm+g} の MAE 平均値は f_{bgm} の約 69%, $f_{b/m+g}$ の MAE 平均値は $f_{bg/m}$ の約 58% となっていることが分かる。以上のことから、提案した二段階推定手法が有効に働いていることが確認できる。

次に、部屋配置属性に対する属性値の違いについて考察する。推定手法別支持度毎の MAE 平均値の上位 3 件は、 $S_{BIN}(p, g)$ を採用した XGBoost (MAE 平均値 20,590), $S_{MRN}(p, g)$ を採用した XGBoost (MAE 平均値 21,469), $S_{NOO}(p, g)$ を採用した XGBoost (MAE 平均値 21,977) であった。また、推定手法別モデル毎に MAE 値が最良となった回数を支持度毎に集計すると、 $S_{BIN}(p, g)$ は 3 回, $S_{MIS}(p, g)$ は 6 回, $S_{MRN}(p, g)$ は 2 回, $S_{NOO}(p, g)$ は 2 回, $S_{ROV}(p, g)$ は 3 回となり、差が小さいとは言え、 $S_{MIS}(p, g)$ が優れていることが示唆された。加えて、評価値が高かった XGBoost を用いたモデル $f_{b/m+g}$ に着目すると、精度が高い支持度は順に $S_{MIS}(p, g)$, $S_{NOO}(p, g)$, $S_{MRN}(p, g)$, $S_{ROV}(p, g)$, $S_{BIN}(p, g)$ となり、属性値として出現を考慮した支持度が有効に働いていることが確認できる。

その一方で、支持度毎の MAE 平均値を算出すると、 $S_{BIN}(p, g)$ は 31,797, $S_{MIS}(p, g)$ は 31,506, $S_{MRN}(p, g)$ は 31,579, $S_{NOO}(p, g)$ は 32,234, $S_{ROV}(p, g)$ は 32,232 となり、平均という観点からは大きな差は確認できなかった。このことは、推定手法別モデル毎の $S_{BIN}(p, g)$ に対する各支持度の勝率 ($S_{MIS}(p, g)$ は 0.50, $S_{MIS}(p, g)$ は 0.56, $S_{MIS}(p, g)$ は 0.44, $S_{MIS}(p, g)$ は 0.44) から確認ができる。以上の結果より、今回の実験では、属性値として出現数を考慮した支持度を採用することは一定の設定においては有効であるが、必ずしもすべての場合において推定精度の向上に寄与するとは限らないことが確認された。

5.3 利用する部分グラフの制限

利用する部分グラフの影響を考察するため、部屋配置属性をサイズ 2 以下の部分グラフに限定した上でモデルの推定を行った。各モデルに対する平均絶対誤差を表 4 に示す。

今回の場合の MAE 最良値は、XGBoost を用いて推定した属性 $S_{ROV}(p, g)$ を採用したモデル $f_{b/m+g}$ における 3,767 であり、すべての頻出部分グラフを利用した場合とほぼ同様の値が得られた。また、推定手法別モデル毎の MAE 平均値の上位 3 件もすべての頻出部分グラフを利用した場合と同じであり、XGBoost を用いた $f_{b/m+g}$ (MAE 平均値 3,967), XGBoost を用いた f_{bm+g} (MAE 平均値 10,994), Random Forest を用いた f_{bm+g} (MAE 平均値 11,757) であった。さらに、MAE 平均値を用いた二段階推定の効果について考

表 3: 実験結果：テストデータに対する平均絶対誤差（すべての部分グラフを利用した場合）

支持度 \ 関数	f_{bgm}	$f_{bg/m}$	f_{bm+g}	$f_{b/m+g}$	平均	f_{bgm}	$f_{bg/m}$	f_{bm+g}	$f_{b/m+g}$	平均
	Linear Regression					Random Forest				
$S_{BIN}(p, g)$	59,537	34,655	58,328	41,419	48,485	27,425	42,791	14,134	14,471	24,705
$S_{MIS}(p, g)$	59,894	35,936	53,382	44,831	48,511	26,189	44,585	10,497	12,074	23,336
$S_{MRN}(p, g)$	64,725	32,609	59,530	42,970	49,959	27,871	41,188	12,149	12,776	23,496
$S_{NOO}(p, g)$	55,978	35,076	58,101	46,906	49,015	33,075	46,919	11,064	12,523	25,895
$S_{ROV}(p, g)$	59,191	36,301	60,075	43,989	49,889	27,476	46,292	10,639	11,705	24,028
平均	59,865	34,915	57,883	44,023	49,172	28,407	44,355	11,696	12,710	24,292
	XGBoost					LightGBM				
$S_{BIN}(p, g)$	30,025	36,904	11,004	4,429	20,590	39,449	38,846	23,441	31,898	33,409
$S_{MIS}(p, g)$	33,568	43,632	10,590	3,768	22,890	30,461	38,706	23,947	32,033	31,287
$S_{MRN}(p, g)$	29,522	41,319	10,950	4,084	21,469	31,640	38,273	23,588	32,071	31,393
$S_{NOO}(p, g)$	28,276	44,402	11,214	4,016	21,977	31,540	39,054	25,454	32,143	32,048
$S_{ROV}(p, g)$	32,426	49,212	11,133	4,104	24,219	30,686	36,630	24,019	31,829	30,791
平均	30,763	43,094	10,978	4,080	22,229	32,755	38,302	24,090	31,995	31,786

察すると、線形重回帰モデルにおける $f_{bg/m}$ と $f_{b/m+g}$ 以外はすべて二段階推定の方が精度が高いことが確認できる。

属性値の違いに関しても、すべての頻出部分グラフを利用した場合と同様の傾向が確認できる。推定手法別支持度毎の MAE 平均の上位三件は、すべての頻出部分グラフを利用した場合と一致し、 $S_{BIN}(p, g)$ を採用した XGBoost (MAE 平均値 19,401), $S_{MRN}(p, g)$ を採用した XGBoost (MAE 平均値 20,939), $S_{NOO}(p, g)$ を採用した XGBoost (MAE 平均値 21,251) であった。また、精度が最も高い XGBoost を用いたモデル $f_{b/m+g}$ においては $S_{BIN}(p, g)$ の精度が最も低く、 $S_{MRN}(p, g)$, $S_{NOO}(p, g)$, $S_{MIS}(p, g)$, $S_{ROV}(p, g)$ の順に精度が良くなっていることが確認できる。一方で、推定手法別モデル毎に MAE 値が最良となった回数を支持度毎に集計すると、 $S_{BIN}(p, g)$ は 1 回、 $S_{MIS}(p, g)$ は 4 回、 $S_{MRN}(p, g)$ は 5 回、 $S_{NOO}(p, g)$ は 2 回、 $S_{ROV}(p, g)$ は 4 回となり、出現を考慮した支持度がより有効に働いていることが確認できる。

5.4 影響力の大きな部分グラフの抽出

賃料に大きな影響を与える部屋配置やその組み合わせを特定するために、推定したモデル f_g に対して、樹状モデルアンサンブルから近似ルール集合を抽出する手法である defragTrees[22]³ を適用した。モデル f_g は部屋配置属性のみを用いた残差調整モデルであり、予測値の大きいルールに現れる部屋配置の組み合わせが賃料に

与える影響が大きいと考えられる。 f_g から得られる近似ルール集合において、予測値が最大であるルールを図 1 と図 2 に示す。なお図 1 は、頻出部分グラフのサイズ上限を 2、支持度 $S_{MIS}(p, g)$ を採用した XGBoost によるモデル f_{bm+g} に対する結果である。一方図 2 は、同じく頻出部分グラフのサイズ上限を 2、支持度 $S_{ROV}(p, g)$ を採用した XGBoost によるモデル $f_{b/m+g}$ に対する結果である。

図 1 と図 2 より、各部屋配置の支持度に関して上限と下限が設定されていることが分かる。また図 2 における最後の条件 [(収納) $\xleftrightarrow{\text{収納}}$ (玄関)] < 1.00 は、玄関に収納がないことを表している。これらのことより、特定の部屋配置が多ければ多いほど良いというわけではなく、適当な量が存在することが伺える。

図 1 には、“収納と窓のある居室”や“収納がある廊下”など、収納に関する条件が多く、収納が賃料を決定する一つの要因となっていることが考えられる。このことは図 2 にも当てはまる。さらに図 2 では、“居室”を含む部屋配置グラフが数多く含まれていることが分かる。両ルールにおいて共通する部屋配置は

$$\begin{aligned}
 & (\text{窓}) \xleftrightarrow{\text{ガラス}} (\text{ベランダ}) \xleftrightarrow{\text{ガラス}} (\text{窓}), (\text{収納}) \xleftrightarrow{\text{収納}} (\text{廊下}), \\
 & (\text{収納}) \xleftrightarrow{\text{収納}} (\text{居室}) \xleftrightarrow{\text{ガラス}} (\text{窓}), (\text{居室}) \xleftrightarrow{\text{ガラス}} (\text{窓}), \\
 & (\text{収納}) \xleftrightarrow{\text{収納}} (\text{居室}) \xleftrightarrow{\text{ドア}} (\text{廊下})
 \end{aligned}$$

の 5 つであり、この結果からも、“居室”と“収納”が中心的な役割を果たしていることが示唆される。

³<https://github.com/sato9hara/defragTrees>

表 4: 実験結果：テストデータに対する平均絶対誤差（サイズ 2 以下の部分グラフのみを利用した場合）

支持度 \ 関数	f_{bgm}	$f_{bg/m}$	f_{bm+g}	$f_{b/m+g}$	平均	f_{bgm}	$f_{bg/m}$	f_{bm+g}	$f_{b/m+g}$	平均
	Linear Regression					Random Forest				
$S_{BIN}(p, g)$	59,425	33,785	58,155	38,297	47,415	29,061	43,169	16,310	14,081	25,655
$S_{MIS}(p, g)$	55,475	32,955	60,620	42,233	47,821	25,551	45,511	10,173	12,467	23,425
$S_{MRN}(p, g)$	60,391	31,764	56,938	39,097	47,048	33,045	38,541	12,162	13,349	24,274
$S_{NOO}(p, g)$	58,077	34,964	58,326	37,405	47,193	25,290	44,376	10,443	12,479	23,147
$S_{ROV}(p, g)$	60,115	33,547	58,401	41,122	48,296	26,990	44,927	9,699	12,333	23,487
平均	58,697	33,403	58,488	39,631	47,555	27,987	43,305	11,757	12,942	23,998
	XGBoost					LightGBM				
$S_{BIN}(p, g)$	22,658	39,505	11,125	4,316	19,401	32,777	38,584	23,862	31,794	31,754
$S_{MIS}(p, g)$	30,024	41,988	10,547	3,793	21,588	29,657	37,206	24,506	31,912	30,820
$S_{MRN}(p, g)$	29,271	39,227	11,231	4,025	20,939	31,826	38,399	24,032	31,610	31,467
$S_{NOO}(p, g)$	29,256	40,669	11,099	3,980	21,251	31,521	37,535	27,113	31,752	31,980
$S_{ROV}(p, g)$	30,075	46,901	10,969	3,767	22,928	31,269	38,129	23,772	31,897	31,267
平均	28,257	41,658	10,994	3,976	21,221	31,410	37,971	24,657	31,793	31,458

$$\begin{aligned}
 +8,828 &\Leftarrow [(窓) \xleftrightarrow{ガラス} (ベランダ) \xleftrightarrow{ガラス} (窓)] \geq 1 \\
 &\wedge [(居室) \xleftrightarrow{ガラス} (窓)] \geq 1 \\
 &\wedge [(収納) \xleftrightarrow{収納} (居室) \xleftrightarrow{ガラス} (窓)] \geq 1 \\
 &\wedge [(収納) \xleftrightarrow{収納} (居室) \xleftrightarrow{ドア} (廊下)] < 4 \\
 &\wedge [(収納) \xleftrightarrow{収納} (廊下) \xleftrightarrow{ドア} (水回り)] < 3 \\
 &\wedge [(収納) \xleftrightarrow{収納} (廊下)] < 3
 \end{aligned}$$

図 1: モデル f_{bm+g} から得られたルール例

$$\begin{aligned}
 +18,902 &\Leftarrow [(窓) \xleftrightarrow{ガラス} (ベランダ)] < 2.67 \\
 &\wedge [(窓) \xleftrightarrow{ガラス} (ベランダ) \xleftrightarrow{ガラス} (窓)] < 1.46 \\
 &\wedge [(窓) \xleftrightarrow{ガラス} (居室)] \geq 1.86 \\
 &\wedge [(水回り) \xleftrightarrow{ドア} (居室)] < 1.71 \\
 &\wedge [(収納) \xleftrightarrow{収納} (居室)] \geq 1.79 \\
 &\wedge [(収納) \xleftrightarrow{収納} (居室) \xleftrightarrow{ドア} (廊下)] \geq 1.24 \\
 &\wedge [(収納) \xleftrightarrow{収納} (居室) \xleftrightarrow{ガラス} (窓)] \geq 1.89 \\
 &\wedge [(収納) \xleftrightarrow{収納} (玄関)] < 1.00
 \end{aligned}$$

図 2: モデル $f_{b/m+g}$ から得られたルール例

6 まとめ

本論文では、賃貸物件の賃料に大きな影響を与える部屋配置を特定することを目的に、単一グラフにおける部分グラフの支持度に関する新たな定義を与えると共に、基本的な属性から得られる賃料推定結果を間取り情報に基づき補正する二段階賃料推定手法を提案した。また実験により、提案手法が推定精度の向上に有効であることを確認すると共に、初期的な結果ではあるが、注目すべき部屋配置の特定を行った。

今後の課題としては、多様な部屋配置を考慮するためにより低頻度な頻出部分グラフを利用することがあげられる。加えて、より効果的に賃料への影響が強い部屋配置を特定するため、傾向スコア分析 [23] や対比集合抽出 [24]、樹状モデルアンサンブル分析手法 [25] 等を併用することを予定している。

謝辞： 本研究では、株式会社 LIFULL が国立情報学研究所の協力により研究目的で提供している「LIFULL

HOME'S データセット」を利用した。本研究の一部は、JSPS 科研費 17K00315 の助成を受けたものです。

参考文献

- [1] 阿部成治, 石崎幸司：首都圏における民間賃貸住宅家賃の重回帰分析, 都市住宅学 (19), pp.39-44, 1997.
- [2] 金春愛, 黄嘉平, 住田潮, 盧韶南：マクロ・ミクロ統合に基づく不動産賃料推定モデルの開発, 筑波大学社会学工学コモンズ Discussion Paper Series, No.1182, 2007.
- [3] 山鹿久木, 中川雅之, 齊藤誠：地震危険度と家賃-耐震対策のための政策的インプリケーション, 日本経済研究 (46), pp.1-21, 2002.

- [4] 瀧澤重志, 材木敦史, 加藤直樹, 具源龍: 新橋に立地するオフィスビルの感性評価を考慮した賃料分析, 日本建築学会計画系論文集 (627), pp.1053-1059, 2008.
- [5] A. J. Smola and B. Schoelkopf: A tutorial on support vector regression, NeuroCOLT2 Technical Report NC2-TR-1998-030, 1998.
- [6] 花里俊廣, 平野雄介, 佐々木誠: 首都圏で供給される民間分譲マンション100m²超住戸の隣接グラフによる分析, 日本建築学会計画系論文集 (591), pp.9-16, 2005.
- [7] 瀧澤重志, 吉田一馬, 加藤直樹: グラフマイニングを用いた室配置を考慮した賃料分析: 京都市郊外の3LDKを中心とした賃貸マンションを対象として, 日本建築学会環境系論文集 (623), pp.139-146, 2008.
- [8] 尾崎 知伸, 小黒 淳斗: 頻出部分グラフを用いた賃料分析, 人工知能学会第111回知識ベースシステム研究会, pp.13-16, 2017.
- [9] D. J. Cook and L. B. Holder eds.: *Mining Graph Data*, Wiley-Interscience, 2005.
- [10] T. Washio, J. N. Kok and L. DeRaedt eds.: *Advances in Mining Graphs, Trees and Sequences*, IOS Press, 2006.
- [11] Z. Zeng, J. Wang, J. Zhang, and L. Zhou: FOGGER: An Algorithm for Graph Generator Discovery, *Proc. of 12th International Conference on Extending Database Technology*, pp.517-528, 2009.
- [12] L. Breiman, J.H. Friedman, R.A. Olshen and C.J. Stone: *Classification and regression trees*, Wadsworth, 1984.
- [13] J. R. Quinlan: Learning with Continuous Classes, *Proc. of the 5th Australian Joint Conference on Artificial Intelligence*, pp.343-348, 1992.
- [14] Y. Wang and I. H. Witten: Induction of model trees for predicting continuous classes, *Proc. of Poster papers of the 9th European Conference on Machine Learning*, pp.128-137, 1997.
- [15] M. Kuramochi and G. Karypis: Finding frequent patterns in a large sparse graph, *Data Mining and Knowledge Discovery*, Vol.11, No.3, pp.243-271, 2005.
- [16] B. Bringmann and S. Nijssen: What is Frequent in a Single Graph? *Proc. of the 5th International Workshop on Mining and Learning with Graphs*, pp.1-4, 2007.
- [17] K. Kamihori, T. Shimano, A. Oguro and T. Ozaki: Towards discovery of human's recognition mechanisms for complex structured images, *Proc. of the Third International Workshop on Skill Science*, pp.26-27, 2016.
- [18] X. Yan and J. Han: gSpan: Graph-based Substructure Pattern Mining, *Proc. of the 2002 IEEE International Conference on Data Mining*, pp.721-724, 2002.
- [19] T. Chen and C. Guestrin: XGBoost: A Scalable Tree Boosting System, *Proc. of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp.785-794, 2016.
- [20] L. Breiman: Random Forests, *Machine Learning*, Vol.45, Issue 1, pp.5-32, 2001.
- [21] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, and T.-Y. Liu: LightGBM: A Highly Efficient Gradient Boosting Decision Tree, *Advances in Neural Information Processing Systems* 30, 2017.
- [22] S. Hara and K. Hayashi: Making Tree Ensembles Interpretable: A Bayesian Model Selection Approach *Proc. of the 21st International Conference on Artificial Intelligence and Statistics*, pp.77-85, 2018.
- [23] 星野崇宏: 調査観察データの統計科学—因果推論・選択バイアス・データ融合, 岩波書店, 2009.
- [24] S. D. Bay and M. J. Pazzani: Detecting Change in Categorical Data: Mining Contrast Sets, *Proc. of the 5th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp.302-306, 1999.
- [25] H. Deng: Interpreting Tree Ensembles with in-Trees, *International Journal of Data Science and Analytics*, 2018.

家庭内行動の深掘りを目的としたIoTセンサによる 行動観察手法の提案

In-home Behavioral Observation Method Employing IoT Sensors

服部 俊一^{1*} 三浦 輝久¹ 市川 玲子² 澤井 大樹²
Shunichi Hattori¹, Teruhisa Miura¹, Reiko Ichikawa², Daiki Sawai²

¹ 一般財団法人電力中央研究所

¹Central Research Institute of Electric Power Industry

² 株式会社イデアラボ

²IdeaLab Co., Ltd.

Abstract: This paper proposes in-home behavioral observation method employing IoT sensors for alleviating the burden of users. Proposed method consists of three components: in-home sensing with IoT sensors, visualization of sensor data, and depth-interview. This method is applied to the investigation of barriers regarding energy saving activities in home for improving behavioral intervention methods. The results show in-home activities targeting twenty households are successfully observed and various barriers are extracted and organized.

1 はじめに

現在、日本ではスマートメータの設置が全国で進められており、一般家庭における電力消費データの簡易な収集・活用が可能となる環境が整備されている。また、2016年4月の電力自由化以降、電力会社の販売部門では顧客満足度向上を目的とした価値創出が求められていることから、電力各社も既にスマートメータデータの見える化や家族の見守りなどを目的としたサービスの提供を始めている。

一方で、価値創出に資するサービス開発には家庭によってそれぞれ異なる制約や潜在的ニーズなど、多様なユーザについての理解が必要である。新サービス開発の方法論としてデザイン思考 [1] やユーザ中心設計 [2] が注目されており、中でもユーザ理解の手段として行動観察が効果的とされている。行動観察は、無意識の行動など本人に聞いても分からない事実や、当たり前すぎて見過ごされてきた事実を発見する質的調査手法であり、サービス利用実態やユーザが抱える制約を解明する手法として活用されている [1, 3, 4]。一般的な行動観察では、調査員がユーザの行動を傍で観察し、その記録を用いたデプスインタビューを実施してユーザが持つ課題や制約を聴取する。その他、ビデオカメラによる撮影で観察を行う場合もある [5]。

このような行動観察を家庭内において実施する場合、

従来の訪問や撮影による観察は以下に述べる2つの問題を抱える。1つ目の問題として、ユーザへ与える負担や影響の大きさが挙げられる。調査員の訪問やビデオカメラの設置はユーザのプライバシーを侵害するため、被験者への心理的負担が大きく長期間の観察は難しい。また、訪問や撮影による直接的な観察によって、ユーザの行動が大きく影響を受ける可能性もある。2つ目の問題として、観察にかかるコストの大きさが挙げられる。調査員の訪問や撮影記録の解析作業には多大なコストを要するため、多数の家庭を観察対象とした網羅的な調査は現実的でない。

一方、日本では前述のようにスマートメータの設置が進められており、一般家庭における電力消費データを簡易に収集・活用するための環境整備が進められている。スマートメータに加えて、「モノのインターネット (IoT: Internet of Things)」という言葉の流行が示すように、家庭内の空気・温熱環境などを計測するための安価なセンサ製品も普及しつつある。家庭内の電力消費や温熱環境データからは、冷暖房機器をはじめとする家電利用情報とそれらに付随する様々な行動が読み取れる。従来の訪問や撮影による行動観察をIoTセンサを用いて代替すれば、長期間かつ多数の家庭を対象とした網羅的な調査が実現できる。

そこで、本稿ではIoTセンサを活用した家庭内の行動観察手法を提案する。提案手法は、IoTセンサによる家庭内の行動計測、計測結果の可視化、可視化結果を用いたデプスインタビューから構成される。IoTセン

*連絡先：(一財)電力中央研究所
〒240-0196 神奈川県横浜市長坂 2-6-1
E-mail: shattori@criepi.denken.or.jp

サの活用によってプライバシー侵害や家庭内行動への影響を低減し、長期間の観察が行える。また、IoT センサによる計測は調査員の訪問や撮影と比較して安価に行えることから、多数の家庭を対象とした網羅的な調査が実現できる。従来手法のように目視情報を伴わない点については、計測データの可視化結果を用いたデプスインタビューで補うことにより、直接の観察に近い事実が発見できる。本稿では、構築した手法を家庭の省エネ阻害要因抽出という目的に対して適用し、抽出された要因の整理結果や観察された行動、および提案手法の特長について述べる。

2 家庭内行動観察の課題

行動観察とはユーザの生活行動やサービス利用実態を把握するための質的調査手法である。行動観察以外の、アンケートやインタビューによる聞き取り調査もユーザ理解の手段として用いられるが、ユーザが無自覚・無意識に行っている動きや言語化が困難な行動を聞き出すことは難しい。行動観察はそのような行動実態や、その背景にある潜在的なニーズを深掘りするための手法として広く活用されている [1, 3, 4]。

しかし、行動観察が実施されるのは商業施設や公共空間が多く、家庭内行動を対象とした観察には以下に述べる困難を伴う。従来の調査員の訪問や撮影による観察は人間の認知能力を活かした詳細な記録が可能である一方で、直接的な観察によってプライバシー侵害の懸念が生じ、また被観察者への強い介入効果によって行動を変えてしまうおそれがある。加えて、長期間かつ多数の家庭を対象とした網羅的な調査は調査員や映像解析にかかるコストも課題となる。

上記の問題に対して、行動ログやサービス利用の記録から間接的に行動を観察するというアプローチが考えられる。このような間接的な行動観察の事例として以下の例がある [6]。フランスでは国民の納税申告手段として窓口や電話など4種類のチャネルを用意していたが、Webサイトの利用率が低かった。そこで、窓口の対応記録とWebサイトのログを分析した。その結果に基づき、行動経済学の手法を用いてWebサイトを改善して利用率を上げることに成功した。

この事例のようにユーザの行動ログを収集し分析できれば、多数のユーザを対象としながらも観察による影響を抑えた調査が実現できる。しかし、観察は間接的なものに留まることから、行動ログのみを用いてユーザの行動実態を詳細に把握することは難しい。加えて、従来はこのような行動ログを家庭内において収集することは技術・費用の両面において困難であるという問題も存在した。これに対し、現在ではスマートメータの設置が国内で進んでいることと、IoT センサの低価

格化によって、家庭内の行動ログ収集が安価に行えるようになった。これらの行動ログを家電利用行動や日頃の習慣がわかるように可視化し、それを用いたデプスインタビューを行うことで、訪問や撮影による従来の行動観察をIoT センサを用いて代替できると考える。

3 IoT センサを用いた行動観察手法

提案手法はIoT センサによる家庭内の行動計測、計測結果の可視化、可視化結果を用いたデプスインタビューから構成される。それぞれについて3.1, 3.2, 3.3節で述べる。

3.1 IoT センサによる家庭内行動計測

IoT センサを用いた家庭内計測は、従来の訪問や撮影による記録を代替し、多数の家庭において長期間の調査を実現するものである。そのために提案手法で用いるIoT センサキットが満たすべき要件について述べる。提案手法では、IoT センサで計測されたデータを可視化し、それを用いたデプスインタビューによって直接の観察に近い事実の発見を目的とする。そのための行動観察に用いるIoT センサキットは、以下2つの要件を満たす必要がある。

1. 長期間の計測が行える
2. 低コストで必要最低限の計測が行える

上記1については、家庭内の制約やユーザが抱えるニーズは日頃の習慣行動や室内状態との関連が強いことによる。短期間の観察では習慣的な行動の把握は困難であり、観察できる行動も調査期間中に行われたものに限定される。

上記2は、低コストで長期間かつ多数の家庭を対象とした観察が可能であるという提案手法の長所を活かすため、目的を達成し得る最小限のセンサ構成が適していることによる。計測に用いるセンサの種類を増やすことで詳細な行動計測が可能となる一方で、撮影による従来手法と同様に観察や分析にかかるコストの増大やプライバシー侵害の問題が生じる。加えて、欠測や故障など計測に関するトラブルが発生する可能性も増大する。

これらの要件を満たしうるセンサおよびゲートウェイ端末として、NextDriveCube¹やOpenBlocks²など既にくいつかの製品が発売されている。そのほか、Raspberry Piなどの小型コンピュータを用いて独自のセンサキットを開発することもできる。製品ごとに対応し

¹<https://jp.nextdrive.io/cube/> (2019/2/7 アクセス)

²<https://openblocks.plathome.co.jp/> (2019/2/7 アクセス)

ているセンサや価格は大きく異なるため、上記の観点
を踏まえて行動観察を実施する目的に応じた製品の選
定やセンサキット開発が必要である。

3.2 計測データの可視化

計測されたデータを本節で述べる方針にしたがって
可視化し、デプスインタビュー時の資料および効率的
な聴取を行うための要点整理に用いる。家庭内の行動
観察にあたって、求められる可視化要件は以下に示す
3点と考える。

1. 大まかな生活行動がわかる
2. 行動に関連する家電利用がわかる
3. 生活習慣とその変化がわかる

上記要件に基づき、センサデータを可視化した例を
図1および2に示す。図1は、ある家庭の1日の電力
消費量とリビングの気温を折れ線グラフ形式で表した
ものである。この例から、この家庭では7時前に電力
消費量と気温が同時に上昇して8時からどちらも低下
し、その後19時頃に再度上昇している様子がわかる。
ここから、7時頃に起床して8時に外出し、19時頃に
帰宅という大まかな生活行動（上記の可視化要件1）が
推測できる。本データを計測したのは12月であること
と、起床前のリビング気温が13度前後でありその後上
昇することから、エアコンなど暖房器具を利用してい
ると推定できる（可視化要件2）。加えて、短期間に電
力消費が2000W前後まで増加する瞬間が複数回観測さ
れている。これは電子レンジや電気ケトルなど、調理
家電によるものと推定できる。IHなど保有家電の情報
が得られれば、より細かな推定も可能となる。

図2は、ある家庭における16日間の電力消費量と気
温をヒートマップ形式で可視化したものである。ヒート
マップ形式によって長期間の計測データが概観でき
るため、データの傾向や変化から以下に挙げるような居
住者の生活習慣や通常と異なる行動を記録できる（可
視化要件3）。

1. 朝6時から6時半に起床し、7時半頃外出する日
が多い
2. 朝は電力消費量が小さく、温度変化も緩やかであ
ることからエアコンによる冷房を使うことは少
ない
3. 2,3日間隔で午前中も在宅の日があり、その場合
はエアコンを利用することがある
4. 13時前後に帰宅し、その際にエアコンを利用し
始めることが多い
5. ほぼ毎日、23時頃には就寝する
6. 23時過ぎから翌日の起床まで、間欠的な電力消
費が発生している

7. 数日おきに、午前2時～6時前後まで電力消費が
発生している（8月7日など、その一部はリビング
グのエアコンと推測される）

上記6については、電力の波形からエアコンの利用
である可能性が高いものの、日中に計測されるエア
コン利用時ほどの消費量がないこと、またリビングの温
度変化が緩やかであることから断定はできない。また、
上記7のように普段の生活習慣と異なる行動の場合、行
動内容の特定やその行動に至った理由までは観察でき
ない。これらのように計測・可視化だけでは不十分な
行動については、次節で述べるデプスインタビューに
おいて聴取できることから、この時点では聴取すべき
箇所として記録するに留める。

なお、図1および2では電力消費量と室内気温のみ
を計測した例を示したが、それ以外にも二酸化炭素濃
度や騒音などを計測することでより多彩かつ詳細な行
動を記録できる可能性がある[7]。ただし、センサを増
やすことで欠測など運用上のトラブルや費用が増大す
るため、目的に応じた最小限の構成が望ましい。

3.3 可視化結果を用いたデプスインタビュー

計測データの可視化結果を用いて、行動の特定や意
図の聴取を目的としたデプスインタビューを実施する。

従来の訪問や撮影による手法と、IoTセンサを活用
する提案手法を比較したものが表1である。従来手法
においてもデプスインタビューを組み合わせた事例は
多く、目視情報に基づく聴取を行うことで家庭内の行
動やその意図が抽出できる。しかし、記録や分析にか
かるコストが高いことと、プライバシーや行動への影
響の大きさといった観点から、観察範囲は少数のユー
ザや家庭を対象とした短期間の行動に留まる。一方で、
表1の項目aおよびbに示したように、提案手法では
IoTセンサを用いた計測でその記録を代替することに
よって記録・分析コストおよび行動への影響を低減し、
多数の家庭を対象とした長期間の観察が実現できる。

項目c「収集できる情報」という観点では、従来手法
は目視情報を得られる点において優位性がある一方、イ
ンタビューで聴取する要点の整理や聴取によって得られ
る情報は調査員のスキルに依存する。提案手法で得られ
る情報は計測データであり従来の目視情報とは性質
が異なるが、上述のように長期間の観察が可能である
だけでなく、計測データの可視化によりインタビューで
聴取すべき要点を簡易に整理できるという利点を持つ。

以上に述べた観点から、IoTセンサによる計測と可
視化、デプスインタビューを組み合わせることが網羅
的な調査を目的とした家庭内の行動観察に効果的と考
える。計測データに反映される行動は3.2節で示した
ように多岐にわたっており、デプスインタビューを組

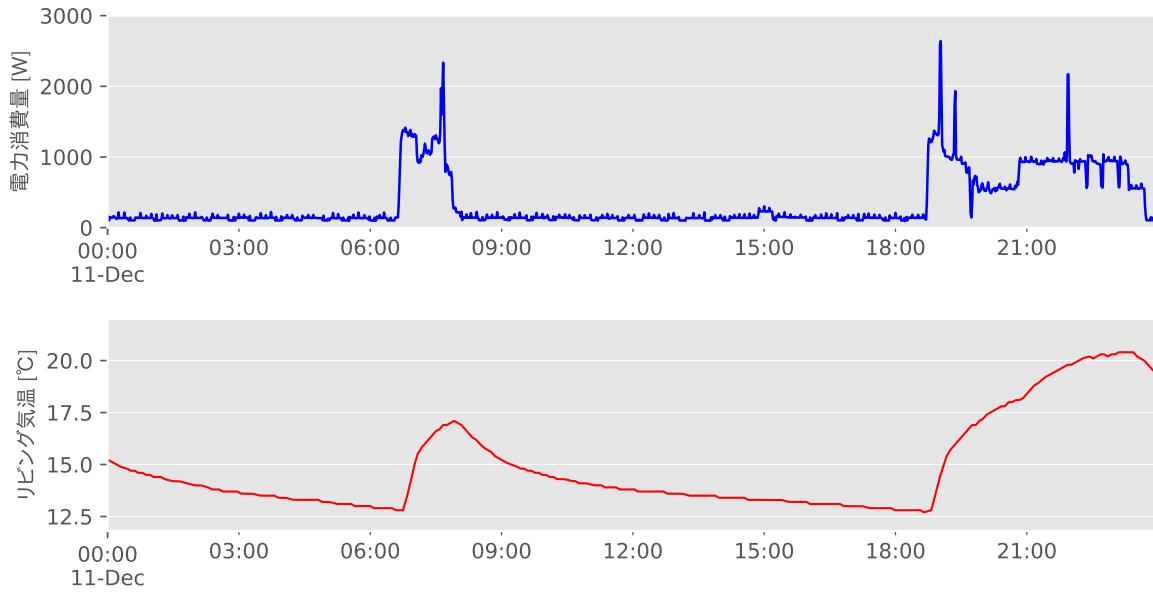


図 1: センサデータ可視化の例 (折れ線グラフ形式)

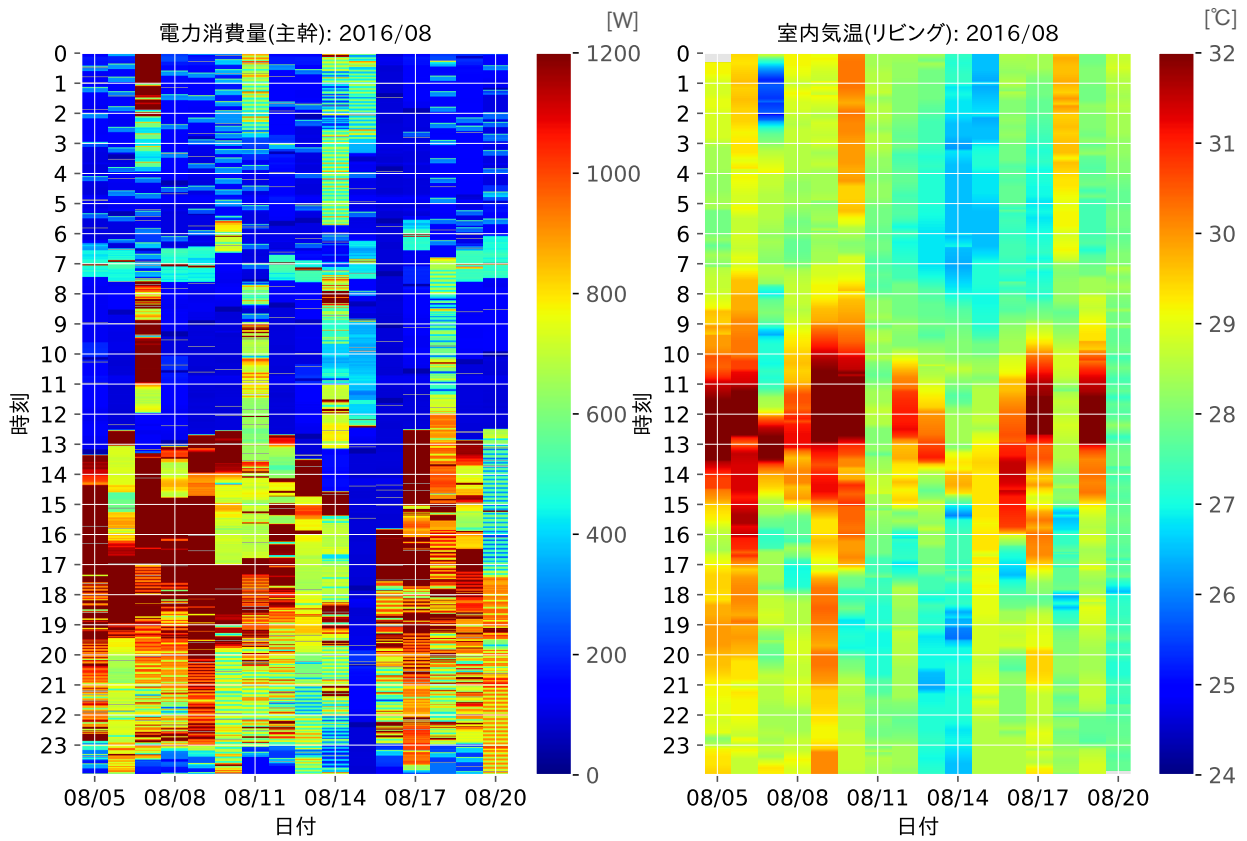


図 2: センサデータ可視化の例 (ヒートマップ形式)

表 1: 従来の家庭内行動観察手法と提案手法の比較

比較項目	従来手法			提案手法		
	1. 訪問/撮影	2. 分析 (要点整理)	3. デプス インタビュー	1. センサ計測	2. 分析・可視化 (要点整理)	3. デプス インタビュー
a. 観察コスト (計測/分析/ インタビュー実施)	× 高コスト (対象家庭は限定的)	△ 映像分析時は 高コスト	△ 調査員の スキルに依存	○ 低コスト (多数の家庭を対 象可)	○ 自動化が容易 なため低コス ト	○ 効率的な聴取 が可能
b. 被観察者への影響 (プライバシー/ 行動への影響)	× 負荷が大きい (長期の観察は困難)	× プライバシー への影響大	△ プライバシー への影響あり	○ 負荷が小さい (長期の観察が可 能)	△ プライバシー への影響小	△ プライバシー への影響あり
c. 抽出できる情報	○ 目視できる行動 の抽出と意図の 推定が可能	△ 調査員・分析 官のスキルに 依存	調査員の認知能 力や主観に基づ く情報	△ 計測可能な行 動のみで意図 推定は困難	○ 自動化により スキルに依存 しない	長期間・多数の 家庭に関する網 羅的な情報



図 3: おうちモニタキット (OMK)

み合わせることで多様な行動やその意図を抽出できる。したがって、本手法によって実現可能な長期間かつ多数の家庭を対象とした行動観察は、家庭内の省エネ行動など日々の習慣とその変化、それに付随する潜在的な課題の調査に適していると言える。

4 IoT センサを用いた行動観察実験

3 節で述べた提案手法を用いて家庭内の行動観察を実施した。本節で述べる実験は家庭向けの省エネサービス考案のための省エネ阻害要因抽出を目的とした。家庭の電力消費傾向に基づく省エネアドバイス [8] など、パーソナライズされた情報提供によって家庭の省エネを促す手法が注目されている。一方で、省エネを阻害する要因は家庭それぞれの習慣行動や環境によって異なることから、これらの要因は十分に考慮されておらず、情報提供が効果的でない場合もある。より効果の高い情報提供の実現には、それぞれの家庭が持つ制約を解明し、それを踏まえたパーソナライズが必要である。本実験はそのための省エネ阻害要因の網羅的な調査を目的としている。

4.1 実施概要

センサ計測には「おうちモニタキット (OMK) [7]」を用いた (図 3)。OMK は多数のセンサに対応しているが、本稿では計測データを電力消費量と室内気温の 2 つに限定した。使用するセンサの種類を増やした場合、コストの増大や故障など設置時・設置後のトラブル増加が懸念される。そのため、今回は省エネ阻害要因抽出という目的を達成し得る最小限の構成とした。

図 3 に示したように、OMK では計測されたセンサデータをリアルタイムに本体ディスプレイ上で表示できる。このようなディスプレイを被験者へ見せることは一種の情報提供に該当し、被験者の行動に影響を及ぼす可能性がある。一方で、情報提供に対する被験者からの反応をデプスインタビューで聴取できれば、省エネ阻害要因抽出という目的に対して有用な意見を得ることができる。デプスインタビューでは計測データの可視化結果を用いることから、データに対するユーザの理解が深まり、より詳細な聴取が行える可能性もある。加えて、何を計測されているかがユーザへ明確に伝わることから、透明性・納得性を確保できるといった利点も期待できる。本実験では目的に対して上述のメリットが大きいと判断し、ディスプレイをそのまま用いることとした。

本実験では環境意識の高い被験者をリードユーザ [9] として募集し、そこから年収幅や家族構成・住居形態・職業といった属性が均等になるよう 20 名を選定した。デプスインタビューは以下の流れに沿って、1 回につき 1 名ずつ、およそ 2 時間実施した。

1. 調査目的、内容、所要時間などを明示し、同意が得られれば調査参加同意書に署名してもらう。
2. 事前課題である「被験者宅の間取り図」と「使用家電型番聞き取り票」を受け取り、記載内容について聴取する。
3. 被験者が「省エネ行動チェックシート」に記入する。回答内容を確認しながら、普段の省エネ行動

の内容や意図を聴取する。

4. OMK の計測データを示しながら、家庭内行動習慣や省エネ阻害要因について聴取する。

上記のうち「被験者宅の間取り図」では、大まかな住宅の間取りと主要家電の設置位置を記入するとともに、OMK の設置箇所を記入してもらった。「使用家電型番聞き取り票」には、所有しているエアコンおよび冷蔵庫について、年式と大きさを記入してもらった。これらの家電は年式・大きさによって消費電力が大きく異なることから、電気利用実態に関する聴取を効率的に行うために用意した。「省エネ行動チェックシート」には、東京都による先行調査 [10] などから、一般的な省エネ行動としてエアコン 3 項目、冷蔵庫 6 項目、キッチン・給湯 7 項目、照明 3 項目、テレビ 4 項目の合計 23 項目を選んで記載した。

4.2 省エネ阻害要因の抽出・整理

計測データの可視化結果を用いてデプスインタビューを実施した結果、様々な習慣行動や、習慣とは異なる行動が抽出できた。その中から省エネに関するものを絞り込み、阻害要因として整理したものが表 2 である。この表では省エネ阻害要因を 10 カテゴリにまとめており、カテゴリごとに省エネ行動が行えない理由としてインタビューにおける被験者の発言を記載した。

実施結果から、同じ家電の利用においても阻害要因は被験者によって異なることが明らかとなった。例として、エアコン利用に関して表 2 のカテゴリ 3 では「寒いのが苦手 (ID1)」と快適性が阻害要因となっているのに対して、カテゴリ 6 では「犬を飼っている (ID17)」とペットへの配慮が省エネを阻む要因となっていることが挙げられる。照明についてもカテゴリ 1 の節約に関する要因やカテゴリ 9 の防犯に関する要因など、同様の家電に対して複数の要因が存在する。

加えて、同じ発言でもその背後に存在する認知プロセスが異なっていることも明らかとなった。デプスインタビューでの発言記録から、省エネができない理由として「面倒くさい」「億劫」といった発言は多くの被験者からなされているが、その理由は様々である。ID10 はトイレの温水洗浄便座の電源が 1 日中入っていることに対して「いちいちやるのが面倒くさくて」と発言しており、文字通り面倒という理由で実行していない。一方で、ID4 は「給湯器の電源を切る」という省エネ行動を当初は実行していたが、使用時に電源の入れ忘れによりお湯が出ないという状況を何度か経験した結果、億劫と感じて切らなくなってしまったと語っている。これらの結果から、家庭内において様々な省エネ阻害要因が存在し、その結果として省エネを行えていないという実態が明らかとなった。

図 4 は、ID1 の被験者に対して日々の習慣行動の抽出を試みた例である³。この被験者の家庭では午前 3 時からほぼ毎日エアコンによる暖房利用が観察されており、その行動は図 4 左側のヒートマップ形式で可視化されたデータから電力消費量と室内気温の上昇という傾向により特定できる。この行動の意図についてインタビューで聴取したところ、飼っている猫が 3 時過ぎに起きるため暖房をタイマーで入れるように設定していることがわかった。長期間の行動観察により抽出できた行動習慣と言える。

デプスインタビューでの発言記録から抽出された家庭内行動の例を、表 3 に示す。これらの行動は日常的な家電利用と密接に関係するものであり、ここから家庭特有の制約や被験者の認知・考え方を深掘りできる。

4.3 考察

本節で述べた実験では、IoT センサによる行動計測、計測結果の可視化、それを用いたデプスインタビューを組み合わせることで、多数の家庭を対象とした長期間の行動観察が実現できることを示した。IoT センサを用いた行動観察は従来手法と比較してユーザのプライバシーに配慮したものとなっており、ユーザへの行動に与える影響も小さい。その結果として長期間の観察が可能となり、図 4 や表 3 に示したように家庭内の習慣行動や普段と異なる行動といった、短期間の観察では抽出困難な実態が判明した。

また、阻害要因抽出という目的を達し得る最小限のセンサ構成とすることで、観察にかかるコストを抑え多数の家庭を対象とした調査が可能となった。表 2 に整理した結果が示すように、省エネ阻害要因は被験者本人の意識や家族・ペットの都合、その他家庭特有の事情によって大きく異なる。このような要因を少数の家庭を対象とした短期間の観察で抽出することは困難であり、提案手法によって多数の家庭を対象とし、長期間の習慣行動を含めた観察が行えたことで得られた結果と言える。

なお、デプスインタビューでの発言記録から、OMK のディスプレイを日常的に閲覧していた被験者は 20 名中 19 名であり、ほぼ全ての被験者が情報閲覧を生活の中に取り入れていることがわかった。本人だけでなく家族、特に子どもが関心を寄せていたとの発言もあり、家族内でのコミュニケーション促進や子どもの省エネ教育に好影響を与える可能性も示唆された。OMK からの情報提供が省エネ行動に与える影響の検証や、OMK を用いてリアルタイムでの情報提供を行う方策の検討なども必要と考える。

³図 4 において生じている日中の欠測は、太陽光発電システムによる。この期間は発電量が消費量を上回ることで逆潮流が生じており、その間はデータ処理上の都合から欠測扱いとなった。

表 2: 抽出・整理した省エネ阻害要因

カテゴリ	カテゴリの内容	発言例 (省エネ行動)
カテゴリ1: 節約	節電や節約に関する発言	スイッチを切り替える時に電気代/負荷がかかる (ID8: 短時間離席時の消灯について)
カテゴリ2: 食品管理	食品や食材に関する発言	腐るのが怖い (ID5: 常温保存可能なものも冷蔵庫に保存)
カテゴリ3: 体調に合わせた空間調整	個人の体調に合わせて空間を調整している旨の発言	寒いのが苦手 (ID1: エアコン利用)
カテゴリ4: 間取り・家具による空間調整	間取りや家具の影響により空間調整をしている旨の発言	裏側にあるので無理 (ID13: テレビのコンセントを抜かない)
カテゴリ5: 生活習慣・心がけ	習慣や個人の考えにより行動をしている旨の発言	忘れてしまう (ID4: 電子レンジ等のコンセントを抜かない)
カテゴリ6: 家族の影 (ペットを含む)	対象者の家族に関する発言	犬を飼っているので (ID17: エアコンの長時間利用)
カテゴリ7: 機器の期待・機能性	利用している機器に対する発言	録画をしているから (ID8: テレビのコンセントを抜かない)
カテゴリ8: 他者からの情報	どこかから得た情報により行動している旨の発言	業者に消さなくていいと言われた (ID5: 給湯器コントローラの電源)
カテゴリ9: 防犯	防犯に関する発言	防犯 (ID18: 照明の点灯時間短縮とテレビの電源オフ)
カテゴリ10: 利便性	利便性を考えて行動している旨の発言	いちいち差すのはストレス (ID10: プリンタやテレビのコンセント)

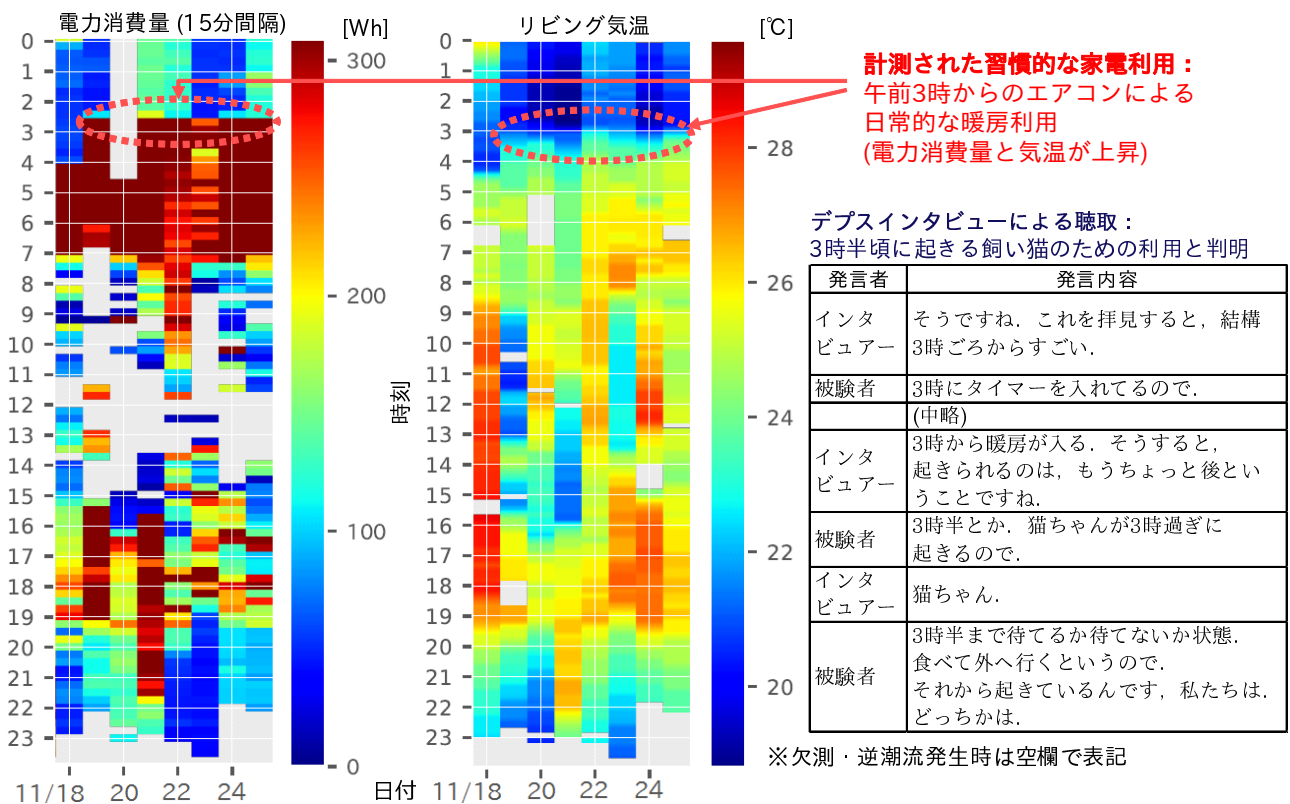


図 4: 行動観察データを用いたインタビュー例 (データは該当箇所のみ抜粋)

表 3: 抽出された家庭内行動の例

ID	抽出された行動
1	猫が3時過ぎに起きるので、3時にはエアコンを入れておかないと寒がる。
1	夫婦が旅行で不在だったが、息子が友人を呼んでエアコンをずっと使っていた。
5	土日は子どもに遊びに行けと言うが、家に友達を呼ぶことが多くてなかなか外出しない。数人でDSを持ち寄ってみんなで充電する。充電するとも言えない。
13	娘夫婦が来て、エアコンやファンヒーターをフル稼働させ、ご飯も山のように作っていた。
13	夫が午前3時から4時のあいだに起きて、アトリエに閉じこもっていることがある。
14	この日は深夜2時までテレビでプロレスを観ていた。
16	この日は娘が休みで、友達呼んで17時まで家でたこやきパーティをしていた。電気代を請求したいくらい。
17	深夜～朝方まで息子？がエアコンをずっと使っていた。怪しい。帰ってから聞いてみる。
18	暖かい日だったので、親子で電気もつけず水も飲まず何も食わず、ひたすらIKEAで買った家具を組み立てていた。
20	猫が出入りするのでもリビングのドアは開けっ放しにしている。

5 おわりに

本稿ではIoTセンサを用いて家庭内の行動観察を実施する手法を提案した。提案手法はIoTセンサによる家庭内の行動計測、計測結果の可視化、可視化結果を用いたデプスインタビューから構成される。家庭内の省エネ阻害要因抽出を目的として行動観察を実施した結果、従来の訪問や撮影による行動観察と比較してユーザへの負担や観察にかかるコストが低減され、多数の家庭に対して長期間の観察が可能であることを示した。

本稿では省エネ阻害要因抽出を目的とした行動観察を実施したが、その結果から省エネに限らない様々な家庭内行動の抽出も期待できる。電力販売分野は競争環境にあり、電力各社は家庭内の課題解決や満足度向上に資するサービスの提供が求められている。独居高齢者の見守りや防犯、その他生活支援などの家庭向けサービス考案や既存サービスの改善を目的とした本手法の活用も進めたい。

謝辞

本研究は、環境省の「低炭素型の行動変容を促す情報発信（ナッジ）等による家庭等の自発的対策推進事業」の助成を受けたものです。また、本研究の遂行にあたってはデロイト トーマツ コンサルティング合同会社および凸版印刷株式会社の協力を得ています。

参考文献

- [1] ティム・ブラウン, デザイン思考が世界を変える, ハヤカワ・ノンフィクション文庫, 2014.
- [2] 山崎 和彦, 松田 美奈子, 吉武 良治, 使いやすさのためのデザイン ユーザーセンタード・デザイン, 丸善株式会社, 2004.
- [3] 松田 千春, 情報探索におけるブラウジング行動: 図書館と書店における行動観察を基にして, *Library and information science*, Vol. 49, pp. 1–31, 2003.
- [4] 畠山 雄豪, 丹羽 由佳理, 佐野 友紀, 菊池 雄介, 佐藤 泰, 立地環境および利用者傾向が行動分布に与える影響 行動観察調査からみたカフェのサードプレイス利用分析 –その1, *日本建築学会計画系論文集*, Vol. 80, No. 711, pp. 1067–1073, 2015.
- [5] 公益社団法人 日本アドバイザーズ協会 Web 広告研究会, おじさんマーケターが知らない若年層のリアルに驚き——「たまった SNS をこなす」「動画を見ながら動画視聴」, 2018 年 5 月 22 日開催 月例セミナーレポート 第 1 部, https://www.wab.ne.jp/wab_sites/general-browse/view/2821/1 (2019/2/7 アクセス)
- [6] E. Singler, F. Waintrop, R. Bordenave and E. Bressoud, “French Government: Nudge Me Tender: how to turn ethnographic insight into more efficient policy-making,” ESOMAR Congress, 2014.
- [7] 服部 俊一, 三浦 輝久, 堤 富士雄, 家庭内センシングを簡易に実現する「おうちモニタキット」の構築とその活用に向けた検討, 第 18 回インタラクティブ情報アクセスと可視化マイニング研究会, pp. 1–6, 2018.
- [8] H. Allcott, “Social norms and energy conservation,” *Journal of Public Economics*, Vol. 95, No. 9–10, pp. 1082–1095, 2011.
- [9] G. L. Urban and E. Von Hippel, “Lead User Analyses for the Development of New Industrial Products,” *Management Science*, Vol. 34, No. 5, pp. 569–582, 1988.
- [10] 東京都地球温暖化防止活動推進センター, 平成 27 年度 家庭の省エネ行動阻害要因等調査結果報告書, 2015. https://www.tokyo-co2down.jp/action/behavioral_inhibition/ (2019/2/7 アクセス)

新しい分析操作を含む操作履歴の提示による テキストマイニングスキルの伝達支援

Skill Transmission Support for Text Mining by Presentation of Operation Histories Including New Analysis Operation

中江 剛士^{1*} 砂山 渡² 畑中 裕司² 小郷原 一智²
Tsuyoshi Nakae¹ Wataru Sunayama² Yuji Hatanaka² Kazunori Ogohara²

¹ 滋賀県立大学大学院工学研究科

¹ Graduate School of Engineering, The University of Shiga Prefecture

² 滋賀県立大学工学部

² School of Engineering, The University of Shiga Prefecture

Abstract: In recent years, text mining skills for analyzing review articles and questionnaire data are required for product development and improvement. However, acquiring the know-how of data analysis is difficult in self-study. Therefore, a text mining software tutorial system supporting text mining beginners and a framework sharing operation history at the time of text mining have been developed, and a framework for learning text mining has been proposed. In this research, we aim to support acquisition of skill of text mining by looking for history including use of tools and button operations which have not obtained valuable analysis results from among many operation histories. We added an operation history presentation function including a new analysis operation to an existing text mining tool and conducted an evaluation experiment of the system of this research. As a result, the interpretation contents of the text analysis tended to improve with the proposed system, and it was found that presenting the operation history including the new analysis operation is effective in supporting the transmission of the skill to the beginner of text mining.

1 はじめに

近年、非常に多くの新製品や新ソフトウェアが登場しており、新たな顧客獲得を目指して熾烈な開発競争が行われている。これにより、新しい製品やアプリケーションソフトウェアの開発を行う企業では、自社の製品やソフトウェアのレビュー記事、およびアンケートデータの分析を行い、得られた改善点から新製品へのアイデアを調査することが求められている。しかし、分析にはテキストマイニングのスキルが必要であり、テキストマイニングによる分析を行える人が少ないことから、テキストデータ分析のノウハウを獲得することは、独学では困難な現状がある。ここで、テキストマイニングのスキルとは、文章テキストを単語で区切り、単語の出現頻度や相関を分析して新たな知見を得る分析のスキルと定義する。

自動車の運転が上手な人の運転を見て真似をするこ

とで自分の運転技術を向上させることができるといったように、他人の行動を模倣することはスキルの獲得に有効である [1]。

これまでにも、テキストマイニングツールの分析操作履歴を他人と共有する枠組み [2] により、他人の分析操作からテキストマイニングのスキルを獲得する枠組みが提案されてきた。しかし、金谷らによって提案された枠組みでは保存される操作履歴の数については考慮されておらず、保存される操作履歴の数が増加するほど、スキルを学びたいと思うユーザにとって参考となる操作履歴の選択が困難になるという点が課題となっている。

そこで本研究では、対象はテキストマイニングを行ったことがない初心者に限定し、テキストマイニングのスキル獲得を目指す人が多数の操作履歴の中から新しい分析操作を含む履歴を探してテキストマイニングのスキル獲得ができる環境の構築を目指すことを目的とする。本研究では、新たな分析操作や文章の絞り込みといった点から評価を行い、システムの利用者へ操作履歴の推薦を行う。

*連絡先：滋賀県立大学大学院工学研究科電子システム工学専攻
〒522-8533 滋賀県彦根市八坂町 2500
E-mail: oh23tnakae@ec.usp.ac.jp

2 関連研究

2.1 履歴の参照によるスキル伝達の研究

他者の学習成果物、成績、学習日記、学習履歴などを蓄積して公開・共有することで学びを誘発するeポートフォリオ推薦システムを提案した研究がある [3]。この研究では、対象と類似の学習履歴を持つ学習者のeポートフォリオを多様に推薦して学習を促す支援を行ったが、本研究では、テキストマイニング時に行った分析の操作内容を書き込んだ操作履歴を蓄積して共有することでテキストマイニングのスキルの伝達支援を行う。

自己および他者の視線行動を可視化・共有できる仮想空間を構築することにより、作業支援を行う研究がある [4]。この研究では、システム利用者の視線が辿った位置の履歴から対象者が注視していた点を算出し、注視点情報をHMD (Head Mounted Display) にて映す映像に出力することでシステム利用者の作業支援を行うが、本研究では、テキストマイニングツールにおける分析操作を再現・共有できる枠組みを構築することにより、テキストマイニングのスキルの伝達支援を行う。

2.2 テキストマイニングのスキル伝達支援の研究

テキストマイニングスキルの伝達支援に関する研究がある [2, 5, 6]。これらの研究では、スキル獲得を支援するチュートリアルの開発により、システム利用者にテキストマイニングの課題を基礎的なものから応用的なものへと順に解かせることでテキストマイニングのスキルの伝達を促している。これらの研究と本研究では、テキストマイニングツールを用いてテキストマイニングのスキルの伝達支援を行うという点で共通している。しかし、本研究ではスキル伝達でテキストマイニングツールを用いるだけでなく、他人の操作履歴を評価し、操作履歴を残した操作者の分析の意図を伝えることでスキル伝達の支援するという点で異なる。

3 スキルの伝達のためのツールの操作履歴の絞り込み支援

3.1 新しい分析操作を含む操作履歴の提示システムの概要

本章では、新しい分析操作を含む操作履歴の提示システムについて述べる。

金谷らにより、TETDM [7, 8] のテキスト分析操作の履歴をデータベースに保存し、他者の分析操作の再現によって知識創発を支援する枠組みは既に出来ている。

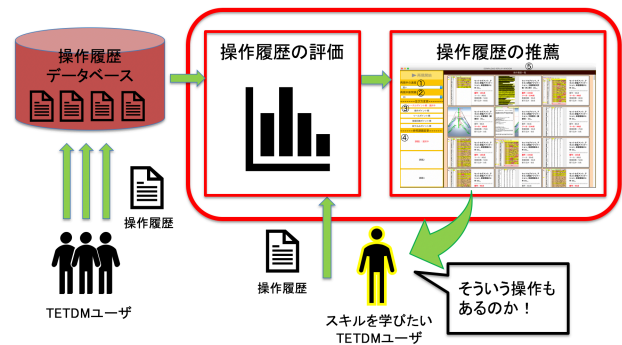


図 1: テキストマイニングのスキルの伝達支援の枠組み

しかし、この枠組みではデータベースに保存された操作履歴の量については考慮されておらず、システムの利用者が増え、データベースに保存された操作履歴の数が增加するほど、テキスト分析の参考になる操作履歴を探し出すことが困難になるという点が課題となっている。

そこで本研究では、テキストマイニングのスキルを学びたいと考えているシステム利用者へ、データベースに保存されている操作履歴の中からシステム利用者にとってスキル獲得に役立つであろうと推測される操作履歴の提示を行うことによってテキストマイニングのスキルの伝達支援を行う。

本研究のテキストマイニングのスキルの伝達支援の枠組みを図 1 に示す。この枠組みには、操作履歴を保存する機能、操作履歴を評価する機能、推薦される操作履歴を提示する機能、操作履歴を再現する機能がある。以下でそれぞれについて触れていく。

3.2 操作履歴を集めるための分析課題

操作履歴の評価を行う前提として、同じ分析の目的を持つものでなければ評価がしにくいということがあがる。同じ分析の目的を持つ操作履歴を集めるため、そのため、操作履歴の評価を行うにあたって、システムで対象とする操作履歴は、共通の分析課題に対する操作履歴とすることにした。

3.3 同じ目的の分析の操作履歴の保存

本研究では、新しい分析操作を含む操作履歴の評価および提示を行うことによってテキストマイニングのスキルの伝達支援を行うので、システムには 2 種類の操作履歴を入力として用いる。

1 つ目は、スキルを学びたい TETDM ユーザの操作履歴である。こちらの操作履歴は、今までその課題で

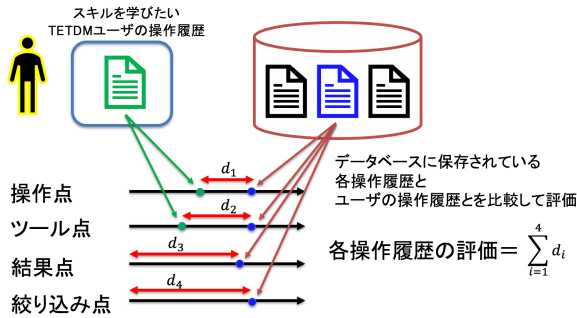


図 2: 操作履歴の評価

行ってきた分析操作，ツールの種類を取得するために用いる。

2つ目は，データベースに保存されている操作履歴である．こちらの操作履歴は，分析操作，ツールの種類を取得に用いる他，同ユーザによって何回目に登録された操作履歴なのか，どのような絞り込み操作を含む操作履歴なのかといった情報を取得に用いる。

3.4 操作履歴の評価

本研究では，操作履歴を評価項目の設定のために予備実験を行った．予備実験の結果，テキストマイニング熟練者の操作履歴では，初心者の操作履歴では使用していないボタン操作とツールでの分析による回答が確認できたため，操作履歴の評価項目として，新しい分析操作を含む事を重視して操作履歴を評価する。

図2に操作履歴の評価のイメージを示す．操作履歴は新しい分析操作，新しいツール使用，分析の結果の登録回数，文章の絞り込みの4項目で評価を行う．各項目ごとに最大100点まで加点し，一つの操作履歴は最大400点でその操作履歴を評価する．ただし，各項目の評価点は予備実験で得られた操作履歴をシステムに入力して，各項目の評価点を均等に設定した．

3.4.1 新しい分析操作による評価

本評価は，テキスト分析の際に押したボタン操作および操作時間に関する評価である．

テキスト分析において，これまで有効に使うことができなかったボタン操作は，そのボタン操作がされている操作履歴の再現を見ることがテキスト分析の参考になる可能性が高い．なお，本研究では，これまで有効に使うことができなかったボタン操作とは，ユーザがこれまでに保存した操作履歴において，「結果と解釈」にまで結びついた分析の操作履歴で使用されていないボタン操作と定義する．

表1に分析操作の評価点を示す．データベースに保存された各操作履歴と TETDM ユーザが保存した操作

表 1: 操作時間の評価点

これまで有効に使えたことがないボタンの数	1個	2個	3個
評価点	30	40	50

表 2: 操作時間の評価点

操作時間	1分未満	2分未満	3分未満	4分未満	5分未満	5分以上
評価点	50	40	30	20	10	0

履歴を比較して，これまで有効に使えたことがないボタン操作が使用されていれば評価が高くなる．評価基準は対象ボタン操作が一つでもあると参考になると考えられるため，最初に大きく加点する．

表2に操作時間の評価点を示す．分析に要した時間が短いほど，分析操作が洗練されていると考えられ，操作時間が短いほど高評価とする．評価基準は予備実験で測定した課題1問あたりの解答にかかる平均の時間より設定した．

システムでは，まずユーザが保存した操作履歴を課題別に参照する．各課題の操作履歴のうち，「結果と解釈」が登録されている操作履歴で使用したボタン操作の抽出を行う．次にデータベースに保存された操作履歴を課題別に参照し，同じ課題において，その操作履歴でユーザの操作履歴から抽出されていないボタン操作がされていれば，数に応じて加点する．

3.4.2 新しいマイニングツールの使用による評価

本評価は，テキスト分析の際に使用したマイニングツールに関する評価である．

テキスト分析において，これまで有効に使うことができなかったツールは，今まで行ったことがないテキスト分析手法を学ぶことができる可能性が高く，そのツールを用いて分析されている操作履歴の再現を見ることはテキスト分析の参考になる可能性が高い．なお，本研究では，これまで有効に使うことができなかったツールとは，ユーザがこれまでに保存した操作履歴において，「結果と解釈」にまで結びついた分析の操作履歴で未使用のツールと定義する．

表3にマイニングツールの評価点を示す．データベースに保存された各操作履歴とシステムの利用者が保存した操作履歴を比較して，データベースに保存された操作履歴に，これまで有効に使えたことがない処理ツールまたは可視化ツールが使用されていれば高評価となる．ただし，「結果と解釈」が登録されていない操作履歴に加点しないものとする．評価基準は，有効に使えたことがないツールが一つでもあることで参考になりやすいため，最初に大きく加点する．

表 3: マイニングツールの評価点

これまで有効に使えたことがない ツール (処理・可視化) の数	1 個	2 個	3 個
評価点	30	40	50

表 4: 分析の結果登録回数の評価点

「結果と解釈」 登録回数	1 回	2 回	3 回	4 回	5 回以降
評価点	50	70	80	90	100

システムでは、まずユーザが保存した操作履歴を課題別に参照する。各課題の操作履歴のうち、「結果と解釈」が登録されている操作履歴で使用したツールの抽出を行う。次にデータベースに保存された操作履歴を課題別に参照し、同じ課題において、その操作履歴でユーザの操作履歴から抽出されていないツールが使用されていれば、その数に応じて加点する。なお、ツールは処理、可視化の各ツールで評価する。

3.4.3 分析の結果登録回数による評価

本評価は、テキスト分析の際に登録した「結果と解釈」の数に関する評価である。

一人のユーザが同じテキスト、同じテキスト分析課題を用いてテキストを分析した操作履歴は、「結果と解釈」が登録されるほど分析操作が洗練化され、操作履歴の再現を見ることがテキスト分析の参考になる可能性が高い。

表 4 に分析の結果登録回数の評価点を示す。一人のユーザが「結果と解釈」に登録した回数が多いほど高評価となる。評価基準は、予備実験において、登録回数が 2 回以上行われた操作履歴の「結果と解釈」の内容が向上していたことより 2 回目登録で他の項目に並ぶような点数としている。

システムでは、データベースに保存された操作履歴を参照する。「結果と解釈」に登録した操作履歴の数を抽出し、その数に応じて加点する。

3.4.4 文章の絞り込みによる評価

本評価は、テキスト分析の際に使用した文章の絞り込みに関する評価である。

テキストを単語や文に注目してテキストの絞り込みを行うことで、その単語や文が出現した理由を探りやすくなる。また、どの単語や文でテキストの絞り込みを行うかはユーザ毎に差が生じるため、テキストの絞り込みが行われている操作履歴は今まで思いつかなかった知識の獲得の補助となる可能性が高い。

表 5 と表 6 に文章の絞り込み方法の評価点を示す。テキストの絞り込みを行なった回数とテキストの絞り

表 5: 文章の絞り込み回数の評価点

絞り込み回数	1 回	2 回	3 回	4 回以降
評価点	50	70	80	90

表 6: 文章の絞り込みツールの種類の評価点

絞り込みをしたツールの種類	1 種類	2 種類
評価点	20	40

込みを行なったツールの種類が多いほど高評価となる。評価基準は、予備実験において 1 種類以上のマイニングツールで複数回絞り込みを行うことで「結果と解釈」の内容が良くなっていったことから、表 5 と表 6 のような評価としている。

システムでは、データベースに保存された操作履歴を参照する。操作履歴から絞り込み機能を使用したマイニングツールの種類、絞り込んだ回数を抽出し、その数に応じて加点する。

3.5 推薦される操作履歴の提示機能

システム利用者が自身の分析操作の参考にする目的の操作履歴をサーバに保存された操作履歴から探しやすくするため、操作履歴の評価により参考になると予測される操作履歴をシステムの利用者に提示する。

推薦される操作履歴の提示は、各評価項目の点数を表示することでシステムで推薦される操作履歴の提示機能を追加した。

図 3 に操作履歴の推薦インターフェースを示す。図 3⑤に操作履歴一覧が表示されており、起動時は評価の合計点が高い順に並んでいる。図 3①②は、操作履歴の再現に関する機能であり、①は操作履歴の再現中の速度を設定することができ、②は操作履歴の再現中に再現を中断したくなつた際に操作ができるように再現を一時停止する間隔を設定することができる。図 3③を操作することで、操作履歴一覧に表示されている操作履歴の並び方を評価項目毎に高い順番で並び替えることができる。また、図 3④を操作することで、操作履歴一覧の表示を異なる課題の操作履歴に切り替えることができる。

4 新たな分析操作を含む操作履歴の提示機能の評価実験

4.1 実験目的

本章では、本研究で開発した新たな分析操作を含む操作履歴の評価を利用することで、テキストマイニングのスキルの伝達に役立てられるかについて評価するため評価実験を行った。以下で実験について述べる。

表 7: レビューテキストの分析課題

課題番号	課題内容
1	レビュー全体を特徴付ける頻度が高い単語を見つけ、その単語がレビュー内でよく使われた理由を調べることで、多くの人がこの映画について抱いている印象を答えて下さい。
2	[低評価]のレビューを特徴付ける頻度が高い単語を見つけて、その単語が[低評価]レビュー内でよく使われた理由を調べることで、多くの人がこの映画について抱いている不満を答えてください。
3	[高評価]のレビューの中から、この映画についてユーザーが不満に思っている点を見つけて答えてください。
4	集めた印象と不満から、[次回作の制作]に向けて、どのような改善案が考えられるか答えてください。

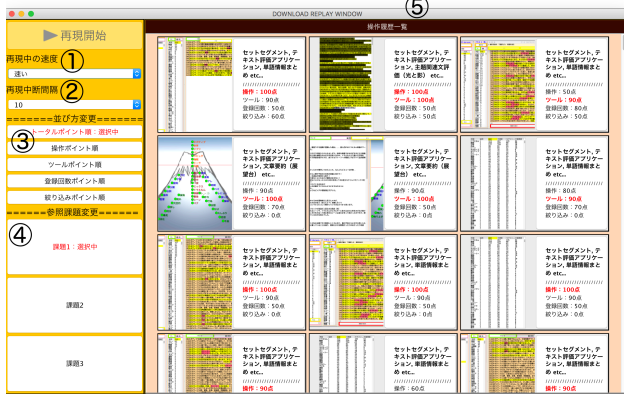


図 3: 操作履歴の推薦インターフェース

4.2 実験方法

本実験では、TETDMのチュートリアル機能を通して初歩的なTETDMの操作方法を学んだ9名の理系の学生を被験者としてシステムの評価実験を行った。2日間を利用して、「被験者がとあるシリーズ映画のプロデューサーである」という立場の仮定のもと、新作映画のアイデアを考える」という課題を出題した。分析に利用するテキストには、映画のレビューサイト[9]に掲載されている子供向けアニメシリーズの映画のレビューを高評価75件、低評価75件ずつ収集し採用した。なお、本実験では、レビューサイトにてレビューを書き込む人がこの映画につけた星の数の評価(1から5の5段階評価)において星4、星5がつけられているレビューを高評価レビュー、星1、星2がつけられているレビューを低評価レビューと定義する。

実験1日目には、表7の分析課題を回答してもらい、レビューの分析結果と解釈を集めた。実験2日目には、提案システムにより他者が行った分析操作の再現を参考にしながら1日目同様の分析課題を回答してもらい、レビューテキストの分析結果と解釈を集めた。ただし、課題1から3がTETDMを用いて行う課題であり、課題4はTETDMを使用せずに課題1から3の結果をまとめた案を解答する課題である。

実験1日目と2日目の課題で登録した「結果と解釈」の違いから、新たな分析操作を含む操作履歴の提示機能があることによってテキストマイニングのスキルが伝達できているか評価を行う。

4.3 結果と考察

4.3.1 実験によって得られたアイデアの比較

新しい分析操作を含む操作履歴の提示を行うことで良いテキスト分析結果が得られるようになったのかについて検証するため、課題4にて被験者に解答してもらった、課題1から3までのテキスト分析結果をまと

表 8: 「新作映画のアイデア」のキーワードの種類数

被験者	1日目のキーワード数	2日目のキーワード数
A	18	21
B	5	15
C	7	18
D	17	49
E	5	12
F	17	25
G	9	9
H	22	31
I	20	30
平均値	13.9	21.6
t検定	t=-2.14, df=16, p=0.047	

めることで得られた「新作映画のアイデア」の内容について比較を行った。

被験者がどれだけ現状の映画について、特徴を捉えることができているかを比較するため、得られた「新作映画のアイデア」の内容について、映画を表す特徴となる名詞・形容詞のキーワードの種類数で比較を行った。

t検定を行ったところ、 $t(16) = -2.14, p = 0.047$ であり、1日目と2日目に得られた「新作映画のアイデア」のキーワードの種類数の平均の差に有意差が見られた。これは、被験者が2日目のテキスト分析で1日目よりも多くの現状の映画の特徴を捉えることができ、1日目よりも詳細な「新作映画のアイデア」を考えることができたためであると考えられる。

4.3.2 上位に提示される操作履歴の閲覧回数と新しい分析操作・ツールの関係

4.4節で述べた評価ポイントによって評価された操作履歴の再現を閲覧することが被験者に新たなテキスト分析手法やマイニングツールの分析の使用につながっているのか検証するため、被験者に提示されている操作履歴のうち、上位12件を対象とし、対象の操作履歴の再現を閲覧した回数と2日目に新しく行った分析操

表 9: 操作履歴の閲覧数と新しいツール数の相関

被験者	課題 1		課題 2		課題 3	
	閲覧数	ツール数	閲覧数	ツール数	閲覧数	ツール数
A	6	2	5	1	5	3
B	7	2	1	1	4	1
C	5	0	0	0	0	1
D	5	3	0	2	2	1
E	0	1	3	1	3	1
F	11	2	2	2	5	3
G	5	0	8	2	1	1
H	5	1	0	2	7	1
I	0	0	0	0	4	0
相関係数	0.48		0.31		0.42	

表 10: 操作履歴の閲覧数と新しいボタン操作数の相関

被験者	課題 1		課題 2		課題 3	
	閲覧数	操作数	閲覧数	操作数	閲覧数	操作数
A	6	0	5	0	5	2
B	7	1	1	0	4	2
C	5	0	0	0	0	1
D	5	1	0	0	2	2
E	0	1	3	3	3	3
F	11	0	2	1	5	4
G	5	1	8	2	1	3
H	5	1	0	0	7	3
I	0	0	0	0	4	0
相関係数	-0.17		0.57		0.29	

作および新しく使用したマイニングツールの数で相関をとった。それぞれの相関をまとめたものを表 9 と表 10 に示す。

表 9 の結果より、4.4 節で述べた評価ポイントによって高評価とされた操作履歴の再現を閲覧した回数と 2 日目に新しく使用したマイニングツールの間には、弱から中程度の正の相関があることが分かった。高評価とされた操作履歴の再現を閲覧する回数が多いほど、今まで有益な結果が得られなかった新たなマイニングツールでテキストの特徴を発見できる傾向があると言える。

また、表 10 の結果より、4.4 節で述べた評価ポイントによって高評価とされた操作履歴の再現を閲覧した回数と 2 日目に新しく行った分析操作の間には、弱から中程度の正の相関があることが分かった。高評価とされた操作履歴の再現を閲覧する回数が多いほど、今までは思いつかなかったテキスト分析方法でテキストの特徴を発見できる傾向があると言える。ただ、課題 1 では二つの間で相関が見られなかった。これは、1 日目の課題 1 のテキスト分析時に行われたマイニングツールで使用されるボタンの操作が少なかったために、再現を閲覧してもツールのボタン操作を行っている操作履歴があまりなく、新たな操作方法をあまり思いつけなかったためであると考えられる。

5 おわりに

本研究では、TETDM を利用してテキストマイニングのスキルの伝達支援を行うことを目的とした。TETDM の分析操作のログが操作履歴として保存される機能を用いて、同一のテキストを使用した同一の分析課題を出題し、課題の操作履歴を収集する枠組みを設計した。また、ユーザがまだ有益な分析結果を得られていないマイニングツールや分析操作によって集められた操作履歴を評価し提示する機能を開発し、有効な操作履歴を探しやすくする枠組みを構築した。

新しい分析操作を含む操作履歴の提示機能がテキストマイニングのスキルの伝達に役立てられるかを検証するために評価実験を行った。映画のレビュー記事について分析を行い、新作映画のアイデアを発見するという課題を操作履歴を評価、提示する機能を実装して回答してもらったところ、提案システムにより操作の参考になる操作履歴を探しやすくなり、テキストマイニングのスキルの伝達を行えた。

今後の課題として、評価されたボタン操作、ツールを表示する、ノイズとなる操作履歴のフィルタリング処理を強化することによって、高評価された理由をより分かりやすくすることが今後の課題である。

参考文献

- [1] 大桑萌：0~2 歳児の仲間関係における模倣の役割，保育学研究 Vol.52, No.2, 172-182, (2014)
- [2] 金谷直哉・砂山渡：ツールの操作履歴の共有によるテキストマイニング時の発想支援，人工知能学会第 12 回インタラクティブ情報アクセス可視化マイニング研究会，pp29-34, (2016)
- [3] 植野真臣：過去の学習者履歴データを利用した e ポートフォリオ・システム，情報知識学会誌，Vol.24, No.47, p.414-423, (2014)
- [4] 藤本武司・砂山渡・山口智浩・谷内田正彦：視線行動の可視化による着眼スキル伝達支援，人工知能学会論文誌，Vol.71, No.1, pp.517-518, (2009)
- [5] 西原陽子・中垣内李菜・川本佳代・砂山渡：TETDM を用いたテキストマイニングのスキル獲得を支援するためのチュートリアルシステムの開発，知能と情報（日本知能情報ファジィ学会誌），Vol.27, No.5, pp.771-783, (2015)
- [6] 中垣内李菜・川本佳代・砂山渡：統合環境 TETDM を用いたテキストマイニング初心者のスキル獲得支援，第 28 回人工知能学会全国大会，1H5-NFC-01b-3, (2014)
- [7] Total Environment for Text Data Mining (テキストデータマイニングのための統合環境)，(URL) <http://tetdm.jp>
- [8] 砂山渡・高間康史・徳永秀和・串間宗夫・西村和則・松下光範・北村侑也：統合環境 TETDM を用いた社会実践，人工知能学会論文誌，Vol.32, No.1, pp.NFC-A-1-12, (2017)
- [9] Yahoo Japan Corporation, (URL) <https://movies.yahoo.co.jp>

引用論文に着目した研究動向把握支援 インタフェースの基礎検討

Basic Study on Interface to Grasp Research Trends Based on Cited Articles

程岡 晃一^{1*} 大杉 隆文² 松下 光範¹
Koichi Hodooka¹ Takafimi Osugi¹ Mitsunori Matsushita¹

¹ 関西大学総合情報学部

Faculty of Informatics, Kansai University

² 関西大学大学院総合情報学研究科

Graduate School of Informatics, Kansai University

Abstract: This goal of this study is to develop a system that supports grasping trends of a research topic. In conducting research, it is an important activity to grasp research trends by investigating prior researches in that field. Search engines are often used in such cases, however, it is difficult to efficiently gather papers in the intended research field because search by keywords or authors is not sufficient. In order to solve this problem, this paper proposes an interface to support comparing the articles cited by research papers. Differences in the cited articles can be assumed to reflect the primary concerns of the paper. Our conducted user observation suggested that the proposed interface contributes to the understanding of the differences between papers.

1 はじめに

近年、様々な分野の論文が執筆され、膨大な数の論文が世の中に存在している [10]. 研究者は研究活動を行う際、膨大な数の論文から過去に行われている研究を把握し、それらを纏めることで考察を行ったり、今までに行われていない新しい研究を行ったりする。そのため、研究者は自身と同じ研究分野の論文を整理し、その分野の研究動向を把握することが重要である。そこで、研究者は過去の研究を把握するために、論文検索のための検索エンジン (e.g., Google Scholar, CiNii) を利用する。その際にキーワードをクエリとして与え論文を検索する。しかし、クエリに用いるキーワードが適切であるとは限らない。そのため、そのクエリを用いた検索で見つかった論文全てが、必ずしも研究者が読みたい論文であるとは限らない。

この問題の解決策として、キーワード検索だけでなく、ある読みたい論文 (以下、起点論文と記す) が引用している論文から、論文を検索するという方法がある。その引用された論文は起点論文に関連した論文である。また、起点論文を執筆した著者 (以下、起点著者

と記す) と同じ著者の論文から読みたい論文を検索するという方法もある。起点著者が執筆した論文が複数存在する場合、起点著者は同じ研究分野に対し、研究を積み重ねていることがある。そのため、起点著者と同じ著者の論文を検索することは、起点論文と同じ研究分野の論文の収集に繋がる。つまり、自身と同じ研究分野の論文を1つ見つけることができれば、その論文に関連する論文を効率的に収集することが期待される。研究者はそれらの収集した論文を読むことで、研究者自身の研究分野や、関連する分野、取り扱われてきた技術、研究動向などを理解する。このような行為は「サーベイ」と呼ばれ、研究活動において、研究者自身の研究の位置付けを明確にするために重要な行為である。従って、サーベイは膨大な数の論文から、研究者が企図した研究分野の論文を探し出し、多くの論文を読むことが望ましい。

サーベイにおいて、研究動向を把握する際は、研究者が企図した研究分野の論文を把握するだけではなく、ある論文の比較対象になる論文や技術的基盤となる論文がどの論文であるかの関係性を理解することが望ましい。しかし、キーワードを用いた検索では、同じ研究分野の論文を全て収集できるとは限らず、著者名を用いた検索では、企図した研究分野の論文のみを収集

*連絡先：関西大学総合情報学部
〒569-1095 大阪府高槻市霊山寺町 2-1-1
E-mail: mat@res.kutc.kansai-u.ac.jp

できるとは限らない。そのため、研究動向の把握が困難である。

そこで、本研究では論文を特徴づける要素の1つである引用論文に着目し、複数の論文間の引用論文の一致や差分を可視化することで、研究動向の把握を支援するシステムの実現を目指す。本稿ではその端緒として、引用論文の比較を支援するインタフェースについて検討する。

2 関連研究

サーベイを効率化するために、論文の引用・被引用関係を可視化する手法が提案されている。井上らは各論文をノードとし、論文間の引用・被引用関係をエッジとするグラフの可視化システムを提案した [5]。可視化システムにおいて、引用・被引用関係にある論文を全て表示すると、論文の数が多く見目が乱雑になってしまう。そこで井上らは、その問題を解決するために、引用理由ごとに論文の表示をフィルタリングできる機能を構築した。実験結果から、フィルタリング機能を使うことでユーザのサーベイの検索効率が向上することが示唆された。しかし、可視化画面と引用理由などの情報が別画面で表示されているため、引用理由を1つにフィルタリングしていない場合、可視化画面のみで論文間の関係性を把握することが負担になる。

Nakano らは、論文の関係性をユーザに理解しやすい形で提示するため、引用・被引用関係に加え、共引用関係を用いた論文関係グラフを提案した [2]。共引用関係とは同一の論文に引用されている論文同士の関係のことである。Nakano らはこれらの論文間の関係を全て提示するのではなく、関係の強さを考慮して提示を行った。2つの論文が引用・被引用関係にある場合の関係性については、1つの論文の文章内で何度も引き合いに出されて引用されている論文ほど関係が強いとした。また、2論文間が共引用関係にある場合の関係性については、共引用されている回数をもとに類似度を計算し、その値が高いもののみを提示した。しかし、共引用されるには時間がかかるため、発行されてもいない論文に共引用関係を利用するのは困難である [3]。そのため、共引用の回数が多い論文同士の類似度が必ずしも高いとは限らない。

榎らは、自然言語には表記揺れが存在し、これが検索や分析の精度低下を招いていると示唆している [6]。この問題を解決するために、専門的な技術用語を含む電子情報通信学会の文献検索システム I-Discover と一般的な日常用語を多く含む Wikipedia のデータベースを用いた表記揺れを補償する辞書 Broadic を提案した。Broadic の有効性を確認するため、I-Discover の文献データを用いた時系列・技術要因分析システムを提案し、

Broadic の有無による影響を評価した。その結果、表記揺れを補償することで、より正確な研究トレンドを可視化することができたと示唆された。しかし、トレンドにある研究がユーザ自身の研究分野に対し、どのような研究動向であるか把握するには、論文を複数収集しなければならない。

共著ネットワークを用いて、萌芽領域の中心研究者を予測するための研究が行われている。森らは著者ごとの発表論文数、被引用数、ダウンロード数などの計量書誌データを各著者のパフォーマンスとみなし、ある研究領域において被引用数が高い著者を領域の中心研究者とした [7]。そして、研究領域を特定するために、シードとなる研究者から共著ネットワークを取得した。そのネットワークの構造に基づく特徴量をもとに、ある研究領域における研究者の被引用数を予測する学習モデルを構築した。これにより、中心研究者の予測を行った。しかし、中心研究者の中でどの論文がユーザが企図した研究分野の論文であるか把握し、その中でも論文同士の差異を把握することは困難である。

3 デザイン指針

前述した関連研究から、論文情報の関係性を可視化することでサーベイを支援できる。論文の引用・被引用関係の可視化を行う研究では、引用・被引用関係にある論文同士の関係性理解に繋がる。しかし、引用・被引用関係にない論文同士の関係性を理解することは困難である。加えて、共引用関係の可視化は、発行されてもいない論文には適さないため、最新の研究動向を掴むことが困難である。

研究動向の把握を支援する研究では、各分野のトレンドの分析を可能にしている。しかし、最も行われている研究の中でも、研究内容の差異を理解するのは困難である。そのため、同じ研究分野の中でどのような研究の動きがあるか理解することは困難である。

著者情報の分析を行う研究では、共著関係を用いることで、中心性の高い著者の把握に繋がると明らかにされている。しかし、その中心性の高い著者が、どのような研究を行っているかまで理解しなければ、研究動向の把握は困難である。加えて、その中心性の高い著者がユーザの研究分野にとって重要な著者とは限らない。そのため、ユーザ自身が選択した論文を執筆した著者の研究動向を可視化することで、ユーザ自身と同じ研究分野の動向把握支援に繋がると考えられる。

そこで本研究では、ユーザが起点とする論文を執筆した著者の研究動向を可視化する。これにより、引用・被引用関係にない論文も把握できる。しかし、その著者の論文の中でもユーザが企図した研究分野以外の論文も存在する。そのため本稿では、引用論文を比較で

きるようにすることで、ユーザが企図した研究分野の論文を把握できるようにする。引用論文は、引用元の論文の根拠や位置付けを示すために用いられる [8]。そのため、複数の論文を比較する際、引用論文は同じ分野の論文の同一性を判定したり論文同士の関係性を理解したりする際の手がかりになる。例えば、起点論文と比較対象論文（以下、比較論文と記す）の共通する引用論文に着目することで、研究分野の同一性や関連性を把握できる。また、起点論文と比較論文の引用論文の差分から、それらの論文間の着眼点や力点の相違といった論文同士の研究の位置付けを把握できる。

また、同一著者の論文は同じ分野の論文である場合が多いため、その著者の論文を継時的に閲覧することで研究動向の把握の一助になる。しかし、論文には複数の著者が関与する場合が多いため、著者ごとに執筆論文を閲覧できるようにする必要がある。

これらを勘案し (1) 引用論文の一致度を把握できるようにすること、(2) 引用論文の差異を判断可能にすること、(3) 著者ごとに論文を閲覧できること、の3つをデザイン指針とする。

4 提案インタフェースの実装

4.1 実装の概要

本稿ではコミック工学分野の論文 350 本を対象として、論文表題、著者名、発行年、引用論文から構成されるデータベースを構築した。そのデータベースを対象として、研究動向把握支援インタフェースを実装した (図 1 参照)。

提案インタフェースは、HTML、CSS、JavaScript、jQuery を用いて実装を行った。可視化部分は、Franzらが提供しているオープンソースの JavaScript ベースのグラフィブラリである Cytoscape.js を用いた。Cytoscape.js は視覚化ソフトウェアコンポーネントであるため、Web ブラウザでインタラクティブなグラフを表現できる [1]。これらを用いて、引用論文の比較を支援するインタフェースを実装した。

4.2 可視化デザイン

グラフ描画は、人間関係や論文の引用関係などのデータセットを直感的に表現するのに有用である [4]。そのため、提案インタフェースはノードとエッジを用いて描画した。ノードは論文を表し、その中には論文の表題を提示した。エッジは起点著者の論文の中で引用・被引用関係があるノード同士を結んだ。図 1 の画面はユーザが選択した論文 (起点論文) を執筆した著者 (起点著者) の動向を示した。赤色のノードは起点論文、青

色のノードは起点著者が執筆した別の論文を表した (図 1 の①参照)。それらのノードの下に表示されている橙色と灰色のノード (図 1 の②参照) は各々の引用論文のリストである。

図 1 の③は、起点著者の論文の中で、引用・被引用関係のある論文同士をエッジで示している。三角形の矢印が引用先であり、丸の印が引用元である。エッジのうち黒色のエッジは起点論文に対し、引用・被引用関係のあるものである。また、灰色のエッジは起点論文以外に引用・被引用関係のあるものである。また、各ノードを時系列順に表示しているため、どちらの論文が引用元で、引用先かはノードが表示されている座標で直感的に理解できる。

デザイン指針 (1) を満たすために、起点論文と比較論文の共通する引用論文のノードを橙色で、異なるノードを灰色で各々表示した (図 1 の②参照)。これにより、起点論文の引用論文と比較論文の引用論文との一致度を直感的に理解できる。例えば引用論文のリストのうち、橙色のノードが多い比較論文は起点論文と目的や手法の類似性が高いと推測できる。反対に、灰色のノードが多い比較論文は起点論文と目的や手法の類似性が低いと推測できる。さらに、起点論文の引用論文のリストと一致する比較論文が存在しなければ、起点論文の研究の方向性が変化したことの推測に繋がる。また、起点論文の引用論文のリストと一致する比較論文が複数存在すれば、起点論文と目的や手法の類似性が高い研究が複数行われていると理解できる。

図 1 の④はユーザが引用論文のノードをクリックすることで、その引用論文と同一の引用論文のノードが黄色に変化する。これによりクリックした引用論文がどの程度、どのような論文に引用されたか確認できる。

また、引用論文のノードの横に表示した数字 (図 1 の⑤、⑥参照) は、起点著者がその引用論文を引用して何回目であるかを示している。数字が 1 の場合はその引用論文の引用元の論文より古い年をユーザが遡っても、その引用論文を引用している論文を収集することはできない。そのため、ユーザはその引用論文が引用元の論文より古い年の論文で引用されているか参照する必要がない。また、数字の背景が深緑色 (図 1 の⑥参照) の場合、その引用論文は引用元の論文以降引用されていないことを表している。そのため、ユーザはその論文を引用している論文を探す際、その論文を引用している論文より新しい年を参照する必要がない。加えて数字が 1 で深緑色のノードの場合、その引用論文を引用したのは、その引用元の論文のみであるということがわかる。これらにより、ユーザが画面移動することの負担を軽減した。

次にデザイン指針 (2) を満たすために、2 論文間で共通する論文と、一方の論文でのみ引用されている論文のノードの色を区別した。起点論文以外の論文 (図 1

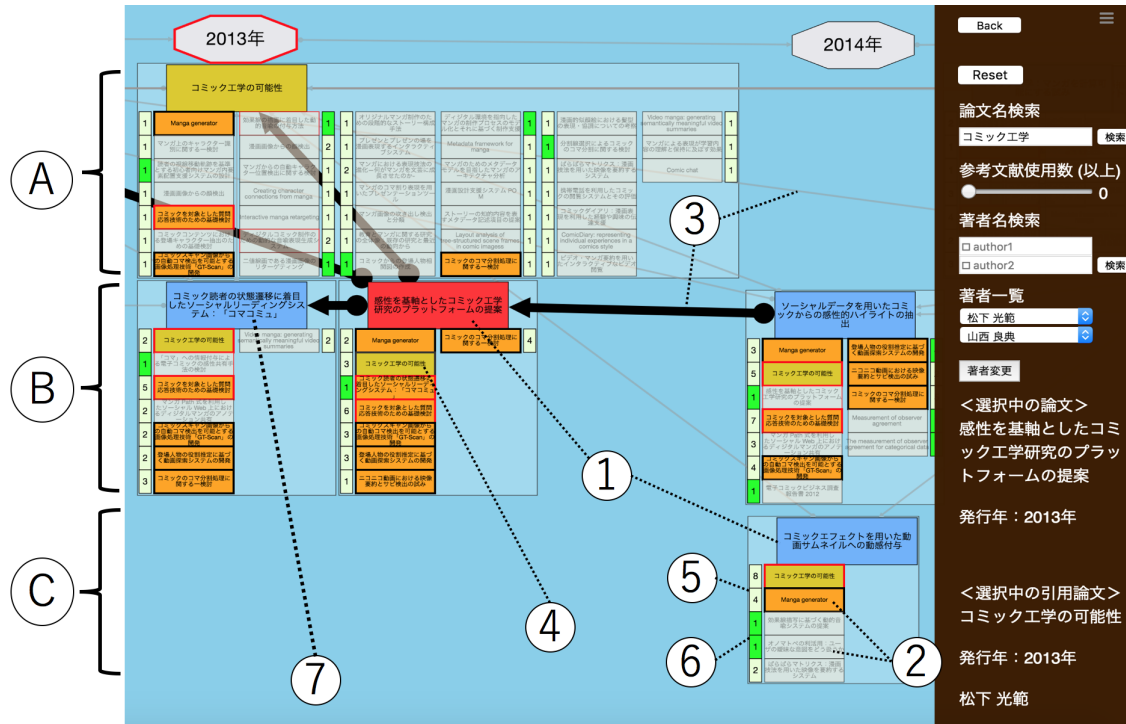


図 1: 提案インターフェース

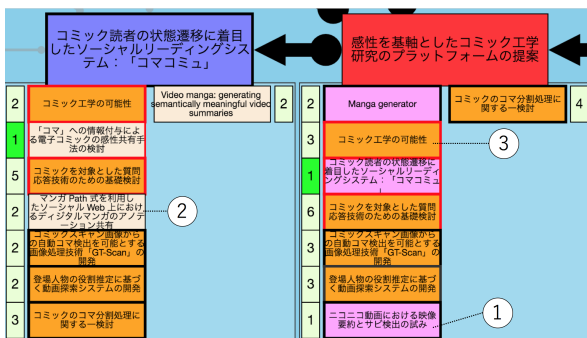


図 2: 引用論文の増減結果

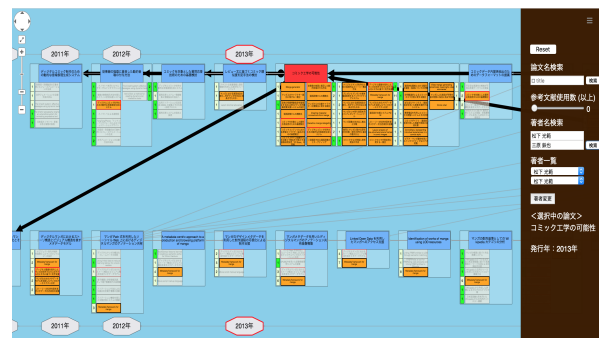


図 3: 共著関係にない著者同士の研究動向

の⑦参照)を右クリックすると、その論文と起点論文の引用論文の色が変わる(図2参照)。桃色のノード(図2の①参照)は起点論文のみで引用された論文であり、黄白色のノード(図2の②参照)は新たに選択された論文にのみ引用された論文である。2つの論文で共通する引用論文は橙色のノードのままである。これにより、直感的に論文同士の差分を理解でき、一方にのみ現れる引用論文がその引用元の論文の特徴を理解する上で手掛かりになる。

また、デザイン指針(3)を満たすために、著者ごとにy座標を統一した。図1では起点著者の1人目が執筆した論文を上段(図1のA参照)、2人目を下段(図1のC参照)に各々表示した。また両者が共同で執筆した論文を中段(図1のB参照)に表示した。これによ

り著者ごとに執筆した論文が容易に理解できる。さらに著者同士の関係性も理解できる。

図3では、ユーザが任意に選択した著者2人の研究動向を表示している。これにより、共著関係にない著者同士の研究動向も比較できる。加えて、両者の研究分野や引用論文の差異の理解に繋がる。

5 ユーザ観察

ユーザ観察では、実装したインターフェースが3章のデザイン指針で企図した利用が確認できるか検証する。そこで、ユーザが引用論文を比較することで、論文同士の関係を理解できているかに注目した。

5.1 ユーザ観察の手続き

まず、実装したインタフェースの説明を行った。その後、インタフェースの評価を行うため、インタフェースを利用し、機能の妥当性に関するインタビューをリードユーザに行った。なお、実装したインタフェースはコミック工学分野の論文を利用しているため、リードユーザはコミック工学分野の論文を執筆した経験のある者にした。ユーザ観察は、リードユーザが2論文間の関係を理解できているか口頭で適宜確認した。そのために、2つの論文を比較させ、その2つの差異を答えるタスクを課した。比較する論文は起点論文と同じ論文を引用しているものや引用・被引用関係がある論文に設定した。

5.2 ユーザ観察の結果

ユーザ観察ではまず2017年の論文Aを起点論文に設定した。その論文Aの引用論文と同じ論文を引用している論文は2016年の論文Bのみであった。そこで論文Bを比較論文に設定し、起点論文との差異を答えるタスクを課した。リードユーザは「論文Aでは、論文Bに扱われていない固有表現の概念を扱うため、その引用論文2本が追加された」と発言した。これは、2つの論文の差分に関する発話であり、デザイン指針(2)で企図した利用が確認された。

次に2017年の論文Cを起点論文に設定した。論文Cの引用論文と同じ論文を多く引用している2015年の論文Dが存在した。そのため、論文Dを比較論文と設定し、起点論文との差異を答えるタスクを課した。リードユーザは「論文Cでは、論文Dで引用されている情報推薦系の論文よりもレビューやコミックを扱う上で必要な論文を引用するようになった」と発言した。これは、2つの論文の差分に関する発話であり、デザイン指針(2)で企図した利用が確認された。

次に2018年の論文Eを起点論文に設定し、他の論文と比較するタスクを課した。論文Eの引用論文と同じ論文を引用している論文は存在しなかった。そのため、リードユーザは「この論文は今までと着眼点が異なるため、他の論文で引用されていたものを引用しなくなった」と発言し、論文Eを新しいテーマの論文であると判断した。これは、起点論文と他の論文の引用論文の一致度に基づく行動であり、デザイン指針(1)で企図した利用が確認された。

一方で、今回の評価ではデザイン指針(3)で企図した利用に関わる行動は観察されなかった。このことから、論文の配置が著者ごとに表示されていることをリードユーザが理解できなかった可能性や、著者ごとに論文を分類することが論文同士の関係の理解に不要である可能性が考えられる。

ユーザ観察終了後に、リードユーザから「情報量が過多であるため、同じジャンルの論文のみを表示してほしい」「広い表示画面を見ることが負担」という意見が得られた。これらのことから、必要な論文を絞り込む機能の必要性が示唆された。

6 議論

5章で行ったユーザ観察から、論文同士を比較したことで、各論文同士の一致度や差分を理解できたと示唆された。しかし、ユーザは普段のサーベイ時、引用論文を比較して論文を検索しないため、提案インタフェースの説明なしでインタフェースを利用するユーザは引用論文を比較せず、論文の表題のみしか見ない可能性が考えられる。そのため、今後はインタフェースの説明なしで利用するユーザでも、ノードの色の意味を理解し、引用論文を自然と比較できるインタフェースのデザインにすることを目指す。本章では、そのインタフェースのデザインと、扱った引用論文の情報、の2つの観点から考察を行う。

6.1 可視化画面の有用性と問題点

可視化画面では、起点著者のうち2人の研究動向を表示している。その2人のうち、1人のみが執筆した論文を可視化画面の上段と下段に表示した。その表示した論文は、その1人の著者が主著者ではない場合も同じ座標に表示している。そのため、起点論文と関連度が低い可能性のある論文も同じ扱いとして表示しているため、今後は起点著者のうち1人しか執筆していない論文の表示は工夫する必要があると考えられる。例えば、その1人が主著者として執筆していない論文を表示する場合は、y座標や色の変化など区別した表示にすることを検討する。

提案インタフェースは起点著者のコミック工学の論文全てを表示しているため、起点著者が行った研究の網羅的な把握や起点論文と類似度の高い論文などの把握を支援した。しかし、論文数の多い著者の場合、全て把握するのは困難になる。また、著者の中にはコミック工学だけでなく、他の分野の研究を行っていることがある。そのため、起点著者全ての論文を表示するだけでなく、ユーザがピックアップした論文のみ表示する機能を実装する必要があると考えられる。これにより、画面を左右に行き来する手間が省け、情報量過多の解消に繋がる。

また、提案インタフェースではサーベイ論文など引用論文数の多い論文を可視化画面で発見することができる。しかし、引用論文数の多い論文は画面の横幅が長くなってしまおうという問題がある。そのため、引用

論文の表示・非表示の切替機能などが必要だと考えられる。

6.2 提示した引用論文の有用性と問題点

実装したインタフェースにおいて、引用論文の表題を各引用元の論文のノードの下に全て表示した。これにより、引用論文から論文を検索する際、引用元の論文を閲覧せず論文を検索することができる。しかし、引用論文の表題全てを表示したため、全体的に文字の情報量が過多になった。そのため、引用論文をジャンルごとに分類するか、キーワードのみを表示し、それらのキーワードをクリックすることで、引用論文の表題が表示されるなどの機能が今後必要だと考えられる。これにより、情報量を減らすことができる。

7 おわりに

本稿では、研究動向の把握支援を目的として、引用論文を可視化することで、論文間の関係理解を促すインタフェースを実装し、その有用性を検証した。ユーザ観察の結果から、論文同士を比較したことで、論文同士の一致度や差分を理解できたと示唆された。しかし、各論文の理解や研究動向の理解を支援できたか検証することはできていない。さらに、ユーザ観察では、リードユーザ1人のみを対象としたため、提案インタフェースが他のユーザにも有効に働くか今後検証する必要がある。

また、提案インタフェースのデザインは6章で述べた通り、改善すべき点が存在する。さらに、本稿で構築したデータベースはコミック工学の論文しか扱っていないため、他の分野の論文でも提案インタフェースが有効に働くか検証する必要がある。また、提案インタフェースは論文の学会情報、ページ数などの書誌情報や論文の内容情報を扱うことができていない。これらを扱うことで、論文同士の差分をより理解できると考えられる。そのため、今後は提案インタフェースのデザインの改善と本稿で扱っていない論文情報を扱うことを検討する。

参考文献

- [1] Franz, M., Lopes, C. T., Huck, G., Dong, Y., Sumer, O. and Bader, G. D.: Cytoscape.js: a graph theory library for visualisation and analysis, *Bioinformatics*, Vol. 32, No. 2, pp. 309–311 (2015).
- [2] Nakano, Y., Shimizu, T. and Yoshikawa, M.: A Visualization of Relationships Among Papers Using Citation and Co-citation Information, *International Conference on Asian Digital Libraries*, Springer, pp. 157–163 (2016).
- [3] Shibata, N., Kajikawa, Y., Takeda, Y. and Matsushima, K.: Comparative study on methods of detecting research fronts using different types of citation, *Journal of the American Society for information Science and Technology*, Vol. 60, No. 3, pp. 571–580 (2009).
- [4] Toeda, N., Nakazawa, R., Itoh, T., Saito, T. and Archambault, D. W.: On Edge Bundling and Node Layout for Mutually Connected Directed Graphs, *Information Visualisation (IV), 2016 20th International Conference*, IEEE, pp. 94–99 (2016).
- [5] 井上絢翔, 韓東力: 参照関係の可視化による論文サーベイの効率化, 第11回インタラクティブ情報アクセスと可視化マイニング研究会, SIG-AM-11-09, pp. 55–61 (2015).
- [6] 槇俊孝, 若原俊彦: I-Discover 文献メタデータを用いた時系列・技術要因分析システム, 電子情報通信学会論文誌 D, Vol. 99, No. 10, pp. 1002–1012 (2016).
- [7] 森純一郎, 原忠義, 榊剛史, 梶川裕矢, 坂田一郎: 大規模学術論文データの共著ネットワーク分析に基づく萌芽領域の中心研究者予測に関する研究, 人工知能学会全国大会論文集, Vol. JSAI2015, pp. 1B21–1B21 (2015).
- [8] 石井仁子, 太田学, 高須淳宏: 引用意図を利用した学術論文閲覧支援のための適切な被引用箇所の特長, 第7回データ工学と情報マネジメントに関するフォーラム (DEIM Forum 2015), F3-5 (2015).
- [9] 棚瀬達央, 大平茂輝, 長尾確: デジタルコンテンツの部分参照・引用関係に基づく論文作成支援, 第75回全国大会講演論文集 2013, 情報処理学会, Vol. 1, pp. 623–624 (2013).
- [10] 井坂徳恭, 中山泰一: 重要論文検索システム Iask の実装と評価, コンピュータと教育研究会, Vol. 2011-CE-109 (2011).
- [11] 林佑磨, 奥野峻弥, 山名早人: 単語の意味概念行列を用いたキーワード生成による関連論文検索システム, 情報処理学研報, Vol. 2014-DBS-159, No. 10 (2014).

キュレーションマップにおけるインタラクティブUIの提案

Proposal for interactive UI on Curation Map

阿部穰太郎^{1*} 渋谷 英潔² 森 辰則²
Jotaro Abe¹ Hideyuki Shibuki² Tatsunori Mori²

¹ 横浜国立大学大学院 環境情報学府 ² 同 環境情報研究院

^{1,2}Graduate School of Environment and Information Sciences, Yokohama National University

Abstract: We define “Curation Map” as information complex that is compiled by iteratively linking “overview texts”, which comprehensively cover multiple viewpoints, to detailed texts written in corresponding individual viewpoints. In this paper, we propose a user interface in which users interactively able to browse dynamically generated Curation Map efficiently. In order to improve visibility and legibility of overview texts and to generate appropriate links to corresponding detailed text, we examined methods to improve the accuracy of dividing overview text according to individual viewpoints. Since gisting links branched out from overview text may have some influence on ranking text in terms of appropriateness as overview text, we also examined how the accuracy of ranking “overview text” changes.

1 はじめに

意思決定に Web を用いることが日常となっている一方で、情報を取捨選択するためには利用者による能動的な評価が必要とされるがその負荷は大きい。このため、近年、「NAVER まとめ¹」など、特定のトピックに沿って複数の情報(文章)を収集・吟味し、分析・判断結果等を付記した文章を作成し、他者と情報共有するキュレーションサービスが注目されている。しかし、現在、その作成は人手により、投稿者の個人技に依存する。一方で、同サービス以外にも、二つ以上の情報間の関係を解説するいわゆる「まとめ文章」が Web 上に存在する。そのため、複数文書間に内容の類似性に基づく参照リンクを自動的に張ることにより、より詳細な観点で記された「詳細文章」を「まとめ文章」で繋いでいくことを繰り返せば、文章を関係付けて理解するための情報複合体が得られる。我々は、この情報複合体をキュレーションマップと呼び、文献[1]で自动生成するための手法を提案している。

文献[1]では、あるトピックに関する文書集合が入力として与えられた場合、まず各観点到に分割されたまとめ文章(文書)を提示し、分割された個々の観点からその観点对する詳細文書と推定された文章に対してリンクが張られたネットワーク構造(キュレーションマップ)を出力する。しかしながら、文献[1]では、観

点に対応した適切な長さのテキストに分割されておらず、まとめ文章の検出精度やキュレーションマップの可読性の低下を招いていた。それゆえ、本稿では、分割されたテキストを適正化することでこの問題の解決を試みる。さらに、動的に生成されたキュレーションマップを利用者との対話により効果的に閲覧するための UI を提案する。

2 キュレーションマップ自動生成手法の概要

キュレーションマップ自動生成手法の基本的な考え方は以下の通りである。2つのテキストを比較して、互いのテキスト中に同じ観点から書かれた記述²がある場合、そのテキスト間にリンクを張る。これをあるテキスト集合に対して繰り返すと、テキストをノードとしたネットワーク構造ができる。リンクが集中しているテキストは、他の多くのテキストと観点を共有していることとなり、それらのテキスト群をまとめているテキストとみなすことができる。したがって、このネットワークに HITS[2] や PageRank[3] などのグラフベースのランキングアルゴリズムを適用することで、まとめ文章としての重要度を計算できる。ここで問題は、ある観点对する記述の範囲をどのように決定するかである。我々は、観点として考えられる最小単位にテキストを一度分割し、その後、同じ観点をもつテキ

*連絡先: 横浜国立大学大学院 環境情報学府 情報環境専攻
神奈川県横浜市保土ヶ谷区常盤台 79-7 森辰則研究室
E-mail: jotaro@forest.eis.ynu.ac.jp

¹<https://matome.naver.jp/>

²互いの記述の長さは問わない。

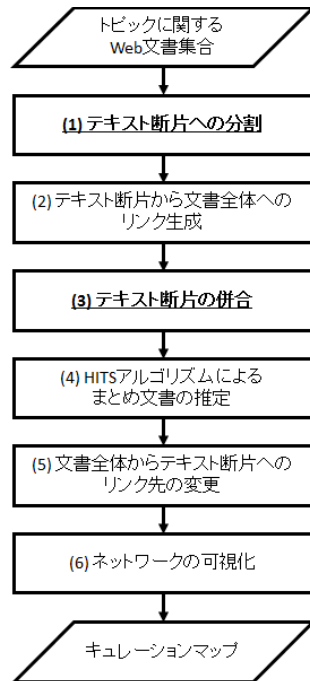


図 1: キュレーションマップ自動生成手法の流れ

ト断片同士をまとめることで、適切な範囲となるテキスト断片を決定した。

図 1 に従来手法 [1] の流れを示す。まず、(1) 入力された各文書を観点を端的に表しうる最小単位（テキスト断片）に分割する。従来手法では、観点を表す端的な表現の最小単位を 1 つの述語項構造と仮定し、文書中のテキストを述語の直後で分割する。(2) 各テキスト断片から、単語の包含性に基づいて同じ観点を持つと判断された文書に対して重み 1 のリンクを張ることで、テキスト断片や文書をノードとしたネットワーク構造を構築する。(3) 連続した 2 つのテキスト断片を比較して、それぞれのリンク先の文書集合が同一または包含関係にある場合、2 つのテキスト断片は同じ観点を持つとして 1 つのテキスト断片に併合する。(4) 併合後のネットワークに対して HITS アルゴリズムを適用し、各ノードのまとめ文章としての重要度を計算する。(5) リンク元のテキスト断片とリンク先の文書内のテキスト断片とを単語の包含性に基づいて比較し、同じ観点を持つと判断された場合にリンク先を文書からテキスト断片に変更する。最後に、(6) ネットワーク構造を可視化することでキュレーションマップを提示する。

3 従来手法の問題点

従来手法 [1] によるキュレーションマップの例を図 2 に示す。図の例では、トピックとして「アイドルマス

ターシリーズ」が与えられている。図の左側には、まとめ文章が観点ごとに区切られて提示され、右側には観点ごとの詳細文章が提示されている。左側の観点ごとに区切られたテキストを見ると、必要以上に小さく断片化されており、人間にとって理解しやすいテキストであるとはいえない。

テキストの過剰な断片化は、図 1 中の下線の処理である「(1) テキスト断片への分割」と「(3) テキスト断片の併合」に原因があるため、本稿ではこれらの処理に修正を加えることでテキスト断片の適正化を行う。

4 テキスト断片の適正化

4.1 テキスト断片への分割

従来手法のキュレーションマップにおいてテキスト断片が読みにくい原因の一つとして、同じ観点について述べている一連の文章が、複数の別のテキスト断片として分割されていることが挙げられる。従来手法では、1 文に複数の観点があることを想定し、述語項構造を最小単位としてテキスト断片に分割している。しかしながら、述語項構造単位に分割されたテキストは可読性の点で問題があるとともに、述語項構造に分割することで、仮定条件の従属節と主節が別のテキストにされるなど、書き手が本来意図する内容と異なった内容になってしまう危険性がある。したがって、テキスト断片の最小単位を述語項構造から文へと変更し、より大きく読みやすい単位に変更する。この処理を文単位化と定義する。

4.2 テキスト断片の併合

ある程度の長さの文章を作成する際、書き手は同じ観点で書いた文を文章中に散在させるのではなく、まとめて記述することが一般的である。そのため、観点が不明なテキスト断片の前後に同じ観点で書かれた 2 つのテキスト断片が存在している場合、不明なテキスト断片を前後のテキスト断片と同じ観点であるとみなし、1 つのテキスト断片として併合する。これにより、同じ観点で書かれた文章中に内容的に関連性が低い文が混ざっていた場合でも、それらの文が 1 つのテキスト断片となることが期待できる。この処理を挟み撃ち併合と定義する。

従来手法でテキスト断片を併合する際、リンク先の文書集合が同一または包含関係であることが併合の条件であったが、その条件では、併合されるべきテキスト断片が併合されない事例があった。それゆえ、リンク先に重複する文書が存在するに条件を緩和する。これにより、例外的に他の文書にもリンクが張られていることで包含関係が成り立たなかったテキスト断片同

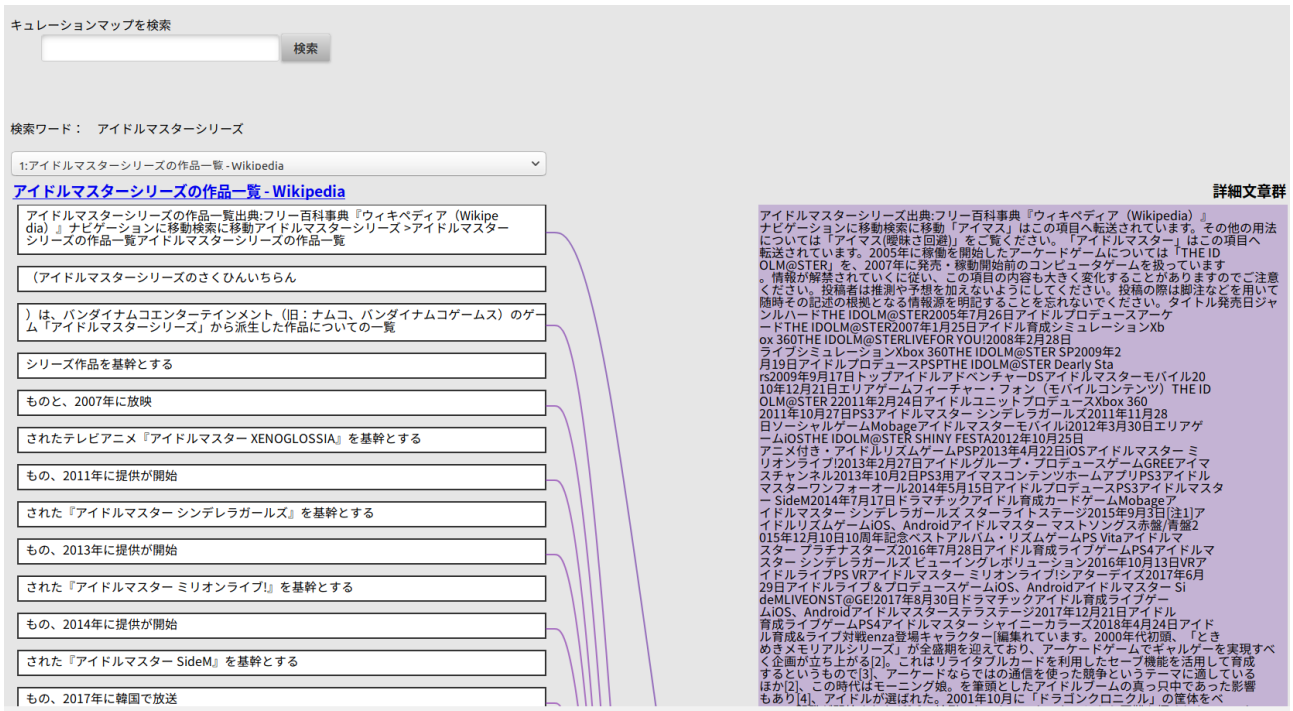


図 2: 従来手法によるキュレーションマップ (適正化前)

表 1: トピック (検索クエリ) とその分野

分野	トピック
人物	コブクロ
スポーツ	ビリヤード
地名	尾瀬
自然言語による質問	なぜ空は青いのか
政治経済	アベノミクス

士が併合されるようになると期待できる。この処理を重複文書併合と定義する。

以上の3つの処理を加えて生成されたキュレーションマップの例を図3に示す。図2と同じトピックであるが、左側のまとめ文章のテキスト断片が内容を理解しやすい長さで分割されており、可読性が向上している。また、ネットワークの構造も大きく変化しており、HITS アルゴリズムによる重要度計算の結果である、まとめ文章にも影響が及ぶことが予想される。

5 実験

提案した3つの手法の有効性を示すために、以下の2通りの実験を行う。1つ目は、まとめ文書の検出精度がどの程度改善されたかを目的とした実験であり、ま

とめ文書検出実験と定義する。2つ目は、分割されたテキスト断片が人間の観点とどの程度一致するかを目的とした実験であり、観点分割実験と定義する。

使用した文書集合は、どちらの実験でも、表1に示す5トピックを検索クエリとして BingSearchAPI³で取得した Web 文書であり、トピックごとに上位 50 文書を入力とした。表1の5トピックは多様な分野から選ぶようにした。3つ全ての処理を適用した手法と比較するために、従来手法 [1]、3つ全ての処理を適用した手法から文単位化を除いた手法 (文単位化なし手法)、3つ全ての処理を適用した手法から挟み撃ち併合を除いた手法 (挟み撃ち併合なし手法)、3つ全ての処理を適用した手法から重複文書併合を除いた手法 (重複文書併合なし手法) の4手法を用いた。

まとめ文書検出実験では、HITS アルゴリズムによる重要度ランキングの上位5件を出力とした。複数の観点が存在する文書を正解文書として第一著者が判定し、評価尺度として MRR (平均逆順位) を用いた。

$$MRR = \frac{1}{|Q|} \sum_{i=1}^{|Q|} \frac{1}{rank_i} \quad (1)$$

Q は全てのトピック、rank は上位5件において初めて正解文書が出た順位である。

³https://azure.microsoft.com/ja-jp/pricing/details/cognitive-services/search-api/

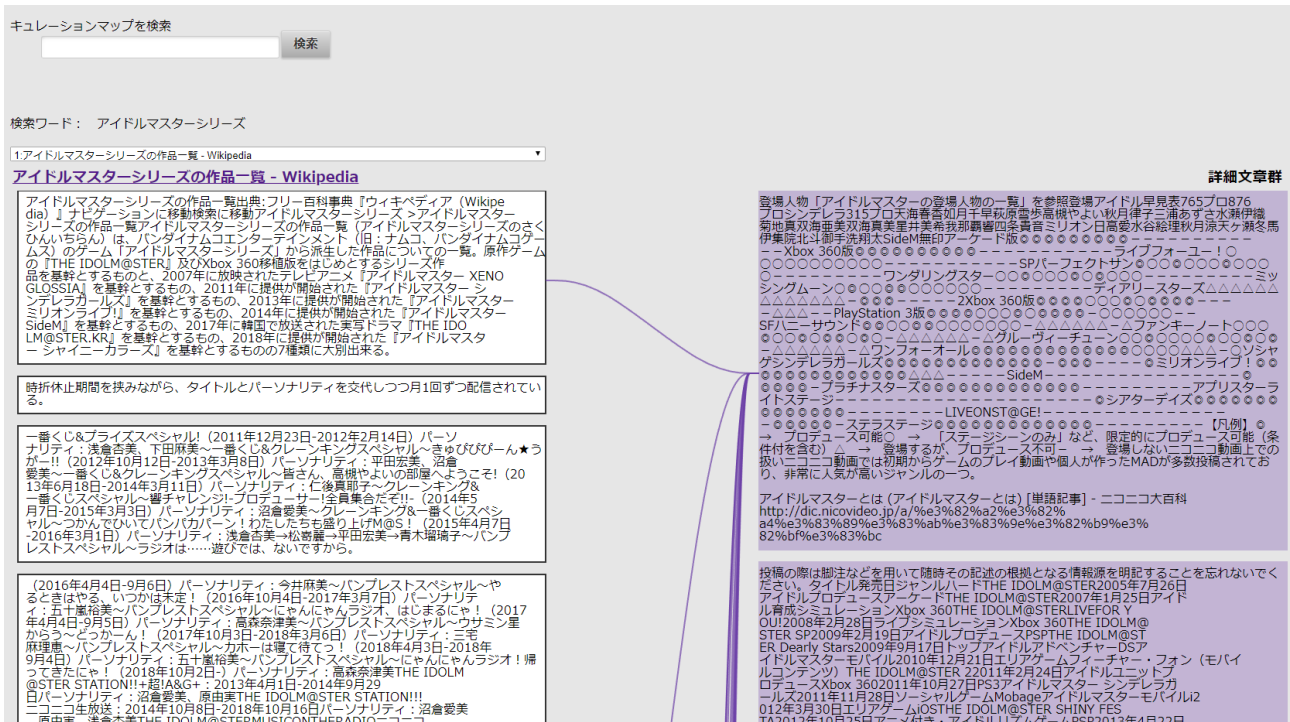


図 3: 全ての処理を適用した手法によるキュレーションマップ (適正化後)

観点分割実験では、まとめ文書検出実験で正解文書と判定された上位 1 件を対象として、第一著者が観点ごとに分割し、その分割点を正解情報とした。評価尺度として、以下の式で適合率 P 、再現率 R 、F 値 F を用いた。

$$P = \frac{\text{一致した分割点の数}}{\text{出力文書中の分割点の数}} \quad (2)$$

$$R = \frac{\text{一致した分割点の数}}{\text{正解文書中の分割点の数}} \quad (3)$$

$$F = \frac{2PR}{P + R} \quad (4)$$

分割点が一貫しているかどうかの判定は、テキストの分割に個人差が存在することを考慮し、正解の分割点との差が 50 字以内であれば正解とした。ただし、正解の分割点に対応する各手法の分割点は 1 つまでとし、1 対 1 対応が保証されるようにした。

6 考察

まとめ文書検出実験と観点分割実験の結果を表 2 と表 3 にそれぞれ示す。

表 2 から、まとめ文書検出実験において、全ての処理を適用した手法から一部の処理を除いた、文単位化なし手法、挟み撃ち併合なし手法、重複文書併合なし手法の 3 手法は、従来手法の精度を上回った。したがっ

表 2: まとめ文書検出実験の結果

手法	MRR
従来手法	0.65
全ての処理を適用した手法	0.61
文単位化なし手法	0.75
挟み撃ち併合なし手法	0.68
重複文書併合なし手法	0.76

て、それぞれのテキスト断片の適正化処理は、まとめ文書の検出精度の改善に一定の効果があったと考えられる。その一方で、全ての適正化処理を行った手法は従来手法よりも精度が低下した。この理由について考察する。ネットワーク構造を構築する際、単語の包含性に基づいてリンクを張るため、長い文章の方が多くのリンクを張られやすい。また、多くの観点を含むまとめ文章は、その性質上、文章全体が長くなる傾向にある。結果として、従来手法においてテキスト断片が適切に併合されていなくとも、ある程度の精度でまとめ文書を検出できていたと思われる。一方、全ての処理を適用した手法では、併合が促進された結果、テキスト断片の数が従来手法と比較してトピック平均で 1915.2 から 394.4 に減少した。その結果、ネットワークのリンク数が 1313.0 から 148.4 に減少し、検出精度が低下し

表 3: 観点分割実験の結果

手法	P	R	F
従来手法	0.09	0.95	0.16
全ての処理を適用した手法	0.25	0.60	0.33
文単位化なし手法	0.18	0.67	0.28
挟み撃ち併合なし手法	0.17	0.91	0.29
重複文書併合なし手法	0.22	0.62	0.32

たと思われる。したがって、リンクの重みを文章全体の長さに基づいて正規化したり、併合の際にリンクの重みを変更したりすることで改善できるのではないかと考えている。

表 3 から、観点分割実験において、全ての処理を適用した手法は従来手法と比較して、再現率が減少 (0.95 → 0.60) したが適合率が大きく増加 (0.09 → 0.25) し、 F 値が 0.16 から 0.33 に向上した。3 つの比較手法の F 値から、文単位化、挟み撃ち併合、重複文書併合の順に効果が高かったことが分かった。挟み撃ち併合なし手法の再現率が 0.91 と従来手法と比較して減少度合が小さいことから、挟み撃ち併合が再現率を大きく低下させている。しかしながら、適合率の改善には挟み撃ち併合の貢献が最も大きかった。また、向上した全ての処理を適用した手法の F 値も 0.33 とあまり高いものではないため、改善の余地があると思われる。重複文書併合なし手法の F 値は 0.32 と全ての処理を適用した手法とあまり差がなく、まとめ文書検出実験で最も高い値 (0.76) であったことを考慮すると、重複文書併合なし手法が総合的に最も良い結果だったといえる。

7 インタラクティブ UI の提案

ここでは、動的に生成されたキュレーションマップを効率的に提示するためのユーザインタフェースを提案する。

基本的には図 2、図 3 のように画面左側にまとめ文書の本文をテキスト断片区切りで表示し、画面右側に詳細文章の一覧を並べて表示する。しかし、このままではテキスト断片のリンク先の詳細文章が遠く離れた位置に描画されている場合、リンクを目で追いながら画面をスクロールする必要があるなど、テキスト断片と詳細文章の対応がわかりづらいため、互いの関係をわかりやすく描画する必要がある。そこで、まとめ文書のテキスト断片をクリックすると、そのテキスト断片から出ているリンク先の詳細文章一覧を表示させるようにする (図 4)。これにより、各テキスト断片に対応する詳細文章が簡単に発見できる。

また、その詳細文章のさらに詳細を知りたい場合に対応するため、詳細文章をクリックした場合、その詳細文章が含まれる文書を新たにまとめ文書として、図 5 のように表示する。この操作は、まとめ文書内の気になった詳細文章をクリックするたびより詳細な内容を見ることができるようにし、粒度の粗い内容が書かれた文書から細かい内容が書かれた文書への誘導を行う目的がある。この誘導を行うことでユーザは、最初から読解するために様々な前提知識が必要な詳細な文書を見せられることなく、様々な内容がまとめられたページから順により詳しいページへと読み進めることができる。

8 まとめ

本稿では、キュレーションマップ自動生成手法において、文単位化、挟み撃ち併合、重複文書併合の 3 つの処理によりテキスト断片の適正化を行うことで、まとめ文書の検出精度向上とキュレーションマップの可読性の向上を試みた。まとめ文書検出実験において、それぞれの適正化処理では一定の精度向上があり、重複文書併合なし手法が最も良い結果を示した。観点分割実験において、 F 値が 0.16 から 0.33 に向上し、文単位化、挟み撃ち併合、重複文書併合の順に効果が高かったことを確認した。今後、リンクの重みを変更するなど改善していきたいと考えている。また、効率的にキュレーションマップを閲覧するためのインタラクティブ UI の提案も行った。

謝辞

本研究の一部は、JSPS 科研費 16K00296 の助成を受けたものである。

参考文献

- [1] 小林隼人, 小笹哲哉, 渋谷英潔, 森辰則. 観点毎の詳細度を考慮したネットワーク構造の発見に基づく Web 文書群の関係の可視化. 言語処理学会第 22 回年次大会発表論文集, pp. 1073–1076, 2016.
- [2] Jon M. Kleinberg. Authoritative sources in a hyperlinked environment. Proceedings of the Ninth Annual ACM-SIAM Symposium on Discrete Algorithms, pp. 668–677, 1998.
- [3] L. Page, S. Brin, R. Motwani, and T. Winograd. Proceedings of the 7th International World Wide Web Conference, pp. 161–172, 1998.

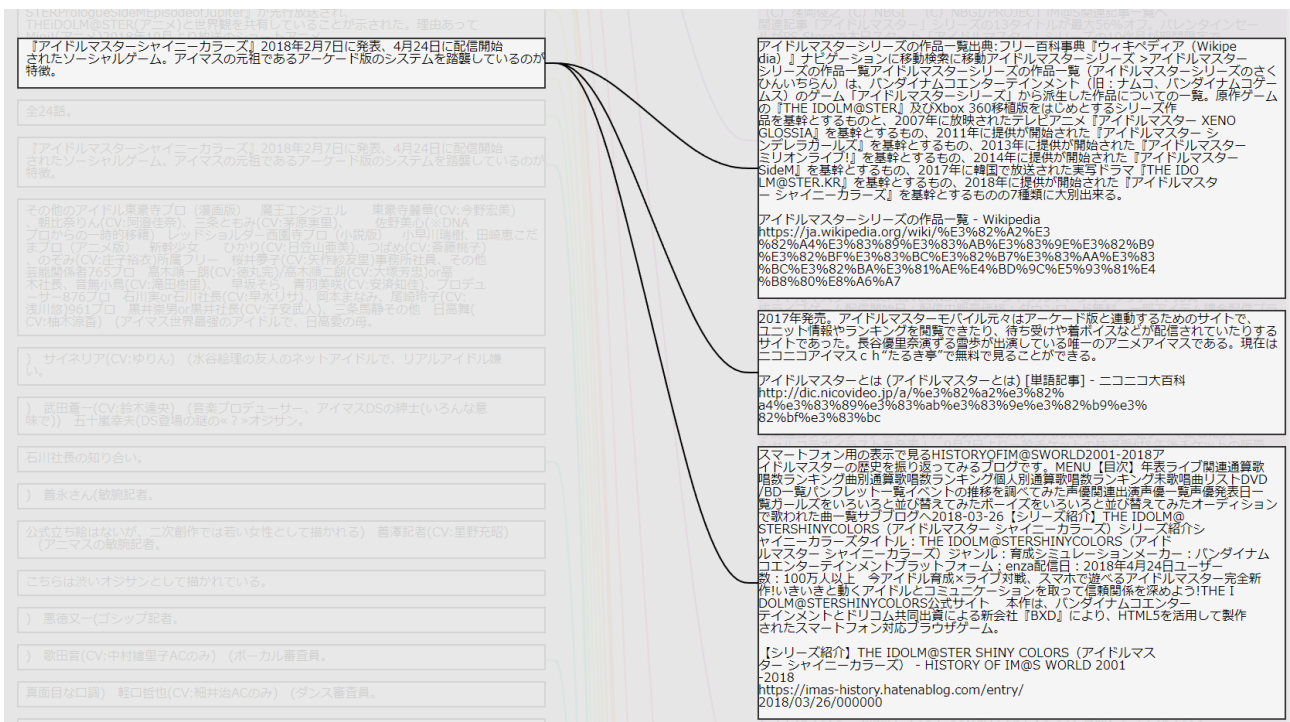


図 4: テキスト断片をクリックしたときの画面表示



図 5: 図 4 の 3 番目の詳細文章をクリックしたときの画面表示

宿泊施設のレビューの時系列分析による季節を表す特徴語の抽出

Extracting Feature-Words Depending on Seasons by Time-Series Analysis of Web Reviews for Accommodations

佐藤裕次郎 山西良典 西原陽子*

立命館大学情報理工学部

College of Information Science and Engineering, Ritsumeikan University

Abstract: The users sometimes concern some keywords which represent the seasonal events/features such as “fireworks” and “colored leaves,” when they look for the accommodations. Such keywords can be also helpful for the accommodation sides; those keywords can be the good advertisements appealing the features of that place. To find those keywords, we propose a time-series analysis of Web reviews for accommodations in this paper. The review for the accommodation is divided into monthly units, and the nouns characteristically appear in the monthly time-series are extracted as the keywords depending on the seasons. Through the evaluation tasks, it was confirmed that the proposed method extracted the nouns appearing in some specific seasons well.

1 はじめに

宿泊施設をユーザが決める際、ウェブレビューが決定要因になることが多い。このとき、特に、宿泊する季節の内容が含まれたレビューを参考にすることがある。例えば、「花火」や「紅葉」などの季節のイベントや、時期によって混雑具合が変わる周辺施設の情報のレビューは、旅行者にとっては充実した滞在をする上で有用な情報となりうる。また、宿泊施設側の観点からも、これらの情報は効果的な広告・宣伝として活用できる。宿泊施設のレビューを対象として、宿泊施設の戦略構築 [1] や宿泊施設サービス改善のための情報抽出 [2, 3] などの研究も報告されている。

ウェブレビューの解析、情報の抽出や比較の研究は数多く存在しており、様々な方法で分析が行われている。中山らは、レビューテキストから条件付き意見の抽出 [4] と評価条件の抽出 [5] を行っている。これらの研究では利用者の宿泊施設の選択の意思決定支援に有益な情報の抽出を目的としている。しかし、同一のレビュー対象についての時系列変化に着目した分析とはなっていない。松尾らは、レビューのユーザの評価の根拠を提示しており [6]、ユーザの評価の根拠の中には本研究で着目する季節によって変化する特徴が含まれる可能性がある。旅行は時期と場所が重要な要素となるアクティビティである。宿泊施設のレビューを対象とする本研究では、ユーザの評価の根拠の中でも特に

時期によって変化する特徴の抽出を目指す。乾らはテキストから評価辞書を使用した評価情報の抽出を行っており [7, 8]、評価を記述するもの、要求、提案、認識、印象などの意見の分類を問題としている。

本稿では、レビュー対象としては宿泊施設を扱い、時系列変化する評価情報 [9] の中でも特に一定の季節において盛り上がりを見せるような季節を表す特徴語の抽出を目指す。提案手法によって抽出された単語が、主観評価実験によっても季節を表す単語であるかを評価する。

2 提案手法

提案手法では、ある宿泊施設のレビュー文から季節の特徴を表す単語の抽出を目的としている。単語は名詞とする。レビューデータとして、楽天トラベルのデータ¹を使用する。宿泊施設に対して書かれたレビューを月ごとにまとめ一つの文書として扱い、12ヶ月分の文書に出現する名詞の1ヶ月毎の出現回数を算出する。その出現回数の時系列変化の割合が大きい名詞を、その宿泊施設の季節の特徴をあらわす名詞として抽出する。月単位の文書内に出現する名詞の出現回数を全単語の出現回数で割った tf 値と、全文書数を単語 t の出現する文書数で割り、その値の対数である idf 値を掛け合わせて $tf-idf$ 値を計算する。 $tf-idf$ 法を用いるこ

*連絡先: 〒 525-0072 滋賀県草津市野路東1-1-1
E-mail: {is0309he@ed, ryama@fc, nisihara@fc}.ritsumeikan.ac.jp

¹楽天株式会社 (2016): 楽天トラベルデータ. 国立情報学研究所情報学研究データリポジトリ. (データセット). <https://doi.org/10.32130/idr.2.2>

とにより、出現した名詞のその宿泊施設の特定の月における重要性が計算できる。また時系列変化の割合が大きいものを出力するために、正規化を行い比較できるように処理する。レビューに出現した名詞の月単位の変化として、12ヶ月分の $tf-idf$ 値の平均値と最大値の差が大きいものから順に出力する。この処理により、特定の季節にのみ $tf-idf$ 値が上昇した名詞が出力される。

処理手順をまとめると、以下のようになる；

1. レビューの分割を行う。月単位でレビューを分割し、1つの宿泊施設に12個の文書が存在するとする。
2. レビューの分ち書きの際に形態素解析器 MeCab と辞書に NEologd を用いて出現した名詞を取得する。
3. 出現した名詞の月単位の $tf-idf$ 値を計算する。
4. 解析中の宿泊施設のレビューに1回でも出現した名詞に対し、12ヶ月分の $tf-idf$ が計算され、 $tf-idf$ 値の変化の割合の大きい順に名詞を出力するため、正規化を行い比較できるようにする。最大値を0、最小値を1とし、12ヶ月分の $tf-idf$ 値に対し正規化を行う。
5. 正規化された12ヶ月分の $tf-idf$ 値の平均と最大値の差が大きい順に名詞を出力する。
6. 正規化の最大値が1なので、平均の小さい順に並び替え出力する。

2.1 $tf-idf$ 値の計算

$tf-idf$ 値は、その文書内で使われている単語の重要性を表す。対象の文書内での出現頻度が高い単語ほど重要であるという考えに基づく tf 値と、いくつもの文書で横断的に出現する単語は重要でないという考えに基づく idf 値の掛け合わせで $tf-idf$ 値を計算する。

以下、本稿における tf 値と idf 値の計算方法を記す。式(1)に tf 値、式(2)に idf 値の計算式を記す。

$$tf(t, d, month) = \frac{n_{t,d,month}}{\sum_{s \in d} n_{s,d,month}}, \quad (1)$$

$$idf(t, month) = \log \frac{N}{df(t, month)}, \quad (2)$$

ここで、 $n_{t,d,month}$ は、ある名詞 t の解析する月 $month$ の文書 d 内での出現回数を示し、 $\sum_{s \in d} n_{s,d,month}$ は、解析する月 $month$ の文書 d 内での全ての単語の出現回数を示す。また、 N は分析対象とする全文書数を示し、

$df(t, month)$ はある単語 t が解析する月 $month$ に出現する文書の数を示す。つまり、 df 値を計算する文書集合を12ヶ月のレビューとすることで、各月に特徴的に現れる単語の抽出を図る。

2.2 正規化と単語の抽出

出現した名詞の $tf-idf$ 値の変化の割合を比較するために正規化を行う。データの最大値を1、最小値を0にする、式(3)に従って正規化を行う。

$$SX = \frac{X - x_{(min)}}{x_{(max)} - x_{(min)}}, \quad (3)$$

ここで、 SX 、 X はそれぞれ正規化後の値と正規化前の値を示し、 $x_{(min)}$ と $x_{(max)}$ はそれぞれデータの最小値と最大値を示す。

この正規化された数値グラフから、特定の月においてのみ、その $tf-idf$ の値の割合が大きくなっている名詞を出力する。特定の月における値の上昇を検知するために、12ヶ月分の平均値を参照とする。12ヶ月分の $tf-idf$ の平均値と最大値の差が大きい名詞を宿泊施設の季節を表す特徴語として抽出する。

ただし、以下の条件で分析を行うものとした；

- レビュー数の上位200件の宿泊施設を使用する
十分にレビュー数が存在するデータを分析対象として使用する
- 数字や、記号など宿泊施設の季節の特徴を明らかに表さない名詞を除外する
- レビュー内で出現回数が9回以下の名詞を除外する
出現回数の少ない名詞では、相対的に12ヶ月分の $tf-idf$ の平均値と最大値の差が大きくなる傾向にあるが、このような名詞は特定のユーザーのみが記述した単語である可能性が高く季節の特徴を表すとは言えない
- 「1月」や「2月」などそれぞれ単体で特定の月を表す名詞はあらかじめ除外する

3 季節を表す特徴語の抽出結果

表1に、「神戸ポートピアホテル」「横浜桜木町ワシントンホテル」「天然温泉 劔の湯 ドーミーイン富山」「ホテル阪急インターナショナル」「グランドホテル浜松」のレビューから、抽出された名詞の上位10件を示す。抽出された名詞のうち、「年末」「お正月」「お盆」「GW」「盆」「クリスマス」といった名詞は、特定の季

表 1: 5 件の宿泊施設のレビューから抽出した名詞の上位 10 件

	神戸ポートピア ホテル	横浜桜木町ワシ ントンホテル	天然温泉 劔の 湯 ドーミーイン 富山	ホテル阪急イン ターナショナル	グランドホテル 浜松
1 位	ルミナリエ	花火	盆	年末	浜松まつり
2 位	年末	要望	システム	名前	祭り
3 位	お正月	程度	煙草	椅子	ふぐ
4 位	お盆	開港	枕	レート	総合的
5 位	不要	不備	雪	クリスマス	空室
6 位	GW	不自由	風	クーポン	若干
7 位	屋外プール	込み	知人	こと	誕生日
8 位	キャンセル	近隣	一緒	タイミング	今年
9 位	渋滞	電車	朝ごはん	みたい	基本的
10 位	不愉快	ランク	おかげ	パノ ラマバス ルーム	交換

節と時期を明確に表す名詞であることがわかる。また、「ルミナリエ」や「屋外プール」「花火」「浜松まつり」は開催される時期があり、通年で催されるとは考えにくい。そのため、特定の季節のレビューに特徴的に出現すると考えられる名詞の抽出において、一定の有効性が確認されたと考えられる。

4 評価実験

主観評価実験により、抽出された名詞が宿泊施設のレビュー文の特定の季節に特徴的に出現するかどうかを評価した。比較対象として提案手法によってランク付けされた下位 5 件の名詞も評価した。実験は以下の手順で行った;

- 手順 1 : 宿泊施設のレビュー数上位 200 件のうち、任意に 5 件の宿泊施設を選択
- 手順 2 : その宿泊施設から抽出された名詞の上位 5 件と下位 5 件を選択
- 手順 3 : 被験者には上位と下位を伝えず、研究の内容を伝え、抽出された名詞が「特定の季節に特徴的に出現すると思うか」を問う
- 手順 4 : 被験者 44 人に答えを「はい」、または「いいえ」の二択で回答してもらう

4.1 実験結果

表 2~6 に、「神戸ポートピアホテル」「横浜桜木町ワシントンホテル」「天然温泉 劔の湯 ドーミーイン富山」「ホテル阪急インターナショナル」「グランドホテル浜松」の結果をそれぞれ示す。

表 2: 「神戸ポートピアホテル」についての結果。表中の人数は、「特定の季節に特徴的に出現する単語であるか?」という質問への回答者数。

	名詞	YES	NO
上位 5 件	ルミナリエ	26 人	18 人
	年末	44 人	0 人
	お正月	44 人	0 人
	お盆	44 人	0 人
	不要	2 人	42 人
下位 5 件	一流	0 人	44 人
	用意	2 人	42 人
	電話	0 人	44 人
	対応	1 人	43 人
	南館	3 人	41 人

また、図 1, 2, 3, 4, 5 に、縦軸が $tf-idf$ 値、横軸が $month$ となる月単位の $tf-idf$ 値の推移グラフをそれぞれ示す。

4.2 考察

「宿泊施設の季節の特徴を表す名詞である」と抽出された名詞の 5 つの宿泊施設の上位 5 件ずつ、計 25 件の名詞のうち、被験者の過半数が「宿泊施設のレビューに特定の季節に特徴的に出現する名詞である」と思うと回答したのは 13 件 (26%) であった。この割合が十分に高くない理由として、ホテルチェーンが経営する宿泊施設は季節の特徴がレビューに書かれにくいこ

表 3: 「横浜桜木町ワシントンホテル」についての結果. 表中の人数は、「特定の季節に特徴的に出現する単語であるか?」という質問への回答者数.

	名詞	YES	NO
上位 5 件	花火	43 人	1 人
	要望	2 人	42 人
	程度	0 人	44 人
	開港	14 人	30 人
	不備	0 人	44 人
下位 5 件	時間	4 人	40 人
	出張	2 人	42 人
	バイキング	3 人	41 人
	朝食	4 人	40 人
	立地	3 人	41 人

表 5: 「ホテル阪急インターナショナル」についての結果. 表中の人数は、「特定の季節に特徴的に出現する単語であるか?」という質問への回答者数.

	名詞	YES	NO
上位 5 件	年末	43 人	1 人
	椅子	0 人	44 人
	レート	3 人	41 人
	クリスマス	44 人	0 人
	名前	1 人	43 人
下位 5 件	窓	3 人	41 人
	綺麗	9 人	35 人
	友人	3 人	41 人
	夜景	17 人	27 人
	リクエスト	1 人	43 人

表 4: 「天然温泉 剣の湯 ドーミーイン富山」についての結果. 表中の人数は、「特定の季節に特徴的に出現する単語であるか?」という質問への回答者数.

	名詞	YES	NO
上位 5 件	盆	43 人	1 人
	システム	0 人	44 人
	煙草	2 人	41 人
	枕	0 人	44 人
	雪	44 人	0 人
下位 5 件	料金	10 人	34 人
	フロント	0 人	44 人
	快適	5 人	39 人
	サウナ	11 人	33 人
	バス停	0 人	44 人

表 6: 「グランドホテル浜松」についての結果. 表中の人数は、「特定の季節に特徴的に出現する単語であるか?」という質問への回答者数.

	名詞	YES	NO
上位 5 件	浜松まつり	38 人	6 人
	祭り	42 人	2 人
	ふぐ	30 人	14 人
	総合的	0 人	44 人
	空室	5 人	39 人
下位 5 件	金額	9 人	35 人
	豪華	1 人	43 人
	デラックス	2 人	42 人
	親切	0 人	44 人
	駅	0 人	44 人

とも関係すると考えられる. 評価実験を行った宿泊施設の「神戸ポートピアホテル」以外は全てチェーンの宿泊施設であった. ホテルチェーンが経営する宿泊施設は対象のユーザに年中一定の品質を提供するため季節の名詞が抽出されにくいと考えられる.

また出現回数が低いことによって, 誤って季節を表す語句として抽出されている名詞が見られた. 出現回数が少なく $tf-idf$ 値が年間を通じて小さい場合, 現在の提案手法では抽出される順位が高くなってしまふ. 例えば, 「要望」や「程度」は意味として季節を表す名詞として考えることは難しいが, これらの単語についての全レビューでの総出現回数が 12 回しかなく, 出現回数が小さいことによって偶然に出現回数が増加した月があったため, 抽出対象として選出されたと考えられる. 今後は, 評価式の再構成や例外処理の条件を設けることで, 抽出の精度向上を目指す.

一方で, 「ルミナリエ」や「クリスマス」といった他

の宿泊施設に出現していないような名詞や, 「年末」や「お正月」といった特定の季節を表す名詞の抽出には成功した. 図 1 の「お正月」と図 4 の「クリスマス」を見ると抽出された名詞の $tf-idf$ 値が特定の月や, 前後の月で上昇しているのがわかる. このことから特定の季節を表す名詞が, レビューに特徴的に出現する季節の $tf-idf$ 値が高くなっていることがわかる.

現段階では, 提案手法によって抽出した名詞すべてについてを, 宿泊施設の季節の特徴を表す単語とすることは難しい. 抽出された名詞を含むレビュー文を解析することで, 本研究の目的に活用できる季節の情報を抽出する可能性を探る.

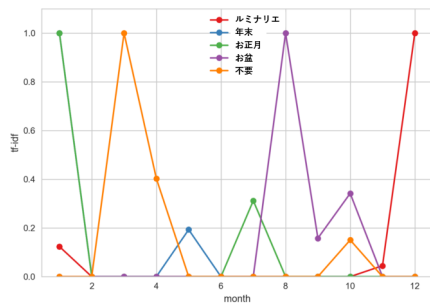


図 1: 「神戸ポートピアホテル」についての tf-idf 値の推移グラフ。

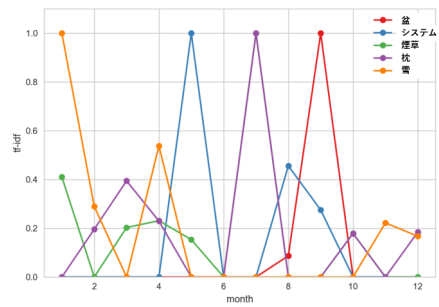


図 3: 「天然温泉 剣の湯 ドーミーイン富山」についての tf-idf 値の推移グラフ。

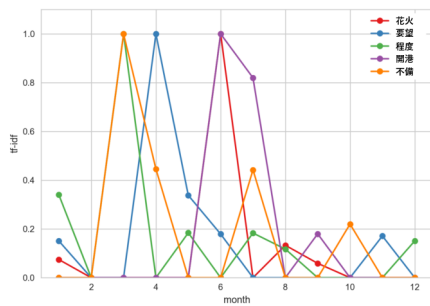


図 2: 「横浜桜木町ワシントンホテル」についての tf-idf 値の推移グラフ。

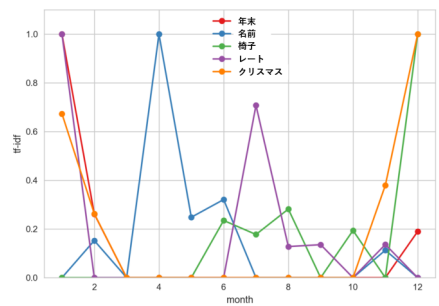


図 4: 「ホテル阪急インターナショナル」についての tf-idf 値の推移グラフ。

5 おわりに

本研究では宿泊施設のレビュー文から、ユーザーが宿泊施設の決定要因になりうる情報、または宿泊施設の広告や効果的な宣伝に活用するために、季節の特徴を表す名詞の抽出を行った。レビューを月単位に分割し、 $tf-idf$ 値の変化の割合が大きいものを抽出し、宿泊施設のレビューに出現した名詞に、季節の特徴を表しているかどうかの順位をつけた。評価実験の結果、抽出した名詞の中に、宿泊施設の季節の特徴を表していると考えられる名詞があることがわかった。

しかしながら、抽出方法の問題点もいくつか見つかった。今後は、提案手法の問題点の改善とともに、抽出された名詞を含むレビュー文を分析することによって、より詳細な季節の特徴抽出を目指す。また、抽出された情報について、ユーザーが宿泊施設の決定要因になりうる情報、または宿泊施設の広告や効果的な宣伝に利用可能な情報といった情報の評価も行っていく。

謝辞

本研究では、国立情報学研究所の IDR データセット提供サービスにより楽天株式会社から提供を受けた「楽天データセット」を利用した。記して謝意を表す。

参考文献

- [1] 田邊亘, 後藤正幸. 宿泊施設の戦略構築を支援するユーザーレビュー分析に関する一考察. メディアセンタージャーナル, Vol. 9, pp. 91-101, 2008.
- [2] 辻井康一, 津田和彦. 宿泊レビューを用いた宿泊施設サービス改善のための情報抽出. 人工知能学会全国大会論文集, pp. ROMBUNNO.2E1-5, 2013.
- [3] 辻井康一, 津田和彦. テキストマイニングを用いた宿泊レビューからの注目情報抽出方法. 情報処理学会デジタルプラクティス, Vol. 3, No. 4, pp. 289-296, 2012.

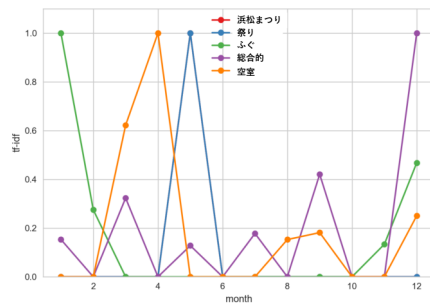


図 5: 「グランドホテル浜松」についての tf-idf 値の推移グラフ。

- [4] 中山祐輝, 藤井敦. レビューテキストを用いた条件付き意見文の抽出. 言語処理学会 第 20 回年次大会発表論文集, pp. 888-891, 2014.
- [5] 中山祐輝, 藤井敦. レビューテキストを対象とした評価条件の抽出手法. 言語処理学会第 19 回年次大会発表論文集, pp. 248-251, 2013.
- [6] 松尾哉太, 新妻弘崇 1, 太田学. レビュー解析に基づくユーザ評価の根拠提示の一手法. 情報処理学会研究報告, Vol. 2014-DBS-160, No. 14, pp. 1-6, 2014.
- [7] 乾孝司, 奥村学. テキストを対象とした評価情報の分析に関する研究動向. 自然言語処理, Vol. 13, No. 3, pp. 201-241, 2006.
- [8] 乾孝司, 梅澤佑介, 山本幹雄. 評価表現と文脈一貫性を利用した教師データ自動生成によるクレーム検出. 自然言語処理, Vol. 20, No. 5, pp. 683-706, 2013.
- [9] 打田裕樹, 吉川大弘, 古橋武, 平尾英司, 井口浩人. Web ユーザレビューにおける評価情報の時系列変化の可視化. 日本知能情報ファジィ学会誌, Vol. 22, No. 3, pp. 377-389, 2010.