

# 地域情報提供サービスのための行動目的抽出

横山 法子<sup>1,a)</sup> 船越 要<sup>1</sup> 戸田 浩之<sup>1</sup> 小池 義昌<sup>1</sup>

受付日 2014年9月20日, 採録日 2015年1月5日

**概要:** 人の行動を促すためには、「動機」と「行動のしやすさ」が情報提示の際に必要なといわれている。しかし、今までの地域情報提供サービスは、場所や金額等を伝え、「行動のしやすさ」を判断する材料を提供することにフォーカスしており、「動機」付けはあまり行われてこなかった。そのため、明確な動機を持たないユーザに対して、行動を起こさせることが難しかった。本研究では、動機付けを行うことで、元々の動機の有無にかかわらず人の行動を促すことができる情報提供を目指す。このために、他人の経験から行動にあたっての目的や、TPO 情報や手段情報といった行動に付随する情報を収集する。本稿では、この研究方針に従い、ソーシャルメディアから手がかり表現と係り受け解析を用いて行動にあたっての目的を抽出する方法を提案する。また、抽出した目的を用いてユーザ評価実験を行い、目的の提示が動機付けに有効であることを確認した。

**キーワード:** テキストマイニング, ソーシャルメディア, 行動目的抽出, 行動支援

## Extracting Purpose-for-Action to Enhance Local Information Service

NORIKO YOKOYAMA<sup>1,a)</sup> KANAME FUNAKOSHI<sup>1</sup> HIROYUKI TODA<sup>1</sup> YOSHIMASA KOIKE<sup>1</sup>

Received: September 20, 2014, Accepted: January 5, 2015

**Abstract:** To suggest a target action and encourage the user to perform it, “motivation” and “ability” are necessary when a trigger is offered. However, traditional local information services focus on giving information related to the ability to perform the behavior (cost, place, etc.) and motivation is not considered; without effective motivation the target action is unlikely to be performed. Our goal is to identify an effective motivation approach. To achieve this goal, we collect purpose-for-action. In this paper, we propose a method to extract purpose-for-action from social media texts by using clue expressions and modification structure. Moreover, we conduct a user experiment whose results confirm that showing purpose-for-action yields effective motivation.

**Keywords:** text mining, social media, purpose-for-action extraction, action support

### 1. はじめに

インターネット上には、行動の参考になる様々な情報提供サービスがある。たとえば外食の際は「食べログ」<sup>\*1</sup>等のグルメサイトから、旅行の際は「じゃらん」<sup>\*2</sup>等の旅行サイトから、施設等の情報を得ることができる。また、近年は「周辺ガイド」<sup>\*3</sup>や「Foursquare」<sup>\*4</sup>のような位置情報サービスが普及してきており、周辺の様々な行動に関する

情報を、横断的に検索できる。

このような情報提供サービスは、ユーザに適した情報を提供し、行動を起こしてもらうことが1つの目標である。従来のサービスは、施設等の情報をジャンルや場所等で分類し、また、施設に関する詳細な情報を提示することで、行動を選択しやすいよう工夫されている。しかしこのようなサービスは、「外食」「美術館」のようにユーザの行動が具体的に決まっている場合には有効であるが、ユーザが漠

<sup>1</sup> 日本電信電話株式会社 NTT サービスエボリューション研究所  
NTT Service Evolution Laboratories, Yokosuka, Kanagawa  
230-0847, Japan

a) yokoyama.noriko@lab.ntt.co.jp

\*1 食べログ, <http://tabelog.com/>

\*2 じゃらん, <http://www.jalan.net/kankou/>

\*3 周辺ガイド, <http://local.smt.docomo.ne.jp/g.php>

\*4 Foursquare, <https://ja.foursquare.com/>

然と「暇つぶしに何かしたい」等と思ったときに、選択したい行動が見つからない場合には不向きである。

この原因を心理学やマーケティングの分野で用いられる行動モデルを用いて考えてみる。行動モデルの1つに、Foggが提案するFogg Behavior Model [1]がある。このモデルは、人に期待する行動をとってもらうためには、「動機」、「行動のしやすさ」、「トリガ（行動を知らせる情報提供）」の3つの要素を同時に満たす必要があると説明している。つまり、情報提供サービスにおいては、推薦情報や検索結果がトリガとしてユーザに与えられたときに、「①動機」「②行動のしやすさ」が必要である。

たとえば、サービスが推薦したい行動が「散歩」であるとき、気分転換をしたいと思っているユーザは、動機を持ち、かつ簡単な行動であるので、「散歩」をしてもらえる可能性が高い（図1）。反対に、動機がない、または行動への障壁があると、情報提供が失敗しやすい。つまり情報提供において、①動機付けと②行動障壁を下げるのが重要な課題といえる。

②行動のしやすさは、時間やお金、身体的負担のなさ等がその代表的要素であり、これらの障壁自体を下げることは難しい。しかし、ユーザが障壁を高く見積もっている場合もあり、こういったユーザに対して、情報を整理し提示することで、正しい判断を助け、行動に導ける可能性もある。また情報を収集するという行動への障壁を下げることも可能となる。「周辺ガイド」等の従来のサービスでは、分かりやすく距離や値段等の施設情報が提示されており、②行動への障壁を下げる工夫がなされている。

一方、①動機は、喜びや希望、社会的受容等がその代表的な要素であるが、①行動への動機付けに関するサービスは限定的な方法でしか検討されてこなかった。クーポンにより、経済的な得をすることを動機付けに用いる方法は頻繁に行われており効果も認められている。しかし、クーポンを使える範囲は限られているため、幅広い行動に対応できる動機付けの方法が必要となる。

そこで、本研究は、ユーザにトリガを与える際、動機付けも同時に行い、納得して具体的な行動につなげる情報提供に着目する。動機付けを行う方法には、行動にあたって

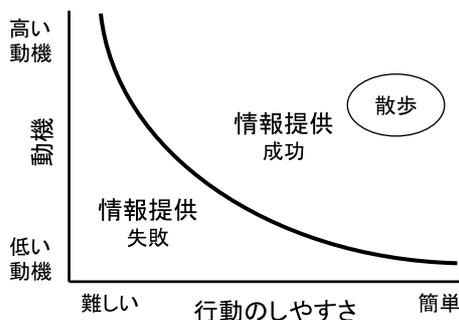


図1 Fogg behavior model  
Fig. 1 Fogg behavior model.

の目的（以下、単に「目的」と呼ぶ）を提示する方法や、他者の経験情報を提示することで、information cascade（他者と同じ行動をとろうとする心理）[2]を起こさせる方法等がある。このような目的や他者の経験情報は、幅広い行動に対して収集できるものであり、従来限定的であった動機付けの課題を解決できる。

そして具体的な行動への動機付けができれば、ユーザの行動支援が可能である。たとえば、施設情報が記述された情報サイトにおいて、単に施設を羅列するだけでなく、目的と関連付けて施設を紹介でき、施設を選びやすくすることが可能である。

情報提供方法として様々な形式が考えられるが、たとえば、TPO情報を考慮した地域情報提供サービス [3]を想定した場合、情報提供は図2の流れで行う。また、アプリケーション例を図3に示す。カテゴリごとに情報が整理されたサービスを想定し、大きく2つのステップから構成されている。第1ステップでは、サービスは選択された場所で行える行動と目的、他人の経験情報をユーザに提示する。これにより、行動への動機付けを行う。そして第2ステップでは、選択された行動に関連する施設情報等の詳細情報をユーザに提示する。この第2ステップは、従来の情報提供サービスと同じである。本研究が目指す動機付けはこの第1ステップであり、図2の太枠で囲んだ部分に該当する。提示する行動への動機付けを行い、行動に関連する

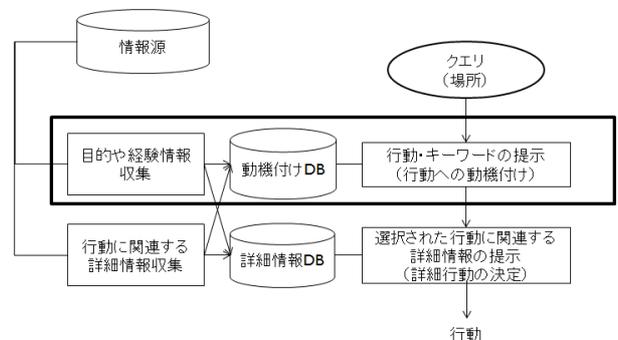


図2 本研究が対象とする領域  
Fig. 2 Target area of this paper.



図3 アプリケーション例  
Fig. 3 Example of application image.

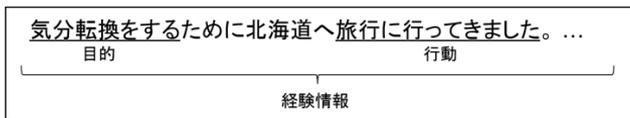


図 4 経験情報からの行動目的抽出

Fig. 4 Extracting purpose-for-action from experience information.

詳細情報への取得要求を起こさせる，または実際に行動を起こさせることを目的とする。

そこで本稿では，この目的実現に向けた第 1 歩として，目的の収集に取り組み，大量の経験や目的が記述されたソーシャルメディアから，指定した行動の目的を抽出する手法について述べる (図 4)。また，ユーザ評価実験を行い，提案手法を用いて抽出した目的の提示が行動への動機付けに有効であるかを評価した結果を報告する。

以下，本稿の構成を示す。まず，2 章で関連研究について述べ，3 章で問題設定，4 章で予備分析，5 章で提案手法について述べる。6 章で目的抽出実験，7 章でユーザ評価実験，最後に 8 章でまとめと今後の課題を述べる。

## 2. 関連研究

### 2.1 テキストからの目的抽出

テキストからある行動の目的を抽出する方法として，2 つのアプローチが考えられる。

#### (1) 目的を表す特徴語に注目するアプローチ

(例)「達成感」等，目的を表す特徴語をあらかじめ辞書に登録し，これを含む文を目的として抽出する。

#### (2) 目的が記述される文のパターンに注目するアプローチ

(例)「～ため」等の表現で行動に接続されている文を目的として抽出する。

(1) の手法は，抽出したい目的が事前に明らかな場合に有効であり，「○○のために○○する」のように行動と目的の関係の手がかりとなる表現が明示的に記述されていなくても，目的を抽出できる可能性が高い。しかし，抽出したい目的が未知の場合には抽出できない。(2) の手法は，抽出したい目的がはっきりしていない場合にも抽出可能であるが，接続助詞等で行動との関係が記述されていない場合には抽出が難しい。

(1) の手法には，旅行ブログサイトから旅行に関わる目的を抽出する研究がある [4]。この研究では，目的が旅行に関わるものに限定されているため，様々な行動の目的を抽出するためには，辞書の拡張が必要である。また，様々な行動に関係する目的辞書の作成については，長期的な行動目的の辞書を構築する研究がある [5]。この研究では，「新年の決意」を含む twitter の投稿から LDA を用いて目的トピックを抽出している。目的を単体で抽出しているが，我々は行動-目的ペアの抽出を目指す。行動-目的ペアの関係は，要因と結果からなる因果関係に含まれる概念である。

(2) の手法は，ニュース記事から因果関係を抽出する研究で多数用いられている [6], [7], [8], [9]。たとえば，「～ため，」や「～を背景に」等の手がかり表現と，構文パターンを用いて，因果関係を含む文を抽出している [9]。これらの研究では，ニュース記事から因果関係を抽出しており，正しい文法で記述される文を対象としている。また，原因・理由の用法と目的の用法を区別せずに，両者を要因として抽出している。我々が対象とするのはソーシャルメディアであり，文法的な誤りや砕けた表現を多く含むため，構文解析だけでは目的ではないものを抽出してしまうことがある。また，原因・理由の用法を含まない目的のみを抽出する。そのため，抽出した目的のフィルタリングが必要となる。

我々は (1), (2) の 2 つのアプローチを組み合わせ，まず (2) のアプローチにより，手がかり表現等を用いて目的を抽出する。次に，(2) で取得した目的を辞書として用いる (1) のアプローチにより，抽出事例を増やすことを目指す。

本稿ではまず (2) の手がかり表現を用いて目的を抽出する方法を提案する。機械学習を用いて目的を含むパターンを学習すること等も考えられるが，まずはルールベースの手法を試みたくうで，今後利用を検討する。

### 2.2 地域情報提供サービス

外出時の参考になる地域情報提供サービスは多数存在する。たとえば，「Foursquare」を利用すれば，指定したエリア周辺の店舗やスポット情報を取得できる。こういったサービスは，場所や入場料等の施設情報をユーザに提供することで，情報を収集するという「行動への障壁を下げること」ができる。また，他人の経験情報が記述されたレビューが掲載されていれば，information cascade を起こすことができ，「動機付け」の効果が期待できる。

しかし，行動に対する目的が端的に提示されておらず，また経験情報がどの施設に対しても記述されているわけではない。本技術を用いて行動目的を提示し動機付けを強化することにより，より施設を選びやすくなる等の効果が期待できる。

また，別の観点では，ユーザの嗜好や状況に応じた情報を提供しようとする研究がある。ユーザの嗜好に合った情報提供の例として，「Google」\*5 はユーザの検索履歴を利用したり，「Amazon」\*6 は類似ユーザの情報を利用したりし，コンテンツや商品の推薦を行っている。また，状況に応じた情報提供を目指す研究として，時間や同伴者等をもとに飲食店情報の推薦を行う研究がある [10]。個人の嗜好や状況に合った行動が推薦できれば，動機付けを行う必要はなく，行動を実行してくれる可能性がある。しかし，これらを実現するためには，ユーザ情報や状況を取得する必要がある。

\*5 <https://www.google.co.jp/>

\*6 <http://www.amazon.com/>

ある。本研究では、ユーザ情報や状況を必要としない一般的な状況を想定した情報提供技術を提案する。また、これらの技術は、どういったアイテム（たとえば店舗）を推薦するかということにフォーカスしており、どういった情報（行動目的等）を付加して提示するかということには着目していない。今後、ユーザの嗜好や状況等を考慮した情報提供を目指す際には、従来の技術と本技術を組み合わせることで、個人の嗜好や状況に応じた、より効果的な情報提供を目指す。

### 3. 問題設定

本稿では、文法的な誤りやくだけた表現を含むテキストから、原因・理由の用法を含まない、指定した行動の目的を抽出し、目的抽出の精度向上を目指す。ここで抽出したい目的とは、指定した行動を行うことで書き手が達成したい行動や状態、または指定した行動を行うことで得られる効果であると読み手が判断できるものと定義する（図 5）。目的は行動に対して相対的に定義されるものであり、たとえば、「気分転換するためにサイクリングをする」という文章があった場合、「気分転換する」が、「サイクリングする」ことで達成したい行動であるので、目的となる。また、「走ったら気分転換になった」という文章の場合、「走る」ことで「気分転換」という効果が得られたと判断できるので、「走る」ことによって「気分転換」という目的が達成可能と考える。指定する行動は、行動の具体化支援という本研究の観点から、行動を行うか行わないか、また何を行うかの選択の余地が大きい行動とする。その中の1つとして、本稿では市場の大きいレジャー活動（特に外出をともなうもの）にフォーカスして検討を行う。以下本稿で使用する言葉の定義をする。

#### 定義 1 (手がかり表現)

指定した行動の目的を抽出するためには、文章の中での行動と目的との関係を把握する必要がある。「ために」のように、行動と目的を接続する表現を手がかり表現と呼び、目的を抽出する際に利用する。

#### 定義 2 (目的クラス)

抽出した目的のうち、内容が類似しているものがあればまとめ、目的クラスと呼ぶ。たとえば「健康のために」「健康の為に」といった文であれば、「健康のために」といった目的クラスにまとめる。また、同一の目的クラスに属する

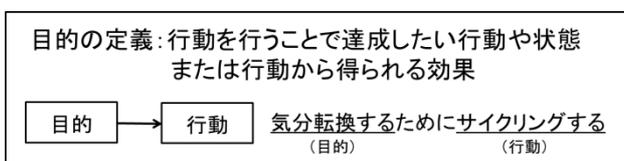


図 5 目的の定義

Fig. 5 Definition of purpose-for-action.

目的の個数を「目的クラスの要素数」と定義する。

## 4. 予備分析

本章では、分析対象として選択したデータセットの説明と、そこでの目的の記述のされ方（使用される手がかり表現）を分析した結果を述べる。

### 4.1 分析対象のデータ

使用するデータセットの条件として、記述された経験が大量に含まれていること、行動の目的がはっきりと記述されていることが必要である。経験が記述された公開データとして、マイクロブログ、Q&A サイト、ブログが考えられるが、その中でまずは1件あたりの記述量が多いブログを対象にする。対象データは今後拡張予定である。

### 4.2 データの分析 (手がかり表現の選定)

まず、ブログにおける行動の目的の記述のされ方を確認する。既存の因果関係抽出技術においては、「ために」「を背景に」等の手がかり表現を使用し、因果関係の抽出を行っているが、分析するデータによって使用される手がかり表現が変化することが指摘されている [9]。そこで、ブログというニュース記事よりもくだけた文体で記述される文章において、使用されている手がかり表現の種別、および行動に対する目的が記述される割合を調査する。

まず、レジャー活動の中から10種類の行動を選び、それらの行動を綴ったブログ（100件×10種類）から手動で行動の目的を抽出した。レジャー活動のサブカテゴリとして、スポーツ、趣味・創作、娯楽、観光・行楽等があげられる [11] ので、この中から10種類の行動を選定した。目的を抽出する際には、目的かどうか判断がつきにくい場合がある。そこで、内閣府の生涯学習世論調査 [12] で使用されている「生涯学習をしてみたい理由」12個を参考にし、12個の理由に関連した記述は、無条件で目的であるとした。たとえば、「体力がついてきた」という文章は、「健康・体力づくりのため」という理由と類似しているため、目的と判断する。この結果、430個の行動目的を抽出した。

次に、抽出した目的と行動を結ぶ文字列を、手がかり表現として抽出した。目的と行動が同一文中にない場合は、抽出対象外とした。結果として抽出された手がかり表現を表 1 に示す。目的と行動の記述方法は、目的と行動の出現位置によって大きく以下の2つに分類できる。

- (1) 目的, 手がかり表現, 行動の順番で記述  
(例) 「ダイエットするために運動する」
- (2) 行動, 手がかり表現, 目的の順番で記述  
(例) 「走ったら気分転換できた」

表 1 では、(1), (2) のどちらかの用法でのみ使用されていた手がかり表現と、(3) 両方の用法で使用されていた手がかり表現とで結果を分けている。目的と行動を同一文中

に含む事例のうち、表 1 の (1), (2) の手がかり表現を含む事例数はほぼ同数であった。本稿では、まず (1) の用法で記述される目的の抽出を行う。

そこで、(1) の用法についてさらに実験を行った。「目的 ⇒ 行動」の順序で使用されていた手がかり表現のうち、正答率が低いと予想される 1 文字からなるものを除いた 15 個の手がかり表現を対象とする。「手がかり表現+行動 (例：を兼ねてサイクリング)」が記述されている文を抽出し、直前の文章に目的が記述されているか評価した。データは、2013 年 2 月 1 日の 1 日分のブログ約 100 万件を対象とした。行動は前述の 10 種類の行動のうち、外出をともなっている 5 種類の行動 (旅行, 外食, 美術館, 遊園地, サイクリング) とした。3 章で行った目的の定義より、たとえば「雨が降ったためサイクリングは中止」のように原因が記述されているものは目的ではないと判断する。評価の結果を表 2 に示す。事例数が 0 の手がかり表現は表から削除した。

ここで対象とすべき手がかり表現は、事例が多く、高い確率で目的を抽出できるものである。その観点では、「には、なので」は、事例数は多いが、正答率が低い。その次に事例数が多い「ために (為に)」は、正答率も高い。以上の分析から、本稿ではまず手がかり表現「ために (為に)」

表 1 ブログから抽出した手がかり表現

Table 1 Clue expressions extracted from blogs.

(1)目的⇒行動	をかねて (or を兼ねて), ため (or 為), ため (or 為) に, のために (or の為に), を求めて, 思うから, と 思って, には, なので, たくて, という方には, と思い, ということになり, からってことで, ようと, が, を, に,
(2)行動⇒目的	は〇〇のため, って, は〇〇が目的, たら, でも, が, 時の, も, てみせます!, ことで, したり, だけでも, も, のが, けど,
(3)どちらも	で, から, (直後の動詞にかかる), (直後の名詞にかかる), は, ので, (文が連続), して, て, と, し, の, 事で

表 2 手がかり表現の出現頻度

Table 2 Frequency of clue expressions.

行動	事例数	正解数	正答率
をかねて or を兼ねて	2	2	1.00
ため or 為	8	4	0.50
ために or 為に	15	12	0.80
のために or の為に	9	7	0.78
を求めて	1	1	1.00
と 思って	1	0	0.00
には	69	2	0.03
なので	24	1	0.04
たくて	3	3	1.00
と思い	1	1	1.00
ようと	2	2	1.00
合計	137	35	0.26

を用いて目的抽出を行う。

## 5. 提案手法

本章では、「ために (為に)」を手がかり表現として用いた目的抽出手法の詳細を述べる。

提案手法は、目的抽出、目的の絞り込みの 2 つの技術から構成される。以下それぞれの技術について述べる。

### 5.1 目的抽出

手がかり表現と係り受け解析を用いて指定した行動の目的を抽出する。係り受け解析を用いた目的抽出は乾らの研究 [6] と類似しているが、本研究では指定した行動の判定を自動で行うために、文節数に工夫を加えた。また、誤り除去や絞り込みを行い、原因・理由の用法やその他の誤抽出を排除する点でも乾らの研究 [6] と異なる。図 6 に示すとおり、目的抽出は 4 つのステップから構成される。

**ステップ 1.** 手がかり表現と指定した行動を表すキーワードを含む文をデータセットから抽出する。

**ステップ 2.** 抽出した文に係り受け解析を適用する。解析には係り受け解析器 Jdep [13] を用いた。

**ステップ 3.** 手がかり表現によって修飾される文節 (第 1 文節と呼ぶ) を抽出し、さらに第 1 文節を修飾する最も近い文節を第 2 文節として抽出する。そして、第 1, 第 2 文節の 2 文節に、指定した行動を表すキーワードが記述されているかを判断する。

**ステップ 4.** ステップ 3 で指定した行動が記述されていれば、手がかり表現を修飾する文節を第 1 文節として抽出し、第 1 文節を修飾する文節を第 2 文節とする。そして、第 1, 第 2 文節の 2 文節を、目的として抽出する。

指定した行動を表すキーワードとは、たとえば「サイクリング」のように、「サイクリングする」という行動を端的に表すキーワードである。ステップ 3, 4 において、第 1 文節を修飾する文節が複数ある場合には、第 1 文節に最

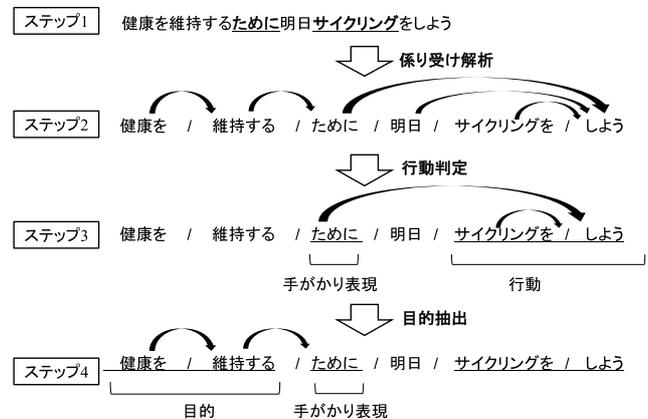


図 6 手がかり表現と係り受け解析を用いた目的抽出

Fig. 6 Purpose-for-action extraction using clue expression and modification structure.

も近い文節1つを第2文節として採用する。文節数に関して、行動や目的として、「動詞+目的語」のセットを抽出することを想定し2文節とした。文節数は検討の余地があり、6章で再度検討する。

また、2.1節で述べたとおり、目的のフィルタリングが必要である。次節は、特に文法的な誤りやくだけた表現を含む文章から誤って抽出した目的を取り除くフィルタリングの一手法である。実験結果をふまえ、原因・理由の用法を排除するルールベースのフィルタリング手法も6.1節で提案する。

## 5.2 目的の絞り込み (適合率の向上)

5.1節で抽出した目的候補のうち、出現頻度の多い目的を採用することにより、ノイズを軽減させる。文法的な誤りやくだけた表現を含む文章から誤って抽出した目的は、複数回出現する可能性は低いと考えられるため、これらの誤りが排除できると考えられる。絞り込みの手順は下記のとおりである。

**ステップ1.** 目的を形態素に分割し、名詞・動詞のみを抽出する。

**ステップ2.** 名詞・動詞が完全に一致している目的どうしを同じ目的クラスとしてまとめる。

**ステップ3.** 各目的クラスに含まれる目的のうち、最も頻度の多い目的、または文字列が最も長い目的をその目的のクラス名とする。目的が1種類しか存在しない目的クラスについては、その目的をクラス名とする。

**ステップ4.** 目的クラスの要素数が2以上の目的クラス名を目的として採用する。

## 6. 目的抽出実験

提案手法の妥当性を確かめるために実験を行った。適合率と再現率を総合的に評価するF値を用いて提案手法の精度を評価する。データセットは2013年2月の1カ月分のブログであり、行動は、予備分析で用いた5種類の行動(旅行、外食、美術館、遊園地、サイクリング)を指定し、指定した行動を表すキーワードも上記を使用した。

### 6.1 目的抽出

指定した5種類の行動の目的を抽出する実験を行った。

提案手法として、上述のブログから「ために」と係り受け解析を用いて目的を抽出する。行動抽出範囲は、5.1節に記述した2文節に加え、1文節、3文節、全文節とした場合も試す。

ベースラインとして以下の2つの手法を用いた。

BL1: 「ために」の直後に行動が続いている場合に、係り受け解析[13]を適用し、5.1節のステップ4と同様の処理により目的を抽出する方法

BL2: 直後を含め「ために」の後方に行動が記述されて

いる場合に、係り受け解析[13]を適用し、5.1節のステップ4と同様の処理により目的を抽出する方法  
目的抽出範囲は、提案手法・ベースラインともに、「動詞+目的語」を抽出することを想定し、2文節とした。

それぞれの手法で抽出した結果に対し、人手で指定した行動の目的かどうかの判定を行い、適合率、再現率、F値を算出した。判定の基準としては、指定した行動・目的ともに正しい場合を正解とした。指定した5種類の行動(旅行、外食、美術館、遊園地、サイクリング)は、旅行に行く、外食をする、美術館に行く、遊園地に行く、サイクリングするという意味で記述されている場合を正解とし、たとえば「健康のために外食を控える」と記述されていた場合には不正解とする。目的は、行動を行うことで書き手が達成したい行動や状態、または行動を行うことで得られる効果であると読み手が判断できるものを正解とした。また、適合率、再現率、F値はそれぞれ以下の式で求めた。

$$\text{適合率} = \frac{\text{正解事例数}}{\text{抽出した事例数}}$$

$$\text{再現率} = \frac{\text{正解事例数}}{\text{全事例数}}$$

$$F \text{ 値} = \frac{2 \times \text{適合率} \times \text{再現率}}{\text{適合率} + \text{再現率}}$$

ただし、BL2は指定した行動を表すキーワードが後方のどの位置にある場合も抽出するため最も基準が甘く、BL1、提案手法(2文節)、文節数を変化させた手法を包括する手法である。そこで、BL2が抽出した正解事例数を、再現率の分母にあたる全事例数とした。

抽出結果を表3、表4、表5に示す。表5が示すとおり、提案手法(2文節)がベースラインに比べ高い適合率を

表3 正解事例数

Table 3 The number of correct examples.

行動	BL1	BL2	1文節	2文節	3文節	全文節
旅行	58	245	55	177	180	194
外食	18	31	19	27	28	29
美術館	9	44	6	29	31	35
遊園地	8	15	3	11	12	15
サイクリング	3	20	6	15	15	16
合計	96	355	89	259	266	289

表4 抽出した事例数

Table 4 The number of extracted examples.

行動	BL1	BL2	1文節	2文節	3文節	全文節
旅行	108	700	81	298	322	421
外食	30	112	30	55	62	76
美術館	13	94	6	35	40	55
遊園地	11	23	3	13	14	17
サイクリング	3	30	8	18	18	20
合計	165	959	128	419	456	589

表 5 抽出精度

Table 5 Extraction accuracy.

	適合率	再現率	F 値
BL1	0.58	0.27	0.37
BL2	0.37	<b>1.00</b>	0.54
1 文節	<b>0.70</b>	0.25	0.37
2 文節	0.62	0.73	<b>0.67</b>
3 文節	0.58	0.75	0.66
全文節	0.49	0.81	0.61

示した。BL1 の結果を見ると分かる通り、「ために」の直後に指定した行動を表すキーワードがある場合でも、その行動に関する目的となっているとは限らない。係り受け解析を用いることで、たとえば、「この日のために旅行の時にのみやげを買った」といった文を排除できた。適合率の平均は、1 文節が最も高い値を示したが、再現率は低い。適合率と再現率とを合わせた F 値は提案手法 (2 文節) が最も高く、2 文節とした提案手法の効果が確認できた。

目的の抽出において、目的でないものを目的として抽出してしまった誤りがある。これを本稿では抽出誤り (FP: False Positive) と呼ぶ。提案手法 (2 文節) の抽出誤り (FP) を分析した結果、以下の原因が明らかになった。

(1) 「ために」の問題

- ・原因・理由が抽出されてしまう (例)「雨のために」
- ・定型文の一部として使用されているものが抽出されてしまう (例)「何のために」

(2) 行動判定の問題

- ・否定の意味で使われている (例)「外食を控える」
- ・行動が違う (例)「旅行保険を申し込む」

(3) 目的抽出の問題

- ・目的が一部しか抽出されていない (例)「見るために」

(4) 行動判定/目的抽出の問題

- ・係り受け解析ミス (括弧がある場合や目的が並列で並べられている場合等) (例)「欲しいものを買うためにも旅行に行くためにも」

(5) 前処理の問題

- ・スパム

そこで、5.1 節で述べたとおり、これらの抽出誤り (FP) を除去するために、時制や品詞等で除去手法を検討した。大まかな改善方針は表 6 に示す。

たとえば、原因・理由の除去に関しては、以下のとおり、時制判定と言い換え判定により誤り除去を行う。

ステップ 1. 行動目的部の時制を調べ、過去形のは「原因・理由」と判断し削除する。

ステップ 2. 「～せいで」等の手がかり表現を連結し、利用例を調査。「～ために」の出現頻度よりも多ければ、「原因・理由」と判断し削除する。

提案手法 (2 文節) により抽出した目的に対して、除去

表 6 誤り除去方針

Table 6 Initial improvement policies.

誤りの要因	誤り除去方針
原因・理由	時制判定, 述語選好性 (「～せいで」に連結し, 事例数の多いものを排除)
何のために	Stop word (「何のために」)
否定	辞書 (否定行動に対する目的を辞書に格納)
行動が違う	名詞への連結 (クエリの直後に名詞がある場合を削除)
並列	Stop word (ためにが 2 回以上出現)
目的が一部	目的語の有無を判定 (目的の前に目的語がくるものを冗長な文として排除)

表 7 誤り数の変化 (クローズドテスト)

Table 7 The change in the number of errors (close test).

誤りの要因	誤り数	誤り数 (除去後)
原因・理由	15	<b>4</b>
何のために	3	<b>0</b>
否定	30	<b>15</b>
行動が違う	64	<b>21</b>
係り受け解析	14	<b>8</b>
目的が一部	23	<b>12</b>
スパム	11	<b>2</b>
(誤り合計)	160	<b>62</b>

表 8 誤り除去 (オープンテスト)

Table 8 Removal of errors (open test).

	適合率	再現率	F 値
BL2(誤り除去前)	0.31(189/611)	<b>1.00</b>	0.47
2 文節(誤り除去前)	0.57(166/289)	0.88	0.69
BL2(誤り除去後)	0.38(166/442)	0.88	0.53
2 文節(誤り除去後)	<b>0.75(145/194)</b>	0.77	<b>0.76</b>

手法を適用した結果、誤り数は表 7 に示すとおりとなり、適合率の平均は 62%から 78%に、F 値は 67%から 70%に改善された。

また、誤り除去方法がデータセットに依存せず、一般的にも有効に機能するかどうかを確認するために、2014 年 2 月の 1 カ月分のブログを用いて、目的抽出、誤り除去を行った。目的抽出には、提案手法 (2 文節) と BL2 を用いた。BL2 を用いた理由は、BL2 の再現率が最も高く、誤り除去による適合率向上により、F 値が提案手法を超える可能性が高いと考えたからである。

それぞれの抽出精度を表 8 に示す。誤り除去による適合率、F 値の向上が確認できた。また BL2 に比べ提案手法 (2 文節) の適合率、F 値が高く、提案手法の有効性が確認できた。

提案手法の適合率は 8 割近くあり、また事例数も多いこ

表 9 目的の絞り込み  
Table 9 Clustering.

行動	誤り除去後の適合率	$P@Nc \geq 2$
旅行	0.80(154/193)	<b>0.94(15/16)</b>
外食	0.62(23/37)	<b>1.00(2/2)</b>
美術館	0.81(25/31)	<b>1.00(3/3)</b>
遊園地	0.82(9/11)	0
サイクリング	0.92(12/13)	0

とから、実用に耐えうる精度と考えられる。行動の判定の仕方や目的の長さを検討することで、さらなる適合率向上が期待できる。

### 6.2 目的の絞り込み (適合率の向上)

6.1 節で見たように、抽出誤り (FP) がある。これらを除去し、さらなる適合率向上を目指して、目的の絞り込みを行う。6.1 節の提案手法 (2 文節) を用いて目的を抽出し、誤り除去を行った表 7 の結果をクラスタリングし、6.1 節で付与した評価結果を用いて、 $P@Nc \geq 2$  (クラスの要素数が 2 以上の目的クラスの適合率の平均) を算出した。評価結果を表 9 に示す。

遊園地、サイクリング、外食、美術館に関しては、事例数が少なく、効果の判断が難しい。今後データ数を増やし再実験する。事例数が比較的多かった旅行に関しては、適合率の改善が確認できた。また、抽出された旅行の目的は、「食事」「人に会う」「お祝い」「気分転換」等であり、様々な目的が抽出できた。しかし、絞り込み前の結果と比べると、事例数は極端に少なくなってしまう。これらの結果より、適合率と再現率を考慮した F 値の観点では、誤り除去後の結果をそのまま用いる方が良いといえる。そして、事例数を犠牲にしてもかまわず、適合率を重視する場面においては、適合率の高い本絞り込み手法が有効であるといえる。事例数を落とさない適合率の向上は今後の検討課題である。

## 7. ユーザ評価実験

6 章の方法で抽出した目的を用いてアンケート調査を行い、本手法により抽出した目的の提示による動機付けの効果を確認した。動機付け効果の高い目的をどのように選択するかは今後の検討課題とし、まずは複数提示する目的の中で 1 つでも行動への意欲を高めるものがあれば、動機付けの効果があると判断する。

### 7.1 実験方法

#### 質問内容

##### (質問 1)

場所を入力すると、その場所でのおすすめ行動が推薦される情報検索サービスから、行動を単体で推薦された状況

表 10 質問 2 の例

Table 10 Example of question 2.

ランニング	行動意欲が高まるか				
	かなり高まる	高まる	変わらない	下がる	かなり下がる
↓目的					
体力をつける	<input type="checkbox"/>				
健康維持	<input type="checkbox"/>				
考え事をする	<input type="checkbox"/>				
...	<input type="checkbox"/>				

(ランニングはどうか?) を想定してもらい、提示される行動を起こしたいかを判断してもらう。

##### (質問 2)

同様のサービスにおいて、「行動+目的」を推薦された場合 (ex. ダイエットのためにランニングはどうか?) に、行動意欲が増すかを 5 段階で評価してもらう (表 10)。

(1) で単体の行動を提示された場合に、行動を起こしたいと答えた人には、(2) において、仮に (1) で行動を起こしたくないと思った場合を想定して答えてもらった。これは、我々が対象としているのは、行動単体では行動を起こしたいと思わない人を、動機付けすることを目指しているからである。

#### 提示する行動・目的

6 章で用いた 5 つの行動を含めた 11 行動 (ランニング、サイクリング、映画館、美術館、外食、カラオケ、旅行、遊園地、ショッピング、バーベキュー、エステ) について質問した。これらの行動は、レジャー活動白書 [11] に記述されていた 5 つのサブカテゴリ (スポーツ部門/趣味・創作部門/娯楽部門/観光・行楽部門/その他部門) から均等に選択した。

提示する目的は、2013 年 1 月から 2013 年 12 月までの 1 年分のブログ約 4 億件から、6.1 節の方法で抽出した。6.2 節の絞り込みの結果、要素数 1 となった目的クラスも含め、できるだけ多くの目的について動機付け効果を検証するために、抽出した目的を手動でクラスタリングした。クラスタリングは、意味の近い目的を徐々にまとめる親和図法 [14] を用いて行った。評価をするユーザの負担を考え、各行動に対する目的クラス数は最大 50 クラスとし、最終的に行動ごとに 18~50 個 (平均 36 個) の目的クラスを作成し、アンケートに用いた。

#### 被験者

1 行動あたり 400 人の被験者に協力してもらった。

### 7.2 結果と考察

行動意欲が増す (かなり高まる or 高まる) と回答した目的の数  $x$  を 6 つの区分に分け、行動ごとに各区分の人数を算出した結果を表 11 に示す。1 つ以上の目的によって行動意欲が増したことを示す  $x > 0$  の区分の平均人数は 342 人であり、約 85% の人に、いずれかの目的による動機付けの効果が確認できた。これより、ブログから抽出した目的を

表 11 情報提供システムの評価

Table 11 Evaluation of the effect of system.

	x=0	1<x<10	10≤ x<20	20≤ x<30	30≤ x<40	40≤x
ランニング	69	102	72	67	49	41
サイクリング	76	130	157	37	0	0
映画館	51	121	134	85	9	0
美術館	58	107	156	79	0	0
外食	47	133	131	65	24	0
カラオケ	61	132	89	64	28	26
旅行	32	73	92	81	72	50
遊園地	50	188	162	0	0	0
ショッピング	50	105	108	63	42	32
パーベキュー	56	189	155	0	0	0
エステ	87	86	56	59	54	58
平均	59	123	119	55	25	19

表 12 目的スコアのランキング例 (ランニング)

Table 12 Examples of purpose-for-action scores (Running).

順位	目的	行動意欲が高まるか		
		かなり高まる +高まる	変わらない	かなり下がる +下がる
1	健康維持	0.54	0.43	0.03
2	体力をつける	0.52	0.45	0.03
3	基礎代謝を上げる	0.51	0.45	0.04
4	カロリー消費	0.47	0.48	0.05
5	気分転換	0.46	0.47	0.07
⋮				
41	体力を消耗する	0.22	0.60	0.18
42	暇つぶし	0.21	0.63	0.17
43	イベント(デート、 結婚式、旅行等)	0.19	0.67	0.14
44	孤独を埋める	0.16	0.68	0.17
45	役作り	0.12	0.71	0.17

提示することにより、行動意欲を高め、行動を促すことができることが分かった。

ただし、動機付けの効果を示す指標となる「かなり高まる + 高まる」の割合を「目的スコア」と定義すると、目的ごとの目的スコアは7%~75%の範囲をとり、提示する目的間で動機付けの効果が大きな差が見られた。ランニングに関する目的スコアの例を表 12 に示す。高い目的スコアを示したものは、「健康維持」「気分転換」「楽しむ」等であり、これらの目的は多くの人々が抱く目的に合致するものと考えられる。一方、目的スコアの低かった目的は、「役作り」「ブログを書く」等であり、これらの目的は限定したユーザのみがいただく目的と考えられる。この結果から、動機付けの効果が高い目的は、ユーザの目的や状況に合致したものであることが分かった。各ユーザに提示した目的のうち、行動意欲が高まった目的数は平均 13 個であり、目的提示数に対する割合は約 36%であったため、提示する目的を絞った場合でも、効果的な目的を提示できる可能性は高いと考えられる。ユーザの嗜好や状況等も考慮した効果的な目的の選定は、今後の検討課題とする。

## 8. まとめと今後の課題

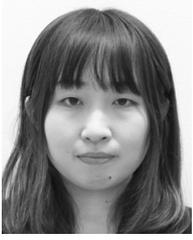
本稿では、ユーザを動機付けし、納得して具体的な行動につなげることができる情報提供サービスを目指し、ソーシャルメディアから目的の抽出を行う方法と、それを用い

た地域情報提供サービスを提案した。評価実験を行い、提案手法が目的抽出と動機付けに有効であることを確認した。

今後は、他人の経験から 5W1H 情報の抽出を行い、同伴者等の状況に応じた目的や行動の選択を行う。また、図 2 で示したとおり、行動に関連する詳細な情報提示も動機付けや行動障壁を下げるために必要である。既存のサービスで提供している情報とあわせて、行動に関連する詳細な情報を提供することで、行動を促す地域情報提供システムの確立を目指す。

## 参考文献

- [1] Fogg, B.J.: A Behavior Model for Persuasive Design, *Proc. 4th International Conference on Persuasive Technology*, Article No.40 (2009).
- [2] Bikhchandani, S., Hirshleifer, D. and Welch, I.: Learning from the Behavior of Others: Conformity, Fads, and Informational Cascades, *Journal of Economic Perspectives*, Vol.12, No.3, pp.151-170 (1998).
- [3] 藤田尚樹, 数原良彦, 片岡良治: 地域特有の話題発見につながるスマートフォン向け検索サービス: 発見探地図エリアダス, *DEIM Forum 2012*, F8-2 (2012).
- [4] 若木裕美, 有賀康顕, 中田康太, 藤井寛子, 住田一男, 鈴木優: 口コミ情報からの目的情報抽出, *FIT2011* (第 10 回情報科学技術フォーラム) (2011).
- [5] Zhu, D., Fukazawa, Y., Karapetsas, E. and Ota, J.: Intuitive Topic Discovery by Incorporating Word-Pair's Connection into LDA, *Proc. 2012 IEEE/WIC/AIM International Conferences on Web Intelligence and Intelligent Agent Technology*, pp.303-310 (2012).
- [6] 乾 孝司, 乾健太郎, 松本裕治: 接続標識「ため」に基づく文書集合からの因果関係知識の自動獲得, *情報処理学会論文誌*, Vol.45, No.3, pp.919-933 (2004).
- [7] Sakai, H. and Masuyama, S.: Cause information extraction from financial articles concerning business performance, *IEICE Trans. Inf. Syst.*, E91-D(4), pp.959-968 (2008).
- [8] Sakaji, H., Sakai, H. and Masuyama, S.: Automatic extraction of basis expressions that indicate economic trends, *Proc. PAKDD 2008*, pp.977-984 (2008).
- [9] 坂地泰紀, 竹内康介, 関根 聡, 増山 繁: 構文パターンを用いた因果関係の抽出, *言語処理学会第 14 回年次大会*, pp.1144-1147 (2008).
- [10] Adomavicius, G. and Tuzhilin, A.: Context-aware recommender systems, *Recommender systems handbook*, pp.217-256 (2011).
- [11] 公益財団法人日本生産性本部: レジャー白書 2013 (2013).
- [12] 内閣府大臣官房政府広報室: 生涯学習に関する世論調査 (平成 24 年 7 月調査) (2012).
- [13] Imamura, K., Kikui, G. and Yasuda, N.: Japanese dependency parsing using sequential labeling for semi-spoken language, *Proc. ACL*, pp.225-228 (2007).
- [14] 川喜田二郎: 「知」の探検学, 講談社 (1977).



横山 法子 (正会員)

2010年早稲田大学理工学部応用物理学専攻卒業。2012年同大学大学院先進理工学研究科物理学及応用物理学専攻修士課程修了。同年日本電信電話株式会社入社。現在、NTT サービスエボリューション研究所所属。行動支援の

研究に従事。



船越 要 (正会員)

1995年図書館情報大学図書館情報学部卒業。1997年北陸先端科学技術大学院大学情報科学研究科博士前期課程修了。同年日本電信電話株式会社入社。以来、分散人工知能、コミュニティ支援、情報検索等の研究に従事。

ACM, 電気情報通信学会各会員。



戸田 浩之 (正会員)

1997年名古屋大学工学部材料プロセス工学科卒業。1999年同大学大学院工学研究科材料プロセス工学専攻博士課程前期課程修了。同年日本電信電話株式会社に入社。以来、情報検索、データマイニングの研究開発に従事。

2007年筑波大学大学院システム情報工学研究科コンピュータサイエンス専攻博士後期課程修了。博士(工学)。ACM, 電子情報通信学会, 日本データベース学会各会員。



小池 義昌

日本電信電話株式会社サービスエボリューション研究所所属。1989年東北大学大学院材料化学専攻博士前期課程修了後、日本電信電話株式会社に入社。以来、パターン認識の研究、遠隔教育システムの研究開発、地域情報検索サービスの研究に従事。

研究に従事。

(担当編集委員 山田 泰寛)