

状況に依存した行動パターン抽出手法の検討

高橋 公海^{1,a)} 佐藤 進也^{1,b)} 松尾 真人^{1,c)}

概要: Web 上の blog 記事や Twitter などのテキストには、日常生活における人々の行動や出来事と、その繋がり（行動パターン）が記述されている。それらのテキストデータを元に行動パターンをモデル化することが出来れば、人間の行動や思考の予測が可能となり、ユーザに対して状況に即した行動提示（ナビゲーション）や推薦、といったコンテキストウェアサービスへの応用が見込まれる。とり得る行動は状況により異なるため、モデル化するにはどのような状況において成立する行動パターンであるかを考慮することが望ましい。そこで本稿では、blog 記事集合を対象として、キーワードに関連する状況毎に文書集合をクラスタリングし、文書集合全体から推定した行動パターンの頻度と、クラスタ内の文書集合から観測される頻度の差を利用して重み付けすることにより、状況に依存した行動パターンを抽出する手法について検討を行った。実験では、本手法を用いて blog 記事集合から抽出した状況ごとの行動パターン例についても報告する。

キーワード: テキストマイニング, 行動モデル, トピックモデル, 残差 IDF

Mining Models of Context Dependent User Behaviors

Abstract: The ability to understand our daily behaviors has long been regarded as enabling a variety of useful applications(e.g. activity-based actuation, recommendation). In this paper, we present an approach to extract context-dependent human behavior models from weblogs. At first, we apply clustering method to weblogs, and then extract important behavior patterns from each cluster using word frequency. We estimated our method by experiment and made sure the effect of it.

Keywords: Text Mining, Human Behavior Modeling, Topic Model, residual idf

1. はじめに

ユーザが発信する情報を集約することでコンテンツを形成する CGM(Consumer Generated Media) の普及や、センサ技術の発達により、人間の日常生活の情報を観測・蓄積することが容易になりつつある。特に Web 上では、blog やマイクロブログを通じて、人々の興味や主観的な意見、その起こった出来事や行動などが記述されたテキストが大量に蓄積されている。これらのデータを元に、日常生活に潜在する人間の行動や出来事の繋がり（行動パターン）をモデル化することが出来れば、行動や思考を予測し、ユーザに対して状況に即した行動提示・ナビゲーションや推薦

を行う、といった応用が見込まれる。

行動や行動パターンは状況により異なるため、どのような状況において成立する行動パターンであるかを考慮し、モデル化することが望ましい。先行研究として、季節毎に成立する行動パターンを抽出する中島ら [1] の手法が挙げられる。これに対し本研究では、季節のように期間で区切ることが難しい状況にも依存した行動パターンを抽出する問題に取り組む。

例えば、「かぼちゃ」に関連する状況として、「冬至」「お菓子作り」「離乳食」「ハロウィン」などが存在する。そして「冬至」という状況で成立するパターンとして「かぼちゃを食べる → ゆず湯に入る」、「離乳食」という状況では「野菜を煮る → お粥を作る」といった行動パターンが挙げられる。冬至やハロウィンは特定の日付と対応しているが、お菓子や離乳食を作る状況は、特定の期間に依存せず発生する状況であると考えられる。本研究では、後者のように時

¹ NTT 未来ねっと研究所
180-8585 東京都武蔵野市緑町 3-9-11

a) t.masami@lab.ntt.co.jp

b) sato.shinya@lab.ntt.co.jp

c) matsuo.masato@lab.ntt.co.jp

期に依存しない状況における行動パターンも抽出することを目指している。

行動パターン抽出には大量の blog 記事集合を用いるが、記事全体で出現頻度の高い行動パターンを採用するという手法では、どのような状況に依存している行動パターンであるのかが明確ではない。また、特定の状況において出現頻度が高い行動パターンが、記事全体を対象とした時にも頻度が高いとは限らず、埋もれてしまう可能性がある。そこで、文書集合を状況毎にクラスタリングし、クラスタに属する文書集合内で頻度の高い行動パターンを出力することで、状況を考慮した行動パターン抽出を実現する。頻度の算出方法は (1) 行動と行動パターンの文書頻度、(2) 文書集合全体から推定した行動パターンの頻度とクラスタ内の文書集合から観測される頻度の差、という特徴を組み合わせることにより、クラスタ内特有の行動パターンに重みを与える。実験では、いくつかのキーワードに対して状況の分類と状況毎の行動パターン抽出を行い、人手で抽出結果の評価を行った。

本論文は次のように構成される。2 章で、本研究が対象としている行動パターンの定義と状況依存性について述べる。3 章では提案手法を説明し、4 章で実験とその結果を示す。5 章で関連研究について述べ、6 章はまとめと今後の課題である。

2. 本研究で扱う行動パターン

2.1 行動

本研究で扱う行動と行動パターンの定義は、先行研究 [1] と同様である。行動は、係り受け関係にある名詞と格助詞、動詞の 3 つの要素で構成する。文を係り受け解析し、同じ分節中に名詞と格助詞が連続して出現しており、かつ動詞を含む分節に直接係る場合のみ行動を取得する。例えば、「かぼちゃを食べます。」という文に対して係り受け解析を行い、図 1(1) のような結果が得られた時、「かぼちゃを食べる」という行動が得られる。また、「かぼちゃを食べて、ゆず湯に入ります。」という (2) の文からは、「かぼちゃを食べる」「ゆず湯に入る」という 2 つの行動が得られる。

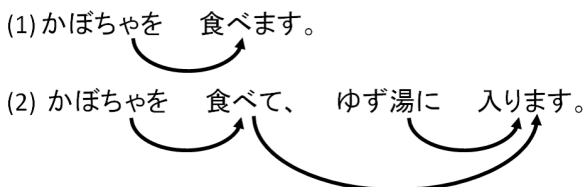


図 1 文中から行動を抽出する例

2.2 状況に依存した行動パターン

日常生活において、ある状況で何らかの行動をとった後、関連する別の行動をとることがある。例えば、冬至の日に

「かぼちゃを食べる → ゆず湯に入る」、離乳食を作る際に「野菜を煮る → お粥を作る」、就職活動で「面接を受ける → 内定をもらう」といった行動の連鎖が挙げられる。このように、「行動の組 $\langle X, Y \rangle$ があったとき、ある状況において X が起きた際、それに続いて Y が生じる確率が高い行動の組」を状況に依存した行動パターンであると定義する。

なお、冬至という状況における「かぼちゃを食べる → ゆず湯に入る」というパターンは定義に該当するため本研究で抽出対象としているが、「家に帰る → 風呂に入る」という行動パターンは、連鎖する可能性の高い行動の組であるが冬至という状況に依存せず生じる行動パターンであると考えられるため、冬至に依存するパターンとしては獲得の対象ではない。

3. 提案手法

状況に依存した行動パターンを抽出するため、本研究では対象の文書集合を状況毎に分割し、分割された文書集合内で頻度の高い行動パターンを出力するというアプローチをとる。次節以降で、文書集合の分割と行動パターン候補の生成方法、そして各行動パターンに対して行う出現頻度をベースとした重み付けの算出方法を説明する。

3.1 文書集合のクラスタリング

「かぼちゃ」という物が関連する状況として、「冬至」「お菓子作り」「離乳食」「ハロウィン」などが存在するように、1 つの語には複数の状況が関連していると考えられる。そのため、文書集合全体からキーワードを含む記事のみに絞り込み、更にそれらの記事をトピックモデル [2] を用いてクラスタリングすることにより、キーワードに関連するより粒度の細かい文書集合に分割する。

トピックモデルとは、様々な離散データに隠れた潜在的なトピック (話題・分野など) を推定するモデルである。文書集合を用いてモデルを学習することにより、各文書が含むトピックや、各トピックの特徴的な単語を発見することが出来る。実際にキーワード「かぼちゃ」を含む blog 記事約 2000 件を対象として、トピックモデルの代表である LDA を利用して特徴語抽出を行った場合、表 1 のような結果が得られる。

トピック番号 0 は冬至、1 番は夏野菜のカレー、2 番はハロウィン、3 番は離乳食といったように、各トピックが状況に対応しているのではないかと考え、今回は LDA を利用した。実験では、キーワードを含む各文書が最も関連するトピックをその文書が属するトピックとして、文書のクラスタリングを行った。

また、今回は行動パターンの候補集合を、ある 1 つの文書内に出現する n 個の行動が $Action = [A_1, A_2, A_3, \dots, A_n]$ という順で出現する時、ある行動が出現した以降の行動と組

表 1 抽出されたトピックの例

トピック番号	特徴語
0	冬至, 柚子, 風呂, ゆず湯, 風邪, 太陽, 冬, 野菜, 昼
1	夏野菜, カレー, トマト, スッキーニ, 夏, じゃがいも, 旬, ピーマン, 人参, とうもろこし
2	ハロウィン, お菓子, 仮装, パーティー, 魔女, おばけ, ランタン, 子供, 夜
3	離乳食, 大さじ, 粥, うどん, ミルク, 豆腐, ブロッコリー, さつまいも, 野菜, 小さじ
4	お弁当, サラダ, ご飯, トマト, 煮物, 卵焼き, 野菜, きゅうり, キャベツ, 豚肉
...	...

にすることで生成した. 例えば $Action = [A_1, A_2, A_3, A_4,]$ という 4 つの行動が順に出現している場合, 生成される行動パターンの候補は $ActionPattern = \{A_1 \rightarrow A_2, A_1 \rightarrow A_3, A_1 \rightarrow A_4, A_2 \rightarrow A_3, A_2 \rightarrow A_4, A_3 \rightarrow A_4\}$ という 6 つのパターンが候補として生成される.

3.2 出現頻度による行動パターンの重み付け

本稿では, 「行動の組 $\langle X, Y \rangle$ があったとき, ある状況において X が起きた際, それに続いて Y が生じる確率が高い行動の組」の抽出を目的としている. そのため, 次の 2 つの仮定をもとに行動パターンに重み付けを行う.

- (1) 行動の組 $\langle X, Y \rangle$ が文書集合中の多くの文書で $X \rightarrow Y$ という順に出現している場合, X に続いて Y が生じる確率が高い.
- (2) 文書集合全体よりもクラスタ内の文書集合に偏って出現する行動パターンは, クラスタが対応する状況に依存している.

仮定 (1) を基に, 文書集合全体における行動パターンの文書頻度 (Document Frequency: df) を指標として利用する. 行動は文中から係り受け関係にある名詞・格助詞・動詞の 3 つを自動的に抽出しているが, ブログに記述される文章は文法的に正しくない文も多いため, 日本語として不自然な行動も抽出されてしまう. 例えば, 「お粥を作る」は行動として生じるものだが, 「食事に組立てる」は不自然であり生じることは滅多にない. 後者のような行動は多くの文書集合で出現するものではないため, 行動パターンの文書頻度を指標とすることで, 日本語として不自然な行動を含むパターンはフィルタリングすることができる.

次に, 仮定 (2) に基づき残差 idf[3] を行動パターンの重み付け指標として利用する. 残差 idf とは, ポアソン分布を利用して単語の idf 値を推定し, 実際の idf 値との差を測ることで, 一般語よりも内容語に重み加わらせた重要語抽出技術である. 今回は文書中の単語ではなく, 行動パターンを 1 つの単語とみなし残差 idf 値を算出した. また, 実際の idf 値は状況に対応するクラスタ内における idf 値, 推定した idf 値はキーワードを含む文書集合全体か

ら算出することで, 状況に対応するクラスタ内に偏って出現する行動パターンに高い重みを与えるようにした.

具体的な算出方法を以降に示す. ポアソン分布は文書において語がランダムに生起する場合の生起回数を確率的に表現するモデルである. ポアソン分布は k を行動パターンの生起回数, λ を期待値として次の式で表される.

$$P(k; \lambda) = e^{-\lambda} \frac{\lambda^k}{k!} \quad (1)$$

ここで, F_i を行動パターン i の大域的頻度, n をキーワードを含む全文書数とすると, ある文書中で行動パターン i が 1 回以上出現する確率 p は次の式で表される.

$$p = 1 - P(0; \frac{F_i}{n}) = 1 - e^{-\frac{F_i}{n}} \quad (2)$$

さらに idf_{ij} を, クラスタリング後のクラスタ j 内での行動パターン i の文書頻度の逆数とし, n_j をクラスタ j に属する全文書数とする. このとき, 残差 idf 値 $ridf_{ij}$ は次の式で求められる.

$$ridf_{ij} = idf_{ij} - \log(1 - e^{-\frac{F_i}{n}}) \quad (3)$$

4. 実験

前章で述べた提案手法を収集した blog 記事集合に適用し, 複数のキーワードに関して状況に依存した行動パターンを抽出することが出来るか人手で評価を行った. 今回対象としたデータは, goo ブログ*1の「日々の出来事」というカテゴリに属する記事約 166 万 5000 件である. あらかじめ, 本文を形態素解析器 Macab*2と構文解析器 Cabocha*3で解析し, 行動を抽出した. なお, その際に軽動詞を含む行動は削除し, 動詞は原型の平仮名に統一することで動詞の表記のゆらぎを吸収した.

4.1 文書集合のクラスタリング

崩壊型ギブスサンプリングを利用し, トピック数は予備実験により設定した. 「就職活動」というキーワードの場合, 対象となる記事数は約 1800 件あり, トピック数は 20 と設定した. 就職活動は行動の順序が比較的決まっているため人手での評価が行いやすいと考え, キーワードとして選択した.

キーワードを含む各 blog 記事を Macab で形態素解析し, 名詞と未知語の Bag-of-Words(BOW) を作成し, LDA を適用した. トピックを尤度順に並べた結果を表 2 に示す.

トピック番号 0 は新卒での就職活動, 1 番は新卒もしくは転職における就職活動, 2 番は官僚の天下り, 3 番は就職活動の準備, 4 番は海外の方の就職活動という状況にそれぞれ対応していると考えられる.

*1 <http://blog.goo.ne.jp/>

*2 <http://mecab.googlecode.com/svn/trunk/mecab/doc/index.html>

*3 <https://code.google.com/p/cabocha/>

表 2 「就職活動」に関して抽出されたトピック

トピック番号	特徴語
0	就職, 学生, 企業, 内定, 新卒, 活動, 採用, セミナー, 説明会
1	面接, 会社, 履歴書, 採用, 試験, 応募, 志望, 転職, 選考, 説明会
2	官僚, 政治家, 公務員, 再就職, 次官, 退職金, 省庁, キャリア, 役人
3	スーツ, 写真, リクルード, サイズ, 色, 顔, スタイル, 髪, 靴, 靴, 撮影
4	外国人, 日本語, 講座, 就職, 英語, 支援, 事業, 農業, 交流, 再就職
...	...

別の例として、キーワード「札幌」を含む記事を対象とした場合の結果を表 3 に示す。地名をキーワードとして抽出して行動パターンを抽出出来れば、観光ナビゲーションへの応用が可能であると考え、キーワードとして選択した。約 12,000 件の記事を同様に形態素解析し、名詞と未知語の BOW を作成して LDA を適用した。トピック数は予備実験により 50 に設定した。

表 3 「札幌」に関して抽出されたトピック

トピック番号	特徴語
0	北海道, 車, バス, ホテル, 駅, 写真, 小樽, 温泉, 旅行, 列車, 函館
1	ラーメン, 店, 北海道, スープ, カレー, 麺, キャラメル, ビール, ジンギスカン, 味噌
2	ドーム, 選手, 日本ハム, 試合, 楽天, ファイターズ, 巨人, 観戦, プロ野球
3	アイヌ, 文化, 民族, 調査, 地域, 支援, 事業, 生活, 歴史, 音楽, 演奏
4	選手, サッカー, コンサド札幌, 監督, 節, ゴール, ホーム, 得点, 試合
...	...

トピック番号 0 は北海道内の旅行, 1 番は札幌の食べ物, 2 番は野球, 3 番はアイヌ民族の文化, 4 番はサッカーという状況にそれぞれ対応していると考えられる。今回は尤度の高いトピックを対象に、行動パターンの抽出を行う。

4.2 行動パターンの重み付け

あらかじめ、キーワードを含む各文書がどのトピックに関連しているかを LDA の結果から算出し、文書集合のクラスタリングと行動パターン生成を行う。例えば、「就職活動」のトピック 0 番に最も関連する文書集合からは約 84,000 件の行動パターン候補が生成された。

まず、各行動パターン候補について、キーワードを含む文書集合における行動と行動パターンの文書頻度を算出する。いくつかのキーワードについて 1 つの行動パターンが何件の記事に出現しているか調査した結果、1~数件の文書にしか出現しない行動パターンが全体の 8 割以上存在することが分かった。そこで今回は、文書頻度順に行動を並べ上位 20%にあたる頻度の行動パターンを残し、それ以外をフィルタリングした。

次に、各行動パターン候補について残差 idf を算出し、それぞれの状況において特徴的な行動パターンに重み付けを行う。各トピックについて、トピックの特徴語を含んでおり、かつ残差 idf 値の高い上位の行動パターンを表 4 に示す。なお、下線が引いてある行動パターンは、キーワードを含む文書集合全体では頻度が上位にくるものではないが、各クラスターでみた時に残差 idf 値の高いパターン（状況に依存している可能性が高いパターン）である。

表 4 「就職活動」に関して状況毎に抽出された行動パターン例

トピック番号	特徴語
0	<u>内定をもらう</u> → 就職活動を始める, <u>内定を取り消す</u> → 内定をもらう, <u>来春に採用</u> → 内定を取り消す, <u>就職を目指す</u> → 資格を取る ...
1	<u>会社に入る</u> → 気を引き締める, <u>転職面接で貰う</u> → 内定を貰う, <u>履歴書を書く</u> → 志望動機を書く, <u>選考に参加</u> → 内定をとる ...
2	<u>定年前に天下り</u> → 退職金を貰う <u>外から抜擢</u> → 退職金を貰う <u>官僚に聞く</u> → 退職金を貰う, <u>歳を取る</u> → 省庁に再就職 ...
3	<u>鏡で見る</u> → 顔を見る, <u>床屋に行く</u> → 髪を切る, <u>スーツを着る</u> → 写真を撮る, <u>自分で買う</u> → スーツを着る ...
4	<u>就職を目指す</u> → 資格を取る <u>日本語を学ぶ</u> → 職を探す <u>外国人労働者が増える</u> → 職を失う <u>職を失う</u> → 就職を目指す
...	...

同様に、キーワード「札幌」の各トピックにおいて特徴語を含む行動パターンのうち残差 idf 値の高い上位の行動パターンを表 5 に示す。

4.3 結果と考察

表 4~5 に示した合計 10 のトピックについて、それぞれ上位 10 件 (計 100 件) の行動パターンについて人手で評価を行ったところ、トピックごとに若干差はあるが正解率は 68%であった。不正解となった行動パターンについては、「内定をもらう → 就職活動を始める」のように行動の順序が逆のもの、「札幌」におけるトピック 1 番で抽出した「散歩に行く → 店に入る」のように状況に依存せず生じると考えられるものなどがあつた。また、「ホームランで先制 → ヒットを打つ」のように元々どちらの行動が先に出現しても不自然ではない行動パターンについては、今回は正解とカウントした。

今回は行動の前後関係をつけるために blog の記事に書かれた順序を利用しているが、文中の言語的な手掛かりや、blog 記事以外のデータ (例えばセンサーデータなど) を用いて行動の前後関係を推定することにより、行動の順序が逆

表 5 「札幌」に関して状況毎に抽出された行動パターン例

トピック番号	特徴語
0	峠を越える → 写真を撮る, 温泉に入る → ホテルに戻る, 電車に乗る → 札幌に向かう, 小樽に向かう → バスに乗る, 自転車に乗る → 写真を撮る ...
1	ジンギスカンを食べる → ガイド本に載る, キャラメルを買う → お土産に配る, 札幌に行く → スープカレーを食べる, 店に入る → 味噌ラーメンを注文, ビールを飲む → ガイド本に載る ...
2	札幌下二ツ木に行く → 応援に行く 本拠地を移る → 北海道に来る, 日本シリーズを見る → 日本ハムが勝つ, 北海道に来る → 野球を見る, 家に帰る → 試合を見る ...
3	地域で暮らす → ルールを守る, アイヌ民族を認める → 先住民と認める, アイヌ文化を発信 → 理解を得る, 生活を支える → 制度を確立 ...
4	選手を集め → 対策を打つ コンサドーレ札幌に移籍 → 穴を埋める, 試合を見に行く → 試合を見る, ボールを奪う → 試合を決める ...
...	...

であるという理由で不正解となるパターンを減らすことが出来ると考えられる。

また、実験では各トピックに対応する行動パターンの上位を対象に精度の評価を行ったが、興味深いのが今回の指標では上位にこない行動パターンも存在する可能性がある。例えば札幌という地名に関する行動パターンは観光ナビゲーション等の応用を検討していたが、「北海道を訪れる → 札幌を訪れる」というパターンよりは「キャラメルを買う → お土産に配る」「大通り公園を歩く → とうもろこしを食べる」「ビール園に行く → 生ビールを飲む」というパターンの方が興味深く有用であると考えられる。「札幌」を含む文書集合において「北海道を訪れる → 札幌を訪れる」というパターンは出現頻度が高いため、各トピックで抽出した行動パターン集合からそれらを除くことで、ナビゲーション等に有用なパターンを抽出できる可能性がある。この点の検証については今後の課題としたい。

5. 関連研究

5.1 状況を限定し人手で行動プロセスをモデル化する技術

あらかじめいくつかの状況を想定し、人手で行動プロセスをモデル化した技術としてタスクオントロジ [4] が挙げられる。タスクオントロジは、ユーザが実世界で認識する問題（タスク）に対して、どのような行動プロセスで解決していくかを示した知識ベースである。「移動」「食事」「遊ぶ」「買い物」「宿泊」「旅行」という 6 つの状況において、起こり得る問題を洗い出し、それぞれの問題を解決する行動プロセスを人手で記述している。しかしタスクオントロ

ジでは、状況や記述されている行動プロセスが非常に限定的であり、かつオントロジの構築には多大な労力を要するという 2 つの課題がある。

5.2 限定的な状況における行動プロセスを自動的に抽出する技術

料理レシピサイトのように、行動のプロセスが順を追って明確に書かれている形式化された文書を対象とし、大量の行動プロセスを自動的に生成する技術として Perkowitzm らの手法 [5] が挙げられる。しかし、シンプルなテキスト解析で行動プロセスを生成しているため、「step1, step2 ...」といったように、形式的に書かれた料理レシピなどの文書にしか対応しておらず、抽出可能な行動プロセスは限定的である。

5.3 多様な状況における行動プロセス抽出技術

大量かつ多様な人間の行動プロセスを抽出するためには、形式化された文書だけでなく、個人が実世界においてどのような行動をとったかが記述された非構造的な文書（blog や Twitter, 質問応答サイトなど）から行動プロセスを抽出する技術が必要である。アプローチとしては、Nguyen らのように言語的な手掛かりを利用した手法 [6] や、倉島らの相関の高いイベントや行動・主観をマイニングする技術 [7] が挙げられる。抽出した行動プロセスを人間への行動提示などに応用する場合、状況によって取り得る行動が変わるため、提示する対象の人間がどのような状況下にあるかを考慮に入れることが望ましいが、上記の先行研究ではどのような状況で成立する知識なのかは意識されておらず、前後関係など行動間の関連性も考慮されていない。

状況を考慮するという意味で最も関連する先行研究は、季節という状況において成立する行動パターンを抽出する中島ら [1] の手法である。これは、時刻情報を利用して文書集合を分割し、各季節における単語やイベントの出現回数・イベント間の共起情報・接続標識などの言語的手掛かりを特徴量として、機械学習を用いて春夏秋冬それぞれに成立する行動パターンを獲得する手法である。例えば夏という季節に成立するものとして、「窓を閉める → クーラーをつける」という行動パターンが得られる。これに対し本研究では、季節のように期間で区切ることが難しい状況にも依存した行動パターンを抽出するという問題に取り組んでいる。

6. おわりに

本稿では、blog 記事を対象としてトピックモデルと残差idfを利用することにより、状況に依存した行動パターンを抽出する手法について検討と実験を行った。実験により、本手法がいくつかのキーワードにおいてある程度有効な結果を得られることを示した。

今後は、より大規模なデータを収集し検証を行う他、行動パターン候補を生成する際に言語的な手掛かり等を利用して候補を絞り込むことや、blog 記事以外のデータを利用した順序関係の特定方法を見出すことで精度向上を図りたいと考えている。

参考文献

- [1] 中島直哉, 吉永直樹, 鍛冶伸裕, 豊田正史, 喜連川優, ”時期依存性を有するイベント連鎖の獲得”, DEIM Forum, 2013.
- [2] Blei, D. M., Ng, A.Y. and Jordan, M.I., “Latent Dirichlet Allocation” Journal of Machine Learning Research 3, pp.993-1022, 2003.
- [3] 北研二, 津田和彦, 獅々堀正幹: 情報検索アルゴリズム, 共立出版 (2002).
- [4] 笹嶋 宗彦, 古谷 孝一郎, 來村 徳信, 深澤 佑介, 長沼 武史, 倉掛 正治, 溝口 理一郎, ”実規模モバイルサービス向けタスク指向型メニューの開発と評価” 人工知能学会全国大会, 2009.
- [5] Perkowitzm M., Philipose, M., Fishkin, K., Patterson, D. J., ”Mining models of human activities from the web” Proc. 13th international conf. on World Wide Web, pp. 573-582, New York, USA, May 2004.
- [6] ゲンミンテイ, 川村隆浩, 中川博之, 田原康之, 大須賀昭彦, “Web からの自己教師あり学習を用いた人間行動マイニング” 電子情報通信学会人工知能と知識処理研究会, AI2009-22, 2010.
- [7] 倉島健, 藤村考, 奥田英範, “大規模テキストからの経験マイニング” 第 19 回データ工学ワークショップ (DEWS2008), 2008.