

リンク航行に着目した行動ターゲティングによるユーザ興味抽出

酒井 まり穂[†] 高田 秀志[‡] 島川 博光[‡]

[†]立命館大学理工学部 〒525-0058 滋賀県草津市野路東 1-1-1

[‡]立命館大学情報理工学部 〒525-0058 滋賀県草津市野路東 1-1-1

E-mail: [†]rs016031@se.ritsumei.ac.jp, [‡]{htakada, simakawa}@cs.ritsumei.ac.jp

あらまし 近年、インターネット利用が急速に増加し、新たな広告媒体として注目されている。広告主は、サイトの多様な特性に応じて、自社の広告を掲載する。一方ユーザは、より限られた目的を持ってサイトを訪れるため、ユーザの興味と広告特性のずれを解消しきれていない。

そこで、本研究では、Internet Explorer ブラウザのキャッシュに記憶された、ユーザが閲覧した Web ページ履歴を用い、HTML パス上から単語を抽出し、それぞれの単語に重要度を設定することで、ある目的を持ってネット・サーフィンをするユーザの興味の対象を見つける手法を提案する。導出された単語に適合する広告を提供できれば、広告の利用率を高められると考える。

キーワード ユーザ興味、Internet Explorer ブラウザのキャッシュ、HTML パス、行動ターゲティング

1. はじめに

コンピュータの発展と共に、インターネット利用は、ここ十数年の間に急成長してきた。以前は現実世界で提供されていたサービスも、インターネット上で提供されるようになってきている。そのひとつに、インターネット広告の配信がある。

当然ながら、広告主は効率的に自社の広告を配信したいと考える。そのため、Web ページに広告を訴求力高く掲示するさまざまな手法が提案されてきた。代表的な手法として、文脈ターゲティングと行動ターゲティング[4]が存在し、これらの手法を用いた既存のサービスもすでに開始されている。しかし、前者を用いた場合、Web ページの特性と広告の内容の一致をもって広告を掲示するため、ユーザが得たい情報が Web ページの特性と異なる場合、掲示された広告へのユーザの関心が薄くなる可能性がある。それに対して後者を用いた場合は、Web 上でのユーザの行動からユーザの興味対象を導き出すため、ユーザの持つ興味に適した広告を掲示することができる。しかし、既存サービスでは、ユーザの行動履歴が限られた Web サイト間のみであるため、ユーザ行動のすべてを把握できていない。さらに、ユーザ興味生成を目的とした研究が多く行われ、多角的な観点から評価されている[6]が、どの既存研究においても、ユーザの負担が大きい、ユーザの行動履歴に含まれたノイズがユーザ興味抽出に影響を与えるといった問題点を解消しきれていない。

そこで本研究では、ある目的を持ってネット・サーフィンを行っているユーザの行動履歴から、そのユーザが持つ興味の対象を導出する手法を提案する。本手法では、ユーザの行動履歴として、Web ブラウザのキャッシュを用いることで、ユーザが閲覧した Web ペ

ジを把握することができる。また、ユーザの興味対象として、本研究ではユーザのリンク航行という動作に着目した。さらに、ユーザがリンク航行するさい、ある 1 つの Web ページから、複数の異なる Web ページへリンク航行するという行動にも着目した。提案する手法は、まず、ユーザが閲覧した Web ページの URL 履歴を基に、その Web ページを構成する HTML ファイル内からユーザがリンク航行に用いた文字列とタイトルにあたる文字列を抽出する。次に、抽出した文字列を品詞ごとに単語に分解する。最後に、それぞれの単語に重要度を設定し、その値が高い単語ほど、ユーザの興味を強く表すと判断する。

本手法の有効性を評価する目的で、本手法を自動化したツールを実装し、実験を実施した。被験者が、一定時間ある目的を持ってネット・サーフィンを行ったさいの URL 履歴を自動化ツールに入力し、抽出された単語と、被験者があらかじめ設定していた目的との比較を行った。導出単語の正確性と完全性というトレードオフの評価基準を用いて比較した結果、ばらつきがみられたものの、一部の被験者の実験内容からは、高い一致性がえられた。本論文では、提案手法と自動化ツールについて論じたのち、実験結果を説明する。さらに本手法と既存研究を比較し、既存研究で提案されていた手法における、ユーザの行動履歴の少なさ、ユーザへの負荷の多さ、Web ページの内容への依存の高さという問題点を、本手法が解決することを示す。

2. 背景

2.1. Web 広告の躍進

近年、世界のインターネット利用者数は、増加の一途をたどっている。国際電気通信連合 ITU

(International Telecommunication Union)が公表している推計によれば,2005年末のインターネット利用者数は,約8.63億人に達している[1].もちろん日本も例外ではなく,インターネットの利用人口はおよそ8,529万人と,前年よりも581万人増となった[1].

これにより,企業が自社のサービスや製品にかかる広告費も,インターネット広告にかかる費用が年々増加している[2].各企業は,自社のインターネット広告の利用率を高めることに注目しており,これまでにさまざまな手法が考案されてきた.

2.2. 既存手法と問題点

既存手法は,文脈ターゲティングと行動ターゲティングに大別される.前者はWebページに書かれている内容を読み込み,関連の強い広告を掲示する仕組みのことで,既存サービスとしては,AdSense[3]が相当する.この手法を用いて広告を掲示した場合,時にはWebページの内容とユーザの欲している情報とが一致しない状況が起こる.たとえば,あるユーザが「季節の献立」という目的を持ってWebページを閲覧していたとき,「ダイエット」という特性を持ったWebページにたどり着いた場合,表示される広告の特性は「ダイエット」に関連するものとなる.これでは,ユーザの持つ目的とは異なった広告を表示してしまう.このように,文脈ターゲティングでは,表示された広告の持つ特性と,ユーザの目的との間に,ずれが生じる可能性が高いという問題がある.

ユーザの目的に適した広告を表示させるためには,ユーザの目的や興味が何であるのかを知る必要がある.そこで新たに行動ターゲティング手法が発案された.これは,ユーザの行動履歴を追っていき,どんなWebページを閲覧したのか,どんな広告サイトを訪問したのかによって,配信する広告を変化させる仕組みである.Webページの特性ではなく,ユーザの興味に合わせて広告を掲示するため,閲覧しているWebページとは関連のない広告が表示されることも少なくないが,文献[4]は,文脈ターゲティングを用いて掲示した広告より,行動ターゲティングを用いて掲示した広告の方が,約17%多く見られていると報告している.この手法を用いた既存サービスの例として,ac Target[5]がある.このサービスでは,サイト内におけるユーザの閲覧履歴や商品の購買状況から,ユーザの興味の対象を導き出し,適した商品を推薦する.しかし,現行のサービスの多くは,サービスに登録しているWebサイト内のみでユーザの行動を特定するため,他サイト内でのユーザの行動を把握できていないという問題点がある.

Web上でのユーザの行動履歴から興味を抽出する研究はいくつか行われており,ユーザの手間という観点

から,それぞれの手法の分類,比較もなされている[6].それぞれの手法において,ユーザの手間が大きい,クライアントの手間が大きい,興味のないWebページの特性に影響を受けるなどの問題があることがわかっている.

3. リンク航行時のユーザ興味の導出

本研究では,Web上のユーザが,リンクを航行しながらWebページ間を移動していくという行動に着目し,ある目的を持ってネット・サーフィンを行っているユーザの行動履歴から,そのユーザが持つ興味の対象を導き出す手法を提案する.

3.1. ユーザの行動履歴

本研究では,Web上でのユーザの行動履歴として,Webブラウザのキャッシュを用いる.

キャッシュとは,ユーザがアクセスしたWebページ上のデータを,クライアント側で一時的に保存しておく領域である.同じWebページを再度閲覧するさいには,サーバに接続せずにクライアント上の一時データを表示するため,Webページの表示速度を上げる効果を持っている.Internet Explorer や Netscape, Opera など,ほとんどすべてのWebブラウザがキャッシュ機能を持つ.

図1が示すように,クライアント側が保存している一時データは,ユーザがアクセスしたWebページを閲覧した順に全て保存しているため,Web上でのユーザの行動そのものであると考えられる.

3.2. HTML パスによる興味の抽出

現在,ほとんどのWebページは,HTML言語で記述されたHTMLファイルである.本研究では,あるHTMLファイルの内部を木構造とみなしたときの,根ノードから葉ノードにいたるまでの枝の連なりを,HTMLパ

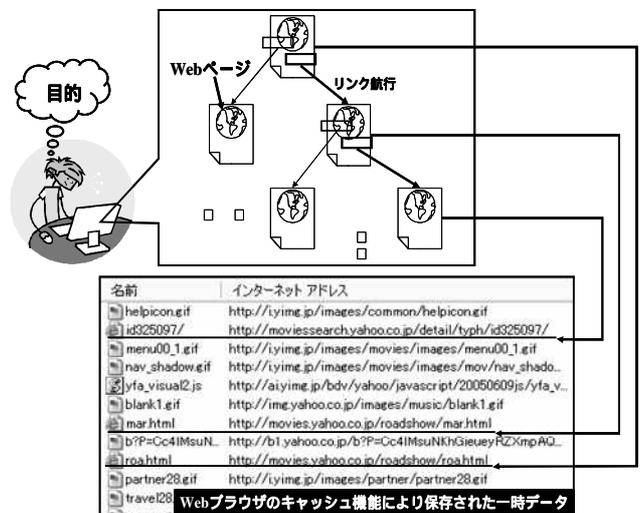


図1: ユーザの行動履歴

スと定義する．HTML パスのイメージを図 2 に示す．

図 3 に示すように，ネット・サーフィンをしているユーザは，Web ページ内で興味のあるリンクを航行することで，次々と Web ページを閲覧していく．つまり，リンクの貼られた文字列が，ユーザの興味に相当している可能性が高い．そこで著者が提案するユーザ興味抽出手法では，ユーザがアクセスした Web ページの HTML パス内に記述された，リンクに着目する．キャッシュ上の Web ページ・アクセス順序において，ある Web ページ P_s の直後に別の Web ページ P_d があるとし， P_s の HTML パス内に P_d に相当する URL がリンク先として記述されているとしよう．この様子を図 4 に示す．図 4 において，上部はキャッシュ上の URL の履歴であり，下部は P_s の HTML パスである．このとき， P_s 内のリンクが航行されて P_d が訪問されたと考えられる．本研究では， P_d を航行先ページと呼ぶ．さらに，この P_d へのリンクを示す P_s の HTML パス内の文字列

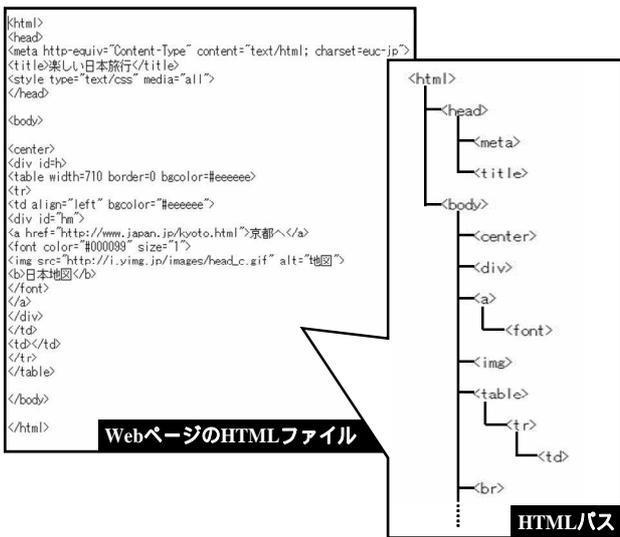


図 2：HTML パス

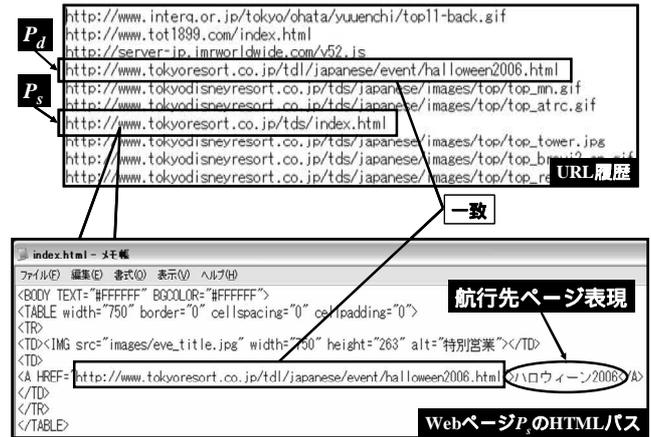


図 4：航行先ページ表現の抽出

を航行先ページ表現と呼ぶ．リンクが HTML の a-タグで示されているときは，タグ内の文字列がこれにあたり，リンクが画像で示されているときは，alt-タグ内の文字列がこれにあたる．

3.3. リンク航行元からの興味抽出

ほとんどの Web サイトは，ある Web ページをトップページとし，トップページ内に複数の Web ページへのリンクが貼られ，さらにそれらから異なる Web ページへのリンクが貼られるという構造になっている．本研究では，トップページを親，トップページからリンクの貼られた複数の Web ページを子とした木構造として Web サイトを捉える．

ユーザ興味抽出手法では，前節で述べた航行先ページ表現以外に，ユーザがアクセスした Web ページのタイトルを表す文字列を抽出する．なぜなら，アクセスした時点で，ユーザはその Web ページへの興味を持っているため，Web ページ全体の特徴を表現したタイトルは，ユーザ興味を把握する上で有用と考えられるからである．図 5 のように，ユーザは，親ページ P_a



図 3：ユーザのリンク航行

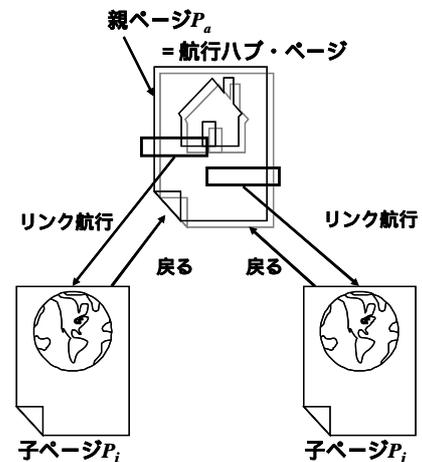


図 5：親ページから複数の子ページへのリンク航行

から子ページ P_i へリンク航行した後，再び親ページ P_a に戻り，さらに異なる子ページ P_j へと移動する場合がある．このとき，親ページの内容をより詳しく知るために子ページを次々と閲覧したと考えられる．このような親ページ P_a のタイトルを示す文字列は，ユーザの興味対象が強く表れていると考えられる．本研究では，何度も航行の元となる親ページを，航行ハブ・ページと呼ぶ．そして，航行ハブ・ページのタイトルを表す文字列を，航行ハブ・ページ表現と呼ぶ．航行ハブ・ページ表現の抽出モデルを図 6 に示す．キャッシュ上に， P_a を訪問したあとで P_i, P_j を訪問したことが記録されており， P_a の HTML パス内に P_i, P_j へのリンクが存在するならば， P_a は航行ハブ・ページである可能性が高い．

3.4. ユーザ興味抽出手法

本手法は，以下に示す 5 段階で構成される．

1. ネット・サーフィン中のユーザの Web ブラウザのキャッシュに随時保存された，ユーザがアクセスした Web ページの URL 履歴をある一定時間分取り出す．
2. 取り出した URL に対応する Web ページの HTML パス内から，タイトルを表す文字列を抽出する．さらに，キャッシュ上の履歴を用いて，航行先ページ表現にあたる文字列を抽出する．
3. 抽出されたすべての文字列を，品詞ごとに単語に分解する．その後，名詞，動詞，形容詞の単語のみを取り出す．
4. それぞれの単語に重要度を設定する．
5. 重要度の高い単語ほど，ユーザの興味を的確に表現していると判断する．

3.5. 単語の重要度

抽出した単語には，単語が属していた文字列の抽出

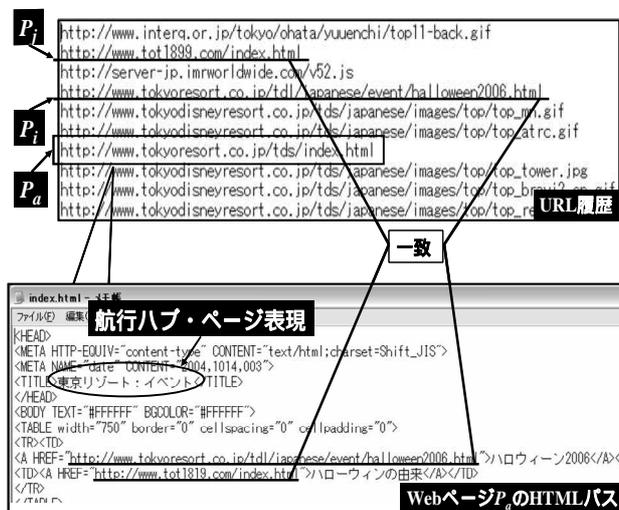


図 6：航行ハブ・ページ表現の抽出

場所と出現頻度によって，異なる重要度を設定する．本研究では，単語への重要度の設定について，以下の 3 点を基準とする．

- ユーザが閲覧した Web ページにおいて，航行先ページ表現に含まれる単語の重要度を高く設定する．1 つの Web ページには，さまざまな情報が混在していることが多い．よって，全体的な特徴よりも，ユーザが航行したリンクに相当する文字列内の単語の方が，よりユーザの興味に近いと考えられる．
- ある Web サイト内において，航行ハブ・ページ表現に属していた単語の重要度を高く設定する．なぜなら，ユーザが 1 つの親ページから複数の異なる子ページへアクセスした場合，親ページは，ユーザの興味に適した文字列を，リンクとして複数個持っている判断できるからである．
- それぞれの単語の出現回数が多いほど，その単語の重要度は高く設定する．単語の出現回数が多ければ多いほど，その単語を含む文字列がよりユーザの興味をひいたと考えられる．

本研究では，単語の重要度の設定手法として，TF・IDF[6]を用いた．TF・IDF は，単語の重み付けを次のようにモデル化している．

$$w_{t,d} = tf_{t,d} \cdot idf_t$$

$$tf_{t,d} = f_{t,d} / \sum_{t=1}^M f_{t,d}$$

$$idf_t = \log N / df_t$$

$f_{t,d}$ はある閾値以上の出現回数を持つ単語 t の出現回数， M はある閾値以上の出現回数を持つ単語の種類の総数を表している．今回，出現回数 2 回を閾値とした． df_t は単語 t の出現回数， N は全単語の出現回数の総和を表している．

4. 自動化ツールの実装

著者らは，本手法の有効性を検証する実験を行う目的で，本手法を自動化したツールを実装した．本ツールの構成を図 7 に示す．本ツールの処理手順は，以下の 8 段階で構成される．

1. ユーザが閲覧した Web ページのデータは，Web ブラウザのキャッシュ機能により，自動的に index.dat ファイルとして保存される．本ツールでは，既存のツール pasco[7]を使用し，index.dat ファイルを text ファイルに変換する．
2. 変換されたキャッシュのデータから，ある一定時間内に閲覧された URL 履歴のみを取り出す．
3. URL 履歴を元に，それぞれの Web ページの HTML パス内から，タイトルを表す文字列，航行先ページ表現にあたる文字列を抽出する．
4. 抽出した文字列を，品詞ごとに単語に分解する．

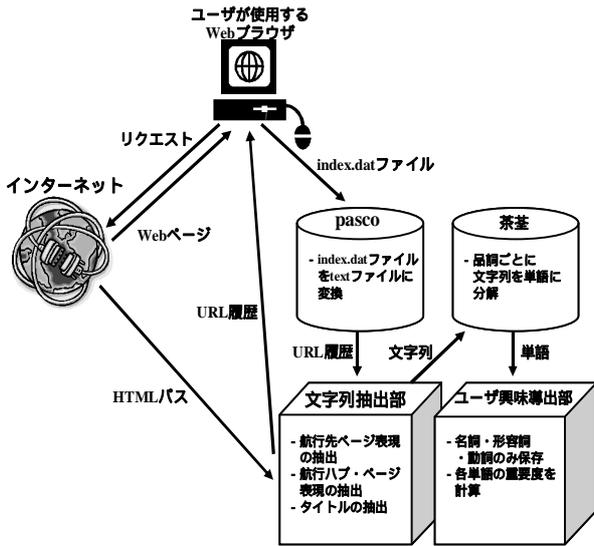


図 7: ツールの構成

そのために、著者らは既存の形態素解析ツール茶釜[8]を用いた。

5. 分解された単語のうち、名詞、動詞、形容詞のみを取り出す。
6. 各単語の頻出回数を算出する。
7. それぞれの単語の重要度を計算する。
8. ある閾値以上の重要度を持つ単語を、ユーザの興味であると判断する。

5. 有効性の検証

5.1. 実験による検証

著者らは、本研究で提案するユーザ興味の抽出手法の有効性を検証するため、4章で述べた自動化ツールを用いて実験を行った。実験の手順は、以下の4段階で構成される。

1. 被験者は、ある1つの目的を持ってネット・サーフィンを行う。WebブラウザとしてInternet Explorerを使用している。
2. 収集したURL履歴を自動化ツールに入力する。自動化ツールは、ユーザが閲覧したWebページのタイトルと航行先ページ表現に現れる単語を抽出する。本研究では、抽出された単語の集合を、抽出単語集合 U とする。さらに各単語の重要度がある閾値以上の単語を、ユーザの興味を表す単語として導出する。本研究では、導出された単語の集合を、ユーザ興味表現集合 N とする。今回の実験では、ユーザ興味表現集合 N を求めるさい、閾値として各単語の重要度の平均値を用いる。
3. 被験者に対するアンケートを実施する。アンケートでは、被験者が、抽出単語集合 U の要素である単語から、あらかじめ設定しておいた目的に適

合する単語を選ぶ。本研究では、ユーザにより選出された単語の集合を、ユーザ選出単語集合 C とする。

4. ユーザ興味表現集合 N の要素である単語とユーザ選出単語集合 C の要素である単語を比較し、両集合に属する単語を、ユーザの興味を真に表現した単語とみなす。本研究では、これらの単語の集合を、ユーザ興味集合 R とする。

ここで、本実験において被験者があらかじめ設定しておく目的は、「旅行」という枠組みに限定した。これは、根本的に異なる目的に沿ったネット・サーフィンを行っても、抽出単語に違いが表れることは明白であり、それでは本手法の有効性の検証には不適と考えたためである。ネット・サーフィンの大きな枠組みは「旅行」であり、詳細設定として、「旅行先」、「旅行の目的」を被験者自身が設定した。各被験者が設定した詳細事項を表1に示す。

本実験によって求められた、単語の集合 U, N, C, R を用いて本手法の評価を行った。これらの集合のベン図を図8に示す。本研究では評価の指標として、適合率、再現率、 F 値、さらに平均適合率を用いた。適合率 p とは、ユーザ興味表現集合 N 中に、どれだけユーザ興味集合 R に属する単語を含んでいるかという正確性の指標であり、次の式で表される。

$$p = R / N$$

また、再現率 r とは、ユーザ選出単語集合 C 中に、ユ

	被験者1	被験者2	被験者3	被験者4	被験者5	被験者6
旅行先	近畿	ラスベガス	近畿	日本	北海道	カナダ
旅行の目的	新入生歓迎	観光	温泉	テーマパーク	観光	オーロラ観賞

表 1: ネット・サーフィンの目的

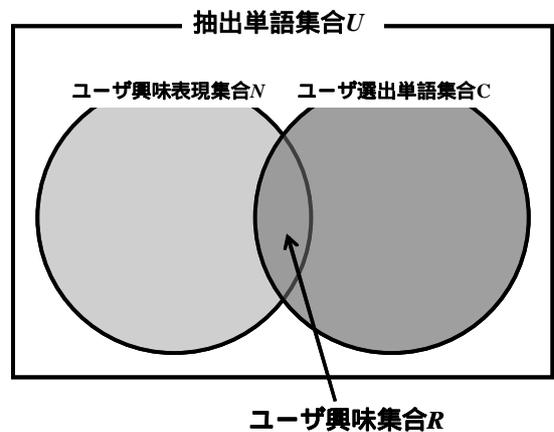


図 8: 各単語の集合のベン図

ユーザ興味集合 R に属する単語を含んでいるかという完全性の指標であり、次の式で表される。

$$r = R / C$$

適合率と再現率はトレードオフの関係にあるため、さらに F 値 f を用いて、本手法の性能の良さを表現する。 F 値は、次の式で表される。

$$f = 2 \cdot p \cdot r / (p + r)$$

さらに、平均適合率 a とは、各再現率レベル k での適合率の平均値であり、次の式で表される。

$$a = (1 / C) \cdot \sum_k c r_k \cdot p(k)$$

以上の指標を用いて本手法を評価した。評価結果は表 2 に示されるように、各被験者によって結果にばらつきが生じた。被験者 2 の結果に基づいた評価のようにすべての指標において高い値を示した結果もあれば、被験者 4 の結果に基づいた評価のように低い値も存在する。また、全体的に適合率は高く、本手法によって導出された単語がユーザ興味を正確に表現していると評価できる。しかし、導出された単語がどれだけユーザ興味を網羅できているかを示す再現率は、ほとんどの実験結果において低い値であった。この理由として、2 つの原因が考えられる。1 つ目の原因は、被験者への依存である。本実験で行ったアンケートでの、ユーザ選出単語集合 C に属する単語の選出は、各被験者によって選出基準が異なる。これにより、評価結果の元となる実験結果に違いが生まれたと考えられる。2 つ目の原因は、実験で用いた自動化ツールのプログラム内容である。本ツールではユーザの興味を表す単語を、Web ページを構成している HTML パス中の HTML タグ内から抽出する。しかし、近年の Web ページは、JavaScript 言語で書かれたものや、Flash 技術を多用したものが増加している。よって、それらの Web ページからは単語が抽出されていないと考えられる。

5.2. 既存研究との比較

2.2. で述べたように、Web 上でのユーザの行動履歴から、そのユーザの興味の対象を導き出す研究は、これまでにいくつも行われている [6][9][10][11]。ここで

	被験者1	被験者2	被験者3	被験者4	被験者5	被験者6
適合率 p	0.750	0.800	0.188	0.143	0.900	0.250
再現率 r	0.473	0.571	0.500	0.250	0.529	0.300
F 値 f	0.581	0.667	0.273	0.182	0.667	0.273
平均適合率 a	0.787	0.858	0.371	0.371	0.766	0.461

表 2：評価結果

は、それら既存研究と著者らが提案する手法との比較を行った。それぞれの比較結果を表 3 に示す。

既存研究で提案されているユーザ興味抽出方法としては、たとえば以下の 4 種類があげられる。

既存手法1：ユーザが Web ページ閲覧中に行う、なぞり読み、リンクの上にマウスポインタを置く、リンクをクリックする、テキストを選択するといったマウス操作から、ユーザが興味を持ったと思われるテキスト部分を抽出する [6]、

既存手法2：ユーザが、Web ページ閲覧中に見つけた興味のある箇所に、現実世界で書籍に書かれた文章にカラーペンで線を引くようにマーキングを行い、それに対してシステムが表示する情報の中から必要な情報を選択する [9]、

既存手法3：あるユーザの年齢、性別、PC の利用状況などの個人データと、そのユーザが約 1 ヶ月の間に閲覧した Web ページのアクセス時間、タイトル、URL などの Web 視聴データとから、興味の対象を導き出す [10]、

既存手法4：ユーザが閲覧した Web ページから、そのユーザが過去に閲覧した Web ページから抽出した頻出語を基準に、新たに頻出する単語を抽出する [11]。

既存手法 1, 2 では、ユーザ自身が興味のあるなしを判定するため、当然ユーザの興味との適合率は高くなる。しかしその反面、ユーザがその動作を行わない限り興味抽出が的確に行われないため、ユーザへの負荷が大きく、ユーザの動作への依存度が高くなる。また、既存手法 1 では、ユーザがなぞり読みした文章に、ユーザの興味が強くと表れているが、その動作には個人差があり、興味のない動作も認識してしまう恐れがある。そのため、抽出されたユーザの興味信憑性は薄くなる。既存手法 3, 4 では、自動的にユーザの行動履歴を保存するため、ユーザへの負荷は全くない。しかし、Web ページのタイトルや頻出語からコ

	本手法	既存手法1	既存手法2	既存手法3	既存手法4
ユーザの行動履歴の多さ					
抽出単語の洗練度					
クライアントの負担の軽さ					
ユーザの負担の軽さ			x		
Web ページの内容への依存度の低さ					x
ユーザへの依存度の低さ			x		

表 3：既存手法との比較結果

文 献

ユーザの興味を導出しているため、Web ページの特性に依存するという問題がある。また、この場合も興味のない Web ページの内容がユーザ興味に含まれてしまう可能性が高く、その正確性は低くなる。

これらに対して著者らが提案する手法では、Web ブラウザのキャッシュに自動的に保存されたデータをユーザの行動履歴とするため、ユーザの負担は全くない。自身の興味に合った広告・サービスの提供を望むユーザは、本手法を実現したツールを自身のコンピュータに組み込む手間がかかるものの、Web ページ閲覧履歴を用いて興味の対象を導出するという処理はすべて自身のコンピュータ内でのみ自動で行われるため、普段のネット・サーフィンを行う場合と全く同じ環境での動作が可能となる。また、どの既存手法においてもあげられている、ユーザ興味を導出するさいに、ユーザが興味を持たない Web ページの影響を受けるという弊害も、本手法では最小限に留めることができる。なぜなら、Web 上でのリンク航行は、ユーザが興味を持っているからこそ行う行動であるため、航行先ページ表現にあたる文字列は、ユーザが強く興味を持った文字列であることを示しているからである。

6. おわりに

本研究では、ある目的を持ってネット・サーフィンを行っているユーザの、リンク航行という行動に着目したユーザ興味抽出手法を提案した。本手法は、Web ブラウザのキャッシュに保存された、Web ページの閲覧履歴と HTML パスを用いて、その Web ページのタイトルと航行先ページを表す文字列に出現する単語のうち、高い重要度を持つものを、ユーザの興味の対象として抽出する。

提案手法の有効性の評価を目的として、本手法の自動化ツールを用いた実験を行った。結果としては、それぞれの被験者の実験データによってばらつきが生じたが、一部の被験者の実験データから、評価の高い値が検出された。結果にばらつきが出た原因としては、実験後のアンケート結果が、ユーザ個人によって捕らえ方が大きく異なっていたこと、本手法の自動化ツールで実装されている単語抽出方法の欠点があげられる。また、従来手法との比較を行い、ユーザの行動履歴数の少なさ、ユーザへの負荷の重さ、Web ページの内容への依存度の高さという問題点を、本手法は解消できることを確認した。

今後は、本研究で提案したユーザ興味抽出手法の有効性を高めるために、自動化ツールの文字列抽出部に改良を加え、よりの確にユーザの行動を把握する。

- [1] 平成 18 年版 情報通信白書 第 2 章第 6 節 1.3 インターネットの普及状況
<http://www.johotsusintokei.soumu.go.jp/whitepaper/ja/h18/index.html> .
- [2] 株式会社 電通 2005 年日本の広告費
http://www.dentsu.co.jp/marketing/adex/adex2005/_media5.html#internet .
- [3] Google 株式会社 AdSense
http://www.google.co.jp/services/adsense_tour/ .
- [4] Bill Harvery, Next Century Media, “The Advantage of Behavior Targeting Increases Dramatically with Frequency,” TACODA, 2006 .
- [5] 株式会社 activecore ac Target 機能概要
<http://www.activecore.jp/target/index.html> .
- [6] 土方嘉徳, “情報推薦・情報フィルタリングのためのユーザプロファイリング技術,” 人工知能学会誌, 19 巻 3 号, pp.365-372, 2004 .
- [7] McAfee, Inc. pasco v1.0
<http://www.foundstone.com/index.htm?subnav=resources/navigation.htm&subcontent=/resources/prod/sc/galleta.htm> .
- [8] 形態素解析ツール茶釜 chasen-2.3.3
<http://chasen.naist.jp/hiki/ChaSen/> .
- [9] 松岡有紀, 坂本竜基, 伊藤禎宣, 武田英明, 小暮潔, “Web 文書に対するマーケティングからの個人知識の獲得,” The 20th Annual Conference of the Japanese Society for Artificial Intelligence, pp.1-4, 2006 .
- [10] 山田和明, 中小路久美代, 上田完次, “インターネットユーザ間の長期にわたる興味遷移パターンの抽出と比較,” The 19th Annual Conference of the Japanese Society for Artificial Intelligence, pp.1-4, 2005 .
- [11] 松尾豊, 福田隼人, 石塚満, “ユーザ個人の閲覧履歴からのキーワード抽出によるブラウジング支援,” 人工知能学会論文誌 18 巻 4 号, pp.203-211, 2003 .