

リアクションサーチ

リアクションをクエリとするウェブ情報検索

莊司 慶行[†] 田中 克己[†]

[†] 京都大学 大学院 情報学研究科 社会情報学専攻 〒606-8501 京都府 京都市 左京区 吉田本町

E-mail: †{shoji,tanaka}@dl.kuis.kyoto-u.ac.jp

あらまし 本研究では、“泣ける”や“行ってみたい”など、発見した文書を読んだ際に読み手がどのように感じるかをクエリとして入力可能なウェブ情報検索を実現する。現状で、ウェブ上で“C言語に関する分かりやすい文書”など、読み手がどう感じるかに基づいた検索を行った場合、本当に望んでいる文書を発見することは難しい。既存の検索エンジンは文書中に出現する語に基づきインデックスを作成するため、書き手が本文中に記述した語に検索結果が大きく左右され、また読み手がどう感じるかは本文中に現れないため、望んだ文書が検索結果に現れない。一方、TwitterやSNSに代表されるウェブ2.0サービスにおいては、ユーザが紹介されたウェブページに対して“泣ける”や“分かりやすい”など、コミュニケーションの一環としてリアクション（反応）をとっている場合がある。そこで、本研究ではウェブコミュニケーションデータに含まれるリアクションをウェブページと対応付けし、クエリとして入力可能にすることで、読み手がどう感じるかに基づく検索が可能なウェブ検索システムについて提案する。本論文では、提案するシステムを実現する手法を提案し、実際にシステムを実装し実験を行うことでその有用性を検証した。

キーワード ウェブ検索, Twitter, Web2.0

1. はじめに

ウェブ2.0が叫ばれてから5年が経過したが、近年のウェブの有りようは、かつてのウェブに対して大きく変貌している。知識文書が主体だったウェブページは多様性を増し、商業サイトやウェブコミュニケーションサービス、個人の日記など、ウェブ上のリソースは今までになく多種多様になってきている。また、コンピュータやネットワークの知識に明るくない人たちも気軽にインターネットを利用するようになり、ウェブ利用者は利用者層、利用者数ともに増してきている。このように多様性を増している現代のウェブにおいて、従来の検索エンジンでは不十分な点が出てきている。

数も多様性も増した現状のウェブの中で、ユーザが望んだ文書を発見することは容易ではない。ことさら近年増加した利用者の中で、コンピュータに詳しくないユーザは、検索エンジンの仕組みを理解しておらず、自分の持っている検索要求を検索エンジンに満身に伝えることができない。またこういったユーザは先行知識に乏しく、既存の検索モデルの要求するような文中語形式のクエリがうまく作成できない。こういったユーザでも、望んだ文書が発見可能な検索システムの必要性が高まっている。

こういったユーザが普段しているであろう検索モデルを考える上で、インターネット上から離れて、現実世界における検索行為について目を向ける。現実世界における文書検索の具体的な例として、詳しくない分野についての書籍を探す場合を挙げる。ある分野に詳しくないユーザが、書籍を紹介してもらった際、例として“マヌエルとマイクロフトという登場人物が登場するSF小説を紹介してほしい”や、“C言語に関して、ポイントと

いう語を含む本を教えてください”というような聞き方をすることは稀である。登場人物名や具体的な内容など、本文中に含まれる要素に基づいた検索は、対象の文書の内容をある程度理解していなければならないため、現実世界ではあまり行われず、通常、このような検索を行う場合には、“読んだら面白いと感じる、SFの本を教えてください”や、“C言語に関する本で、わかりやすい本を探してほしい”というような聞き方をする。現実世界でこういった検索をする際には、本文の内容や含まれる語に基づく検索でなく、“わかりやすい”、“スカッとする”、“泣ける”など、文書を実際に読んだ際に読み手がどのように感じるかに基づいた検索をしている。

このような“読み手がどう感じるか”に基づくような検索は、現状のインターネット検索においては、満足に行えない。先の例のような検索をしようとして、従来の検索エンジンにクエリとして“C言語 わかりやすい”と入力した場合、望むような検索結果は得られない。“わかりやすいC言語の本を探しているが、見つからない”という個人の日記や、“我が社のわかりやすいC言語の本を買ってください”という通信販売のページなどが検索上位に並ぶ一方、本当に必要なページは検索結果に現れない。このようなことが起こる原因は2点ある。一点目は、従来型の検索モデルの問題点に起因する。現状の検索エンジンでは、検索対象とする文書中に含まれた単語に着目した検索を行う。そのため、“わかりやすい”という語が一度も登場しないがわかりやすい文書があった際など、その文書を発見できず、検索結果の再現率が低下する。もうひとつの原因は、文書に現れている語は、あくまで書き手の意図によるもので、読み手がどう感じるかとの間に相違がある点である。書き手が面白いと思ったものが読み手にとって面白いとは限らないほか、商業的

なサイトにおいては、ポジティブなイメージをもつ語を売り文句として記載している場合がある。そのため、読み手が面白いと感じない文書が検索結果に現れ、適合率が低下する。従来の検索エンジンの、文中語をクエリとする情報検索は、すでにある程度知っている内容を掘り下げたい場合や、以前に一度読んだ文書を探している場合には有用であるが、新規文書の検索に向かない。この問題を解決するため、“読み手が実際に文書を読んだ際にどのように感じるか”をクエリとして入力可能な検索システムを提案する。

提案システムでは、あるウェブページを読んだ際、読み手がどのように感じるかに基づく検索を可能にするために、ウェブ 2.0 コンテンツに含まれるユーザのリアクション（反応）に着目する。Twitter^(注1)や SNS といったウェブコミュニケーションを目的とするウェブ 2.0 サービスの中では、コミュニケーションの一環として、ウェブページを紹介したり、されたりしている場合がある。そして、紹介されたウェブページに対して、ユーザが“おもしろい”や“泣ける”などといった反応（リアクション）をとっている場合がある。このリアクションは、“読んだ時にこう感じる”というものが、具体性を伴って記述されたものである。具体例として、猫の画像の掲載された“かわいい”ウェブサイトを紹介されたユーザたちは、“キュートだ”、“プリティ”、“抱きしめたい”といった文をリアクションとしてコミュニケーションサイト投稿している。これらウェブコミュニケーション中のリアクションを収集し、ウェブページと対応付けることで、

- (“かわいい”, “猫”)
- (“行ってみたい”, “イベント”)
- (“よくある”, “大学生活”)
- (“賛成”, “法案”)

という形式で、通常の検索エンジンにおけるクエリに加え、文書を読んだ際にとられるリアクションをクエリとした検索を可能にする。これにより、上述したような、その文書を実際に読んだ際に読み手がどのように感じるかに基づいた検索が可能になると考えられる。

本稿では、このようなウェブ検索システムを実現する手法を提案し、プロトタイプを実装したうえで、実験、評価を行った。以下で本稿の構成について述べる。2 節で本稿で提案するリアクションサーチについて、3 節で提案手法の有用性を検証する実験とその結果についてそれぞれ述べ、4 節で結果が何を示しているのかについて解説する。また、5 節で、本研究と関連する先行研究について述べ、6 節で今後の課題について、7 節でまとめについて述べる。

2. リアクションサーチ

通常のトピッククエリに加えて、リアクションをクエリとして入力することで、発見した文書を読んだ際にどのように感じるかに基づくウェブ検索ができる検索エンジンを実現するための手法について説明する。この検索システムの概要を図 1 に示

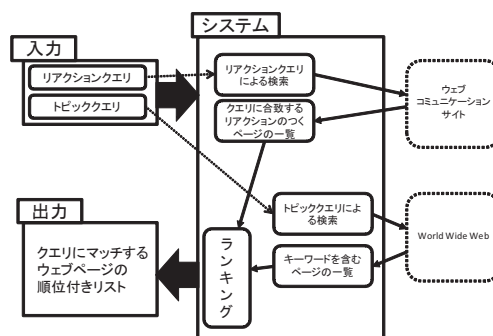


図 1 ウェブ検索システムの概要

す。システムへの入力、読んだ際にどう感じるかを表すリアクションクエリと、何に関する文書かを表すトピッククエリである。これら二つのクエリは文字列として入力される。検索対象はすべてのウェブページであり、スコアリングにはウェブコミュニケーションのログおよびウェブページの内容を利用する。システムの出力は、クエリにマッチするウェブページの順位付きのリストである。ページの順位は、ページとリアクションクエリ、トピッククエリとの適応度によって決定される。

システムの受け付けるリアクションクエリは、読み手が文書を読んだ際にどう感じるかを表す語である。ここで言うリアクションとは、文書を読んだ際にどう感じたかがコミュニケーションを通じて記述されたものである。この記述は具体性を伴って自然言語で行われるため、表現の形態に多様性がある。例として、読み手が“分かりやすい”と感じるような文書を読んだ際に実際にとられるリアクションは、

- “簡単だ”, “難しくない” などの 換言表現
- “おかげで理解した” などの 因果に基づく表現
- “へえ!”, “なるほど” などの 感嘆表現

など、さまざまである。提案するシステムでは、これらのうちどのような形式のリアクションでも、リアクションクエリとして入力可能である。また、“なるほど”というクエリ入力で“わかりやすい”と感じる文書を検索するなど、別の表現で記述されたリアクションに関しても検索可能にする。

2.1 モデル

すべての検索対象のウェブページ集合 P 、全単語を V とした際、入力されるクエリ q は $q = (q_r, q_t)$ で表される。ただしこの際、 $q_r \subset V, q_t \subset V$ である。入力されたクエリ q を構成するリアクションクエリ q_r 、トピッククエリ q_t について、リアクションサーチは q_r というリアクションのつく q_t に関するウェブページ p の集合をランクづけして返す。

あるウェブページ p がリアクションクエリ q_r およびトピッククエリ q_t とどれだけ適合するかを表すランキング関数 $rank(p, q_r, q_t)$ を式 (1) のように定義する。

$$rank(p, q_r, q_t) = score_T(p, q_t) \cdot score_R(p, q_r) \quad (1)$$

ここで登場する関数 $score_T(p, q_t)$ はあるウェブページ p がトピッククエリ q_t に対してどれだけ適合するかの度合いであり、 $score_R(p, q_r)$ はウェブページ p がリアクションクエリ q_r に対してどれだけ適合するかの度合いである。これらの関数に

(注1): <http://twitter.com/>

関してはそれぞれ 2.2 項, 2.3 項で述べる.

このようなモデルを実現するうえで, 解決すべき技術的課題点を 2 点挙げる.

- リアクションクエリ q_r とウェブページ p との間の適合性計算を行う. この際, 同じ意味を持つ別表現のリアクションに対しても検索対象とする.
- リアクションのひとつもついていないウェブページ p_0 に関してリアクションクエリ q_r との適合度を推定し検索対象化する.

本論文においては, 特に前者に注力する.

2.2 トピッククエリに対するウェブページの適合性

あるページ p がトピッククエリ q_t に対してどれだけの適合性を持っているの計算を, 以下の式 (2) で求める. 本稿では, もっとも単純化したクエリ尤度モデル [1] を用いた.

$$score_T(p, q_t) = Pr(p|q_t) \propto Pr(p) \prod_{t \in q_t} Pr(t|M_p) \quad (2)$$

$Pr(t|M_p)$ は, トピッククエリ q_t に含まれる語 t が, ウェブページ p 中に登場する回数を $tf_{t,p}$, ウェブページ p の長さを L_p としたとき, 式 (3) で推定できる.

$$Pr(q_t|M_p) = \prod_{t \in q_t} Pr_{mle}(t|M_p) = \prod_{t \in q_t} \frac{tf_{t,p}}{L_p} \quad (3)$$

2.3 リアクションクエリに対するウェブページの適合性

あるウェブページ p についているリアクションが, どれだけリアクションクエリ q_r に対して適合性が高いかをページごとに取りまとめたスコア $score_R(p, q_r)$ について説明する.

リアクションとは, “ある文書を読んだ際にどう感じたか” が具体性を伴って自然言語で記述されたものであり, 文書を読んだ際に感じたことそのものではない. そのため, わかりやすい文書に 実際につくりアクションは, “わかりやすい” のほかに, “なるほど”, “へえ!” など, 多彩な表現を伴う. そこで, リアクションクエリ q_r に対するウェブページ p の適合性を計算するためには, 同一の感動を表すが別の表現で記述されたリアクションについて考慮する必要がある.

ウェブページ p につくりアクション集合を $R(p) = \{r_1, \dots, r_j\}$ としたとき, それぞれ個別のリアクション $r \in R(p)$ の q_r に対する適合度に関して, 平均をとったものをページ p とリアクションクエリ q_r の適合度 $score_R(p, q_r)$ として用いる. これを以下の式 (4) で表す.

$$score_R(p, q_r) = \frac{\sum_{r \in R(p)} s_r(r, q_r)}{|R(p)|} \quad (4)$$

これら個別のリアクション r の入力されたリアクションクエリ q_r に対する適合度 $s_r(r, q_r)$ は, リアクション r を構成する複数の語 $W(r) = \{w_1, w_2, \dots, w_i\}$ についてそれぞれリアクションクエリ q_r との適合度をとり, 少なくとも一語が q_r と同義であるかを表す式 (5) により合算することで決定する.

$$s_r(r, q_r) = 1 - \prod_{w_i \in W(r)} (1 - s_w(w_i)) \quad (5)$$

ここで, リアクション r を構成するそれぞれの語 $W(r) = \{w_1, w_2, \dots, w_i\}$ について, その語が入力されたリアクションクエリ q_r とどれだけ近い意味を持つかについて考える.

ひとつのウェブページに対して, 複数のユーザから, 複数のリアクションが付けられる場合がある. 通常, “かわいい” というリアクションのつくページには, “かわいい” と同義のリアクションが付きやすいと考えられる. その一方で, 同じ文書を読んだ場合でも, 人によって, どのように感じるかは異なるため, 同じページに対する複数のリアクションが, 常に同種の感動を表すものとは限らない. そこで, リアクションに含まれるある語 w_i がリアクションクエリ q_r と同じ感動を表す語である度合い $s_w(w_i, q_r)$ を, 式 (6) によって定義する.

$$s_w(w_i, q_r) = s_a(q_r, w_i) \cdot s_c(q_r, w_i) \quad (6)$$

式中の $s_a(q_r, w_i)$ は q_r と w_i の観点の同一性を, $s_c(q_r, w_i)$ は q_r が w_i とどれだけ特徴的に同じ対象に言及しているかをそれぞれ表す. ある語 w_i が同一な観点に基づいており, なおかつ同じ対象に言及している場合, 同じ感動の別表現である可能性が高い. 以下でこの 2 つのスコアに関して説明する.

同一のウェブページに対して, 複数の個人がそれぞれ別の観点に基づいたリアクションをとっている場合が考えられる. 例として, 動物の形をした家具を紹介するページを閲覧して, 動物の形に注目して “かわいい” と感じる人もいれば, 家具としての性質に注目して “便利そう” と感じる人もいる. そこで語 w_i がリアクションクエリ q_r と同一の観点である度合い $s_a(w_i, q_r)$ を式 (7) で定義する.

$$s_a(w_i, q_r) = \frac{|P_{reaction}(q_r) \cap P_{reaction}(w_i)|}{|P_{reaction}(q_r)|} \quad (7)$$

$P_{reaction}(w_i)$ は語 w_i を含んだリアクションが少なくともひとつ付いたウェブページの集合を返す関数である. ある語 w_i がリアクションクエリ q_r と同一観点に基づく語であった場合, リアクションクエリ q_r を含むリアクションのついた他のウェブページに対するリアクションにもまんべんなく含まれていることが予想される. 先の例において, “かわいい家具” へのリアクション, “かわいい菓子” へのリアクション両方に含まれている語は, 同一観点に属することが期待できる. “便利そう” という別の観点に基づくリアクションは前者の “かわいい家具” にしかつかない一方, “キュートだ” という同一観点に基づくリアクションは “かわいい家具” と “かわいい菓子” の両方につきうる. 式 $s_a(w_i, q_r)$ はリアクションクエリ q_r を含むリアクションが付いたすべてのウェブページに, 語 w_i を含むリアクションがどれだけまんべんなく付いているかを表す.

リアクションクエリ q_r がある語 w_i とどれだけ特徴的に共起するかを表す尺度 $s_c(w_i, q_r)$ は, 以下の式 (8) で表される.

$$s_c(w_i, q_r) = \frac{|\{(r, p) | r \text{ include } w_i, p \in P_{reaction}(q_r)\}|}{|R_{reaction}(w_i)|} \quad (8)$$

$R_{reaction}(w_i)$ は語 w_i を含むリアクションが少なくともひとつ付いたウェブページに対するリアクションすべての集合を返す関数である. ある語 w がリアクションクエリ q_r とどの程度似

た意味を持つかを判断するうえで、語 w_i がリアクションクエリ q_r とどの程度特徴的に共起するかを利用する．先の例で、“かわいい” ウェブページに対するリアクションのみに含まれ、“かわいくない” ウェブページへのリアクションには含まれない語は、“かわいい” の同義語であることが期待できる． $score_R(w_i, q_r)$ は、ある語 w_i が、リアクションクエリ q_r を含むリアクションが付くようなウェブページにつくりアクションにどれだけ特徴的に登場するかを表す．

以上の計算で、少なくともリアクションのひとつ以上ついたすべてのウェブページについて、付いたリアクションにリアクションクエリ q_r が含まれていなくても、適合性 $score_R(p, q_r)$ を算出することができる．しかし、リアクションがひとつも付いていないウェブページ p_0 があった場合、ページの内容にかかわらず、この適合性スコア $score_R(p_0, q_r)$ は 0 となる．そこで、リアクションのついていないウェブページについて、 $score_R$ を推定する手法について、2.4 項にて述べる．

2.4 リアクションのついていないウェブページのリアクションクエリに対する適合性の推定

インターネット上に公開されているすべてのウェブページに対して、ウェブコミュニケーションサイトにて紹介されリアクションがとられているウェブページは、ごく一部にすぎない．提案する検索システムでは、リアクションがひとつも付いていないウェブページについても検索可能にするため、リアクションクエリとリアクションのついていないページとの適合度についても定義する．内容の類似したウェブページは、読み手に対して同じ感動を与えると考えられる．そこで、リアクションのついていないウェブページと付いていないウェブページの類似度をとることで、つくであるリアクションを推定する．リアクションが付いており、なおかつトピッククエリ q_t にマッチするウェブページ集合 $P(q_t)$ についてすべてのウェブページ $p \in P(q_t)$ を連結した文書 p_{qk} について、各次元が語 w_i で、各次元の重みを語 w_i の p_{qk} における登場頻度 $tf(w_i, p_{qk})$ とした特徴ベクトル $v_{p_{qk}}$ を作成する．そして、トピッククエリ q_t にマッチするウェブページのうち、リアクションのついていないウェブページ p_0 に関して、各次元を語 w_i 、各次元の重みを語 w_i の p_0 における登場頻度 $tf(w_i, p_0)$ とする同様のベクトル v_{p_0} を作成した．これら二つの特徴ベクトル $v_{p_{qk}}, v_{p_0}$ について、式 (9) の計算を行うことで、リアクションクエリ q_r とリアクションのついていないウェブページ p_0 の適合度 $score_R(p_0, q_r)$ を設定した．

$$score_R(p_0, q_r) = \cos(v_{P(q_t)}, v_{p_0}) \cdot \min(\{score_R(p, q_r) | p \in p_{qk}\}) \quad (9)$$

ここで、リアクションクエリ q_r に適合するリアクションのついたすべての文書を結合してひとつの大きな文書にしたのは、リアクションのついた特定のページに極めて良く似た結果が検索上位に現れても情報が重複するだけで、検索結果の向上につながらないためである．また、リアクションクエリ q_r との適合度を計算する際に最小値をとっているのは、上位に似た結果が固まって表示されるのを避けるためである．



図 2 実装したシステムのスクリーンショット



図 3 取得したリアクションの例

3. 実験

提案するウェブ検索手法について、その有用性に関して評価を行うため、実際に動作するウェブアプリケーションシステムを実装し、そのシステムを使って被験者実験を行った．

3.1 実装

実験のため、2 節で提案したアルゴリズムに則って動作するプログラムを、ウェブブラウザからアクセス可能なウェブアプリケーションプログラムとして実装した．作成したウェブアプリケーションのスクリーンショットを図 2 に示す．二つ並んだテキストボックスの左側にリアクションクエリを、右側にトピッククエリをそれぞれ入力することで、上述のランキング関数にしたがって順序づけされた検索結果が得られる．この際、発見したウェブページのタイトル、URL、スニペットのほかに、実際にそのウェブページに対してなされたリアクションも表示される．

実装を行う上で、リアクションの収集は Twitter から行った．Twitter 内の発言の中で、あるユーザが他人の URL を含む発言を引用しつつ、その URL についてコメントしている形式のものを抽出し、コメント部分のみをリアクションとして利用した．収集した Twitter 内の書き込みの具体例を図 3 に示す．この例において、検索の対象となる URL は “http://www.dl.kuis.kyoto-u.ac.jp” であり、このウェブページを読んだ際のリアクションが “こんなに面白い Web サイト見たことない www” である．

また、リアクションのついてない Web ページの検索には Yahoo ウェブ検索 API^(注2)を利用した．これらのウェブページについてリアクションクエリに対する適合性の推定を行う際、特徴ベクトルの生成にスニペットを用いた．

(注2): <http://developer.yahoo.co.jp/webapi/search/websearch/v2/websearch.html>

実際にアプリケーションとして実装する際、処理の流れは以下のようにした。

- (1) 前処理としてウェブコミュニケーションサイトからリアクションを収集し、URL とリアクションの対応付けを行い利用可能な状態にする
- (2) ユーザの入力したリアクションクエリとトピッククエリを受け取る
- (3) リアクションクエリを用いて、2.3 項の手法を用いて類似した意味の別表現を探し、単語のスコアを決定する
- (4) リアクションのついているウェブページの一覧に対して、トピッククエリで文中語による検索を行う
- (5) トピッククエリを用いて通常のウェブ検索を行う
- (6) 手順(5) で取得した結果について、リアクションのついていないものに関して、リアクションのついていないものと類似度比較を行う
- (7) 手順(3), (6) よりリアクションクエリとの適合度を、手順(4), (5) よりトピッククエリとの適合度を算出し、ランキングを行い出力する

実装上の都合として、検索結果の上限を 3000 件と定めた。この制限は手順(3) において課され、本文中に含まれるトピッククエリ語の登場回数による順位が 3000 位より下のウェブページに関しては、リアクションクエリとの適合度計算、検索結果への出力は行われない。

実装にはスクリプト言語である Ruby を使用し、作成したシステムは debian linux 上の apache2 により CGI として動作する。リアクションおよびウェブページ本文の形態素解析には MeCab を、MeCab の辞書には IPA 辞書を用いた。この際、品詞情報などは考慮せず、いくつかのストップワードを除いたすべての語をスコア計算に利用した。

3.2 実験

提案手法について、システム全体を通して結果の精度に関して、ユーザ評価に基づく比較実験を行った。

比較のためのベースラインとして、従来の情報検索のアルゴリズムとして、クエリ尤度モデルに基づく情報検索システムを用意した。比較用システムでは、受け取ったリアクションクエリ、トピッククエリの両クエリについて、対象文書の本文中に含まれるかどうかに基づいた AND 検索の結果を返す。ベースラインシステムのスコアリングアルゴリズムは式(2)に則る。これら両システムに同一の検索対象ページ群とクエリの組を与え、検索結果のランキングに含まれる各ウェブページについて、クエリとの適合度を人手で評価することで、検索システムの評価を行った。

評価のために 20 組のクエリをあらかじめ作成した。クエリの形式はリアクションクエリ、トピッククエリひとつずつの組み合わせである。実際に利用したクエリについて、表 1 に記す。

被験者の人数は 4 名であり、全員が男性の情報学への知識を有する大学生および大学院生である。それぞれの被験者は 10 件ずつのクエリについて提案手法、ベースライン手法により算出された検索結果上位 20 件について適合度合いを評価した。評

表 1 実験に使用したクエリ

リアクションクエリ	トピッククエリ
行ってみたい	ラーメン屋
わかりやすい	経済
泣ける	コピペ
共感	体験談
お買い得	ノートパソコン
笑える	日記
おいしそう	レシピ
分かりやすい	Ruby
面白い	SF 小説
役に立つ	豆知識
意外	芸能ニュース
ためになる	プレゼン
なるほど	郵政民営化
えげつない	商売
おすすめ	京都
かわいい	猫
やってみたい	アルバイト
難しい	C 言語
素敵	年賀状
美味しくない	ラーメン屋

価用のシステムでは、被験者がどちらの検索結果を評価しているか分からなくするため、与えられる評価用タスクの順序は被験者ごとにランダムとし、画面にクエリと URL だけが表示されるようにした。被験者はリンクをクリックして実際にページを閲覧したうえで、クエリとウェブページの適合度を 1 から 5 の 5 段階で評価した。

実際に実験を行う際に用いたデータについて述べる。2010 年 11 月より 2010 年 12 月にかけて、Twitter ストリーミング API で無作為に取得した URL を含むツイートのうち、日本語で記述され、なおかつ先述した形式に則ったものをすべて利用した。収集した日本語ツイート総数は 30,134,604 であり、このうち 2,370,222 件がリアクションとして利用可能であった。リアクション中に含まれた URL の指すウェブページは、短縮 URL による重複を取り除いたうえで 1,286,752 件であり、本実験ではこれらのウェブページとそれについてリアクションを対象として取り扱った。今回実験で利用した上位 20 件までの検索結果の中には、リアクションが付いておらず本文類似度からリアクションクエリとの適合度を推定したウェブページは含まれていない。そのためベースライン手法においても、ここで取得したリアクションに含まれる URL のみを検索対象とした。

3.3 結果

実験における各評価項目ごとの結果を表 2 に示す。

各被験者の付けたウェブページへの適合性評価について、同一タスクごとに被験者間の平均をとり、ウェブページとクエリの適合度とした。

$P@k$ は適合度が 3 より大きいものを適合、そうでないものを不適合とした場合の検索結果上位 k 件目までの適合率である。両システムの検索結果について、1, 3, 5, 10, 20 位までの結果について、各クエリについて算出したのち、全クエリを平均

表2 実験結果

評価項目	提案手法	ベースライン手法
Precision at 1	0.50	0.50
Precision at 3	0.45	0.38
Precision at 5	0.44	0.35
Precision at 10	0.36	0.34
Precision at 20	0.33	0.28
MAP	0.55	0.51
取得結果数平均	2397.8	675.9

して求めた。

MAP@20 は全クエリに対する適合度の総合尺度であり、両手法とも 20 位までの結果について計算している。

取得結果数平均は、入力されたクエリによって何件の結果が返ってきたかを表す。ただし、各検索結果において、最大件数を 3000 件と設定したため、それ以上結果が発見されたものに関しては考慮しない。

検索結果の K 件目まで閲覧した際の適合率を表す $P@k$ では、1 件目から 20 件目まですべての k において、ベースライン手法に対して提案手法の有用性が認められた。

4. 考 察

提案手法がベースライン手法に対して特に良い結果を返した 2 件のクエリについてそれぞれ個別に考察および解説を行う。

q_r = “役に立つ”, q_t = “プレゼン” という検索タスクにおいて、提案手法がベースライン手法を適合率で大きく上回った。このクエリで得られた検索結果の上位 10 件のウェブページについて、2 名の被験者のつけた 5 段階での適合度の平均と、ページの内容の概要を表 3 に記す。提案手法ではプレゼンテーション技法について詳細に語ったページが多数発見されたのに対し、既存手法では他人のプレゼンテーションを聞いたという報告や、会合や書籍の紹介記事のようなウェブページが多く検索結果に現れている。これは “ためになる” という記述が、“ためになるプレゼンテーションを聞いた” と記述されているが具体的な内容に触れていないページや、宣伝文句として “ためになるイベントに参加してください” というページに多く含まれている反面、プレゼンテーション技法について解説したページでは淡々と技法のみ紹介しており、“ためになる” という記述が本文中に含まれないためである。提案手法による検索結果において、上位 20 件中に “ためになる” という語を直接含むようなリアクションが付いたページは 1 件も存在しなかったが、“必読”、“なるほど”、“勉強”、“興味深い” などの語が、“ためになる” というリアクションクエリと近いと判断され、これらの語を含むリアクションのついたウェブページが上位に表示された。

q_r = “やってみたい”, q_t = “アルバイト” というタスクに関して、両手法とも適合率は高くないが、ベースライン手法の適合率が極めて低い結果を返した。両手法に共通する適合率を下げる要因として、企業のウェブサイトが紹介された際、多くの企業ではメニュー項目にアルバイトという語が入っていることがあり、検索結果上位にこれらのページが出てきた際に適合率

表3 検索結果例 (“ためになる”, “プレゼン”)

従来手法	適合度
ベンチャービジネスの紹介記事	1.0
勉強会の参加受付	2.0
ビジネス本の紹介	2.0
勉強会の参加受付	2.0
Web デザイナー講演会のレポート	4.0
プレゼン本の広告のある blog	1.0
プレゼン練習会の Twitter まとめ	1.0
政治ニュースへの Twitter まとめ	1.0
プレゼン本の広告のある blog	1.0
就職活動の心構え	1.0
提案手法	適合度
プレゼン手法の提案	4.5
プレゼン配信ソフトの紹介	2.0
プレゼン手法の提案	4.0
プレゼン本の解説	4.0
プレゼン本の解説	4.5
プレゼンソフトの解説	3.5
スティーブ・ジョブズの記事	1.0
プレゼン本の解説	3.0
プレゼンテーション協会のサイト	2.0
プレゼンのここのまとめ記事	4.0

が下がった。これは文中語によるトピック検索の欠点であり、本研究で提案する “やってみたい” と感じるかどうかというリアクションに基づく検索以前に、“アルバイト” というトピックと関係ないページが結果に現れたためである。その中でもベースライン手法のスコアがより低いのは、アルバイト斡旋系サービスのウェブサイトがうたい文句として “やってみたい” という語を多用するためである。アルバイトを紹介するポータルサイトでは “あなたのやってみたいアルバイトが、きっと見つかる” というような宣伝文句がページ内に数ヶ所書き込まれている。そのため文中語に基づくスコアリングをした際、リアクションクエリに対するスコアが高くなるが、そのページには実際に読み手が “やってみたい” と感じるアルバイトの紹介情報は書かれていない。

一方で、いくつかのクエリで、提案手法がベースライン手法に劣る結果を返した。このうちもっとも適合率の差が顕著であった q_r = “上げつない”, q_t = “商売” の検索タスクについて、考察および解説を行う。このタスクでは、ベースライン手法、提案手法ともに検索結果の適合率が高かったが、ベースライン手法の結果がより高い評価を得た。ベースライン手法の適合率が極端に高かった理由として、“上げつない” というリアクションクエリがネガティブな意味合いが強く、通常ウェブサイトを書き手が記述したがない点が挙げられる。ベースライン手法においてランキング上位に登場した “上げつない” および “商売” の両方を含む 20 件のウェブページのうち、11 件が Twitter や掲示板の意見を集約したまとめサイトであった。これらのサイトは、ニュースやトピックについてウェブコミュニケーション上で行われたやり取りの結果をそのままページのコンテンツとしているため、ニュース記事と読み手の声が一体になったよう

表 4 語スコア例 (えげつない)

w_i	$s_a(w_i, q_r)$	$s_c(w_i, q_r)$	$s_w(w_i, q_r)$
えげつない	1.0000	0.8923	0.8923
酷い	0.1250	0.0210	0.0026
ひどい	0.0893	0.0091	0.0008
すぎる	0.1071	0.0038	0.0004
ない	0.3571	0.0011	0.0004
悪	0.0714	0.0053	0.0004
現実	0.0714	0.0050	0.0004
レベル	0.1071	0.0032	0.0003
やめ	0.0714	0.0042	0.0003
だろ	0.2143	0.0014	0.0003

なページ構成をしている。そのため、文中に読み手の意見が含まれており、ベースライン手法でトピック、リアクションともに適合率の高い記事が発見されたと思われる。一方で、提案手法で発見したページには、ニュース記事や企業、事業主のウェブサイト、商売繁盛の祭事が多く含まれた。提案手法で発見されたウェブページの中で、直接的に“えげつない”という語を含むリアクションのついたページは1つもなく、意味的類似度の高いと計算された語を含むウェブページが表示された。このときのリアクションクエリ“えげつない”と語の関連度の上位10件とスコアを表4に表す。提案手法では、“商売”に関するニュースなどの通常のウェブページのうち、“酷い”や“悪”などの語を含むリアクションを含むサイトを検索結果上位に表示した。提案手法においてクエリに適合していると判断された7件のウェブページのうち本文中に“えげつない”という語が含まれているページは1件のみであった。

5. 関連研究

本研究を行う上で、テーマ的、技術的な関連をもつ先行研究について説明する。

5.1 ソーシャルアノテーション

読み手がウェブページを閲覧した際にどう感じたかを付与する別の研究として、ソーシャルブックマークやタギングに代表されるソーシャルアノテーションが挙げられる。ソーシャルアノテーションの情報検索への利用は Heyman [2], 山家 [3] らによって研究がなされている。本研究におけるリアクションも広義のソーシャルアノテーションの範疇に含まれるが、ソーシャルアノテーションサービスを通して付与されたタグとリアクションは性質が異なる。1日当たりに付与されるアノテーションの量に関して、形式の固定されるこれらサービスのタグ付けに対して、Twitterをはじめとするウェブコミュニケーションから発見可能なリアクションは多い。また、アノテーションが行われる Web ページの質に関しても、わざわざタグ付けするまでもないような瑣末なページに対して、リアクションという形でならアノテーションが発見可能である。一方で、タグが単語からなり構造化が容易であるのに対し、リアクションは短文で表されることが多く、より多様な表現が許される。そのため、ユーザの生の声が反映されやすい反面、検索に利用する際には、2.3 で述べたように、表現の揺らぎを吸収する処理が必要

となる。

5.2 評判情報検索

本研究においては、ウェブページを読んだ際にどう感じるかというユーザの記述をコミュニケーションサイトに含まれるリアクションからマイニングした。これと同様に、商品を購入、利用した際にどう感じるかというユーザの記述をブログや個人のウェブページなどから抽出する研究として、評判情報検索がある [4]。評判情報検索では、多くの場合、ある商品に関する評判がどのようなものであるかをウェブから収集することを目的とする場合が多く、また結果は観点と、ネガティブ / ポジティブによる極性からなる。評価表現をクエリとする商品検索の研究として杉木ら [5] の研究などがある。

5.3 センチメント分析

文書の持つ感情情報を利用した検索手法として、センチメント検索 [6] がある。ウェブページやニュース記事を内容から分析し、“明るい 暗い”、“承認 許否”などのいくつかの定められた軸ごとに感情の度合いを算出する [7]。ある文書の持つセンチメントの算出には文中の要素を利用しているため、現れるセンチメントは読み手ではなく書き手によって大きく左右される。この点が本稿における、読み手がどう感じるかに基づいたウェブ検索と大きく異なる。

6. 今後の課題

本論文において提案した手法について、実装および評価を行った結果、以下のような課題が見出された。

今回はプロトタイプとして、もっともシンプルなランキングアルゴリズムを適用した。そのため、すべてのケースに適用できているわけではなく、一部で不適当なランキングが行われている。アルゴリズムの洗練化が必要である。

また、本稿においては、実験用データセットのため、リアクションの収集もとを Twitter に限定し、特定の形式のもののみを用いた。これは本来ウェブ上に存在するリアクションのうちごく一部であるばかりか、不適当なものも含まれている。より精度の高いリアクションの収集、ならびに匿名掲示板 2ch^(注3)や Yahoo 掲示板^(注4)など収集対象の拡大が課題である。

類似したリアクションのトピック非独立性、ページ間類似度のリアクション非独立性についても考慮する必要がある。今回の手法では、リアクションクエリと語の類似度は、トピッククエリとは独立に算出している。しかしながら、実際にはトピックによって同一の感動を表すリアクションは異なる。例として、カレーについて“おいしい”と“辛い”は類似したリアクションといえるが、トピックがケーキだった場合には“辛い”は“おいしい”と同義ではない。同様に、ページ間類似度についても、今回はリアクションの種類によらずすべて単純なコサイン類似度に依っているが、検索に用いたリアクションの種類によっては、コサイン類似度では不適当な場合がある。例として、“簡潔だ”というリアクションに基づいた検索を行っている際、本

(注3): <http://2ch.net/>

(注4): <http://messages.yahoo.co.jp/>

文のコサイン類似度が高くても、長く冗長なページは結果としてふさわしくない。これら二つの点については今後対応してゆく必要がある。

本稿で提案したシステムでは、クエリとして入力可能なリアクションは一種に限られ、検索の自由度が十分でない。複数のリアクションを入力可能であることが望まれる。この際、通常のトピック検索と異なり、単純なアンド検索、オア検索では不十分である。また、“賛否両論”や、“好みが分かれる”など、リアクションのつき方の特徴による検索も考えられる。

先行研究として挙げたソーシャルアノテーションや、QA サイトに現れるユーザの入力したデータは、リアクションと似た性質を持っている。これらのデータを統合し利用可能にすることで、より多彩な意図を汲むことができ、より多くのページを網羅することが可能になる。

同じ文書を読んだ時にその文書から何を感じるかは、人によって大きく異なるため、検索のパーソナライゼーションが必要である。本稿においてはすべてのリアクションを同列に取り扱っているが、検索者と似た感性の持ち主のリアクションを重視するようなランキングにすることで、より読み手の立場に立った検索が可能になると考えられる。

今回提案した手法では、リアクションのスコアリングをする際、使われている語や対象となっているウェブページのみを考慮し、発信者情報などは利用していない。ウェブページによっては一部の熱狂的なファンだけが好意的なリアクションをとっている場合など、バイアスがかかっていることが考えられる。ウェブコミュニケーションサイト上における発言者の関係性やユーザごとの嗜好などのソーシャルファクタを考慮することで、より万人受けするランキングにすることができると考えられる。

また、今回はリアクションのスコアリングに自然言語的な処理を施しておらず、否定語を含むリアクションが検索結果の精度を下げる要因となっている。リアクションのランキングに関しても、クエリとして入力されたリアクションとのマッチ度のみに基づいている。“とても”や、“少し”などの程度を表す表現があった際に、正しいランキングが行われない問題がある。スコアリング時に言語パタンを用いることで、否定語や程度などを考慮した検索が可能になる。

本稿においては、手法の提案を目的としているため、もっとも基礎的で単純な形式をとっている。実際にウェブ検索エンジンとして実用するためには、さらなる改良が必要である。

7. ま と め

本稿では、文書を読んだ際にどのように感じるかに基づいた読み手目線の検索を実現する、リアクション検索について提案を行った。提案手法ではウェブコミュニケーションデータを利用し、クエリとしてリアクションを入力可能にすることでこのような検索を可能にした。

提案した手法について実際にウェブアプリケーションを実装し、被験者による評価実験を行った。結果としてクエリに適合するウェブページがおおむね意図にしたがった順位で表示されることを確認した。また、クエリ尤度モデルに基づく既存の検

索手法との比較を行い、提案した検索手法の有用性について明らかにした。

謝 辞

本研究の一部は、グローバル COE 拠点形成プログラム「知識循環社会のための情報学教育研究拠点」、文部科学省科学研究費補助金特定領域研究「情報爆発時代に向けた新しい IT 基盤技術の研究」、計画研究「情報爆発に対応するコンテンツ融合と操作環境融合に関する研究」(研究代表者: 田中克己, 課題番号 1809041) によるものです。ここに記して感謝の意を表します。

文 献

- [1] Christopher D. Manning, Prabhakar Raghavan and Hinrich Schtze “Introduction to Information Retrieval,” Cambridge University Press (2008).
- [2] P. Heymann, G. Koutrika and Hector Garcia-Molina: “Can social bookmarking improve web search?,” Proceedings of the international conference on Web search and Web Data Mining, pp. 195–206 (2008).
- [3] Y. Yanbe, A. Jatowt, S. Nakamura and K. Tanaka: “Can Social Bookmarking Enhance Search in the Web?,” Proceedings of the 7th ACM/IEEE-CS Joint Conference on Digital Libraries (JCDL2007), pp.107–116 (2007).
- [4] 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).
- [5] 杉木健二, 松原茂樹: “消費者の意見に基づく商品検索,” 情報処理学会論文誌 49 7 pp.2598–2603 (2008).
- [6] Na J.C. and Khoo C.S.G., Chan S. and Hamzah N.B.: “Sentiment-based search in digital libraries”, Digital Libraries, 2005. JCDL’05. Proceedings of the 5th ACM/IEEE-CS Joint Conference, pp.143–144 (2007).
- [7] Tadahiko Kumamoto and Katsumi Tanaka: “Proposal of Impression Mining from News Articles,” Lecture Notes in Computer Science, Volume 3681/2005, pp. 901–910 (2005).