5. 評価実験

本章では,4 章で提案したクエリ拡張手法が,元の入力 クエリで得られた上位 k 件のウェブページから得られるタ スクに対して,どの程度新しいタスクを含んだウェブペー ジを発見できているのかを評価した.

5.1 ベースライン手法と実験設定 提案手法の有効性を検証するため,検索連動型広告を用

いず,通常のウェブ検索の検索結果を用いてクエリ拡張を 行いウェブページを収集する手法をベースライン手法と して用意した.具体的には,5.1 節で示した手法のうち, {a1 , . . . , an } を検索連動型広告ではなく,通常のウェブ検 索結果上位 n 件とした手法である.

実験にあたりウェブ検索結果の取得には Google Custom

情報処理学会研究報告

IPSJ SIG Technical Report

Search API\*1を用い,検索連動型広告の取得には,Yahoo! Japan が提供するスポンサードサーチ\*2を利用した.また, 形態素解析には MeCab\*3を用いた.

節で述べた手法のパラメータについては,今回の実験で は,n = 15,k = 20,k′ = 5 と設定した.すなわち,クエ リ拡張を行い最終的に 20 件のウェブページを取得し,そ のウェブページにどの程度,元のクエリで得られる 20 件 のウェブページに含まれていないタスクが含まれているの かを評価した.

5.1.1 実験に用いたクエリ 本稿では,“花粉症の対策をする”,“部屋を掃除する”,

“結婚する”,“コーヒーを淹れる”,“ダイエットする” とい う 5 種類のクエリを用いて実験を行った.これらのクエリ を選択した理由は,これらの目的を達成する手段が豊富に 存在すると考えたためである.

5.2 評価尺度 提案手法の性能を評価するに当たって,以下の 2 種類の

評価基準を設け,人手による検証を行った.

* 行動数:クエリ拡張で得られた上位 20 件のウェブペー ジ集合から得られた,目的の達成に必要な行動の数の うち,元のクエリで検索して得られた上位 20 件のウェ   ブページには存在しなかった行動の数.
* 詳細度:新しく得られた行動のうち,その行動を達成 するために必要な情報が得られたウェブページに書か れている場合はその行動に対して 2 点,行動が列挙さ れているだけでその行動に達成するために必要な情報 が得られない場合は 1 点として上記の行動数を重みづ   けた尺度.
* たとえば,クエリ拡張を行い新しく取得したウェブペー
* ジの中から,“加湿器を使う” という行動が得られたとき, 行動数を 1 として計算する.その際,もしウェブページに 加湿器を購入するための情報や良い加湿器に関する情報な どが含まれていたら詳細度を 2 点,そうでない場合は 1 点 として計算する.  5.3 実験結果 得られた結果を表 1 に示す.まず,新しく得られた行  動数に着目してみると,表 1 から分かるように,“部屋の 掃除をする”,“結婚式する”,“コーヒーを淹れる” という クエリでは,目的を達成するために必要な行動として,提 案手法によりクエリ拡張を行うことで新しいものがベース ラインよりも多く取得できていることが分かる.また,詳 細度について着目してみると,表 1 から分かるように,“ ダイエットをする” 以外の 4 つのクエリについて,提案手
* \*1  https://developers.google.com/custom-search/?hl=ja
* \*2  http://search.yahoo.co.jp/search/ss
* \*3  https://code.google.com/p/mecab/

表 1 Table 1

花粉症の対策をする 部屋の掃除をする 結婚する コーヒーを淹れる 5 9 4 4 ダイエットする 5 9 5 10

法がベースライン手法を上回っていることが分かる.これ

は,検索連動型広告を用いてクエリ拡張を行うことで,よ

りタスクの達成に特化した,タスクの達成に必要な情報が

記載されたウェブページが取得できたことを示していると

考えられる.

クエリ拡張語としてどのようなものが得られていたのか を分析すると,“部屋の掃除をする” というクエリについて は,提案手法では “クリーニング”,“配送”,“洗浄” とい た単語が得られていた.一方ベースラインでは,“上がる” や “維持” といった単語が得られており,検索連動型広告 を利用することで,タスクに関連した単語がクエリ拡張語 として得られていることが分かる.また,“部屋の掃除を する” というクエリでは,提案手法は “掃除の代行業者に 頼む” というタスクを含んだウェブページが得られており, 検索連動型広告に出現するクリーニングや洗浄といった単 語が,こうしたタスクに関連したウェブページの発見を可 能にしたと考えられる.

一方で,“ダイエットする” というクエリについては,ど ちらの評価尺度ともベースラインと比較して,高い値は得 られなかった.提案手法では,“ダイエットをする” に対し て,“吸引”,“体験” といった単語が拡張語として得られ ており,実際に “吸引” という語を追加することで脂肪吸 引に関するウェブページを発見することができていた.一 方で,ベースライン手法では “食べる” という語がクエリ 拡張語として得られており,元のクエリでは既にダイエッ トに有効な食事に関するウェブページが多く出現していた ものの,“食べる” を元のクエリに追加することで,より多 様な食事に関するダイエット方法を記載したページが得ら れ,結果として提案手法よりも得られる行動数が多くなっ ていた.