

クエリ推薦に基づくモバイル協調検索支援手法の評価

大重 智志¹ 中村 聡史¹ 田中 克己¹

概要: 我々は、モバイル環境におけるグループでの意思決定を効率化することを目的として、グループとしての意図を汲み取り、グループの検索を発散または収束させるクエリを推薦することによって検索行動を支援するモバイル協調検索を提案してきた。本稿では、提案手法を用いた評価実験を実施することにより、手法の有効性を明らかにする。また、詳しい分析を行うことによりその特性を明らかにするとともに、手法の改良を行う。

キーワード: 協調検索, モバイル協調検索, クエリ推薦

Evaluation of Mobile Collaborative Search with Query Suggestion

SATOSHI OHSHIGE¹ SATOSHI NAKAMURA¹ KATSUMI TANAKA¹

Abstract: We have proposed a method to support mobile collaborative search by suggesting queries for promoting divergence of a topic and ones for promoting convergence of a topic so that group members can make decision smoothly using their smartphone. In this paper, we make an experiment using proposed method and clarify usefulness of our method. Additionally, we analyze in detail to reveal features of our method and improve our method.

Keywords: collaborative search, mobile collaborative search, query suggestion

1. はじめに

爆発的な人気を博すスマートフォンによって、人々は時と場所を選ばずに Web 検索を行うことができるようになった。VisionMobile によると 2011 年上半期時点での全世界での携帯電話保有者のうち、スマートフォン利用者は 27%に達し [1]、アメリカでは、スマートフォン利用者は 46%にもなる。特に、25 歳から 34 歳までの利用者は 71%にも達する [2]。

Morris[3] は、複数のユーザが協調して Web 上で情報検索を行うという行為が多く行われていることを明らかにしている。例えば、グループ旅行の計画のために必要な情報を皆で検索したり、ネットショッピングで様々な候補の中から最善のものを選択し購入したりするなどの状況が挙げられる。本稿では、このような複数ユーザでの検索行動を

協調検索と呼ぶ。スマートフォンの普及により、時と場所を選ばずに Web 検索を行うことができるようになったため、PC を利用した協調検索だけではなく、小型のモバイル端末を用いて状況に応じた協調検索（モバイル協調検索）を行う機会も増えてきた。例えば、グループでの旅行中にその周辺の観光スポットを探すような状況、パーティーの後で急遽行くこととなった 2 次会の店を探す状況、災害直後に被災したメンバーと帰路を探す状況などがある。一般的な PC を用いた協調検索とモバイル協調検索の大きな違いは、後者は突発的に検索を開始することが多く、前者に比べて情報検索に十分な時間を掛けることができない点である。また、後者はディスプレイサイズが小さいため表示および共有できる情報量にも制限がある。

モバイル協調検索では、グループとして満足のいく結論に達する必要があることが多く、最後に意思決定が行われる。しかし、現在の何の支援もないモバイル協調検索では意思決定までスムーズに進めることは容易ではない。この

¹ 京都大学大学院 情報学研究科 社会情報学専攻
Department of Social Informatics, Graduate School of Informatics, Kyoto University

原因として、我々は大きく分けて2つあると考えている。1つ目は、グループの複数ユーザが重複またはとても類似した検索行動を行ってしまう点である。グループとしての検索行動の方針が決まっていない検索開始初期などにおいて、似たようなクエリで検索すると、グループ全体での探索範囲が狭まってしまう。そのため、グループ全体として網羅的な探索を行うことができず、質の高い良いページに辿り着けないなど、満足のいく結論に至らないことがある。2つ目は、各ユーザが個人の好みに合わせすぎて自由に探索を行ってしまうという点である。各ユーザの好みだけで行動すると、いつまで経っても探索範囲がバラバラでグループとしての的を絞られなくなってしまう。結果として、最終的なゴールに辿り着けず目的を果たせなくなってしまううえ、時間の浪費となってしまう。なお、この両者は探索範囲が狭くなっていることと探索範囲が広がっていることとでトレードオフの関係にある。

我々は上記の問題を解決するため、クエリ推薦によってモバイル協調検索における円滑な意思決定を支援する手法を提案してきた[4]。本稿では、提案および実装した手法について簡単に述べた後、システムを用いた実験を実施し、その結果とともに詳細な分析を行う。

2. 関連研究

SearchTogether[5]は、遠隔地にいるユーザがPCを用いて同期的・非同期的に協調検索を行うためのシステムである。各ユーザがそれぞれのPCで入力したクエリが他の協調ユーザと共有されたり、気に入ったページを推薦したり、ページに対して評価やコメントなどを付加できるなどの機能がある。ユーザはこうした機能を利用することで、効率的に協調検索を行える。

CoSearch[6]では、複数のユーザが1台のPCと複数のマウス・モバイル端末を用いて協調検索を行う。各ユーザにはカーソルが1つずつ割り当てられ、ディスプレイ上に色で区別されたカーソルが表示されている。各ユーザは自身のカーソルをマウスで操作し、クエリや閲覧ページをキューに貯めて協調検索を行う。モバイル端末では、マウスの機能に加え、クエリ入力やページの閲覧、ページに対するメモなどが行える。これにより、1台のPC上で効率的に協調検索を行うことができる。

伊豆らの研究[7]では、ユーザの探索に同期してシステムが有用なWebページをグループの共有データ内から選び提示することで、探索を支援する。検索結果に至るプロセスやその行為を含めて共有することで、ユーザの探索目的や探索範囲を把握する。ユーザの行動に沿ったページを提示するために、ユーザの検索行動を、いろいろなページを見ている発散フェーズ、類似ページを続けて見ている収束フェーズ、ユーザが考える範囲を調べ終えた補完フェーズに分け、ユーザがどのフェーズにいるかを判定し、それに

応じて何を提示すべきかを判断している。それらのフェーズは、探索目的に沿ったページをどれ程多く閲覧したかを表す指標である達成度と、どの程度他の類似Webページと吟味されて推薦ページは選ばれたのかを表す指標である精選度の2つの尺度を用いて決定している。

これらの研究は1台以上のPCが使えることを想定している。本稿は、モバイル端末のみの状況を想定しているという点でこれらの研究とは異なる。また、伊豆らの研究は既に閲覧されたページを提示するのに対し、本稿ではクエリを推薦するという点で異なる。さらに、本稿でも発散・収束という語を用いているが、発散・収束がどの程度の強さであるかという考え方をしている我々とは異なる。

奥らの研究[8]では、ユーザの意見を反映したウェブコンテンツを共有し、それらをユーザの評価で整理および提示することで、ウェブコンテンツの比較・検討を支援するインタフェースを提案している。

小谷らの研究[9]では、気に入ったウェブページを他のユーザの検索結果一覧に割り込ませることにより通知する機能や、各ユーザの探索空間を共有する機能により、ユーザ間の情報共有を支援している。

これらの研究は我々の研究同様モバイル端末を用いた協調検索を想定しており、情報の比較・共有に注目している。本稿では比較や共有という観点ではなく、話題の発散や収束を促すことで意思決定をスムーズに進めるよう支援するものである。

3. 提案手法

3.1 モバイル協調検索の問題点と解決案

モバイル協調検索はグループ全体として満足する必要がある。全員が満足する意思決定のためのページを発見しなければならない。しかし、モバイル協調検索では意思決定までスムーズに進めることは容易ではない。1章で述べた通り、我々はグループの重複した検索行動と的を絞れない検索の仕方が原因であると考えている。

そこで我々は、話題を発散させることで探索範囲を広げてグループ内での重複した検索行動を防ぎ、話題を収束させることでグループ内での探索範囲を狭めて的を絞った情報検索を行えるようにする。具体的には、システムが各ユーザの入力クエリからグループ全体としての意図を推定し、話題の発散を促すクエリと話題の収束を促すクエリを推薦することで、グループ全体としての話題の発散・収束をコントロールし、意思決定における問題の解決を図る。ここでは、発散を促すクエリを「発散クエリ」、収束を促すクエリを「収束クエリ」と呼ぶ。なおここで、発散クエリとは、ある入力されたクエリとは関連があるが別の観点によるもののことである。また、収束クエリとは、ある入力されたクエリと関連があり、さらに詳細なものを指すものとする。

3.2 推薦クエリの生成

各ユーザの入力したクエリからグループ全体としてのおおまかな検索の背景を推定し、その後、その背景に基づいて発散クエリと収束クエリを生成する。ここで背景とは、検索の文脈を表すもので、そのような背景を表した語を背景語と呼ぶ。

3.2.1 発散クエリ

発散クエリとしては、あるクエリをシードとし、そのクエリと関連はあるが別の観点であるものを生成する。以降、シードとなるクエリのことをシードクエリと呼ぶ。

例えば、“沖縄 観光”をシードクエリとして考える。このクエリには2つの背景が考えられ、それは「沖縄へ旅行するときに、観光地について検索する」という背景と「暖かく海の綺麗な場所を観光したいので、特に沖縄について検索する」という背景である。前者の場合、沖縄で観光だけでなくショッピングのことも考えられるため、“沖縄 ショッピング”が発散クエリとして考えられる。後者の場合、暖かく海の綺麗な観光地は他にも奄美大島があるため、“奄美大島 観光”が発散クエリとして考えられる。

ここで、各クエリに含まれる語に関して注目すると、前者の場合「沖縄」、後者の場合「観光」という語が共通して現れており、それぞれが背景に対応している。また、前者の場合、「観光」と「ショッピング」という語が、後者の場合、「沖縄」と「奄美大島」という語が、それぞれ同位概念のような関係になっている。以降、このように同位概念のような関係になる語を兄弟語と呼ぶ。

そこで、あるシードクエリ“ $t_1 t_2$ ”から背景を表す語 t_1 を選び、残りの語 t_2 の兄弟語 t_2' を求め、新しい発散クエリ“ $t_1 t_2'$ ”を求める。このようにして発散クエリを生成する。

3.2.2 収束クエリ

収束クエリとしては、あるクエリをシードとし、そのクエリと関連があり、さらにそのクエリを詳細化するものを生成する。

以下、“沖縄 観光”をシードクエリとして考える。“沖縄 観光”で「観光」に注目して詳細化する場合、「沖縄での観光、特に海に近いところ」という意味で“沖縄 観光 ビーチ”が収束クエリとして考えられる。また、「沖縄」に注目すると、「沖縄、特に那覇での観光」という意味で“沖縄 那覇 観光”などが収束クエリにあたると考えられる。

ここで、クエリに含まれる語に関して注目すると、「ビーチ」は「沖縄」という背景で「観光」を詳細にするもので、「那覇」は「観光」という背景で「沖縄」を詳細にするものである。

また、グループ全体としての的を絞るという意味で、シードクエリ自身もそのシードクエリを入力したユーザ以外のユーザにとっては収束クエリとなりうる。

そこで、あるシードクエリ“ $t_1 t_2$ ”から背景を表す語 t_1

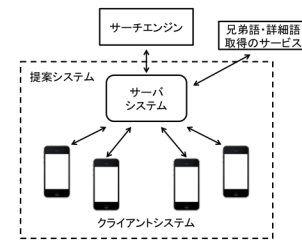


図1 システムの構成

を選び、残りの語 t_2 をより詳細にする語 t_3 （詳細語）を求め、収束クエリ“ $t_1 t_2 t_3$ ”を求める。また、シードクエリ“ $t_1 t_2$ ”自身も収束クエリとする。このようにして収束クエリを生成する。

3.2.3 背景語とシードクエリの発見

クエリに多く現れる語はそれだけ背景を意味していると考えられるので、背景語は全クエリに含まれる単語のうち頻出のものとする。

また、シードクエリは、クエリ履歴の中から、求めておいた背景語が含まれるクエリを取得する。ここでは、簡単化のために、シードクエリはある単語と背景語の2つだけからなるものに限る。

3.2.4 兄弟語と詳細語の発見

クエリ“ $t_1 t_2$ ”から背景を表す語 t_1 を選んだ後は、残りの語 t_2 の兄弟語もしくは詳細語を見つける必要がある。

兄弟語や詳細語を求める研究は数多くあるが、本稿では大島ら[10]、野田ら[11]の研究に着目した。大島らは本稿での「兄弟語」を「同位語」と呼んでおり、野田らは本稿での「詳細語」を「話題語」と呼んでいる。これら同位語と話題語を取得する手法が実装されたAPI^{*1}を利用し、本稿では兄弟語と詳細語を取得する。

3.3 システムの実装

我々は、上述の手法を基にサーバ/クライアントモデルを採用したプロトタイプシステムを設計した。サーバはPythonを用いてCGIとして動作し、クライアントはObjective-Cを用いてiOS5.0を搭載した端末上で動作する。システムの構成は図1のようになっており、サーバは検索エンジンや兄弟語・詳細語を取得するためのサーバとやりとりしている。

図2は検索を行うメイン画面である。上部にクエリを入力する検索バーがあり、それを選択すると図3のようにキーボードが現れる。キーボードの検索ボタンを押すと、中央に検索結果がランキングで表示される。また、検索バーと検索結果の間に、推薦クエリを表示するエリアがある。本システムでは、6つのクエリを推薦する。

^{*1} <http://www.dl.kuis.kyoto-u.ac.jp/~ohshima/wiki/index.php?YayalaAPI>

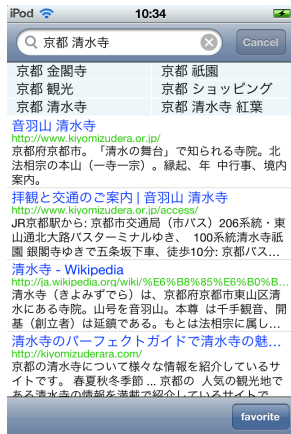


図 2 メイン画面

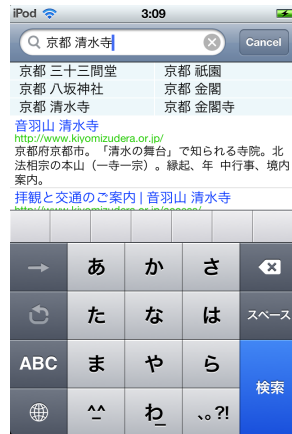


図 3 メイン画面のキーボード

4. 評価実験

提案手法の有用性を示すため、協調検索によって被験者にあるタスクを達成してもらうという実験を実施した。実験によって、発散クエリによってグループとして網羅的に検索し様々な候補を検討できているか、収束クエリによってグループとしての的を絞った検索ができているか、それらによって意思決定がスムーズに進んでいるか、これら3つを明らかにする。以下では、比較システム及び実験の手続き、その結果について述べる。

4.1 比較システム

我々は、提案手法との比較のために、3.3 節のプロトタイプとほとんど同じではあるが、一部機能が異なるシステムを作成した。

提案手法によるプロトタイプシステム（以下、提案システム）と比較のためのシステム（以下、比較システム）とでは、推薦されるクエリの生成・推薦の仕方のみが異なっている。提案システムでは、推薦されるクエリを3章の手法により生成・推薦しているが、比較システムでは、Yahoo!の関連検索ワードのAPI^{*2}を用いており、ユーザの入力したクエリからAPIによって関連検索ワードを取得し、それらをクエリとして推薦する。

比較システムは、一般的な関連キーワードによるクエリ推薦を模したものであり、大量のログデータから関連して検索されるものをクエリとして推薦する仕組みで、システムを通して協調することは想定していない。それに対して、提案システムでは、グループの意図を汲み取ってその意図にちなんだクエリを推薦することを目指している。そのため、これら2つのシステムを比較することで、大量のログデータに頼ったものでなく、そのグループの意図に沿った推薦を行うことが重要であるということを明らかにできる

^{*2} <http://developer.yahoo.co.jp/webapi/search/assistsearch/v1/webunitsearch.html>

と考えている。

4.2 実験内容

我々は、大きく分けて2つの実験を行った。1つ目は協調検索による実験とその被験者による主観的な評価であり、2つ目は全ての協調検索の実験が終わった後の他者による客観的な評価である。

4.2.1 協調検索による実験

我々はまず、システムを利用しモバイル協調検索のシチュエーションを模した実験を行った。この実験では先述の通り、網羅的に検索できているか、的を絞った検索ができているか、意思決定がスムーズに進んでいるか、これら3つを明らかにする。また、システムが推薦したクエリのクエリとしての適合性と、ユーザの感じたそのクエリの発散・収束具合を調べるために、個人の判断によるクエリ推薦の精度の評価も行う。

まず、実際にグループ単位でタスクを与え、協調検索を行ってもらう。その後、どのような方針で協調検索を行ったかをディスカッションしたうえで、最終的にアンケートに回答する。この流れを1グループにつきそれぞれの手法で1回ずつ行う。なお、協調検索を開始する前には、システムに依存しない事前アンケートに回答してもらう。内容は、日常的に協調検索を行うか、スマートフォンの入力に慣れているかなどである。

被験者は32名の男女で、日常的に会話を行うメンバー4人をひとつのグループとして、合計8グループに分けた。

我々が与えたタスクを、提案システムと比較システムとでそれぞれ1タスク10分程度で行ってもらった。実験は、20m²程度の個室で立った状態で、経過時間が分かるようにタイマーを被験者に見えるようにしたうえで、以下の手順で行った。

- (1) 実験の内容と目的を説明する。
- (2) システムの使い方を説明し、操作に慣れてもらう。
- (3) 事前アンケートに回答してもらう。
- (4) 1つ目のタスクを与え、協調検索を行ってもらう。
- (5) 実際に推薦されたクエリの精度を評価してもらう。
- (6) 協調検索の方針をディスカッションしてもらう。
- (7) 5段階のリッカート尺度で問うアンケートを行う。
- (8) 2つ目のタスクに対して、(4)～(7)を繰り返す。

実験で使用したタスクは、タスク0：京都で飲みに行く場所、タスク1：博多で観光する場所、タスク2：札幌で観光する場所、タスク3：金沢で食事する場所、タスク4：仙台で食事する場所、それぞれを探すというもので、タスク0はシステムに慣れるための練習用タスクである。

さらに、我々は協調検索のタスクを終えた後に、上記の(5)のように被験者にクエリの精度を主観的に評価してもらった。実際にユーザに推薦された全てのクエリを、そのクエリが推薦されたときにユーザが入力していたクエリと

表 1 掛かった時間と発話数の平均

	結論に達するまでの秒数	発話回数
提案	589.1 ($\sigma = 54.5$)	25.2 ($\sigma = 12.9$)
比較	604.3 ($\sigma = 37.4$)	29.8 ($\sigma = 14.7$)

ともに表示し、それぞれを2つの尺度で評価してもらった。2つの尺度とは、そのクエリが推薦されるクエリとして適切であるかという尺度である「適合度」と、そのクエリがより広範囲を検索するのに役立つかまたは的を絞って狭い範囲を検索するのに役立つかという尺度である「収束度」で、それぞれ1~5の5段階である。適合度は1の場合が不適切、5の場合が適切であり、収束度は1の場合がより広い範囲を検索するのに役立ち、5の場合がより狭い範囲を検索するのに役立つということである。

4.2.2 協調検索後の客観評価

我々は、全ての実験を終えた後、推薦された全てのクエリの精度を客観的に評価した。協調検索の直後に個人によるクエリの精度の評価は行ったが、それはそのユーザ自身の視点からの主観的な評価であり、この評価は第三者による客観的なものである。被験者にはグループの検索意図(タスク)を見せ、実際に各時点でユーザが入力したクエリを提示した。それらを考慮したうえで、先述の適合度と収束度の2つの尺度で、推薦されたクエリを全て評価してもらった。

4.3 実験結果

どのタスクについてもほぼ10分程度で結論に達した。実験の様子を撮影した映像と操作ログ、アンケート、クエリの精度の評価について分析を行った。

グループ毎の結論に達するまでに掛かった時間と発話数の平均を表1に示す。結論に達するまでに掛かった時間は提案システムの方が比較システムよりも短くなっており、10分を切っている。つまり、提案システムの方が短時間でスムーズに意思決定に至ったことがわかる。また、グループ毎の平均発話数は、提案システムの方が比較システム比べて少なくなっている。

操作ログからのユーザの入力したクエリ数や閲覧ページ数などを表2に示す。ユーザの入力したクエリのグループでの総数の平均は18.6個($\sigma = 6.0$)、比較システムでは平均19.6個($\sigma = 4.2$)であり、ユニーククエリ数は提案システムでは平均14.0個($\sigma = 3.2$)、比較システムでは平均16.0個($\sigma = 3.5$)であった。提案システムでは各ユーザに同じ時点では同じクエリを推薦するので、クエリ数もユニーククエリ数も比較システムの方がわずかに多くなっている。しかし、入力クエリのうち推薦クエリから選択したものの占める割合は、提案システムで31.5%、比較システムでは13.4%と、提案システムの方が圧倒的に多く、より興味を示すクエリが提示されたということがわかる。グルー

表 2 クエリ数や閲覧ページ数など

	クエリ数	ユニーククエリ数	入力クエリのうち推薦クエリの割合
提案システム	18.6 ($\sigma = 6.0$)	14.0 ($\sigma = 3.2$)	31.5%
比較システム	19.6 ($\sigma = 4.2$)	16.0 ($\sigma = 3.5$)	13.4%
	閲覧ページ数	ユニーク閲覧ページ数	-
提案システム	64.0 ($\sigma = 11.4$)	38.4 ($\sigma = 8.4$)	-
比較システム	74.5 ($\sigma = 8.4$)	42.9 ($\sigma = 4.4$)	-

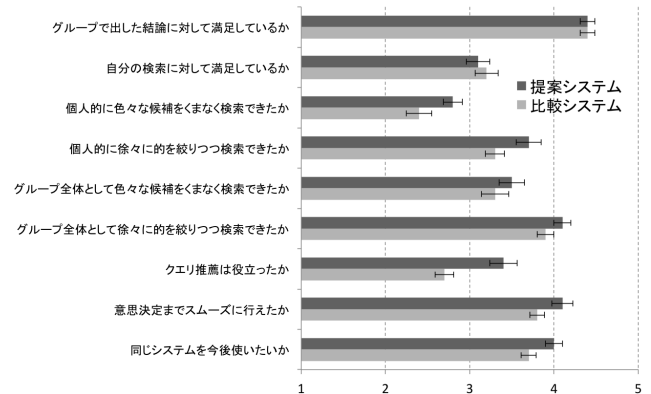


図 4 アンケート結果

プ毎の閲覧ページの総数の平均は提案システムで64.0ページ($\sigma = 11.4$)、比較システムで74.5ページ($\sigma = 8.4$)であり、ユニーク閲覧ページ数の平均は提案システムで38.4ページ($\sigma = 8.4$)、比較システムで42.9ページ($\sigma = 4.4$)となっている。いずれも提案システムの方が少なく、ばらつきは大きい。つまり、提案システムは比較システムよりも見ているページが少ないということになる。

アンケートの結果を図4に示す。「グループで出した結論に対して満足しているか」、「意思決定までスムーズに行えたか」といった質問に対しては、提案システムと比較システムで有意差はなかった。しかし、「クエリ推薦は役立ったか」という質問に対しては、提案システムと比較システムで有意な差が見られ($p < 0.05$)、提案システムの方が良い結果となった。また、「個人的に色々な候補をくまなく検索できたか」という質問に対しても有意差があり、提案システムの方が良い結果となった。

クエリの精度を2つの尺度で評価した結果を表3に示す。上部が、協調検索を行ったユーザによる主観の評価で、下部が第三者による客観的评价である。全推薦クエリについての平均、発散クエリとして推薦したクエリについての平均、収束クエリとして推薦したクエリについての平均を載せている。ただし、主観的评价に関しては、推薦クエリのうち、ユーザが実際に選択したもののみの結果である。全体の適合度は、主観的评价でも客観的评价でも高めになっている。また、客観的评价における発散クエリの収束度は低く、収束クエリの収束度は高くなっていて、すなわち、発散クエリは発散クエリとして適切で、収束クエリは収束クエリとして適切であったということになる。しかし、主観的评价における収束度は、発散クエリでも収束クエリで

表 3 クエリの精度

主観的評価	全クエリ	発散クエリ	収束クエリ
適合度	4.20	4.22	4.18
収束度	-	3.28	3.54
客観的評価	全クエリ	発散クエリ	収束クエリ
適合度	3.33	3.19	3.58
収束度	-	2.51	3.18

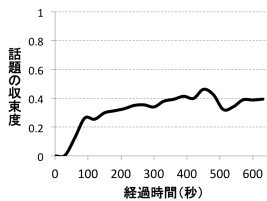


図 5 各グループの収束度の平均

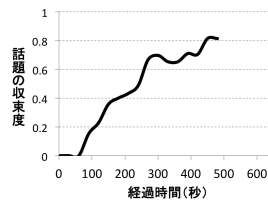


図 6 グループ 8 の収束度

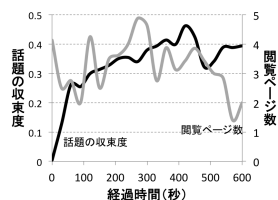


図 7 各グループの収束度と
閲覧ページ数の平均

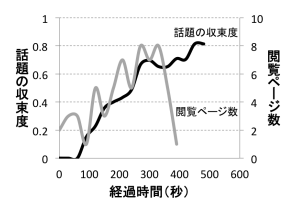


図 8 グループ 8 の収束度と
閲覧ページ数

も高めで、すなわち、ユーザの選択したクエリは、発散クエリか収束クエリによらず収束クエリとして認識されている傾向にあることになる。ただし、発散クエリよりも収束クエリの方が収束度がより高めであるので、両者に少しの差異はみられることになる。

また、協調検索を行った際にシステムが推定した各グループの収束度の平均の推移(図5)と、特に変化が顕著であったグループ8の収束度の推移(図6)の2つを示す。どちらのグラフも右肩上がりであり、網羅的に検索し、徐々に的を絞っていったということがわかる。

さらに、話題の収束度とグループでの閲覧ページ数の推移を重ね合わせた図を図7と図8に示す。閲覧ページ数の推移とは、30秒間隔で区切ったときに、その区間内でいくつかのWebページを閲覧しているかというものである。グラフから、話題の収束度が大きくなると、閲覧ページ数も大きくなる傾向にあることがわかる。特に、図8にはその様子が顕著に表れている。

5. 考察

5.1 結論に達するまでの検索行動と意思決定に対する評価

グループで出した結論や自分の検索に対する満足度は、提案システムと比較システムでそれほど違いはない。しかし、個人の判断とグループ全体としての判断によらず、色々な候補をくまなく検索できたか、徐々に的を絞りつつ検索できたかという質問に対しては、提案システムの方が

高く、比較システムを下回ることにはなかった。また、表1からわかるように、意思決定に至るまでの時間は提案システムの方が短く、10分を切っている。つまり、提案システムの方が、ある程度の満足度を保ったまま、短時間で、色々な候補を検索した上で徐々に的を絞って検索できたということがわかる。

両者の満足度に差は出なかったが、この理由としては、時間が十分だったことや、グループの人数が多くなかったことがあげられる。そこで、今後は人数を増やしたり、タスクを難しくしたり、時間をさらに制限することで、満足度に差が出るような実験を行おうと考えている。

閲覧ページ数、ユニーク閲覧ページ数は表2のように、いずれも比較システムの方が多くなっている。ユニーク閲覧ページ数が少ないということは、多くのページをくまなく検索できていないということになるが、アンケート結果(図4)を見る限り、提案システムの「色々な候補をくまなく検索できたかどうか」という項目で、悪い結果になっているわけではない。つまり、提案システムでは、より少ないページの閲覧で、色々な候補をくまなく検索しつつ、的を絞って検索できる傾向にあるということになる。逆に、比較システムでは、提案システムに比べて、ページ自体は異なるが「色々な候補をくまなく検索する」のには役に立たないページが多かったのだと考えられる。

5.2 クエリ推薦の有用性の評価

表2のようにクエリ数、ユニーククエリ数ともに提案システムの方が少なくなっている。これは、他のユーザが入力したクエリを元に推薦クエリを生成しているもので、当然のことである。しかし、入力クエリのうちユーザが推薦クエリから選択したものの割合は提案システムの方が圧倒的に多くなっている。これは、推薦されたクエリに有用なものが多かったということが理由として考えられる。また、クエリ推薦は役に立ったかどうかという質問に対して、提案システムの方が良い結果が得られたことから、推薦クエリは有用であったと考えられる。実際にクエリ推薦が有用にはたっているかどうかを調べるために、撮影映像と操作ログの分析を行った。事前アンケートでスマートフォンの入力に全く慣れていないと答えたユーザは、協調検索開始直後に入力に手間取っていた。そのとき、提案システムではクエリが推薦され、そのクエリを選択して検索を開始していた。つまり、入力の手間自体を解消するということがクエリ推薦によって実現されているということである。また、このユーザは、グループ内で最もクエリ推薦を利用しており、アンケートの全ての項目において、提案システムの方が高いか比較システムと変わらないかであった。

また、実際のクエリログの例を2つ示す。表4は、札幌旅行で観光する場所を探すというタスクで実験を行ったときのログで、表5は、金沢旅行で食事する場所を探すとい

表 4 「札幌旅行で観光する場所を探す」というタスクのクエリログ

経過時間	User A	User B	User C	User D
0:00			札幌 観光	札幌 名所
0:02				
0:13	札幌			
0:20		札幌駅 観光		
0:32			札幌 観光 おすすめ	
0:37	札幌 観光			
1:28				札幌 観光
1:48		札幌 観光 起爆剤 (c)		
1:52				札幌 観光 起爆剤 (c)
2:06				札幌 観光 (c)
2:19		札幌 観光 (c)		
2:34		札幌 観光 人気		
3:17	札幌 観光 起爆剤 (c)			
3:53		札幌 観光 (c)		
4:05		札幌 観光コース		
4:11	札幌 観光 (c)			
4:31				札幌 観光 魅力 (c)
5:03				札幌 観光 (c)
5:47				札幌 スポット
5:59				札幌 スポット 写真 (c)
6:20		札幌 スポット (c)		
6:28				札幌 スポット (c)
7:27	札幌 スポット 情報 (c)			
8:20				札幌 時計台
8:32				札幌 テレビ塔 (d)
8:49			札幌 テレビ塔 (c)	
8:52				札幌 大通公園 (d)

表 5 「金沢旅行で食事する場所を探す」というタスクのクエリログ

経過時間	User A	User B	User C	User D
0:00	金沢 名産			
0:12		金沢駅 名物	金沢 名物	
0:15				金沢 グルメ
0:21			金沢 グルメ 情報 (c)	
1:14				金沢 海鮮
1:42	金沢 ハントんラーメン			
1:59			金沢 海鮮 (c)	
2:07		金沢 のどぐろ		
2:44		金沢 のどぐろ (c)		
2:45				金沢 のどぐろ 塩焼 (c)
2:54	金沢駅 ハントんライス			
3:19		金沢 寿司		
3:27				金沢 のどぐろ 駅前
3:34			のどぐろ	
4:04		金沢 海鮮丼 (d)	金沢 海鮮丼 (c)	
4:18			金沢 海鮮丼	
4:21	金沢 海鮮丼 (c)			
4:37			金沢駅 海鮮丼	
6:05	金沢 半トライス 有名			
8:01		金沢駅 海鮮丼 (c)		
8:23				金沢駅 海鮮丼 (c)
8:38	金沢 寿司			
9:08				金沢 寿司 店舗情報 (c)
9:12				金沢 寿司
9:16			金沢駅 海鮮丼 ひらい	
9:39	海鮮丼 平井			
9:39		金沢駅 海鮮丼 ひらい		

うタスクで実験を行ったときのログである。表中の (d) はユーザが発散クエリとして推薦されたクエリを入力したということを表し、(c) は収束クエリとして推薦されたクエリを入力したということを表し、無印はユーザ自身がクエリを考えて入力してことを表す。

表 4 のグループの User D は、札幌で観光する場所として、「札幌 時計台」(8:10) を検索していた。その後、発散クエリとして推薦された「札幌 テレビ塔」(8:32) というクエリで User D は検索し、そこで「テレビ塔もあるのか」という発話がなされ、候補を広げることができた。さらに、発散クエリとして推薦された「札幌 大通公園」(8:52) というクエリで User D は検索し、「大通公園があるのか」という発話がなされ、さらに候補を広げることができていた。最終的に、「時計台」も「テレビ塔」も「大通公園」から近いということが判明し、そのグループは「大通公園」を結論とした。このように、推薦クエリによって、ユーザの思いつかなかった候補を提示できていた。ほぼ 1 人に対して

の推薦であるが、これは他のユーザがウェブ閲覧中で、推薦クエリを見なかったということに起因する。このような場合でも、会話により情報は広がる可能性があるので、推薦クエリを必ずしも全員が見ておく必要はなく、しっかりと候補を広げることができる。比較システムの場合、「時計台」、「テレビ塔」、「大通公園」のいずれをクエリとして入力しても、残りの 2 つが推薦されることはない。

さらに、表 5 の別のグループでは、金沢で食事する場所として「金沢 寿司」(3:19) というクエリで User B が検索していた。そのとき、「金沢 海鮮丼」(4:04) という発散クエリが推薦され User B はそのクエリで検索し、それに伴い、他のユーザも「金沢 海鮮丼」(4:18, 4:21) で検索を始め、最終的に「海鮮丼」に的を絞る、「海鮮丼が美味しそう」という理由で海鮮丼が有名なある店に決定した。もともと、「金沢 海鮮」というクエリは User D などが入力している (1:42) が、「海鮮丼」には至っていない。「寿司」というクエリから「海鮮丼」というクエリが生成され、候補を広げられ、さらにそのままの絞られている。比較システムの場合、「寿司」というクエリを入力しても、「海鮮丼」が推薦されることはない。

以上のように、提案システムでクエリ推薦が有効にはたらくことが多々あった。

5.3 クエリの精度

クエリの精度は表 3 のようになった。主観的評価における発散クエリと収束クエリの収束度はいずれも高くなっており、発散クエリの収束度が高いということは、システムの意図と食い違っていることになる。今後は、発散クエリが発散クエリとして認識されるようなクエリの生成・推薦の仕方を考える必要がある。本稿では、推薦クエリにそれが発散クエリなのか収束クエリなのかというラベル付けをしたうえでの推薦は行っておらず、ユーザからはそれが発散クエリなのか収束クエリなのかは分からないようになっている。例えば、ラベル付けをしたうえで推薦することで、発散クエリの収束度を下げることが可能かもしれない。

また、客観的評価における適合度は全体的に低めであるが、例えば「博多 観光」というクエリをユーザが入力したとき、システムは「博多 ビジネス」というクエリを推薦している。「博多に行く目的として観光だけでなくビジネスもありうる」という考えに基づくものであるが、元々の意図としては観光であるので、「博多 ビジネス」は不適切なクエリ推薦となる。現在の手法では、「博多」というおおまかな意図を推定することは可能であるが、粒度の小さい「旅行」、「観光」といった意図は考慮していない。そのような理由で、適合度は低くなってしまったと考えられる。

さらに、適合度に関しては収束クエリの場合でわずかに高くなっているが、これは収束クエリがユーザの入力クエリ自身かそれに詳細語を加えるという生成の仕方であるた

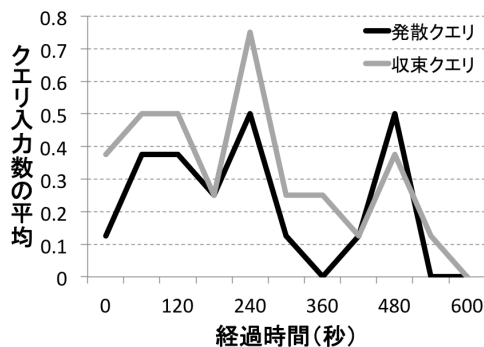


図9 発散クエリと収束クエリの時間的入力数の推移

めに、グループの意図から大きく逸脱したようなクエリは推薦されなかったからだと考えられる。

図9は、グループで推薦されたクエリを入力した数の平均の推移である。90秒で区切って、その区間内で入力された発散クエリと収束クエリの数の平均を表している。理想としては、検索行動における前半ではくまなく検索を行い、後半では的を絞って検索を行うのが良いと考えられる。そのため、グラフとしては、前半は発散クエリ入力数が多く、後半では収束クエリ入力数が増えているのが理想である。しかし、図9はそうになっていない。この理由としては2つ考えられる。1つ目は、ある程度検索が済んだ後に、最後に改めて網羅的にチェックを行うことがあるというものである。これは実際に会話内容からも判断でき、図9の300～360秒当たりにかけて入力クエリ数が減り、480秒にかけて増えていることからわかる。このため、今後は5分などの短い時間での実験を行う必要がある。2つ目は、検索が進むにつれ、推薦クエリの内容が似たようなものになってしまったというものである。終盤は収束度が高めであるので、推薦される収束クエリの割合も増えてくる。そのとき、推薦される収束クエリが既に入力したものや似たようなものばかりであると、発散クエリを選択してしまう可能性が増える。このため、今後は推薦クエリの生成・推薦の仕方を改善する必要がある、クエリに多様性を持たせたり、既に入力したクエリはできるだけ再度推薦しないようにしたりする工夫が必要となってくる。

6. まとめ

本稿では、我々が既に提案してきた、モバイル環境におけるクエリ推薦を用いた意思決定支援手法のより詳細な分析を行った。具体的には、提案システムと比較システムとの間に満足度の大きな差は見られなかったが、結論に達するまでの時間はより早く、クエリ推薦の使用率も高く、より少ないページの閲覧でくまなく検索できる可能性があり、クエリ推薦が有効にはたらくことが多々あったことがわかった。

今後の課題としては、クエリ推薦の精度がそれほど良く

なかったもので、生成の仕方を改善する必要がある。具体的には、グループの意図をより粒度を細かく推定できるようにしたり、兄弟語・詳細語だけでない発散・収束クエリについて検討したりする必要がある。また、本稿では、グループ内の全てのユーザーには同じクエリが推薦されていたが、ユーザー毎に検索を分担していたり、先行知識が違ったりと、全ユーザーが平等であるとは限らないので、ユーザー毎に推薦するクエリを変更するという点についても考えたと思っている。

謝辞 本稿の一部は、文部科学省科学研究費補助金 挑戦的萌芽研究「モバイル協調検索に関する研究」(研究代表者：中村 聡史, #22650018)、若手研究(A)「インタラクティブな再ランキング・再サーチを可能とする次世代検索に関する研究」(研究代表者：中村聡史, #23680006)によるものです。ここに記して謝意を表すものとします。

参考文献

- [1] VisionMobile. The elusive long-tail of mobile shipments. <http://www.visionmobile.com/blog/2011/11/the-elusive-long-tail-of-mobile-shipments/>.
- [2] Pew Research Center. Nearly half of american adults are smartphone owners. http://www.pewinternet.org/media/Files/Reports/2012/Smartphone_ownership_2012.pdf.
- [3] Meredith Ringel Morris. A survey of collaborative web search practices. In *Proceedings of the twenty-sixth annual SIGCHI conference on Human factors in computing systems*, CHI '08, pages 1657–1660. ACM, 2008.
- [4] 大重 智志, 中村 聡史, and 田中 克己. モバイル協調検索のためのクエリ推薦・提示手法. In *DEIM Forum 2012 E1-1*, 2012.
- [5] Meredith Ringel Morris and Eric Horvitz. Searchtogether: an interface for collaborative web search. In *Proceedings of the 20th annual ACM symposium on User interface software and technology*, UIST '07, pages 3–12. ACM, 2007.
- [6] Saleema Amershi and Meredith Ringel Morris. Cosearch: a system for co-located collaborative web search. In *Proceedings of the twenty-sixth annual SIGCHI conference on Human factors in computing systems*, CHI '08, pages 1647–1656. ACM, 2008.
- [7] 伊豆 陸, 中島 伸介, and 田中 克己. グループ支援型 web 閲覧における閲覧履歴の視覚化と共有. *日本データベース学会 letters*, 3(1):121–124, 6 2004.
- [8] 奥 梓, 小牧 大治郎, 荒瀬 由紀, 原 隆浩, 上向 俊晃, 服部 元, and 西尾 章治郎. 携帯端末を用いた協調 web 検索におけるコンテンツ比較支援インタフェース. In *DEIM Forum 2010 D9-4*, 2010.
- [9] 小谷 大祐, 中村 聡史, and 田中 克己. モバイル協調検索におけるユーザー間の情報共有支援. In *DEIM Forum 2011 A6-6*, 2011.
- [10] 大島 裕明, 小山 聡, and 田中 克己. Web 検索エンジンのインデックスを用いた同位語とそのコンテキストの発見. *情報処理学会論文誌. データベース*, 47(19):98–112, 12 2006.
- [11] 野田 武史, 大島 裕明, 小山 聡, 田島 敬史, and 田中 克己. 主題語からの話題語自動抽出とこれに基づく web 情報検索 (情報抽出). *電子情報通信学会技術研究報告. DE, データ工学*, 106(149):239–244, 7 2006.