

グループでウェブの探索を効率化する 検索共有インタフェース

武田達弥[†] 五十嵐健夫[†]

本論文では、グループにおける検索活動を支援するシステムについて提案する。グループ内のユーザの間では興味・関心が共有されているため、検索活動の重複が大きい。そのため、グループのメンバがそれまでに行なった検索履歴を利用することで、次に他のメンバが同じような検索を行ったときに目的となる情報にたどり着くまでの時間と労力を省くことが出来ると考えられる。本稿では、一つの実現例として、検索クエリが入力されたときに、過去に同じクエリで検索され有用と判断されたページを提示するシステムについて説明し、実際にユーザに使用してもらった結果について報告する。

Improving the efficiency of web search by reusing search history of group members

TATSUYA TAKEDA[†] TAKEO IGARASHI[†]

When searching a web using a keyword, the search engine returns many results and the user needs to pick desired one manually by visiting the returned many pages one by one. We try to reduce this burden by reusing the result of this task by other group members. Our basic assumption is that group members are expected to share general interest. For example, if one member picks page A as an appropriate search result for search query X after examining many search results, it is likely that another group member also wants to get page A when submitting search query X. This paper proposes search history sharing system that allows the users to get page A immediately when the second group member submit query X to save the user from checking many search results. Finally, we discuss our preliminary test.

1. はじめに

近年、ウェブ上の情報が爆発的に増え、それらを効率的に入手するためのシステムや研究が急速的に発展し、検索画面や表示される結果のパーソナライズなど各個人の情報活動効率を高めるユーザインタフェースに関する技術が注目を集めるようになってきている。しかし、検索という作業は一人で行うものばかりではない。検索活動には、他の人と協力する場面や、協力出来るような場面がある。例えば、グループで取り組むプロジェクトに関連した情報やグループ旅行の目的地などを探す場合がある。また、明示的に協力はしていても、グループでは興味・関心が共有されることが多いので、検索活動には重複があり、他人と協力できる可能性がある。しかし、それを支援するソフトウェアはあまりないため、検索活動において他人との協力は成し遂げられていない。他人との協力を支援する技術を導入することによって、検索が下手な人でも、検索が上手な人の検索履歴を利用することで有用な情報に簡単にたどり着くことができるようになるなど、検索効率の向上が期待出来る。

本稿では、そのようなグループでの検索活動を支援

する手法の一例として、あるクエリが、ユーザによって使われたときに、過去にそのクエリを用いて他のユーザが手に入れた情報を表示し、有用な情報にたどり着くまでのプロセスを効率化するという方法を提案する。また、その考えに基づいて作成したプロトタイプシステムおよびユーザスタディの結果について報告する。

2. 関連研究

2.1 グループを対象とした検索支援

グループを対象に検索を支援した研究として、Morrisら[1]、田中ら[2]、Romaroら[3]の研究がある。Morrisら[1]と田中ら[2]は、グループのメンバが明示的に協力して探すのを支援した研究である。Morrisら[1]は、グループのメンバが使ったクエリを表示することと、検索結果中のどのページが既に他のメンバに見られているかを表示することによって、ワークスペースにおけるアウェアネスを提供し、重複して同じページを探すという無駄を省く。また、見たウェブページに評価をつけたり、他のメンバにウェブページを推薦したり、あるクエリからの検索を分割して割り当てたりというようなことにより、検索効率を上げる。[2]は、グループで何かを決めるために、検索を使うような場合に、発見したページへグループのメンバがどのようなルー

[†] 東京大学
Tokyo University

トでたどり着いたかを視覚化することで、発見したページがどのような経緯で発見されたかわかるようにし、発見した情報を用いたグループの意思決定を支援する。[3]は、暗黙的な協力を促した研究である。見たwebページに対してアノテーションを残せるようにして、次にそのページを見た人が効率的に情報を得ることが出来るようにした。

2.2 ソーシャルブックマーク

検索を直接対象にはしていないが、ウェブ上の情報を他人と協調して扱うシステムとして、ソーシャルブックマークなどがある。ソーシャルブックマークは、ブックマークをオンライン上で他人と共有することが出来るものである。ブックマークは有効な情報に対してするものであると考えられるので、他人と協力して有効な情報を見つけるといって本研究に似ている。

2.3 協調フィルタリング

GroupLens[4]のような協調フィルタリングを用いたシステムは他人の情報を用いて、有用な情報を推薦するシステムである。協調フィルタリングも暗黙的に他人と協力して情報を見つけられるシステムであるが、協調フィルタリングは、趣向の異なる人たちの中から趣向の近い人を見つけてそれを推薦するという方式であり、本論分は小規模のグループを対象としているので、その点で異なる。

3. 検索履歴の利用

本システムは、グループの検索を支援するために、グループのメンバーの検索履歴を利用する。

グループは、興味・関心の方向性が近いため、ウェブで求める情報も似ていることが予想される。例えば、研究室で共同研究をしているときの関連論文探し、出張や旅行をするときの計画立てなどウェブで同じ対象を求めることがあるだろう。そのようなとき、通常は各メンバー毎に求めている情報を探す、つまり検索エンジンにクエリを打ち、結果として提示されたウェブページを順番に見ていくことになる。この作業は、求める情報によっては、労力と時間のかかる作業になることがある。そこで、グループの誰かが見つけた有効そうな情報を共有することが出来れば、他のグループメンバーはその情報を利用することが出来、作業の効率が上昇することが予想できる。そこで、本システムでは、グループのメンバーの検索履歴を利用することにより、グループの情報取得を支援する。

3.1 発見された情報の判別

グループのメンバーの検索履歴から有用な情報を決める必要がある。検索履歴は、価値のある情報と価値のない

情報からなる。我々は検索するとき、情報を発見するまでいくつかのウェブページを見て、それぞれが求めている情報かどうか判断している。求める情報のないページなら、その検索タスクでは価値がないページであろうし、求める情報があるページなら価値のあるページとなる。この二つのうち、グループのメンバーに提示すべき情報は、価値のあるページである。今回、われわれは、価値のあるページの判定として、ユーザのブラウジング中の行動を用いた。ユーザはブラウジングしているとき、見ているページをブックマークする。印刷する、保存するなどの行動を取る。経験上、これらの行動は、価値のあるページに対してなされることが多いはずである。よって、ユーザのこのような行動が取られたページを価値のあるページであるとみなし、グループのメンバーに提示することとする。

3.2 発見された情報の提示

検索エンジンに打たれたクエリと、価値のあるページを結び付け、次にそのクエリが使われたときに表示するという方法を取る。ユーザが何か検索しようとして、検索エンジンに情報を打ち込んだときに、システムによって共有されている価値のあるページが提示される。こうすることで、ユーザは本システムを使っていないときと変わらず、いつも通り検索することが出来、また価値のあるページが、それを求めている検索をしているときにだけ表示されるので、価値のあるページ郡の中から、今求めている情報を探すという手間も省かれる。

また、システムは価値のあるページを見つけたときに直接利用したクエリと結びつけるだけでなく、そのクエリの前後に使われたクエリとも結びつける。これは、ユーザは検索するとき、いくつかのクエリを試行錯誤しながら検索エンジンに打ち込んで、ページを調べていくので、こうすることで、その検索タスクに使われていたクエリのどれかが次に用いられれば、価値のあるページが表示されるということになる。これにより、グループ内の、ユーザ Alice が、query1 を使ってページを探したが見つからなかったので、次に query2 を使って、page3 という求める情報を発見したという場合に、次にユーザ Bob が query1 を打つだけで、page3 を見つけることが出来ると予想される(図1)。これは、もし Bob の検索能力があまりなくて、query1 を打ったあとに見つけることが出来ず、query2 というクエリを思いつかなかったような場合でも page3 を発見できるということになり、有効であると考えられる。

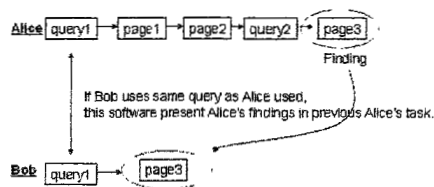


図 1 システムによる検索作業支援

Figure 1 Support by our system

今回、我々はクエリと価値のあるページとの結びつけについて、価値があると判断されたページと時間的に近いクエリを結びつけるという方法を取った。この方法は、検索タスク同士が時間的に離れている場合には、有効であると考えられる。しかし、この方法では、連続してまったく別の検索タスクをこなす場合や、時間的に長い検索タスクには対応できないので、将来的に考える必要がある。

4. インタフェース

本システムの概観を図 2 に示す。本システムは、Firefox ブラウザ [5] の拡張として実装されている。ユーザが検索エンジン(現在は、Google[6]にのみ対応している)のページにクエリを打ち込むと、通常の結果の上、システムが提示する価値のありそうなページが 10 件ランキングされて表示される(図 2 の 1)。表示されるのは一つのページにつき、ページのタイトル、URL、それが選ばれた行動要因である。これにより、ユーザは、普段通りに検索エンジンを使いつつ、グループの他のメンバの行った検索作業で見つかった価値のあるページも受け取ることが出来るようになる。また、ステータスバーのボタン(図 2 の 2)を押すことにより、プライベートモードになり、システムへ履歴が送られないようにすることが出来る。もう一度押せば解除される。



図 2 システム概観

Figure 2 User Interface

5. 実装

本システムはクライアントサーバモデルで実装されている。クライアントは Firefox ブラウザの拡張であり、サーバは、Ruby と MySQL を用いて実装されている。クライアントは、ユーザのブラウザ上での行動を監視し、検索エンジンのページでクエリが打たれれば、そのクエリをサーバに送り、そのクエリと結び付けられているサーバが保持している価値のあるページを受け取る。また、印刷や、ブックマーク、保存の行動も監視し、それらが行われればサーバに送る。サーバ側は、クライアントから、クエリや、ブラウザ上の行動といったデータを受け取り、クエリと価値のあるページの結びつけを行い、それらをデータベースに保存する。

6. 評価

本システムがユーザの検索活動を適切に支援するかを調べ、またこれからシステムをどのように改善すれば良いのかを調べるためにユーザスタディを行った。ユーザスタディの場所は、筆者が所属する大学の研究室で行った。また、参加者は 6 人で、すべてコンピュータサイエンスを研究している学生である。

6.1 実験課題

実験課題として、2 種類の課題を用意した。これらの課題には、われわれが日常行っている検索タスクを模倣したものを 2 種類用意した。それぞれ、課題 A、課題 B とする。

課題 A

課題 A は、波の音が入っている音声ファイルを Google で検索して集めるという課題である。本課題は、ウェブ上で探したいものが明確に決まっている場合に、本システムがユーザの検索活動を支援するかを調べることが目的で行った。また、検索中は、音声ファイルを発見したら、PC に保存を行ってもらった。

課題 B

課題 B は、研究室で行うクリスマスパーティでのプレゼント交換用に、千円から 2 千円のプレゼントの候補を、探すという課題である。本課題は、目的はあるが、探す対象が曖昧で、ウェブを見ながら考えをまとめるというような場合に、本システムがユーザの検索活動を支援し、考えをまとめやすくなるかどうかを調べることが目的で行った。また、見つけたプレゼントの候補や、プレゼントを考える上で役に立ちそうなページを見つけた場合に、ブックマークを取ってもらった。

6.2 手順

まず、一人一台の PC を割り当て、Firefox ブラウザのインストールとクライアントであるブラウザ拡張をインストールしてもらう。そして、システムの簡単な説明と、実験課題の説明を行う。次に、6 人を 3 人一組の 2 つのグループに分け、これをグループ 1 とグループ 2 とする。グループ 1 に課題 B を、グループ 2 に課題 A を割り当て、本システムの支援のもと課題を行ってもらい、これを本番作業とする。しかし、本システムがユーザを支援するには、「以前にユーザが検索したデータ」をあらかじめ保持しておく必要があるため、本番タスクの前に「以前に検索したデータ」を作る必要がある。そのため、本番作業の前に、システムの支援機能なしで、システムへデータの入力だけを行う作業を被験者に行ってもらった。この作業は、本番作業と同じ課題 A と課題 B になるが、同じ被験者が本番作業の前に同じ課題を行ってしまっていると、学習効果などの影響を受けてしまうので、データを入力するための作業では、グループ 1 に課題 A、グループ 2 に課題 B を行ってもらったこととした (図 3)。そして、それぞれのグループに 20 分間データ入力の作業を行ってもらい、10 分間休憩を挟んだ後、20 分間本番作業を行ってもらい、作業終了後に、アンケートを行った。アンケートでは、システムが提示したページは参考になったかとそれらのページ群は通常の検索エンジンが返す検索結果と比べてどうかを尋ねた。




	課題 A	課題 B
データ入力作業 (システム支援なし)	 グループ 1	 グループ 2
本番作業 (システム支援あり)	 グループ 2	 グループ 1

図 3 課題の割り当て

Figure 3 work assignment

6.3 結果

実験中に取られたログとユーザの行動と作業後のアンケートの内容を議論する。図 4 は、システムの提示したページは参考になったかという質問に対する回答で、両課題とも、本番作業を行ったグループの被験者に答えてもらっている。図 5 は、課題 A でユーザが発見した音声ファイルをダウンロード出来るウェブページの数とそのうちでシステムの支援により発見したウェブページの数を表す。

まず、課題 A の音声ファイル探しの結果を分析する。課題 A では被験者によって 92 種類のクエリが打たれ、そのうち 12 種類のクエリで、システムは他人の発見した価値のありそうなページを提示した。システムの提示したページは参考になったかという質問に対し、課題 A に参加した 3 人ともはいと答えた。理由として、「提示されたページから音声ファイルをダウンロードすることが出来、楽であった。」ということが上げられた。実際、被験者は音声ファイルをダウンロード出来るウェブサイトを 15 ページ発見したが、そのうち 4 ページはシステムからの支援で辿り着かれたものであり、システムは被験者の情報の発見を手助けしていた。次に、システムによって提示された結果は、通常の検索結果に比べてどうかという質問には、「洗練されていたと感じた。」という意見があった一方で、提示された結果には「良いときはとても良いが、タスクと関係ない結果が混じっていたと感じた。」という意見もあった。課題と関係ない結果が混じっていた原因として、課題 B を行っている被験者が間違いで課題 A 向けのクエリを打ってしまったということが上げられる。それが原因で一部課題 A と課題 B 向けの結果が混じってしまっていた。今回は間違いのせいで混じったとはいえ、簡単にノイズが多く混じってしまうという点は、改善す

べきである。

課題Bのクリスマスプレゼントを探すという課題の結果について分析する。課題Bでは被験者によって64種類のクエリが打たれ、そのうち9種類のクエリで、システムは他人の発見した価値のあるページを提示した。システムの提示したページは参考になったかという質問に対し、課題Bに参加した3人中2人ははいと答え、1人はいいえと答えた。はいと答えた被験者の意見として、「自分の思いつかなかったプレゼントを発見することが出来、考えが広まった。」「他の人と同じプレゼントを選ばないように気をつけることが出来た。」という意見があった。いいえと答えた被験者の意見は、「システムから他の人の見つけたページが支援されなかった。」ということであった。これは、その被験者の用いたクエリが他のメンバの用いたクエリと重複することがなく、支援を受けることが出来なかったということが原因であり、他のメンバが発見したページをクエリから提示する部分のアルゴリズムの問題点が示唆された。クエリが完全一致しないと他のメンバの発見したページが提示されないというのは、改善すべきであると考えられる。次に、システムにより提示されたページは通常の検索エンジンの結果に比べてどうだったかという質問に対する意見として、「他の人の選んでいる内容を見ることが出来たのは面白く、そればかり見てしまった。」という意見があった。本システムだと、通常の検索結果と違い、人間の作為で選ばれているので面白く感じるのかもしれない。また、「自分の選んだページまで表示されるので、それが邪魔になる。」という意見があった。これは、システムにより、自分がつい最近見ていたページが次のクエリを打ったときも表示されてしまうので、もう見る必要のないそのページが検索タスクの邪魔になるということであり、改善すべきである。

次に、データ入力作業と本番作業のログについて比較することで、システム支援ありとシステム支援なしの比較を行う。支援ありと支援なしの比較によって、本システムの有効性を調べるには、課題達成までに打たれたクエリ数の数で比較し、クエリの実行錯誤の度合いがシステム支援により、どの程度減ったかによって調べるということや、もしくは発見されたターゲットの数によって調べるということなどが考えられる。今回の実験では、時間内に見つかるだけ探してもらうという形式であったので、見つかった数で比較する後の方が適切であると考えられる。比較に、数を集めることが目的である課題Aを用いる。課題Aで見つかった音声ファイルのあるウェブページの数は、図5の通

りであり、システム支援ありと支援なしで有意な差は見られなかった。この原因として、個人差の影響などが考えられるが、アルゴリズムが不十分である可能性もある。

結果をまとめると、他人の発見したページが見られることは、情報検索と思考の支援に繋がる可能性があることが分かった。しかし、アルゴリズムとインタフェースは、このままでは不十分であり、改善の余地があるということが分かった。

	はい	いいえ
課題A	3	0
課題B	2	1

図4 システムの提示したページは参考になったか

Figure 4 Were the pages shown by the system useful?

	発見されたページ数	システムの支援により発見されたページ数
データ入力作業	14	0
本番作業	15	4

図5 課題Aで発見された音声ファイルのダウンロード出来るウェブページ数

Figure 5 The number of web page found in task A

7. まとめと今後の課題

本研究では、グループの検索履歴を用いて、グループでのウェブの探索を効率化する手法を提案した。また、ユーザスタディを行い、提案手法がユーザの検索を適切に支援するかを評価した。結果、通常通りに検索するのに比べて、本手法による支援を受けると、情報を楽に発見できる可能性とユーザの思考を広げる可能性があることが分かった。一方で、アルゴリズムとインタフェースに対するいくつかの改善点が示唆された。それらを踏まえて今後の課題として、大きく4つのことが上げられる。

まず、一つ目に、クエリと価値のあるページとを結びつけるアルゴリズムである。今回のアルゴリズムは、価値のあるページを、そのページが見られた時間からの一定時間内に打たれたクエリに結びつけるという方法を取った。しかし、この方法だと、べつべつの検索タスクが、ばらばらに行われるという前提であれば、うまくいくが、近い時間にまったく別の検索タスクを

行った場合、関係のない検索タスクのクエリに価値があると判定されたページが結びついてしまうという問題がある。今回のユーザスタディでもタスク A 向けのページとタスク B 向けのページが一部混じってしまっていて、被験者に邪魔であると感ぜられた。これを解決する方法として現在考えているのは、検索タスクの範囲を判定するということである。なにかを検索するときに打たれた一連のクエリ群がある検索タスクに使われたクエリとみなすことが出来る。ここで、これらの検索タスクのクエリ群を、そのタスクでクエリと結び付けるとさらに効率的に支援できると考えられる。それを実現するには、検索タスクのはじまりが、クエリが打たれた点であるとして、検索タスクがどこで終了するのかということを考える必要がある。これを考えることが、将来の課題のひとつである。

二つ目に、インタフェースの改善である。被験者の「検索エンジンの検索結果が長くなるので邪魔と感ぜることがある。」という意見があった。さりげなく見せる方法や、クエリとページの結びつけのアルゴリズムと組み合わせて、結果をもっと精選してリストするなど改善することが課題である。

三つ目に、検索履歴の中から価値のあるページを判別する方法である。今回は、ブラウザ上でユーザが取った、ページをブックマークする、保存する、印刷するという行動から価値のあるページを判別した。これは確かに価値のあるページであると予想されるが、これら以外にも検索中に見つかった価値のあるページはもっとあると思われる。今回の判定アルゴリズムでかからなかった価値のあるページを判定することが、三つ目の課題である。ブラウザのフォーカスや、ページを見ている時間などを利用することを検討している。

四つ目は、長期的な評価を行うことである。今回の評価は、タスクと時間を決めて行ってもらうという形式であった。我々は、長期につかってもらって評価する必要があると考えている。

参考文献

- 1) Search Together: An Interface for Collaborative Web Search
Meredith Ringel Morris, and Eric Horvitz, Proceedings of the 20th annual ACM symposium on User interface software and technology
- 2) 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.
伊豆 陸, 中島伸介, 田中克己, 日本データベース学会論文誌 (DBSJ Letters) , Vol.3, No.1, pp.121-124

- 3) Collaborative information retrieval environment: integration of information retrieval with group support systems
Romano, Nicholas C. and Roussinov, Dmitri G. and Nunamaker Jr., Jay F. and Chen, Hsinchun, Proceedings of the 32nd Annual Hawaii International Conference
- 4) GroupLens: An Open Architecture for Collaborative Filtering of Netnews
Paul Resnick, Neophytos Iacovou, Mitesh Suchak, Peter Bergstrom and John Riedl, In Proceedings of the 1994 Computer Supported Collaborative Work Conference, pp.175-186,
- 5) Mozilla Firefox
<http://www.mozilla-japan.org/products/firefox/>
- 6) Google
<http://www.google.co.jp/>