

協調検索における役割と検索行動の関係分析

山本 岳洋† 山本 光穂‡ 田中 克己†

† 京都大学大学院情報学研究科

‡ デンソーアイティラボラトリ

† {tyamamot, tanaka}@dl.kuis.kyoto-u.ac.jp

‡ miyamamoto@d-itlab.co.jp

1 はじめに

複数のメンバが同一の情報要求や目的を共有し、お互いに協働しながら情報検索を行う協調検索 [1][2][3] が広く行われるようになってきている。204名の知識労働者を対象に行った Morris の調査 [5] によれば、54.6%のユーザが、これまでに他者と協調してウェブ検索を利用したことがあると回答しており、38.5%が週に1回、15.6%が月に1回程度、恒常的にそのような経験をしていると回答している。このような協調検索を支援するため、近年多くの研究者がユーザインタフェース [6] や検索アルゴリズム [7][8][10] の両面から盛んに研究に取り組んでいる。

協調検索に関する研究の大きな流れの1つに、グループのメンバに異なる役割 (Role) を与えることで、適合ページの獲得にかかる時間の効率化や、検索タスクの最終成果の質を向上させることを目的とした研究がある [4][9][10]。明示的な役割をメンバに与えることによって、各メンバは行動の指針が定まり、他者とより協調しやすくなると考えられる。たとえば、代表的な役割として、Shar らが提案した Gatherer と Surveyor がある [9]。Gatherer とは適合ページをひたすら収集する目的を持ったメンバであり、Surveyor はできる限り多様なページを収集する目的をもったメンバである。Shar らは、これらの異なる役割を想定したユーザに応じた検索ランキングアルゴリズムを提案している [9]。

しかし、各役割に応じた検索ランキングが提案されてきたものの、そうした役割に基づくユーザがどのような検索行動を取るのかについては、明らかにされてこなかった。既存研究では、役割に基づいたメンバの検索行動の分析には焦点が当てられてこないうまま、役割に適した検索ランキングが議論されてきた [8][9]。たとえば、先述した Gatherer や Surveyor といった役割は、その性質上自身や他のメンバのクエリや閲覧ページに影響を受け、クエリの選択やページの閲覧を行うと考えられる。各役割に基づくメンバが、どのように他者とインタラクションを行い、どのような検索行動を振る舞うのかを明らかにすることができれば、協調検索という、複数のユーザが互いに影響を与え合う複雑な検索を理解するための一助になると考えられる。さらに、役割に応じた検索行動の特徴が明らかになれば、そうした特徴を利用することで、各役割に応じた最適な検索ランキングやクエリ推薦アルゴリズムを実

現することが可能となると考えられる。

本研究の目的は、協調検索における明示的な役割が、クエリの選択や検索結果の閲覧にどのような影響を与えるのかを分析することである。本研究では、Gatherer と Surveyor という2つの役割に焦点をあて、2人組のグループに対して、実装した協調検索システムを使用してもらうことで、協調検索ユーザの検索行動を収集、分析した。本論文では、本研究で行った実験について述べるとともに、検索行動、特にメンバが投入するクエリとメンバの役割との関係に関する実験結果を報告し、結果を考察する。

2 関連研究

本章では、まず協調検索における役割を扱った研究について紹介し、その後、協調検索におけるメンバの行動分析に関する研究について述べる。

1章で述べた、Shar らが提案した、Gatherer と Surveyor という役割 [9] のほかにも、Pickens らは、Prospector と Miner という役割を提案している [8]。Prospector と Miner は、それぞれ Surveyor と Gatherer と類似した役割であり、Prospector は検索ドメインの空間を広げるためにクエリの投入を主に行い、一方で Miner は検索結果の適合性判定を主に行う。Pickens らは、Prospector と Miner それぞれに応じた検索インタフェースを用い、それぞれの役割の行動に基づいた動的なランキングアルゴリズムを提案している。また、Soulier らは、メンバの検索行動から各メンバの役割を動的に推定し、それぞれの役割に応じた検索結果ランキングを行うことで、協調検索の検索効率を向上させることを提案している [10]。

検索行動よりもより大きな視点から協調検索の役割を提案した研究として、Imazu らの研究がある [4]。Imazu らは、検索タスクだけでなく、その検索を通して作成する成果物を意識した、作業タスクが協調検索において重要であると指摘し、Searcher と Writer という役割を提案している。ここで、Searcher は PC を操作し実際に情報を検索を行うのに対し、Writer は検索は行わず、メモを取りタスクの成果物として要求されているレポートを作成することを目的とした役割である。

本研究では、Shar らの提案した Gatherer と Surveyor に焦点をあて、これらの役割とユーザの検索行動の関係分析を行う。Gatherer と Surveyor の役割に

注目した理由は、これらの役割が、メンバの利用するシステムや行動について大きな制約がなく一般的である点、また、その役割の一般性のため、得られた知見が個人による検索にも適用できる可能性を持っている点からである。

協調検索において、クエリや検索結果の閲覧といったユーザの検索行動の分析に焦点をあてた研究は少ない。Yueらは、協調検索におけるクエリ修正がどのような要因から影響を受けるのかを被験者実験を通して分析している [11]。彼らは、2名のメンバで構成される10組のグループを対象に実験を行い、メンバ間のチャット、過去に自分（他のメンバ）が投入したクエリ、過去に自分（他のメンバ）が閲覧したページ、といった要因がクエリ修正に与える影響を分析している。その結果、78%のクエリがチャットの内容と間接的に関連していたと報告している。彼らの実験は、特定の役割をメンバに課さない自由な状況下で検索を行った際の分析である。そのため、本研究で対象とする、明示的な役割の基での協調検索では、メンバの検索行動はチャットだけでなく、自身や他のメンバの検索行動に影響を受けると考えられる。

特定の役割を課した際のユーザ行動の分析として、Imazuらの研究がある [4]。Imazuらは、上述したSearcherとWriterという役割を課すことで、メンバ間の会話が最終的な成果物に関する内容と関連する割合が高くなることを報告している。彼らの研究は、役割を課したグループとそうでないグループにおいて、検索行動や会話にどのような影響が生じるのかを分析しており、個々のメンバの検索行動の違いについては分析されていない。

3 実験デザイン

本章では、本研究で行った被験者実験について述べる。まず、本研究で対象とする、GathererとSurveyorの2種類の役割について説明し、その後、実際に行う実験の詳細について述べる。

なお、本研究では、役割の影響の分析に注力するため、他の既存研究 [4][8][10][11]と同様に、2名のグループによる、互いのメンバがお互いに離れた環境 [11]における協調検索を対象とした。

3.1 GathererとSurveyor

Sharらの定義によれば [9]、Gathererとは適合する情報をひたすら収集する目的、Surveyorとは探索空間を広げ、できる限り多様な情報を収集する目的を持った役割である。

これら2種類の役割は、その検索行動がグループのメンバの行動に影響を受けると考えられる。たとえば、Gathererは適合ページを探し求めるため、すでに得られた適合ページと類似する検索結果や、適合ページが得られたクエリと類似するクエリを選択しやすいと考えられる。一方、Surveyorはすでに得られた適合ページとは異なる検索結果や、すでに投入したクエリとは類似しないクエリを選択しやすいと考えられる。本研



図 1: 実験に用いた協調検索インタフェース。

究の目的は、こうした役割と検索行動の関係性を明らかにすることである。

3.2 検索インタフェース

図1に実験に用いた協調検索インタフェースを示す。本研究では、役割に基づく検索行動の分析に焦点をあてるため、協調検索におけるインタフェースについては既存研究 [11]を参考に実装を行った。インタフェースは、大きく分けて以下の4つの機能を持つ。

- (1) **検索**: クエリを入力し検索結果を表示する。本研究では、Bing¹の上位50件の検索結果および上位8件のクエリ推薦が表示される。また、各検索結果には共有ブックマークに追加するためのボタンが表示されている。
- (2) **チャット**: パートナーと会話を行うためのチャットウィンドウ。送信、受信したメッセージがリアルタイムに反映されるとともに、メッセージ受信時には通知音が再生される。
- (3) **検索履歴**: ユーザおよびパートナーが入力したクエリが最新のものから順に、リアルタイムに表示される。
- (4) **共有ブックマーク**: ユーザおよびパートナーが共有ブックマークに保存した検索結果がリアルタイムに表示される。

3.3 検索タスク

協調検索では、一般的に、ある話題の網羅的な調査や旅行計画といった、探索的検索に関するものが典型的である [4][5]。具体的には、Morrisらが指摘しているように、協調検索の典型的な検索タスクには、情報を網羅的に収集するタスク (recall-oriented タスク) と、観光計画のように意思決定を伴うタスク (decision-making タスク) の2種類があり [5]、これら2種類の検索タスクが既存研究の多くでも用いられている。本研究では、情報を網羅的に収集するタスクに注目し実験タスクを用意した。具体的には、Soulierらの実験

¹<https://datamarket.azure.com/dataset/bing/search>

[10]でも用いられているトピックである，地球温暖化に関する情報収集タスクを用いた．下記に，実験に用いたタスクの説明文の例を示す．

あなた達2人は，現在同じ講義を履修しています．その講義では，「地球温暖化に対する世界の取り組み」というテーマで，A4用紙4枚のレポートを2人でまとめることになっています．実験システムを使い，30分間かけ，レポートを執筆するために役に立つと感じるページを共有ブックマークに保存してってください．30分間で，テーマに関する有用なページを，できる限り，さまざまな観点から多く集めることがあなた達の目的です．

3.4 実験手順

実験のため，京都大学に所属する学生2名を被験者とし実験を行った．グループは以下の手順で，デスクトップPCを用いて実験を行った．

1. 実験趣旨の説明．
2. システムの試用（15分程度）．
3. 3.3節で述べたタスクの実行．この際，メンバの1人についてはGathererの役割を，もう1人についてはSurveyorの役割を明示的に与え，タスクを行ったもらった．
4. タスク終了後，現在行ったタスクに対する満足度に関するアンケートを行うとともに，ユーザが投入したそれぞれのクエリについて，そのクエリを投入して良い結果が得られたかどうかをあらわす主観的な満足度と，クエリを投入した理由をアンケートで回答してもらった．

4 結果

はじめに，タスクから得られたクエリ数や検索結果クリック数といった，タスク中の基本的な検索行動が役割ごとにどのように異なるのかを検証する．表1は，3.4節の実験により得られた，グループの検索行動のログを，各役割ごとにまとめた表である．表中の“総クエリ数”，“総検索結果クリック数”，“総ブックマーク保存数”は，1タスクあたりにシステム投入されたクエリ数，検索結果クリック数，共有ブックマークに検索結果を保存した回数をそれぞれ表している．“平均検索結果クリック数”と“平均ブックマーク数”は，1クエリ当たりのそれぞれの回数を表している．また，“平均滞在時間”は，あるクエリを投入したメンバが，次のクエリを投入するまでにかかった時間の平均である．

表1より，Gathererは検索結果のクリック数や共有ブックマークへの保存数がSurveyorよりも多いという傾向が分かる．これは，Gathererの役割である，適合するページするページを次々へ保存していくという目的がこうした行動に表れているためであると考えら

れる．一方，Surveyorは，Gathererに比べて1クエリあたりの滞在時間が少なく，より多くのクエリを投入していることが表より見てとれる．これも，Surveyorの役割である，できる限り多様な情報を探すとという目的が検索行動となって現れていると考えられる．

次に，それぞれの役割に基づくメンバが投入したクエリが，自身やパートナーのクエリからどの程度影響を受けたのかを分析した．本研究では，メンバが投入したクエリと，それまでに投入されていた自身やパートナーのクエリとの類似度を計算することで，投入されたクエリがどの程度自身やパートナーの過去のクエリに影響を受けているのかを分析した．いま，ある時刻 t においてユーザが投入したクエリを $q^{(t)}$ ，時刻 t までに自身，パートナー，あるいはその両者が投入したクエリ集合を $Q_{\text{past}}^{(t)}$ で表すとき， $q^{(t)}$ と $Q_{\text{past}}^{(t)}$ の類似度 $\text{sim}(q^{(t)}, Q_{\text{past}}^{(t)})$ を以下の式で求める．

$$\text{sim}(q^{(t)}, Q_{\text{past}}^{(t)}) = \max_{q' \in Q_{\text{past}}^{(t)}} \text{Jaccard}(q^{(t)}, q'), \quad (1)$$

ここで， $\text{Jaccard}(q^{(t)}, q')$ は，クエリ中に含まれる単語集合間のJaccard係数である．つまり，投入したクエリが，過去のクエリ中に含まれている単語を含んでいるほど， $\text{sim}(q^{(t)}, Q_{\text{past}}^{(t)})$ は高い値をとる．あるメンバの全てのクエリについて式(1)を計算し，その値の平均値を求めることで，あるメンバのクエリが，自身やパートナーのクエリにどの程度影響を受けたのかを分析した．

表2は，自身の過去のクエリ，パートナーの過去のクエリ，自身とパートナーを合わせたグループの過去のクエリそれぞれについて，上述した，クエリと過去のクエリの類似度の平均値を求めた結果である．表から，自身，パートナー，グループにかかわらず，GathererはSurveyorよりも過去のクエリとの類似度が高いクエリを投入していることが分かる．これは，Gathererの目的が，適合するページを収集することにあるため，適合ページが得られると期待されるクエリを，それまでの検索で得られた情報やクエリから作成した結果，多くのクエリが過去のクエリと類似する傾向にあるという結果になったものと考えられる．一方，Surveyorは，Gathererよりも類似度の平均がどれも低い値となっている．Surveyorはできるだけ多様なページを収集するため，これまでに得られていないようなクエリで情報を探す必要がある．そのため，これまでに検索していないようなクエリを積極的に投入していったのではないかと考えられる．

また，自身の過去のクエリとパートナーの過去のクエリとの類似度の差に注目してみると，Gathererは，パートナーの過去のクエリとの類似度の方が，自身の過去のクエリの類似度よりも高い値となっていることが分かる．これは，あるクエリでの適合ページの収集が終わり，次のクエリを投入する際，Surveyorであるパートナーが既に調べた様々なトピックからクエリを選ぶことが多いためではないかと考えられる．

表 1: 実験より得られたクエリ数, 検索結果クリック数, 共有ブックマーク保存数.

	Gatherer	Surveyor
総クエリ数	19.0	21.0
総検索結果クリック数	37.0	12.0
総ブックマーク保存数	32.0	12.0
平均検索結果クリック数	1.95	0.76
平均ブックマーク保存数	1.68	0.57
平均滞在時間 (sec)	98.2	83.8

表 2: 役割ごとの, 過去のクエリと投入したクエリの類似度の平均.

	Gatherer	Surveyor
自身の過去のクエリ	0.71	0.63
パートナーの過去のクエリ	0.88	0.57
グループの過去のクエリ	0.91	0.63

5 おわりに

本研究では, Gatherer と Surveyor という協調検索におけるメンバの役割に注目し, それぞれの役割がメンバの検索行動に与える影響の分析に取り組んだ. 本研究では, 特にそれぞれのメンバが投入したクエリと, 過去のクエリとの類似度に注目し実験結果を分析した. その結果, Surveyor は Gatherer よりも, 過去のクエリとの類似度が低く, Surveyor の目的であるできるだけ多様なページを収集するという目的が検索行動に現れていること, また, Gatherer は自身のクエリだけでなく Surveyor であるパートナーの過去のクエリとの類似度が高く, Gatherer のクエリ選択は Surveyor の過去のクエリから大きな影響を受けている可能性があることが分かった.

今後は, より多数の被験者を集めて実験を行い, 今回の結果が示唆している知見が妥当であるかどうかを検証する. また, 今回は 1 種類のタスクを用いて実験したが, 知見の一般性を検証するため, 他の種類のタスクについても同様の実験を行う. 具体的には, 観光計画に関するタスクについても実験を行う予定である.

今回得られた知見は, Gatherer や Surveyor の検索行動が, 過去の自身やパートナーの検索行動に大きな影響を与えている可能性があることを示していると考えられる. つまり, Gatherer や Surveyor のクエリ推薦や検索ランキングを改善するために, これまでの自身やパートナーの検索行動が有用である可能性が考えられる. 今後は, より多くのデータを収集し, このようなグループの過去の検索行動情報を利用した, クエリ推薦や検索ランキングの改善手法について取り組んでいく予定である.

参考文献

- [1] Jonathan Foster. Collaborative information seeking and retrieval. *Annual review of information science and technology*, Vol. 40, No. 1, pp. 329–356, 2006.
- [2] Jill Freyne, Barry Smyth, Maurice Coyle, Evelyn Balfe, and Peter Briggs. Further experiments on collaborative ranking in community-based web search. *Artificial Intelligence Review*, Vol. 21, No. 3-4, pp. 229–252, 2004.
- [3] Gene Golovchinsky, Pernilla Qvarfordt, and Jeremy Pickens. Collaborative information seeking. *Information Seeking Support Systems*, 2008.
- [4] Marika Imazu, Shinichi Nakayama, and Hideo Joho. Effect of explicit roles on collaborative search in travel planning task. In *Information Retrieval Technology*, pp. 205–214. Springer, 2011.
- [5] Meredith Ringel Morris. A survey of collaborative web search practices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 1657–1660. ACM, 2008.
- [6] 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*, pp. 3–12. ACM, 2007.
- [7] Satoshi Ohshige, Satoshi Nakamura, and Katsumi Tanaka. Supporting Mobile Collaborative Searches with Query Suggestions. Workshop on Collaborative Information Seeking, part of CSCW’13, 2013.
- [8] Jeremy Pickens, Gene Golovchinsky, Chirag Shah, Pernilla Qvarfordt, and Maribeth Back. Algorithmic mediation for collaborative exploratory search. In *Proceedings of the 31st annual international ACM SIGIR conference on Research and development in information retrieval*, pp. 315–322. ACM, 2008.
- [9] Chirag Shah, Jeremy Pickens, and Gene Golovchinsky. Role-based results redistribution for collaborative information retrieval. *Information processing & management*, Vol. 46, No. 6, pp. 773–781, 2010.
- [10] Laure Soulier, Chirag Shah, and Lynda Tamine. User-driven system-mediated collaborative information retrieval. In *Proceedings of the 37th international ACM SIGIR conference on Research & development in information retrieval*, pp. 485–494. ACM, 2014.
- [11] Zhen Yue, Shuguang Han, Daqing He, and Jiepu Jiang. Influences on query reformulation in collaborative web search. *Computer*, Vol. 47, No. 3, pp. 46–53, 2014.