

K_060

行動検索アプリケーションの評価方法の検討 Examination of Method of Evaluating Action Retrieval Application

倉 恒子† 森田 哲之†† 大浦 啓一郎††† 日高 哲雄†† 田中 明通†† 加藤 泰久††
Tsuneko A. Kura Tetsushi Morita Keiichirou Ooura Tetsuo Hidaka Akimichi Tanaka Yasuhisa Kato

1. まえがき

インターネット上における情報の爆発的な増加や更新頻度の多発化に伴い、以前見た質のよい情報を再度見つけることが困難になりつつある。筆者らは、ページに含まれるキーワードの合致度に加え、コンピュータ上の操作履歴を活用することにより、過去に見た Web ページの中から利用者が重要と考えていたページを表示する方法を提案している[1]。

人が調べ物をするとき、目的のものを探しながら興味の対象が絞り込まれ、より深い情報を探していくといった行動をとる。あるいは、興味が別の方向に向かってしまい、全く違うものを探すとといった行動をとることもある。筆者らはキーワードに合致した Web ページだけを表示するよりも、目的のページを含む Web ページ閲覧の遷移を提示することにより、利用者に調べ物をしてきた背景まで含めて提示することができ、過去に重要だと考えたことが短い時間で思い出すことが可能になると考えた。ここでは、ある内容に関連する Web ページの閲覧遷移を、閲覧シーケンスとして定義する。

過去に見た Web ページの中から検索条件に合致した閲覧シーケンスを提示するアプリケーションを開発し、評価実験を実施したので、その結果を報告する。

2. 行動検索アプリケーションの概要

行動検索アプリケーション(以下、Memory-Retriever)は、Web ページ閲覧の遷移を行動として扱うことにより、利用者が探している情報の提示、調べ物をしてきたときの背景に加え、重要と考えていた情報をできるだけ短時間で把握できるように機能を有する。主な特徴を以下に述べる。

(1) 操作履歴に基づく Web ページおよび閲覧シーケンスの評価

過去の閲覧履歴から、入力されたキーワードに合致した複数の Web ページを抽出する。ある Web ページから次の Web ページへ遷移する場合、最初に入力したキーワードを含まないが、内容に関連しているページを見ている可能性が高いと考えられる。Memory-Retriever では、あるページから次のページへ遷移する時間間隔が一定時間を超えない限り、利用者は継続して調べ物をしてきたと判断して、関連した内容のページの遷移を閲覧シーケンスとして扱っている。

また、利用者が Web ページに対して行なった操作一例えば、閲覧時間、印刷、ページ内に含まれる文章の選択などに基づき、個々のページに対してスコアを求める。キーワードの合致度とあわせて、利用者に重要と考えられるページおよび閲覧シーケンスを評価・表示する機能を有す

† NTT サービスインテグレーション基盤研究所

†† NTT サイバーソリューション研究所

††† 西日本電信電話株式会社

る。

(2) 閲覧シーケンス表示インタフェース

調べ物をする際、人は検索エンジンにキーワードを入力し、該当する Web ページを見つけ出す。検索キーワード、Web ページに含まれるタイトル、アンカーテキストは情報検索において有益な手がかりであることが判明している[2]。

Memory-Retriever では、閲覧した Web ページの遷移を、サムネイルを用いて表示し、上記情報も一緒に表示することにより、利用者に調べ物をしてきたときの考えを提示している。

図 1 は閲覧シーケンスを表示した画面である。個々のサムネイルの下にはタイトルを表示している。サムネイルの上には、検索エンジンに入力したキーワード(図 1(あ))や表示されている Web ページに遷移するために前のページでクリックしたアンカーテキスト(図 1(い))を表示している。

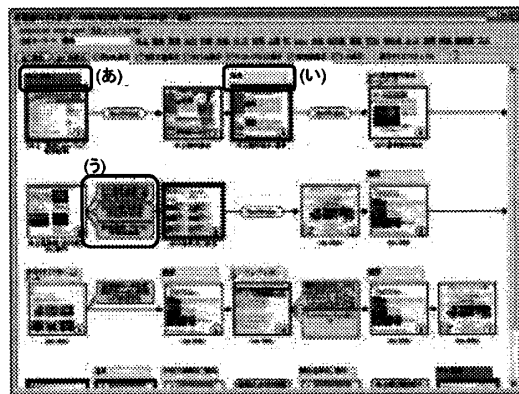


図 1 閲覧シーケンス表示画面

(3) Web ページ内での重要箇所表示

本を読んでいるときに重要な箇所に鉛筆などで線を引く人は多い。同じような操作を Web ページ上で表示した技術を適用した[3][4]。サムネイルの横に表示された吹き出し(図 1(う))は、選択したテキスト情報を表しているのので、この画面から重要なことを思い出すことも可能である。また、サムネイルをクリックすると、選択した部分の背景を黄色で表示するため、よりの確に内容を思い出すことができるようになっている。

3. 評価実験

過去に見た Web ページの中から、重要なものを早く見つける表示方法としての有効性を検証するため、評価実験を行なった。次の 4 つのポイントを用いてアプリケーションの評価を行なった。

- 重要なものと考えた箇所を見つけられるか
- 早く見つけることができるか

- c) 正確に見つけることができるか
- d) 重要なものと判断した根拠となるページを見つけることができるか

3.1 評価実験条件

(1) 評価対象アプリケーション

Memory-Retriever と類似するアプリケーションとして、過去に見た Web ページやファイルを検索するツールで一番広く知られている Google デスクトップ検索β版[5]を選択した。さまざまな機能のうち、Web ページの検索部分の機能を使用して評価実験を行なった。

(2) 被験者の情報

男性4名、女性1名の計5名で行なった。被験者は普段コンピュータを使って仕事をしており、インターネットを使うのに全く抵抗を持っていない。

3.2 評価実験手順

評価実験は調べてもらい、1週間後に思い出してもらい、という2段階で実施した。各々の手順について説明する。

(1) 最初の実験: 調べてもらう

被験者には、与えられた課題に対して興味のある事柄をブラウザで調べ、その結果を紙に記録してもらい、被験者が特に重要な情報と考えた内容には、印をつけて他の内容と区別できるようにしてもらった。

一方、根拠ページの特定のため、被験者が閲覧したページのサムネイルをコンピュータに蓄積しておき、調査後に、記録した内容とそれが書かれている Web ページを対応させることにした。

実施した実験の内容は次の通りである。

- ・課題: 国内旅行(スキー, 東京見物)

候補地を3箇所あげ、被験者が興味ある内容について自発的に調査

- ・調査実施時間: 15分

(2) 次の実験: 思い出してもらう

各々のアプリケーションを用いて、1週間前に調査して記録したことのうち、重要なものをできるだけ早く正確に思い出してもらった。

実施した実験の内容は次の通りである。

- ・課題の割り当て: 1課題1アプリケーション

5人が必ず両方のアプリケーションを使うようにした。課題は、均等になるように割り当てた。

- ・実施時間: 7分

3分、5分、7分と時間を区切り、各々の時間幅で重要なものが見つけられたかどうかを測定した。

(3) 評価方法

各々のアプリケーションを用いて、1週間前に記録したことのうち、重要なものをできるだけ早く正確に思い出してもらった。

記録内容の数値化方法は次のとおりである。

- ・調べたときに記録した内容: 2点
合計点 = 書かれた個数 × 2
- ・思い出すときに書かれた内容
内容が一致かつ根拠ページが一致: 2点
単語は含まれるが根拠ページが不一致: 0点
- ・重要なものを思い出した場合: 10倍

3.3 評価結果と考察

3.2で説明した評価方法に基づき、5名10課題について、記録内容を数値化し、100点換算で値を求めたものを時間

幅で表したものを表1に示す。ここで、MRと書かれているものがMemory-Retriever、GDSと書かれているものがGoogleデスクトップ検索での値である。

表1 評価結果

	3分	5分	7分
MR	17.08	44.34	51.69
GDS	9.81	23.39	30.03
対応のある場合の t検定[6]	0.87	4.73	3.75

有意水準を5%とすると、自由度が4の場合のtの臨界値は2.78であり、5分および7分のときに有意差があると判断される。すなわち、Memory-RetrieverはGoogleデスクトップ検索よりも、過去に重要だと考えたことを見つけやすい方式であることは判明したが、短い時間で見つけられるという主張に対しては有意性を示すことができなかった。どの時間幅においてもMemory-RetrieverがGoogleデスクトップ検索よりも結果がよかった原因として、Googleデスクトップ検索ではキーワードに合致したページだけが表示されること、タイムライン表示にすると文字列の表示になるため、内容の確認には1つ1つクリックして表示しなければいけないのに対し、Memory-Retrieverではキーワードに合致したページに関連して閲覧したページも表示されること、サムネイルで表示されるため、文字列よりも理解しやすかったことが考えられる。

一方問題点として、重要なものを選択する数は被験者によりばらばらであった。今後は選択する数を統一する必要がある。

4. まとめと今後の予定

行動検索アプリケーションを開発し、評価実験を行なった。提案方式がキーワードに合致したWebページだけを表示するよりも、目的のページを含むWebページ閲覧の遷移を提示することにより、利用者に調べ物をしていた背景まで含めて提示することができ、過去に重要だと考えたことを思い出すことを支援する方式であることを確認した。今回の実験で判明した問題点を解決し、再度評価実験を行なう予定である。

参考文献

- [1] 森田他, "Memory-Retriever - 履歴を利用したローカルファイル検索システム," 2005年信学会総合大会, 2005
- [2] 阿部他, "アンカーテキストとリンク構造解析を用いたWeb情報検索の改善," データ工学研究会, 2003
- [3] 大浦他, "NaMiSEN: ウェブ閲覧履歴への個人嗜好情報付加・抽出ツールの開発," オフィスインフォメーションシステム研究会, 2005
- [4] 大浦他, "注目情報付加・利用方式"NaMiSEN"の評価," オフィスインフォメーションシステム研究会, 2006
- [5] <http://desktop.google.com/ja/>
- [6] 森, 吉田, "心理学のためのデータ解析テクニカルブック," 北大路書房, 1990