

PC 操作履歴に基づいた行動検索/提示システム

Retrieval method for user action history

森田 哲之† 倉 恒子‡ 日高 哲雄† 大浦 啓一郎† 田中 明通† 加藤 泰久†

Tetsushi Morita† Tsuneko Kura‡ Tetsuo Hidaka† Keiichiro Ooura† Akimichi Tanaka† Yasuhisa Kato†

1. はじめに

『過去に見た』という体験を思い出し、その体験に含まれる Web ページを再検索することは多いと思われる。Web ページを閲覧するとき、過去に見た Web ページを再度閲覧する割合は、ブラウザの『戻る』操作を含めると 81% であったという Cockburn らの調査報告がある[1]。しかし、過去に見たことがあるにもかかわらず、どうやって見つけたかわからずゼロから探しなおすことや、検索エンジンを利用して結局その Web ページを見つけれないことがある。

そこで、デスクトップ検索技術のように、コンピュータ内のファイルや Web キャッシュを検索するという切り口で、過去に体験した Web ページを検索するツールが公開されている[2]。デスクトップ検索では、ユーザが入力したキーワードを含むコンピュータ内の Web ページのリストを、優先度順あるいは時間順に表示する。Web ページリストの URL をクリックすることで、Web ページの内容を閲覧することができる。

しかし、Web を閲覧するという体験によって得られる情報は、ひとつの Web ページの内容だけではない。その Web ページを閲覧したときに、他にどんな Web ページを見たのか、なぜその Web ページを見たのか、その Web ページをみてどのような知識を得たのか等の情報があると思われる。ひとつの Web ページの内容を発見したとしても、それらの情報を思い出すことは困難である。

本稿では、ユーザがコンピュータ上で行った操作を、単に『行動』と呼ぶこととし、任意のキーワードに関して過去に体験して得られた情報を、『Web ページ』ではなく『行動』という切り口で検索する手法を提案する。

2. 体験獲得情報

Web ページを閲覧するという体験によって得られる情報のひとつは、過去に閲覧した Web ページの記事など、Web ページの内容である。また、ユーザが何かを調べるとき、一つの Web ページだけでなく、同時に複数の Web ページの内容を閲覧することは多い。くちこみのページで推薦されていたので閲覧するなど、Web ページをどのようにして発見したのかといったユーザが辿った Web ページ間の関係も、Web ページを閲覧するという体験によって得られる情報である。この商品のこの機能が魅力的だと理解したことなど、Web ページをどのように解釈したのかといった Web ページから得た知識も、体験によって得られる情報である。このように、体験によって得られる情報は、ひとつの Web ページの内容といった単純なものではなく、多く

の情報が含まれている。Web 閲覧体験によって得られた情報のなかで、本稿の対象とする情報を次のように整理し、体験獲得情報と呼ぶことにする。

- Web ページの内容
- Web ページ間の関係
- Web ページから得た知識

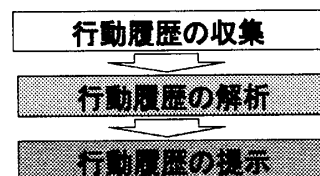
これらの体験獲得情報を効率的に再取得できることを本研究の目的とし、次のように定めた。

- 多くの体験獲得情報を取得できる
- 短期間で体験獲得情報を取得できる
- 重要な体験獲得情報を取得できる

3. 行動検索

キーワードに関連した Web ページを閲覧するなどの行動を行っていた時間には、キーワードに関する体験獲得情報が多く存在すると考えられる。そこで、体験獲得情報をユーザに取得させるための手段として、過去に行ったコンピュータ上の行動を検索する手法を提案する。ユーザは過去に自分自身が行った行動を閲覧することにより、体験獲得情報を想起できる。

行動を検索するためには、行動履歴を収集し、重要な行動履歴データを解析/抽出して、ユーザに効果的なインタフェースで提示する必要がある(図 1)。本章では、アクティブ期間(後述)を抽出し、その検索キーワードに対する重要度を算出する手法、およびユーザに効果的に体験獲得情報を想起させるインタフェースを提案する。



(図 1) 行動検索の流れ

3.1 アクティブ期間

本節では、行動を管理する最小単位として『アクティブ期間』について述べ、ユーザの興味の強さを表すアクティブ期間の注目度、およびキーワードに対するアクティブ期間の重要度を算出する手法を提案する。

現在普及しているコンピュータの多くは、ウィンドウシステム[3]を採用したオペレーションシステム(OS)を採用している。この種の OS を利用する場合、ユーザは Web ページを Web ブラウザのウィンドウ内に表示して閲覧や文字入力などを行っている。アクティブなウィンドウに表示している Web ページを変更する、あるいは別のウィンドウをアクティブすることで、利用する Web ページを切り替えている。たとえば、興味のある Web ページをリンククリックにより探し、いくつか注目すべき Web ページが見

†日本電信電話株式会社 NTT サイバースリユーション研究所

‡日本電信電話株式会社 NTT サービスインテグレーション基盤研究所

つかると、Web ブラウザを複数起動して両者を切り替えながら Web ページを比較することなどを行っている。

そこで、ひとつのウィンドウがある Web ページをアクティブに表示している期間を、『アクティブ期間』と定義する。アクティブ期間はユーザの行動をまとめる特徴的な単位であると考えられる。例えば、インターネットエクスプローラに表示している Web ページを、リンククリック、あるいは別ウィンドウのアクティブ化のたびに、新しいアクティブ期間が発生する。

ユーザはそれぞれのアクティブ期間内に、Web ページを閲覧し、注目した文章をコピーし、特に重要であると判断すれば印刷するなど、さまざまな行動を起こす。そこで、ユーザが多く行動を起こしていたアクティブ期間はユーザが興味を持った可能性が高いと仮定し、アクティブ期間に対するユーザの興味の強さをあらわす指標として注目度を定義する。注目度は、アクティブ期間内にユーザが行った行動の種類をレベルわけして重み付けし、行動の発生回数によって式(1)のようにスコア化した。更に、任意のキーワードに対するそれぞれのアクティブ期間の重要度を式(2)のように数値化した。

$$Att(ap) = \sum_i (E_i \times Fr_i) \quad (1)$$

$$I(k, ap) = Att(ap) \times R(k, ap) \quad (2)$$

$Att(ap)$: アクティブ期間 ap の注目度

$I(k, ap)$: キーワード k に対するアクティブ期間 ap の重要度

$R(k, ap)$: アクティブ期間に表示されている Web ページのキーワード k に対する適合度

E_i : 行動種類 i の重み付け係数

Fr_i : アクティブ期間 ap 内の行動種類 i の発生回数

ap : アクティブ期間

k : キーワード

i : 行動種類

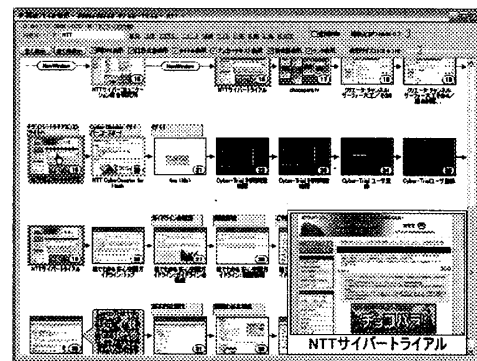
ここで、アクティブ期間に表示されている Web ページのキーワード k に対する適合度は、ユーザが過去に閲覧した時点で収集したすべての Web ページを母集団とする TF/IDF 値[4]とした。行動種類としては、Web ページ閲覧時間、コピー、印刷、マウス操作、キーボード入力、文字列選択を考慮した。また、アクティブ期間を最小単位として管理することで、Web ページの URL でグループ化し、Web ページ毎の重要度を算出することができる。

3.2 行動シーケンス検索インタフェース

図 2 に行動シーケンス表示インタフェースを示す。具体的にはアクティブ期間の遷移が Web ページのサムネイルを用いて時系列に表示される。これにより、ユーザは複数の Web ページの内容を発見できる。

Web ページを閲覧する場合、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ことは多い。本インタフェースでは、個々のサムネイルの下にタイトルを表示する。サムネイルの上に、検索エンジンに入力した

キーワード、あるいは表示されている Web ページに遷移するために前の Web ページ上でクリックしたアンカーテキストを表示する。これにより、Web ページの内容の理解を高めるだけでなく、ユーザにとっての Web ページ間の関係を想起できると思われる。たとえば、検索エンジンに『デジタルカメラ 薄型』と入力して薄型のデジタルカメラを探し始め、ニュースサイトでお勧めの情報を調べた後、メーカーの商品ページで詳細にスペックを調べたといった Web ページ間の関係を想起できると思われる。さらに、短時間で重要な体験獲得情報を想起できるよう、表示されているアクティブ期間のうち、ユーザが指定したキーワードに対するアクティブ期間の関連度の高いものは、赤色、あるいはオレンジ色の枠によって強調表示される。関連度が低いアクティブ期間は、ユーザにとって有用ではない行動であったと想定し、表示しないといったフィルタリング機能を備える。



(図 2) 行動シーケンス検索インタフェース

4. まとめ

本稿では、PC 操作履歴に基づいて、ユーザの過去の PC 上での行動を検索/提示するシステムを提案した。ひとつのウィンドウがある Web ページをアクティブに表示している期間を、『アクティブ期間』と定義し、ユーザが多く行動を起こしていたアクティブ期間はユーザが興味を持った可能性が高いと仮定し、アクティブ期間に対するユーザの興味の強さをあらわす注目度を算出した。また、アクティブ期間を時系列に表示し、閲覧時間やクリックしたアンカーテキストなどのアクティブ期間の属性情報を提示することにより、Web 閲覧体験によって得られた情報を想起させるインタフェースを提案した。今後は、行動シーケンス検索インタフェースの評価実験を行う予定である。

参考文献

- 1) A. Cockburn and B. McKenzie: What Do Web Users Do? An Empirical Analysis of Web Use., Int. J. Human-Computer Studies, 54(6), pp.903-922, (2001).
- 2) <http://desktop.google.com/about.html>
- 3) 尾川浩一, 松山佐和, 岩崎晴美: ウィンドウシステム入門, 朝倉書店(1996).
- 4) 北研二: 情報検索アルゴリズム, 共立出版(2002).