

ページの表示領域を考慮した Web 履歴の拡張

Enhanced Web History with Page Appearance

三浦 元喜* 國藤 進* 佐藤 省吾† 田中 二郎‡

Summary. We propose a method to enhance web browsing history by considering “page appearance.” The page appearance represents properties of web browser window such as size, location, z-order, activated status, and so on. We consider the page appearance can be utilized for both estimating importance of web page and extracting tacit relationship between pages. Using the appearance, users can extract pages which are only visited once but practically used. We developed a logger and a viewer to introduce our method. The tools help users to find informative pages previously visited. We concluded that “visit count” was short for efficient re-visitation by preliminary experiment.

1 はじめに

多くの Web ブラウザは以前閲覧したページへの再訪問を支援するための履歴 (History) 機能を備えている。Cockburn and McKenzie[2] の調査によると、およそ 81% のページ閲覧が「再訪問」である。しかしながら、Catledge and Pitkow[1] の調査でも示されているように、履歴機能による再訪問は一般にはあまり利用されていない。

我々は、履歴機能が再訪問時に使われない理由として、自動的に記録される膨大な記録の中から、ユーザが必要な記録を抽出することが難しい点にあると考えている。既存の Web 履歴機能は、訪問時刻と訪問回数という 2 つの属性によって必要な記録を抽出する機能を備えているが、特に訪問回数は必ずしもユーザが考えるページの重要度と一致する属性とは限らない。

そこで我々は、ユーザが Web ページをどのように扱って閲覧していたかという情報を取得し、Web 履歴の属性として追加する [3]。この情報により、ユーザが重要と考えるページへの再訪問を容易にし、Web 履歴の有用性を高めることが期待できる。

2 ユーザの Web 閲覧行動の記録

我々は、Web ページの表示状態を取得し、ユーザがどのような Web 閲覧行動をとったかを記録するため、Web ブラウザウィンドウの以下の 2 つの属性に着目している。

1. ウィンドウの位置、サイズ、重なり (Z-order)

© 2005 日本ソフトウェア科学会 ISS 研究会。

* Motoki Miura, Susumu Kunifujii, 北陸先端科学技術大学院大学 知識科学研究科

† Shogo Sato, 筑波大学大学院 理工学研究科

‡ Jiro Tanaka, 筑波大学大学院 システム情報工学研究科

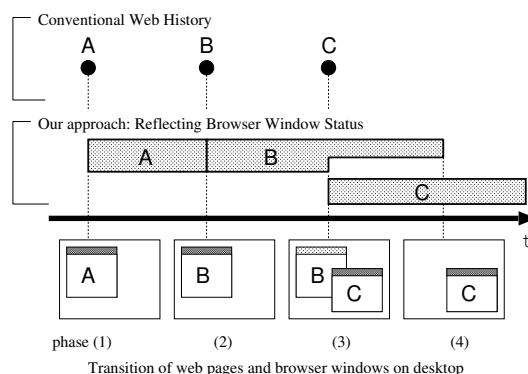


図 1. ページの表出度計算の例

2. ウィンドウの状態 (アクティブ化、最小/最大化)

加えて、ページを読み込んだ時刻 (訪問時刻) に加えて、ページを閉じた時刻を取得し、ページが表示されていた期間 (duration) を得られるようにした。これらの指標により、ユーザがある Web ページをどの程度デスクトップ画面上で表出させていたか、という指標の計算が可能となる。図 1 は具体的な例として、(1) で開いたページ A の領域が (2) でページ B に置き代わり、さらに (3) でページ C が別ウィンドウでページ B を一部覆った状態が表示され、(4) でページ B が閉じられる、という一連の流れにおけるページの表出度を、表示領域と表示時間の積として表現できることを示している。

また複数ウィンドウの情報は「作業時におけるページ間の関連付け」にも利用できる。複数の Web ページを同時に開き、頻繁にアクティブ/非アクティブの状態を繰り返しながらタスクを遂行する行為はしばしば行われる (例えば、会議のページと宿泊施設のページ、交通手段を調べるページを行き来しながら申し込みや予約を行うなど)。このような作業状

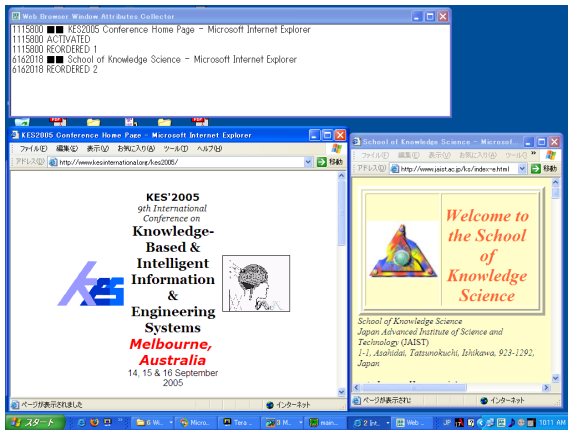


図 2. ログ収集の様子

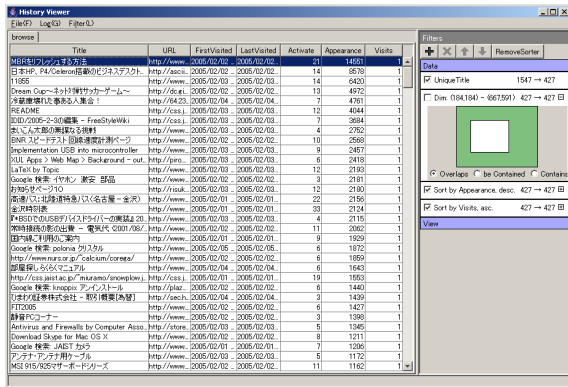


図 3. ログ閲覧システム

況を取得することで、ページのリンク関係からは抽出できない「暗黙的な関連」を抽出する効果が期待できる。

3 システム実装

ログ収集プログラム (タスクトレイに常駐する Windows アプリケーション, 図 2) は、1 秒おきに Web ブラウザウィンドウのサイズと、その内の URL を表現しているウィンドウキャプションを取得し、ファイルに記録する。ログ閲覧プログラム (Java アプリケーション, 図 3) では、ユーザは“Activate (アクティブ化された回数)” “Appearance (表出度)” “Visits (訪問回数)”などをクリックし並べかえたり、ウィンドウ表示領域によって絞り込んだり、URL をクリックしてブラウザを開き再訪問できる。図 3 は、1 回しか訪問していないが、表出度が高いため重要である可能性が高いと判定されたページのタイトルと URL を表示している例である。

4 予備実験

約 2ヶ月間のログ (11,476 訪問) を収集し、その中でタイトルが重複していない 2,505 ページについ

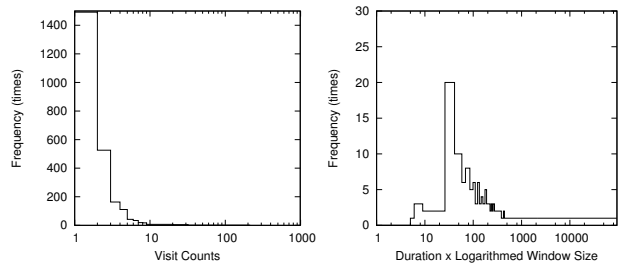


図 4. 訪問回数と、表示指標 (表出度) との分布比較

て、訪問回数と表出度という 2 つの指標での分布を調べた。なお、ページの表出度は、暫定的に表示時間 (秒) $\times \log(\text{ウィンドウサイズ})$ という式で計算している。一度しか訪問していないページは約 59% であり、9 回以下の訪問回数のページが約 96% を占めることから、訪問回数指標でページの絞り込みを行うのは困難であることが確認された。訪問回数に比べると表出度の分布は分散しているため、もし表出度とユーザが再訪問したいページとの関連性が高い場合には再訪問に有効な指標として利用できると考えられる。

5 まとめと今後の課題

Web 履歴機能を強化するため、ユーザの閲覧時のブラウザウィンドウの属性と状態を取得し、Web 履歴の属性として追加する手法を提案した。今後は、表出度の妥当性と暗黙的ページ間の関連の有効性を調査し、利用しやすい Web 履歴機能を提供していきたい。Web 閲覧活動とユーザの作業履歴との関連を詳細に調べるため、個人の作業履歴蓄積システム NecoLogger[4] の利用することも視野に入れながら進めていきたいと考えている。

参考文献

- [1] L. D. Catledge and J. E. Pitkow. Characterizing Browsing Strategies in the World-Wide Web. *Computer Networks and ISDN Systems*, 27(6):1065–1073, 1995.
- [2] A. Cockburn and B. McKenzie. What Do Web Users Do? An Empirical Analysis of Web Use. *Int. J. Human-Computer Studies*, 54(6):903–922, June 2001.
- [3] M. Miura, S. Kunifuji, S. Sato, and J. Tanaka. Capturing Window Attributes for Extending Web Browsing History Records. In *Proc. of 9th Int. Conf. on Knowledge-Based Intelligent Information and Engineering Systems (KES2005)*, Part I, pp. 418–424, Sept. 2005. (LNAI 3681).
- [4] 近藤 秀樹, 三宅 芳雄. Retrospector: 計算機上での履歴を振り返ることによる日常作業支援. In *Proceedings of WISS 2004*, pp. 151–152, Dec. 2004.