

## あしあと記録を用いたウェブコンテンツ探索知識の再利用と共有

## A Footmark record bookmarks for re-using and sharing know-how of finding specific web contents

田村 美帆<sup>†</sup> 佐藤 慶三<sup>‡</sup> 中島 誠<sup>‡</sup> 伊藤 哲郎<sup>‡</sup>

Miho Tamura Keizo Sato Makoto Nakashima Tetsuro Ito

## 1. はじめに

近年、インターネット技術が向上するにしたがい、探索するウェブコンテンツの数も膨大なものとなってきている。また、オンラインの広報誌や電子書籍のように繰り返し参照されるウェブコンテンツも増えてきている。このようなコンテンツのうち、参照時にブックマークされなかったものについては、再探索が行われる[3]。

ウェブコンテンツの再探索には、ブックマークや履歴機能の利用が考えられるが、参照されたウェブページのURLのみが記録されており、それらは、目的のコンテンツについて、どのようにして発見に至ったかという情報が利用できない。また、ウェブコンテンツの情報共有する場合、情報提供者にとって、従来のブックマークや履歴の公開はコンテンツのURLとなるため望ましくない。

本稿では、あしあと記録を提案する。あしあと記録は、ウェブサイトのホームページやポータルサイトから目的のコンテンツに至るまでのユーザの操作を探索知識として収集したもので、これによりURLやキャッシュを利用せずに目的コンテンツの再探索を可能とする。この知識によれば、参照済みコンテンツが増大しても、それらの中から参照したいコンテンツに効率的に辿りつくことができる。辿る過程をユーザに見せることによって、目的のコンテンツ周囲にある他のコンテンツも閲覧することが可能である。実験により、あしあと記録を再利用することで、手動で再探索をした際の操作回数を減らせるとともに、個々のウェブページのデザインの違いにも対応できることを確認した。

## 2. あしあと記録による再探索

あしあと記録では、目的のウェブコンテンツまでの辿り方について、経路上のウェブページのURLではなくリンククリックなどの入力情報を扱う。

図1に、あしあと記録の収集概要を示す。図中、ユーザはあるホームページをあしあと記録の収集のためのスタートページとし、そこから複数のウェブページをたどりながら目的のコンテンツに辿りつく。目的のコンテンツは、ユーザの興味のあるウェブコンテンツを指し、その決定はユーザに委ねられる。ホームページから目的のコンテンツに至るまでの入力情報のあつまりがあしあと記録となる。入力情報として、次のウェブページに進んだ際のリンクの位置座標および、その座標の周囲のテンプレート画像の特徴量の2つをあしあと記録として利用する。

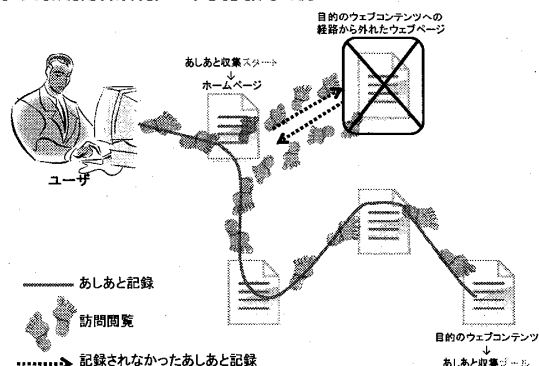


図1 あしあと記録の収集概要

目的のコンテンツを発見するまでに、「戻り」が発生する場合がある。このようなページはコンテンツへの経路としては記録不要であると考えられる。あるウェブページから「戻った」とき、戻る前のウェブコンテンツを記録から削除することで、最適なあしあと記録を作成する。

図中、ウェブコンテンツに至るまでのユーザの操作回数は3回分に相当する。この複数回の操作が、あしあと記録による探索過程では自動的に行われるため、1回の操作で済むようになる。

## 3. 実装

通常のウェブブラウザでは、ユーザの入力情報を記録・再現できないため、CollaboTray[1]の技術を利用する。CollaboTrayでは、アプリケーションを共有した協調作業を行うために、表示されているアプリケーションに対して、ユーザからのマウスやキーボード入力情報などを収集できる。このような機構を用いることで、あしあと記録の収集と再利用が可能になる。

## 3.1 収集と記録

CollaboTrayによるあしあと記録の仕組みを図2に示す。ユーザはディスプレイに表示されているブラウザを操作し、それらの情報をCollaboTrayが収集する。収集されたイベント情報はCollaboTrayを仲介してブラウザへ送信される。ここでイベント情報とは、ブラウザ上でのマウスのクリック情報や、キーボードからの文字入力である。イベント情報とその座標値を収集することで、あしあと記録を生成していく。

クリック対象のリンクの視覚効果にFlashなどの技術を取り入れるウェブページも多く見られる。このようなウェブページでは、クリック対象の位置が頻繁に変動する。動的なリンク位置に対応するため、クリック位置周辺の部分テンプレート画像の特徴量も同時に収集する。図2上部は、CollaboTrayが、ユーザからのイベント情報であるマウスクリックを収集し、あしあと記録を作成、ブラウザにイベント情報を送信する流れを示している。

<sup>†</sup> 大分大学大学院知能情報システム工学専攻  
Dept. of Computer Science and Intelligent Systems,  
Graduate School of Engineering, Oita University

<sup>‡</sup> 大分大学工学部知能情報システム工学科  
Dept. of Computer Science and Intelligent Systems,  
Faculty of Engineering, Oita University

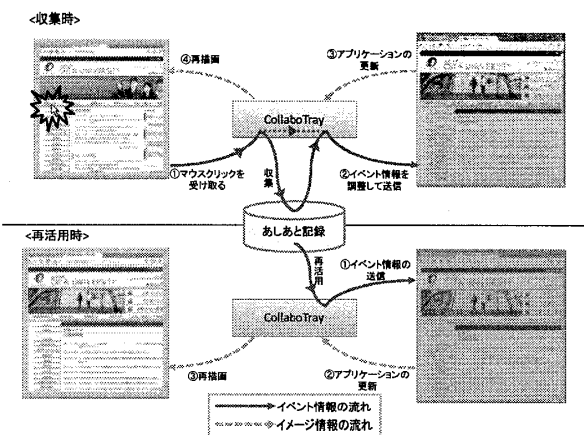


図2 CollaboTrayを利用したあしあと記録の収集・再活用

### 3.2 再活用と共有

ウェブコンテンツを再探索する場合、あしあと記録をCollaboTrayからブラウザに送信する。その際、リンク位置の座標値に加え、座標の周辺のテンプレート画像の特徴量を用いて該当のコンテンツを探索していく。その探索には、アクティブ探索法[2]を使用する。図2下部では、CollaboTrayが作成したあしあと記録をイベント情報としてブラウザに送信し、ユーザがそのブラウザを閲覧することで、あしあと記録が再活用される流れを示している。

あしあと記録は、探索過程の再現という方法により、前回閲覧したウェブコンテンツだけでなく、周辺のコンテンツの探索も容易に行えるという利点がある。ブックマークを使用すると、目的のコンテンツ周囲にあるコンテンツの再探索を行うことは困難であり、履歴を使用すると、その中から目的のコンテンツを探し出さなければならない。

コンテンツのURLであるブックマークを公開する場合、サイトの意向でそのURLの公開が認められない場合がある。一方で、あしあと記録をウェブ上に公開する場合、記録として検索キーワードや座標値、テンプレート画像の特徴量のみを保持しており、URL自体を公開せずに目的のコンテンツへアクセスすることが可能である。

## 4. 検証

本稿で提案したあしあと記録を用いて、公開の妥当性の確認やウェブコンテンツを再探索した際の操作回数の比較を行った。対象として、広報誌BUNDAL.OITA[4]および岡山県立図書館デジタル絵本[5]、へんてことば[6]を使用した。それぞれ対象となるコンテンツ数は26と112、100であった。

### 4.1 あしあと記録の公開

一般的に、ウェブコンテンツ自体のURLの公開は適当ではないため、ホームページのURLのみ公開されている。ブックマーク、履歴を利用した再探索を行う場合は、公開するURLはウェブコンテンツ自体となる。しかしながら、ユーザが閲覧した履歴やブックマークをウェブ上に公開することは不適切である。一方で、あしあと記録を利用した再探索の場合、ウェブコンテンツに至るまでのあしあと記録を公開することになる。

図3 あしあと記録 □ 手動

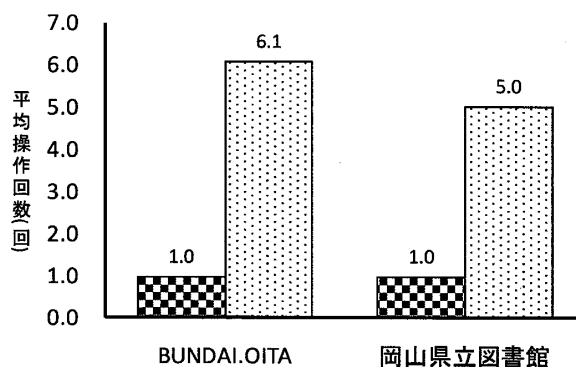


図3 操作回数の変化

図3では、BUNDAL.OITAおよび岡山県立図書館デジタル絵本を対象とした操作回数の変化を示している。従来の手法では、BUNDAL.OITAでは平均6.1回、岡山県立図書館デジタル絵本では平均5.0回の操作回数を必要とした。あしあと記録では対象コンテンツを選択する1ステップのみとなっており、操作回数を軽減できていた。

### 4.2 テンプレートマッチングによる特定率の変化

へんてことばを利用して、テンプレート画像の特徴量によるクリック位置の特定率の変化を調べた。へんてことばのホームページには、Flashで作成された絵本選択画面がある。絵本の表紙画像のクリックで各絵本にアクセスできるデザインとなっており、表紙画像の配置はアクセスするごとに異なる。ここでは、表紙画像配置が異なる3パターンについて、各絵本の表紙画像の特徴量から再参照対象絵本を特定する特定率を調べた。

実験の結果、特定率は平均96.0%であり、要した時間は平均11.3秒であった。ここで、特定できなかった4.0%については、間違ったリンクを抽出したのではなく、該当箇所が発見できなかった場合であった。そのため、テンプレートマッチングによる誤特定はなく、収集するテンプレートのサイズの精度や処理時間が問題になると考えられる。

## 5. まとめ

本稿では、リンクのクリックなどのユーザの操作を元にウェブコンテンツの辿り方を表現するあしあと記録について述べた。今後の課題としては、経路上のウェブページ上でのイベント情報補足をより正確に行うための改良と頑健性についての評価を行う。

### 参考文献

- [1] Yuki, A. et al. "Tolerant Sharing of a Single-user Application Among Multiple User in Collaborative Work," Proc. CSCW2010, Savannah Georgia, USA, pp.555-556, Feb. 6-10, 2010.
- [2] 村瀬 他: 局所色情報を用いた高速物体探索-アクティブ探索法-, 電子情報通信学会論文誌, Vol.J81-D-II, No. 9, pp.2035-2042.
- [3] Suugjoon, S. W. et al, "Contextual Web History: Using Visual and Contextual Cues to Improve Web Browser History," Proc. CHI2009, pp1457-1466. (2009).
- [4] 大分大学広報誌 BUNDAL.OITA: <http://www.oita-u.ac.jp/>
- [5] 岡山県立図書館: <http://www.libnet.pref.okayama.jp/>
- [6] へんてことば: <http://www.fujiigumi.com/hentekotoba/index.php>