

# ブラウジング履歴に基づく Bookmarkless Bookmark の提案

大橋 史武 粕谷英人 大久保弘崇 山本晋一郎 (愛知県立大学大学院情報科学研究科)

## はじめに

近年のインターネット接続環境の普及に伴い、WWW のブラウズによる情報収集は一般的なものとなった。ユーザは、WWW を利用し情報収集をしていく過程で多くのページを見ていくが、その中で自分が興味をもったページを後で参照できるようにするには、そのページをブックマークに登録しておかなければならない。ブックマーク登録しなかったときユーザは、「世界中に存在する全てのページを対象」とした検索エンジンを利用した検索、ブラウザの「履歴情報」を調査、もしくは目的のページまでの経路を思い出しハイパーリンクを辿る、などの作業を強いられるが、それでも目的のページが見つけれられるとは限らない。また手作業によるブックマークのカテゴリ分けは、カテゴリ間に重複部分が生じるため分類に工夫が必要となり、ブックマーク登録し、参照する際にカテゴリを取り違えると、かえって参照の妨げとなる。さらに、ページをブラウズしている段階で、後からどのページが必要になるか正確に判断することはできない。そこで、ユーザが閲覧した全てのページの内容、閲覧状況をデータベースに保存し、必要に応じてこの中から検索、キーワードの取り出しを行うことができる、Bookmarkless Bookmark システムを提案する。そして「あらかじめブックマークする」必要性からユーザを解放することで、情報収集のための快適なブラウジング環境を提供する。

## Bookmarkless Bookmark システム

図 1 に Bookmarkless Bookmark システムの概要を示す。本システムでは、ユーザがページを探索し情報収集をしていく過程で閲覧した全ての HTML 文書に対して、その内容、閲覧状況などのデータを収集しデータベースに保存する。そしてこのデータベースを活用して、閲覧したページの検索・キーワードの取り出しなどの、ブラウジング効率を向上させる諸機能を提供する。

Web 利用の調査 [1] によれば、一般的なユーザが閲覧するページ数は 1 日約 41 ページである。1 日に閲覧する HTML 文書のデータ量は 1~2MByte 程度であり、昨今のハードディスクの大容量化を考えれば、本システムで十分管理可能なサイズである。

次に、本システムが備える機能を以下に挙げる。

- 閲覧済みページの再取得
  - 語のゆらぎを考慮に入れたキーワード検索
  - 閲覧した日付、時間、閲覧場所などのシチュエーションによる検索
- ページの自動カテゴリ分け
  - 閲覧ページからキーワードを抽出し、その相関関係によりページの主題を特定し、概念の包含関係からページの階層的カテゴリ分けを自動的に行う

本システムを利用することで、以下に挙げることが可能となる。

- 閲覧したことのあるページから、再び必要になったページを容易に取り出すこと

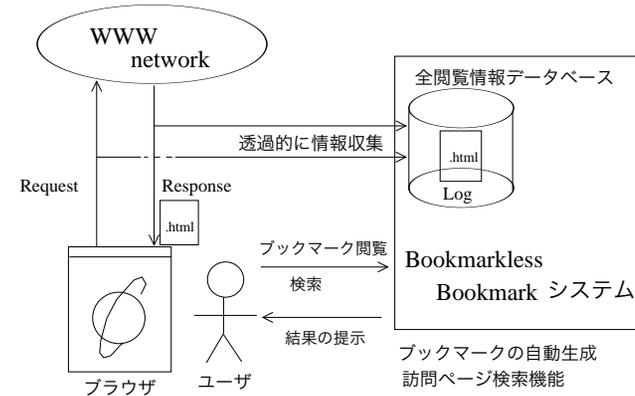


図 1: Bookmarkless Bookmark の概要

ができる。またそのためのブックマークの作成・整理の作業は不要

- ユーザに密着したシステムであり、検索語以外に「ページのリンク元、閲覧した日時、閲覧場所」といった多角的な情報を用いた検索ができる
- 各人の関心の方向性に応じたカテゴリの細分化が自動的に行われる

## 3 まとめ・今後の課題

Bookmarkless Bookmark システムにより、ユーザが過去の閲覧ページを記憶する必要がなくなり、情報収集のための快適なブラウジング環境の提供が可能となる。さらに、本システムのデータベースを活用することで、閲覧したページ他ページからの被参照数・閲覧回数などを基にしたユーザのそのページへの評価、ユーザによるページへの注目度のフィードバック、ページの巡回経路の可視化、ブラウズ状況のリプレイ、などの機能も提供できる。また、本システムの概念を RFC822 形式のメッセージに適用し、メールフォルダの管理に拡大することで、個人の情報管理をさらにサポートするシステム [2] に発展させることも可能である。

## 参考文献

- [1] Bruce McKenzie and Andy Cockburn: An Empirical Analysis of Web Page Revisitation, In Proceedings of the 34th Hawaiian International Conference on System Sciences, HICSS34, Maui Hawaii, (CD-ROM). IEEE Computer Society Press, 2001
- [2] 伊地知宏, 倉部淳: メールでの話題の把握を支援するソフトウェア, 未踏ソフトウェア創造事業, 平成 14 年度