

統合ブックマーク生成システムの提案

045154 渡辺伸一

(指導教員 速水治夫教授)

1 はじめに

近年、インターネット上の Web ページ数は増加の一途を辿っている。しかし情報の質の幅が広がったために、情報量の多さや信頼性の高さが優良な Web ページを探す労力の増加が問題となっている。¹⁾

Web ブラウザにはブックマーク機能がある。これは利用者が再度訪れたいと考えた Web ページを Web ブラウザに保存できる機能である。ブックマークに登録されている Web ページは、その管理者によりフィルタリングされているため、優良な Web ページが多く存在すると推測される。

そこで本研究では、複数利用者のブックマークを統合することにより、集団知を利用して優良な Web ページが集まるリンク集を出力するシステムを提案し、構築した。

2 システム概要

利用者がインターネット上で本システムを利用できるようにするため、Web アプリケーションの形をとる。そこで開発環境として MySQL, Apache, PHP の組み合わせを用いた。

本システムで出力するリンク集である統合ブックマークは、上位下位の二階層構造に分かれる。

統合にはデータベースに登録されたブックマークを利用する。まず順位算出部でフォルダの順位を計算し、構造構築部で構造を作成、構造調整部で余分なリンクを削除し、最終的に統合ブックマークを出力する。

システム構成図を図 1 に示す。

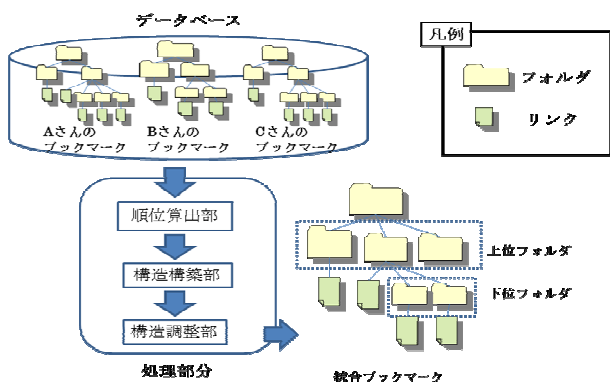


図 1 システム構成図

3 評価実験

インターネット上から取得したブックマークを本システムに登録し、統合ブックマークを出力した。この統合ブックマークを Google, Yahoo!カテゴリと比較した。評価尺度として、出力された Web ページ中のノイズでない Web ページの割合である精度とともに、実験協力者に協力していただき、目的の Web ページを得るまでの時間である速度、検索結果の Web ページが実験協力者にとってどれだけ役に立つ度合であるかを 5 段階評価していただいた有用性、検索結果中の知らない Web ページの割合である新規性を計った。評価結果を表 1 に示す。

表 1 評価結果

	本システム	Google	Yahoo! カテゴリ
速度 (秒)	16.889	27.000	200.333
精度 (%)	86.301%	84.667%	94.231%
有用性・平均 (5 段階)	3.812	4.185	4.136
有用性・標準偏差	1.168	1.074	1.081
新規性 (%)	75.198%	78.586%	78.189%

速度は他のシステムより良い結果となり、精度は Google より良い結果となった。しかし、有用性は他のシステムより劣る結果となった。標準偏差の高さと新規性の低さから、出力された Web ページの質の幅が広がったと推測される。

4 おわりに

実験結果から、一定の情報を持つ Web ページを探しやすい検索システムを構築できたといえる。しかし本システムで出力された Web ページは有用性が低く、質の幅が広がってしまった。これには登録するブックマーク数を増やすことや、インターネットから取得したブックマークでなく実際に利用されているブックマークを登録することで改善が可能だと推測される。今後、この点に関して実験を行うとともに、Web ページのフィルタリングを行うことで精度を更に高めていく。

参考文献

- 1) 森幹彦, 山田誠二: ブックマークエージェント: ブックマークの共有による情報検索の支援, 電子情報通信学会論文誌, J-83-D-I, pp.487-494(2000)