

robot と Hot Sauce を用いた個人 Home Page 環境管理

5 L-8

豊吉雅昭 中津川徳子 芝野耕司
東京国際大学商学部

1. はじめに

今日では、インターネットの爆発的な普及により、インターネット上に膨大な量の情報が氾濫するようになってきている。Webの利用者はこの膨大な量の情報の中から、いかに効率よく、価値の高い情報をタイムリに手に入れることができるかが大きな問題となっている。

ここでは、robot と Hot Sauce[1]を用いて、個人の興味ある Home Page を管理し、情報入手を支援するシステムを提案する。

2 情報の価値とは？

個人にとっての情報価値は、利用者個人が与える価値と情報自体の性質とによって定まる。

利用者個人が与える情報の価値は、利用者が個別情報に対して、重要である、すなわち、価値が高い、とする場合と、ある分野に対して興味を持つというように、ある範囲の情報に対して価値を見いだす場合とがある。

また、こうして利用者個人が与える情報価値にも、高低がある。

一般に情報は、その情報を利用者が入手すると再び読む価値がなくなる一過性の価値と何度も読みたい情報とに分かれる。これは、一般的には新聞や雑誌などのように読み捨てる情報源と書籍のように何度も読み返す情報源があることに対応する。

情報自体の性質によって定まる価値は、情報の鮮度、新規性に依存する。

すなわち、利用者が関心を持つ分野の情報で新規の情報は、価値が高い。

また、特定の情報であっても、その情報の更新度合いによって、利用者個人にとっての価値は異なることになる。

一方、価値の高い情報であっても、その情報に対するアクセスの容易さも重要である。

3 サービスの現状とその問題点

現在、上記のような個人の要求に対して、インターネットで得られるサービスには、検索サービスと巡回・監視サービスとの二種類のサービスが存在する。

(1) 検索サービス

インターネット上にいくつも存在するホームページ検索サービス。一般的な検索方法には、キーワード検索と分野ごとに整理された項目より探し出す方法の二種類である。キーワード検索では、利用者の必要とする情報を表す単語を入力し、同一単語を含んでいるか又はそのキーワードが付与されているホームページを検索する。一方、検索サービスに登録されているホームページは分類分けされているものもあり、その分類のなかから利用者が該当する分野をたどっていき必要な情報を探すサービスもある。これらが現状の検索サービスで用いられている手法である。

(2) 巡回・監視サービス

巡回・監視サービスは、アプリケーションプログラムとして一般に市販されている者が多い。これらは、利用者が頻繁に閲覧するホームページのデータをダウンロードし、オフラインでホームページの内容を見ることができるようにするものである。その際に、目的のページの更新を調べ、更新されている場合にはダウンロードをまた新たに行うようになっている。

このような既存の手法で、前述の価値の高い情報を入手できるかと問えば、そうでもない。検索エンジンでは大雑把な検索結果しか得られないため、最新情報を探し出すためには利用者自身がそのホームページを閲覧する以外に方法がない。更新の度合いを調べることも同様である。また、利用者負担をかけることになるので、情報を見落とす可能性もある。

一方、巡回・監視プログラムにより指定したページの最新情報の入手は可能だが、こちらでもそれ以外の最新情報を見落とす可能性は高い。また、関心分野に該当するホームページのデータが更新されている場合、古いデータは消去されているケースがあるため、利用者自身がそのデータを保存していない限り、再び利用することは不可能となる。

Personal management of Home Page using robot and Hot Sauce

Masaaki Toyoyoshi, Noriko Nakatsugawa, Kohji Shibano

Tokyo International University

4. 問題点解決に向けて

利用者の情報価値と情報自体の価値を活かしたサービスでは次の点を考慮する必要がある。

(1) 個別具体的に対象そのものを利用者が指定する場合

この情報の管理は大きく二つに分けられる。一つ目の問題は一度目を通してしまえば全く必要がなくなる情報である。この情報の管理に最も必要なことは、常に新しい情報が更新されているかどうかを監視することである。二つ目は何度でも目を通したいと、利用者が感じる情報である。この情報管理に必要なことは、利用者が必要だと感じている情報のみをローカルに保存する事で解決する。

この情報群についてはブラウザソフトの Book Mark に利用者自身が保存した URL が相当するので、ここに利用者が登録している BookMark 情報を利用する。

(2) 高い価値がある情報が含まれていそうな情報群について

この情報群は利用者にとって最も高い価値のある情報とまではいかないが、比較的高い価値の情報が当てはまる。この情報群については利用者のブラウザソフトの閲覧履歴 (history) が相当すると思われる。

これらの情報 (URL) を抽出、対比することにより、利用者にとって高い価値のある情報を管理するシステムを構築する。

(3) 利用者が興味を持つ分野

利用者自身が常に興味をもっていて、頻繁に検索を実行している分野が当てはまる。この分野については検索エンジンから得られる情報を利用者が検索エンジン自体に登録し、一定時間ごとに検索を実行させることで、解決の方向を見いだす。

上記の(1)~(3)を統合して、サービスを提供する事が必要である。

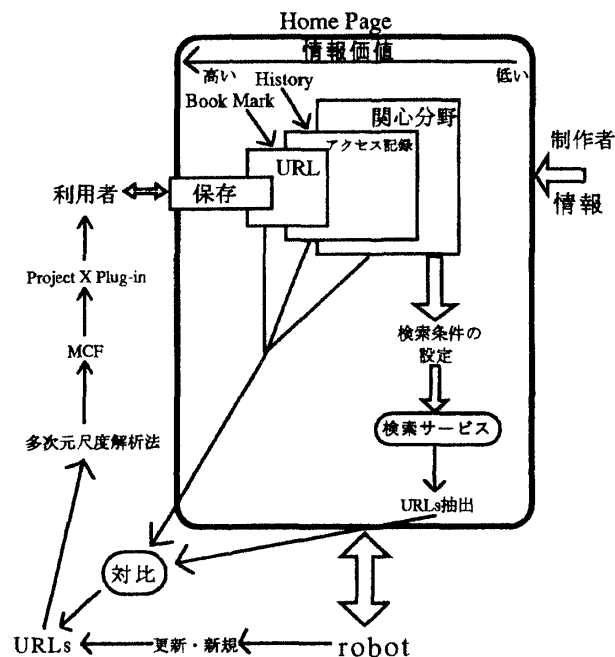
5. システム構成

このシステムは再構成された個人情報をも robot に登録し、インターネット上を巡回、指定されたホームページのファイルを集める。そして指定のホームページは bookmark に登録されたホームページと history ファイルより抽出されるアクセス記録より、閲覧頻度の高いホームページを対象とする。

さらに、関心分野として登録された検索条件を用いて、一般の検索サービスを利用しホームページを検索する。この得られた検索結果と指定されたホームページのデータを比較し、新規、更新された情報を抽出する。

得られたホームページの情報の価値を算出し、Hot

Sauce 上に展開することにより利用者により利便性の高いインターフェースを提供する。アクセス記録などの個人情報と、ホームページの新規・更新具合を元にホームページごとの情報の価値の高さを算出し、この算出された値をもとに多次元尺度解析[3]を利用した構成を疑似三次元上に行い、その結果をブラウザソフトの Project X Plug-in を利用して Hot Sauce 上に配置する。さらに利用者自身が情報をローカル上に保存できる環境を整備する。(下記の図参照)



システム構成図

6. 終わりに

このシステムを使用することにより、個人の利用者にとって情報の検索、抽出が容易になるものと思われる。

参考文献

- [1] Hot Sauce <http://hotsauce.apple.com/>
- [2] MCF <http://mcf.research.apple.com/hs/mcf.html>
- [3] 林 知己夫/鮑戸 弘共編, 「多次元尺度解析法 -その有効性と問題点-」, サイエンス社, 1989