検索機能を付けた個人別履歴の作成

96H068 高徳　哲也

1　はじめに

ここ数年のインターネットの普及により大学内などでも Web ページを閲覧することによって情報を取得する機会が増えてきており、それに伴い Web ページを閲覧する回数も多くなってくる。そこでユーザが過去に見た Web ページを再び見るためによく使われる機能として WWW ブラウザのブックマーク機能がある。しかし、このブックマーク機能には欠点もあり、特にユーザがブックマークを登録すると次からブックマークを利用するためにブックマークを登録したパソコンを利用しなければならなくなる。そのためユーザは使用するパソコンを限定されてしまい、大学の演習室などのようにパソコンが複数台あるような所では使いにくいことがあげられる。

そこで本研究では個人別の履歴表示と検索機能を用いることによってユーザが大学の演習室などで使用するパソコンを限定されず、より簡単に過去に見た Web ページを見つけ出せる環境を提供できるようにした。

2　研究内容

本研究では従来のブックマークの欠点を補うためにキャッシュサーバに保存してある様々なログを用いて、フィルタ処理を行い個人別の履歴を作成する。そのため、キャッシュサーバに保存してあるアクセスログなどのログから履歴に必要なユーザ情報、Web ページにアクセスした時間、URL、Web ページのタイトルや内容を取り出すのにフィルタ処理を行ってその日の履歴を作り、ユーザ個人の履歴表示にはその日の履歴を元に CGI を用いてユーザからの要求があるたびに作成するようにした。ユーザ認証とログ内に保存されているユーザ情報の作成は WWW ブラウザの環境変数を用いて行っている。そのため本研究ではこの環境変数に用いられる情報がクライアントマシン側にあることが前提となるが、これを用いることによってユーザはユーザ認証のためにユーザ名やパスワードを入力すること無く個人の履歴を作成し、見ることができる。

また履歴表示に Web ページに含まれていたキーワードを用いて検索できる機能を付け加えたものを作成する。このキーワードを用いた検索機能には日本語全文検索ソフトである「Namazu」を用いた。この「Namazu」を CGI として動作させることでインデックス作成の高速な検索を実現しており正確な HTML の取り扱いなど様々な利点を持たせることができる。本研究での検索機能にはキャッシュサーバのログに保存している URL のみを検索し、検索する日付を指定することにより、より的確に捜し出すことができ、検索結果から目的の Web ページが簡単に閲覧できるようになっている。

3　まとめ

本研究では、キャッシュサーバとして使われている Squid を利用してログにフィルタ処理を行い、CGI の環境変数を用いることによってユーザ別の履歴を表示させることができた。さらにキーワード検索機能を付けることによってより簡単に過去に見た Web ページを再び閲覧することが可能になった。

しかし、Web ページによってはキャッシュに含まれているバイナリデータによる履歴のタイトル、内容表示での文字化け問題や、大量に履歴が増えると「Namazu」によるインデックス作成に時間がかかるなどの問題がある。また大きな問題としてクライアントマシンの OS によってはユーザ認証のために用いる情報を標準で持たないものもあり、その場合クライアントマシンごとに情報を持たせるよう設定を変えなければならない問題がある。さらに検索機能をより充実するために時間帯指定での URL 検索やユーザによる履歴の保存期間の指定することなどを実現することが今後の課題として考えられる。