

## WWWにおけるグループ経験の共有化による prefetch 機能の効率化\*

6L-3

山田 仁 斎藤 逸郎 石塚 満  
東京大学工学部電子情報工学科  
hitoshi@miv.t.u-tokyo.ac.jp

### 1 はじめに

WWWは、インターネット上においてその主要なサービスの一つであるが、WWWの使用中に問題となるのが、その待ち時間の長さである。この問題を解決するための手段の一つとして、ユーザが情報の転送を要求する前に、あらかじめ情報をとってくるというprefetch機能がある。すでに単純な戦略によるprefetchの機能は実現されているが、トラフィックの増加が大きいなどの問題点がある。そこで、このような問題を解決するために、個人のURLの履歴とグループ内でのURLの履歴をもとにprefetchするページを決定するという、より効率的なprefetchの方法について述べる。

### 2 WWWにおけるprefetch機能

WWWにおける情報は、URLで表されるページを単位とするものである。WWWの使用中は、ユーザは次から次へとページを見ていくのであるが、このとき情報を見ている時間と情報が送られてくるのを待っている時間は別々である。すなわち、ユーザがページを見ている時間は回線が空いている時間であり、この時間を利用してユーザが次に見るであろうページを転送すれば、次に情報を見る時の待ち時間は少なくなる。このように、ユーザからの要求がないうちに、あらかじめ情報をとってくるのがprefetch機能である。

prefetch機能にとって最も重要なのは、prefetchするページの決定である。ユーザからの要求がないうちにページをとってこなければならないので、

ユーザが次に見るであろうページを予測する必要がある。HTMLで記述されるWWWの各ページは、ハイパーリンクによって関連付けられており、一般にユーザはこのリンクをたどってページを見ていくことが多い。従って、ページを予測するときは、今見ているページからリンクが張られているページの中から選ぶようにすればよい。

ここで、一番簡単な方法は、今見ているページの中で上に書かれているリンク先のページから順番に選ぶ方法である。この方法で、リンク先のページのインラインイメージも含めて10個をprefetchするという実験[1]が行われているが、この結果、待ち時間が約2.5分の1になり、同時にトラフィックが約3倍になったことが報告されている。このように単純な方法でも確かに待ち時間は短くなっているが、転送されても利用されない情報が多く、トラフィックの増加が大きいという問題がある。

いま、prefetchのヒット率を、ユーザが実際に要求したページの数に対して、そのページがprefetchしたページと一致する割合とすると、どのような方法でprefetchするページを決定しようと、prefetchするページの数多くすれば、prefetchのヒット率は上がり、待ち時間はより短くなることが期待されるが、同時に利用されない情報が多くなり、トラフィックの増加も大きくなるであろう。トラフィックの増加を抑えたまま待ち時間を短くすることができるような効率的なprefetch機能を実現するためには、prefetchするページを決定する際に、ユーザが次に見るであろうページをより正確に推測する必要がある。

### 3 グループ経験の共有化

そこで、ユーザが次に見るであろうページをより正確に推測するために、まずユーザ個人の好みを判断する。ユーザの好みを判断する方法として、

\*The Efficient Prefetching Technique on the WWW by Sharing Histories of the Group Members  
Hitoshi Yamada, Itsuro Saito, Mitsuru Ishizuka  
Department of Information and Communication Engineering, Faculty of Engineering, The University of Tokyo  
7-3-1 Hongo, Bunkyo-ku, Tokyo, 113, Japan

リンク先の各ページが要求された回数、ハイパーリンク先（あるいはブックマークに登録したページ）を選ぶ割合、アンカー部分のキーワードマッチ、ページの上の方のリンク先を選ぶ割合、などに応じて候補先のページをランク付けすることが考えられる [2]。しかし、この prefetch するページを決定する部分は、ユーザがページを見ている間に行われなければならない、短時間で処理することが重要である。そこで、本研究では、prefetch するページの候補となるリンク先の各ページの URL について、ユーザが要求したページの URL の履歴を調べ、ランク付けすることにした。具体的には、候補となる URL について、要求された回数が多いものや、最近に要求されたものを、ユーザが次に見るであろう確率が高いページと判断して prefetch することにする。

このような個人の好みを判断する方法だけでも、個人が何度も同じページを訪れるような場合には有効である。しかし、初めて訪れたページを見ているときなど、個人の URL の履歴にリンク先のページの URL に関する情報がない場合には利用できない。そこで、そのような場合には、URL の履歴などの情報を個人だけでなく、グループのメンバーから集め、そのグループ内で情報を共有する方法 [3] を用いることにする。これは、ユーザが要求したページの URL の履歴からユーザの興味を調べ、グループ内での URL の履歴の情報の中から似たような興味を持つユーザを見つけ、そのユーザと同じような行動をとるものとして、次に見るであろうページを決定するものである。このように、グループ内で情報を共有することになれば、研究グループなどで各メンバーが同じようなことに興味を持っている場合には、個人の経験だけでなくグループ内の他人の経験も生かされる割合が大きく、特に有効である。

#### 4 prefetch 機能の実現

prefetch 機能は、ブラウザとサーバの間に立つ proxy によって実現される (図 1)。proxy を用いたのは、ブラウザで proxy を指定するだけで利用でき、サーバやクライアントの変更などの手間がかからないためである。

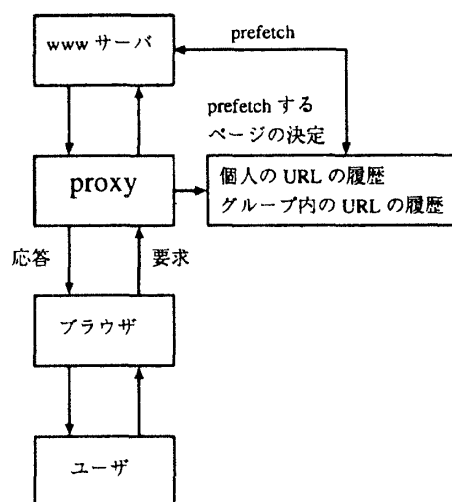


図 1: proxy による prefetch 機能

proxy はブラウザから要求を受け取ってサーバに接続し、サーバからの応答をブラウザに送る。このとき、応答が HTML 文書であればリンク先の URL を取り出し、前述の方法に従ってこの中から prefetch するページを決定する。そして、ユーザがページを見ている回線が空いている時間を利用して決定されたページを取ってきてキャッシュとして保存する。その後に来た要求が、その prefetch したページの URL と一致すれば、サーバに接続することなくキャッシュから読み出せばよく、アクセス時間は大幅に短縮されることになる。

#### 参考文献

- [1] 知念 賢一, 山口 英: 先読みによる WWW アクセスの高速化の可能性, インターネットカンファレンス 96, 1996.7.11.
- [2] Stuart Wachsberg, Thomas Kunz, Johnny Wong: *Fast World-Wide Web Browsing Over Low-Bandwidth Links*,  
URL: <http://ccnga.uwaterloo.ca/~sbwachs/paper.html>
- [3] 斎藤 逸郎, 土肥 浩, 石塚 満: WWW におけるグループ経験の共有を図るメディアエージェントの構築, 情報処理学会第 53 回全国大会講演論文集 2K-7, Sep1996.