

Web アクセス高速化技術の PDA への応用

石丸 大十 庄野 篤司† 佐藤 英昭†

株式会社 東芝 研究開発センター†

1. はじめに

近年、インターネットの普及により、Web 利用者の数が著しく増加している。また検索サイトや EC(Electric Commerce)サイトのよう Web サーバにリクエストが届いた時点で CGI や ASP を利用して動的にコンテンツを作成する技術も多くのサイトで使用されるようになってきている。一方、Web を見る端末も PC に限らず PDA や携帯電話といったモバイル機器に広がっている。モバイル機器は比較的遅い回線を使用していることが多い。データの転送量を減らす手段としてプロキシキャッシュは欠かすことができないが、アクセスの度に生成される動的コンテンツには有効でなかった。そこで、我々は動的コンテンツをキャッシュの対象としてデータ転送量を削減するフィンガープリントキャッシュ[1] (以後 FP キャッシュ)を開発した。本稿では、FP キャッシュを PDA に応用するにあたり、検討すべき問題点について言及する。

2. FP キャッシュ

FP キャッシュは、WEB アクセスを高速化するために開発されたキャッシング技術である。従来のプロキシキャッシュでは、URL とコンテンツを一对一で対応させてキャッシュしている。既にキャッシュされている URL に対するリクエストがあると、その URL に対応するコンテンツをキャッシュから返すことでデータの転送量を削減させている。この方法は、静的コンテンツには非常に有効であるが、URL 対コンテンツが一意に定まらない動的コンテンツには有効でなかった。

一方、FP キャッシュでは URL とコンテンツの対をキャッシュするのではなく、コンテンツを一意に識別するフィンガープリント (以後 FP) と URL の対をキャッシュする。FP とはコンテンツのハッシュ値であり、コンテンツに対して一意に定まる識別子である。

本技術を用いるには、Web サーバ側・Web ブラウザ側にそれぞれ一つずつ FP キャッシュを実装した装置が必要である。Web サーバ側の装置 (装置 1) と Web ブラウザ側の装置 (装置 2) の両方に FP キャッシュを保持し、もし Web サーバから以前取得

したコンテンツが装置 1 に届くとコンテンツの代わりに FP だけを装置 2 に送る。装置 2 は該当 FP と対応付けられているコンテンツを Web ブラウザに返すことで装置 1 ~ 装置 2 間のデータ転送量を削減 (FP 転送) する。また、以前取得したコンテンツと類似したコンテンツの場合は、装置 1 が取得しているコンテンツの FP とそのコンテンツとの差分情報だけを装置 2 に送ることでデータ転送量を削減する (差分転送)。この場合、装置 2 は差分情報と FP からコンテンツを復元して、Web ブラウザにコンテンツを返す。

我々はこれまでに、FP キャッシュの実現モデルとして、Web ブラウザを動作させるクライアントが PC であるという前提で、デュアルプロキシ構成、クライアントモジュール構成、プラグイン構成の開発を行ってきた。現在、FP キャッシュの適用範囲の拡大を目的として、近年普及の著しい PDA 上で動作するクライアントモジュール構成の開発を行っている。

3. FP キャッシュの PDA への適用

3.1 PDA における制限

PDA では PC と比較して以下のような制限があり、FP キャッシュを実装する際の課題となる。

- ・システム全体のパフォーマンスの低さ (貧弱な CPU パワー・実行時メモリ容量の制限)
- ・データメモリ容量 (補助記憶領域) の制限

システムパフォーマンスの制限に関する問題について 3.2 節で、データメモリ容量に関する問題について 3.3 節で述べる。

3.2 システムパフォーマンスによる制限

FP キャッシュを実現するデータ中継転送は、ブラウザとは別スレッドで動作させる必要がある。しかし PDA の CPU は PC の CPU に比べ非常に処理能力に乏しい。このため、クライアントモジュールの処理に CPU パワーの多くを使用するとブラウザやシステム全体の動作が不安定になる恐れがある。また、FP キャッシュ処理を行うためには、データ受信バッファなど、ある程度のメモリ領域を必要とするが、PDA は PC と比べメモリの量も非常に少ない。実行時メモリ容量に制限のある PDA において、メモリの使用量を減らすことも重要となる。実行時メモリでは、動的確保を利用することで瞬間的な最大必要メモリ量を減らすことと、静的確保を利用することで動的確保によるメモリの分散

†Development of Fingerprint-Cache for PDA device

†Dai Ishimaru · Toshiba Corp. Research and Develop Center

†Atsushi Shono · Toshiba Corp. Research and Develop Center

†Hideaki Sato · Toshiba Corp. Research and Develop Center

化や動的確保にかかる時間消費を避けることがトレードオフの関係にあり、効率の良いメモリ管理が必要となる。

Web アクセス高速化という観点から実行時間について FP キャッシュを検討した場合、本技術は装置 1～装置 2 間のデータ転送量を削減することでデータ転送時間を短縮することを特徴とする。しかし一方で、装置 1 および装置 2 においては、FP とコンテンツの対応付けや差分伸長といったデータ処理を行う時間が新たに必要となる。データ処理時間は、PC のようにシステムパフォーマンスの高い機器であれば、短縮されるデータ転送時間に比べると十分に無視できる範囲である。しかし、クライアントマシンに PDA を利用する場合、データ転送の短縮時間に対して装置 2 におけるデータ処理時間が無視できない。PDA 用クライアントモジュールで高速な Web 閲覧を実現するためには高速な処理を行えるアルゴリズムでなければならない。

3.3 キャッシュ容量

FP キャッシュでは、コンテンツを保存するキャッシュが必要となる。PDA はデータメモリ容量（補助記憶領域）の制限がある。PDA での Web ブラウザは Pocket IE (PIE) を想定している。FP キャッシュを PIE のキャッシュ内に実装する場合、PIE ではキャッシュは内部メモリへ保存される為、わずかな領域しか使えない。FP キャッシュを PIE のキャッシュとは別の場所（外部メモリ）へ書き出すことにしても、PC のキャッシュ領域に比べ制限はある。そこで、少ないキャッシュ領域を効率よく使用する仕組みが必要となる。

キャッシュの削除には、参照された日時の古いものから順に削除する LRU (Least Recently Used) アルゴリズムや、利用頻度の低いコンテンツから順に削除する LFU (Least Frequently Used) アルゴリズムなどがある。キャッシュされたコンテンツの総量がキャッシュ可能サイズを超えた場合や何かしらの条件を満たした場合に、上記 LRU や LFU アルゴリズムなどに従ってキャッシュを削除することで、価値のあるコンテンツがキャッシュに残る仕組みとなっている。

通常のキャッシュの場合はキャッシュ容量が小さいとキャッシュヒット率が下がり、転送量が多くなる。一方、FP キャッシュの場合、該当コンテンツがキャッシュから消されてしまっても、直前に取得したコンテンツと似たコンテンツの場合は差分転送によって転送データ量削減とそれに伴うレスポンスタイムの向上効果が得られる。

ここで、差分前のコンテンツサイズ (S) と差分後の差分データサイズ (S') より差分圧縮時の圧縮率を以下のように定義する。

圧縮率 = $100 \times (1 - S'/S)$
圧縮率は圧縮後の差分データサイズ S' が小さく

なるほど高くなる。すなわち、圧縮率が高いコンテンツほど、キャッシュの中に類似度の高いコンテンツが存在していることになる。似たようなコンテンツを数多くキャッシュすることは効率が悪い。コンテンツをキャッシュする時、圧縮率も併せて保存しておく。キャッシュされたコンテンツの総サイズがキャッシュ可能サイズを超えた場合、キャッシュ済みコンテンツの総サイズが予め定められた基準値以下になるまで、圧縮率の高いコンテンツから順に削除をしていく。この方法により、差分圧縮・伸長に対してより効率のよいキャッシュ維持方法が実現できる。

3.4 PDA 用クライアントモジュール

我々は PC 用に開発したクライアントモジュールを基に、PDA で動作するクライアントモジュールの開発を行っている。クライアントモジュール構成はクライアント側にプロキシを設置する必要がないため、比較的柔軟なネットワーク構成を実現できる。PDA 用クライアントモジュール構成を図 1 に表す。

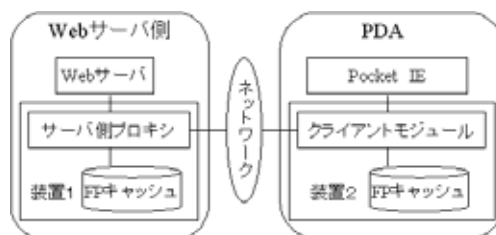


図1 PDAクライアントモジュール構成

装置 1 は FP キャッシュを実装したフォワード型 / リバース型のいずれかのプロキシキャッシュとして Web サーバ側の LAN に設置される。装置 2 はクライアントマシンである PDA 上でアプリケーションとして FP キャッシュモジュールを実装する。

4. おわりに

本稿では、リソースやパフォーマンスが限られている PDA で、FP キャッシュを用いた Web アクセスの高速化について、方法および課題について検討を行った。またキャッシュサイズに制限がある PDA において、差分圧縮における圧縮率を考慮した新しいキャッシュ削除方法により、従来のキャッシュより効率の良いキャッシュができることを示した。さらに PDA におけるクライアントモジュール構成の実装方法についても示した。今後は今回試作した構成において、転送データ削減効果やレスポンスタイム、キャッシュの有効性について定量的に検証する予定である。

参考文献

[1] 吉井謙一郎他, “フィンガープリントキャッシュと動的コンテンツの配信への応用”, 第 4 回インターネットテクノロジーワークショップ