

## プッシュ型放送のためのデータ利用時間と相関性を考慮したキャッシング方式

内田 渉 原 隆浩 西尾 章治郎

大阪大学大学院情報科学研究科マルチメディア工学専攻

{wataru,hara,nishio}@ist.osaka-u.ac.jp

プッシュ型放送を行う情報システムでは、複数のクライアントから発生したデータアクセス要求の全てを一度の放送で満たすことができるため、クライアント数が非常に多い場合に、データアクセスのスループットの向上が期待できる。本稿では、クライアントが相関性をもつ複数のデータアイテムに対して、時間間隔をおいて連続的にアクセス要求を発行するような環境を想定し、データアクセスの平均応答時間短縮のためのキャッシング方式を提案する。提案方式では、データ間の相関性および次のアクセス要求を発行するまでのデータ利用時間を考慮して、応答時間の利得が大きいデータアイテムをあらかじめキャッシュする。さらに本稿では、シミュレーション実験によって、提案方式の有効性を検証する。

## A Caching Strategy for Push-based Broadcast Considering Think-time and Correlation

Wataru UCHIDA Takahiro HARA Shojiro NISHIO

Dept. of Multimedia Eng., Graduate School of Information Science and Technology, Osaka University

{wataru,hara,nishio}@ist.osaka-u.ac.jp

In push-based broadcast information systems, since a server can meet all the requests pending in the system with one broadcast, a larger throughput can be expected even though a large number of clients exist. In this paper, to reduce the response time of data access, we propose new caching strategies assuming an environment where clients consecutively issue access requests for multiple data items with think time. The proposed strategies reduce the average response time by caching data items which have long expected response time according to each client's access characteristics such as correlations among data items and think-time between a data access and the next access request. Moreover, we verify effectiveness of the proposed strategies by simulation experiments.

### 1 はじめに

近年、有線・無線の通信環境の発展に伴い、サーバが放送型通信を用いてクライアントにデータを配送するプッシュ型情報システム（以下では単にプッシュ型情報システムと呼ぶ）に関する研究の関心が高まっている。プッシュ型情報システムでは、クライアントはデータアクセスの際、アクセス要求をサーバに送信せずにサーバの放送帯域を監視し、そのデータが放送さ

れた時点でアクセスを完了する（図1）。クライアントからの要求に応じて個別にデータを配送する1対1のオンデマンド型情報システムとは異なり、プッシュ型情報システムでは、サーバは各クライアントから離散的に発生する要求を一度の放送により満たすことができるため、クライアント数が増加してもシステム全体の負担コストはほとんど変化しない。また、放送内容をクライアントからサーバへ送信されたアクセス要

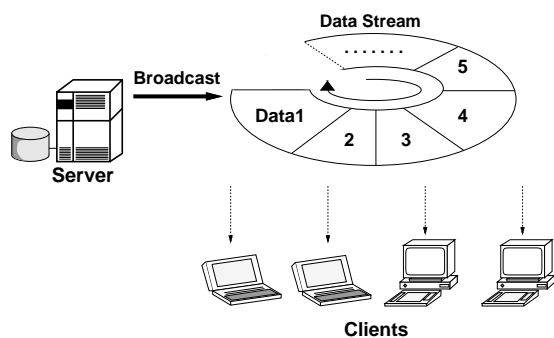


図 1: プッシュ型情報システム

求に基づいて決定する 1 対多のオンデマンド型情報システムと違い、アクセス要求がサーバへ送信されないため、クライアント数が増加した場合の上り帯域での競合も発生しない。したがってクライアント数が非常に多い環境において、データアクセスのスループットの向上が期待できる。

一方、サーバの放送帯域を全てのクライアントが共有するため、多種のデータを放送する必要があり、データアクセスに対するシステム全体の平均応答時間が大きくなる。そこで、プッシュ型情報システムにおける平均応答時間を短縮する方法として、サーバ側でのスケジューリング戦略 [4, 5, 6, 8, 9, 10]、クライアント側でのデータアイテムのキャッシュ戦略 [2, 13, 14]、プッシュ型とオンデマンド型の融合戦略 [3, 7] などがこれまでに提案されている。

これまでに提案されたサーバ側でのスケジューリング戦略の中で最も代表的なものとして、Acharya らによる放送ディスク [1] や Vaidya ら [17]、Su ら [12] が提案した方式がある。これらの方式では、クライアントが頻繁にアクセスするデータを頻繁に放送することによって、平均応答時間を短縮している。また、キャッシュ戦略の中で代表的なものとしては、Acharya らが、各クライアントにおけるデータのアクセス確率や放送予定から、応答時間の利得が大きくなるように効果的なキャッシュの置き換えを行う方法を提案している [2]。

一方、クライアントはあるデータの集合にまとめてアクセスすることが多いといったように、データ間には相関性が存在することが一般的である。クライアントが相関性をもつデータ集合に対して頻繁にアクセス要求するような場合には、スケジューリングおよびキャッシングの際にこれらの相関性を考慮することによって、データアクセスの応答時間を短縮することが可能である。また、相関性をもつデータ集合に対するアクセス

要求は、それらすべてが同時に発行される場合と、アクセス要求が時間間隔において逐次的に発行される場合の二通りが考えられる。

筆者らは、文献 [15]、[20] において、相関性をもつデータ集合に対して同時にアクセス要求が発行されるような環境を想定し、システム全体の応答時間を短縮するためのスケジューリング方式およびキャッシング方式を提案した。一方、後者のアクセス要求が時間間隔において発生する環境におけるスケジューリング方式を、文献 [16]、[18] において提案した。しかし、文献 [16]、[18] において提案したスケジューリング方式では、システム全体のグローバルなアクセス特性に基づいて応答時間の短縮を図っているため、アクセス特性がグローバルなものとは異なるクライアントに対しては、応答時間を短縮することができない。

そこで筆者らは、後者の環境におけるクライアント側のキャッシング方式を文献 [19] において提案した。しかし、文献 [19] では、基本的なアイデアの提案に留まっており、方式の明確な定義や評価は行われていない。そこで本稿では、そのアイデアに基づき、クライアントが相関性をもつ複数のデータアイテムに対して時間間隔において連続的にアクセス要求を発行する環境における、個々のクライアントのアクセス特性を考慮したキャッシング方式の提案および評価を行う。

本稿では、次のようなシステム環境を想定する。

- サーバは 1 つとする。
- サーバで放送されるデータはデータアイテムと呼ばれる単位で放送される。データアイテムは  $M$  種類存在し、1 から  $M$  の識別子を用いて区別する。
- 簡単のため、各データアイテムのサイズは全て等しいものとし、1 アイテムの放送にかかる時間を 1 タイムスロットとする。
- クライアントはキャッシュをもつ。キャッシュに保持しているアイテムに対するアクセス要求の応答時間は 0 とし、保持していないアイテムに対する応答時間は、アクセス要求が発行されてから、そのアイテムの次の放送時刻までの時間とする。
- クライアントは放送プログラムを知っている。これは、例えば、サーバが周期的にプログラム情報を放送することによって実現することができる。
- 各クライアントにおけるデータ間の相関性などのアクセス特性は、各クライアントで異なり、これらは各クライアントにおいて既知とする。
- オンデマンド型配送は行わない。
- データの更新は発生しない。

以下、2章で放送アイテム間の相関性およびアクセス要求の発生に関する本稿の想定について述べる。3章でこれまでに提案されたクライアント側のキャッシング方式について述べ、4章でデータ間の相関性およびアクセス要求発生の時間間隔を考慮したキャッシング方式の提案を行う。5章で提案方式の性能評価を行い、最後に6章で本稿のまとめを行う。

## 2 データアイテム間の相関性とアクセス要求

### 2.1 アイテム間の相関性

一般に、サーバから放送されている様々なデータアイテムは、あるデータアイテム集合がまとめてアクセスされることが多いといったように、互いに相関性をもつことが多い。

例えば、サーバが静的な Web ページを、デジタル放送帯域などを利用して多数のクライアントにプッシュ型配信する場合を考える。つまり、ユーザからの Web ページの要求は能動的に発生するが、クライアントはサーバにアクセス要求を送信せず、そのページが放送されるのを待つ。ここで HTML ファイルや画像ファイルを個別のデータアイテムとして配送する場合、クライアントは、Web ページを構成する複数のデータアイテムに対して同時にアクセス要求を発行する。したがって、これらのアイテムは相関性をもつ。

また、サーバが HTML ファイルや画像ファイルなどの Web ページを構成する要素をまとめて一つのデータアイテムとして、様々な Web ページを放送している場合を考える。クライアントが、ある Web サイトのページを閲覧する場合、まずはトップページにアクセス要求を発行する。トップページを取得すると、クライアントはある時間、そのページを閲覧し、そのページからリンクされる別のページにアクセス要求を発行する。したがって、各ページからリンクされている他のページは、リンク元のページと相関性をもつ。

このように、相関性をもつデータアイテム集合は、同時にアクセス要求される場合と、時間間隔においてアクセス要求される場合の二通りがある。本稿では、後者のように、クライアントが配信されたデータアイテムを利用する時間が、そのデータアイテムの配信に要する時間に比べて充分大きく、相関性をもつ個々のデータアイテムが時間間隔においてアクセス要求される環境を想定する。以下では、この時間間隔をデータ利用時間と呼ぶ。また、本稿では簡単のため、各クラ

アントが相関性をもつデータアイテムに対して同時に発行するアクセス要求は1つまでとする。

相関性の強さは、クライアントがあるデータアイテムにアクセスした後の、他のデータアイテムへの連続したアクセス要求の発生確率として定義することができる。また、データ利用時間は、直前のデータアクセスからの経過時間に基づく確率密度関数によって定義される。

実環境において、各クライアントにおけるアイテム間の相関性や、データ利用時間を示す確率密度関数は、クライアントのアクセス履歴を調べることで決定できる。また、上記の Web ページがデータアイテムとして放送される例のように、各ページ間のリンク関係や、コンテンツの内容などで解析的にアイテム間の相関性及び確率密度関数を決定できる。

### 2.2 連鎖的なアクセス要求

前節の例のように、様々な Web ページが放送されている場合、クライアントはあるページを閲覧した後、そのページと相関性をもつ別のページにアクセスし、さらに別のページにアクセスするといったように、相関性をもつデータアイテムに連鎖的にアクセス要求を発行する。

本稿では、クライアントが、一つ目のデータアイテムに対して、ある確率で無記憶性のアクセス要求を発行した後、それと相関性をもつデータアイテムに、ある確率密度関数に従ったデータ利用時間において連鎖的にアクセスを複数回行う環境を想定する。その一連のアクセスをプロセスと呼び、各プロセスにおいて始めにアクセスされるアイテムを先頭アイテムと呼ぶ。

## 3 従来のキャッシング方式

プッシュ型情報システムの性能向上を目指す研究の一環として、これまでに様々なキャッシング方式が提案されている。これらのうちで、プリフェッチ型キャッシング方式では、重要度が高いと判断したアイテムをあらかじめキャッシュに格納(プリフェッチ)しておく。フィルタリングの手間以外のコストの増加を伴うことなく、重要なアイテムに対するヒット率を高め、平均応答時間を短縮できることから、プッシュ型通信に適したキャッシング方式と考えられている。代表的なプリフェッチ型キャッシング方式として、文献 [2] において PT 法と呼ばれる方式が提案されている。この方式では、データアイテム間の相関性は考慮せず、すべて

のアクセス要求が無記憶に発生する環境を想定している．また、筆者らは文献 [20] において、クライアントが相関性のあるアイテム集合に対して同時にアクセスを要求することを想定して、PT 法を拡張した CB-PT (Correlation-Based PT) 法を提案している．

以下では、PT 法および CB-PT 法の概要について説明する．

PT 法:

1. 各アイテムの放送開始時に、キャッシュ内のアイテムおよび放送されるアイテムに PT 値と呼ばれる値を与える．アイテム  $i$  に与える PT 値  $L_i$  は次式で表される．

$$L_i = p_i \cdot (u_i(t) - t) \quad (1)$$

ただし、 $p_i$  はアイテム  $i$  に対するそのクライアントのアクセス確率、 $t$  は現在時刻、 $u_i(t)$  は時刻  $t$  におけるアイテム  $i$  の次回の放送時刻とする．

2. 放送されるアイテム  $i$  の PT 値  $L_i$  が、キャッシュ内で PT 値が最小となるアイテム  $j$  の PT 値  $L_j$  より大きい場合、アイテム  $i$  と  $j$  を置き換える．キャッシュ内で PT 値が最小となるアイテムが複数存在する場合は、それらの中から無作為に一つのデータアイテムを選択する．

PT 値は、あるアイテムに対してアクセス要求を発行したとき、そのアイテムをキャッシュに保持していない場合に生じる待ち時間を示している．すなわち、PT 法では、各アイテムの放送開始時に、各々のアイテムをキャッシュから追い出すことにより増加する待ち時間を比較し、最も応答時間の利得が大きいようにキャッシュを管理する．ここで、PT 法におけるデータアイテムの置き換えは長期的に見た場合に必ずしも最適とはならない．また、データアイテム間の相関性を考慮した場合、各アイテムのアクセス確率は時間的に一定とはならないため、正確に応答時間の利得を計算することができない．

CB-PT 法:

1. 各アイテムの放送開始時に、キャッシュ内および放送中のアイテムに対して、PT 値を相関性を考慮するように拡張した CB-PT 値を与える．アイテム  $i$  に与える CB-PT 値  $G_i$  は、次式で表される．

$$G_i = (u_i(t) - t) \sum_{k \in C} p_{ik} + \sum_{k \in Q_i} (u_i(t) - u_k(t)) p_{ik} \quad (2)$$

ここで、 $p_{ik}$  はクライアントがアイテム  $i$  と  $k$  に一括してアクセス要求を起こす確率とする．また、 $C$  は放送データのうちキャッシュ内にあるアイテムの集合とし、 $Q_i$  は現時点からアイテム  $i$  の次回の放送時間までに放送され、かつ、 $C$  に含まれないアイテムの集合とする．

2. 放送されるアイテム  $i$  の CB-PT 値  $G_i$  がキャッシュ内で CB-PT 値が最小となるアイテム  $j$  の CB-PT 値  $G_j$  より大きい場合、アイテム  $i$  と  $j$  を置き換える．キャッシュ内で CB-PT 値が最小となるアイテムが複数存在する場合は、それらの中から無作為に一つのデータアイテムを選択する．

式 (2) は、アイテム  $i$  と放送プログラム内の任意のアイテムの 2 つに同時にアクセス要求を起こしたとき、アイテム  $i$  をキャッシュに格納していなかった場合に増加する待ち時間を示す．したがって、式 (2) の値が大きなアイテムをキャッシュに残す置き換えは、相関性をもつアイテムに対して同時にアクセス要求を発行する環境において、その瞬間に生じるアクセス要求の平均応答時間を最短にする．しかし、CB-PT 法における置き換えも、PT 法と同様に、長期的に見た場合に必ずしも最適とはならない．また、相関性をもつアイテムに対するアクセス要求が時間間隔において発生する場合は、正確に応答時間の利得を計算できない．

## 4 アクセス要求発生 の時間間隔を考慮したキャッシング方式

本稿では、文献 [19] のアイデアに基づいて、クライアントが相関性をもつアイテムに対して、データ利用時間において逐次的にアクセス要求を発行する環境を想定したキャッシング方式を提案する．提案方式を RIB-PT (Request Interval Based PT) 法と呼ぶ．

クライアントがアイテム  $i$  にアクセスした後にアイテム  $j$  にアクセス要求を発行する確率を  $c_{ij}$  とする ( $\sum_{j=1}^M c_{ij} = 1$ )．アイテム  $i$  のアクセス後にアイテム  $j$  にアクセスするといった条件の下で、データ利用時間  $t$  で  $j$  へのアクセス要求を発行する確率密度関数を  $f_{ij}(t)$  と表す ( $\int_0^\infty f_{ij}(t) dt = 1$ ,  $f_{ij}(t) = 0$  ( $t < 0$ ))．

現在時刻  $\tau$  において、クライアントが最近にアクセス要求を発行したアイテムを  $i$  とし、アクセス要求の発行時点でキャッシュ内にアイテム  $i$  が存在せず、クライアントが次回の  $i$  の放送を待っているものとする．このとき、次回の  $i$  の放送時刻を  $\zeta$  ( $\tau < \zeta$ ) とすると、アイテム  $i$  のアクセス後にアイテム  $j$  へのアクセ

ス要求が発行されるといった条件の下で、時刻  $t$  に  $j$  に対するアクセス要求が発行される確率密度関数は、 $f_{ij}(t - \zeta)$  となる。

次に、最近のアクセス要求が既にアイテム  $i$  の放送もしくはキャッシュアクセスによって時刻  $\zeta$  に満たされており ( $\zeta \leq \tau$ )、 $i$  へのアクセス後、一つのアクセス要求も発生していないものと仮定する。このとき、アイテム  $i$  のアクセス後にアイテム  $j$  のアクセス要求が発生するといった条件の下で、時刻  $t$  に  $j$  に対するアクセス要求が発生する確率密度関数は次式のように表される。

$$\frac{1}{1 - \int_{\zeta}^{\tau} f_{ij}(t - \zeta) dt} f_{ij}(t - \zeta) \quad (3)$$

時刻  $\tau$  におけるアイテム  $j$  の次回の放送時刻を  $u_j(\tau)$  とすると、 $t$  ( $\tau \leq t \leq u_j(\tau)$ ) に発生する  $j$  へのアクセス要求の応答時間は、 $j$  がキャッシュにない場合、 $u_j(\tau) - t$  となる。したがって、クライアントの最近のアクセス要求がアイテム  $i$  に対するもので、時刻  $u_j(\tau)$  にアイテム  $j$  が放送される場合、 $j$  をキャッシュに格納しないことで生じる、次回のアクセス要求に対する平均応答時間の増加分 (期待値)  $R_j$  は次式のように表される。この値を、アイテム  $j$  の RIB-PT 値と呼ぶ。

$$R_j = S_j \cdot \int_{\tau}^{u_j(\tau)} f_{ij}(t - \zeta) \cdot (u_j(\tau) - t) dt \quad (4)$$

ただし、 $S_j$  は次の式で表されるものとする。

$$S_j = \begin{cases} c_{ij} & (\tau < \zeta) \\ \frac{c_{ij}}{1 - \int_{\zeta}^{\tau} f_{ij}(t - \zeta) dt} & (\zeta \leq \tau) \end{cases} \quad (5)$$

ここで、次に  $j$  をキャッシュする機会である次回の放送時刻以前を考慮するため、積分の範囲は現在時刻  $\tau$  から、 $j$  の次回の放送時刻  $u_j(\tau)$  までとなる。

同様の式により、キャッシュ内のアイテムの RIB-PT 値を定義できる。ただし、この場合は、RIB-PT 値は時刻  $\tau$  にそのアイテムをキャッシュから削除することによって生じるアクセスの平均応答時間の増加分である。

提案する RIB-PT 法では、次のような手順でキャッシュの置き換えを行う。

**RIB-PT 法：**

1. 各アイテムの放送開始時に、キャッシュ内および放送中のアイテムに対して、式 (4) で表される RIB-PT 値を計算する。
2. 放送されるアイテム  $i$  の RIB-PT 値  $R_i$  がキャッシュ内で RIB-PT 値が最小となるアイテム  $j$  の

RIB-PT 値  $R_j$  より大きい場合、アイテム  $i$  と  $j$  を置き換える。キャッシュ内で RIB-PT 値が最小となるアイテムが複数存在する場合は、それらの中から無作為に一つのデータアイテムを選択する。

このように、RIB-PT 法では、最近にアクセス要求を発行したアイテムのアクセス (予定) 時刻に基づいて、そのアイテムと相関性をもつデータアイテムに対して次回に発行されるアクセス要求の応答時間を最短にするようにキャッシュの置き換えを行う。

## 5 性能評価

本章では、提案方式のシミュレーション実験による性能評価を行う。評価では、サーバが 1 つの Web ページを 1 データアイテムとして、様々な Web ページを放送する環境を想定する。サーバが放送する内容は、主に 5 つのサイトからなり、各 Web ページは、同じサイトの Web ページにリンクしており、クライアントは同じサイトの Web ページを続けて要求する可能性が高いものとする。また、クライアントは低い確率で他サイトのページを要求することもあるが、どのページを要求するかを予想するのは困難とする。さらに、各サイトの規模は等しいが、クライアントが頻繁にアクセスする 1 つのサイトのみ、プロセスの先頭においてアクセスされる確率が高いものとする。

実環境では、個々のクライアントのデータアクセスのうち 80% が、放送されている全アイテムの 20% という限られたアイテムに対するものであるという事実が文献 [11] において報告されている。そこで、このような事実に基づいて、想定環境を表現するパラメータ設定を以下のように行った。

まずデータアイテム数を 500 とし、全データアイテムをアイテム数 100 の 5 つのグループ (グループ  $G_1, G_2, \dots, G_5$ ) に分割した。 $c_{ij}$  (データアイテム間の相関性の強さ) は、乱数を用いて与え、 $i, j \in G_k, (1 \leq k \leq 5)$  の場合は 10%、 $i \in G_k, j \in G_l, (1 \leq k, l \leq 5, k \neq l)$  の場合は 40% の確率で  $c_{ij} > 0$  とした。つまり、各データアイテムは同じグループのデータアイテムとはそれぞれ 10%、別のグループのデータアイテムとはそれぞれ 40% の確率で相関性をもつ。ここで、それぞれの  $c_{ij}$  の値は、グループ内、グループ間でそれぞれ等しいものとし、 $\sum_{i,j \in G_k} c_{ij} = 0.8$ 、 $\sum_{i \in G_k, j \in G_l, k \neq l} c_{ij} = 0.2$  とした。このように  $c_{ij}$  を与えることによって、クライアントはあるデータアイテムにアクセスを行った後、0.8 の確率で、同じグルー

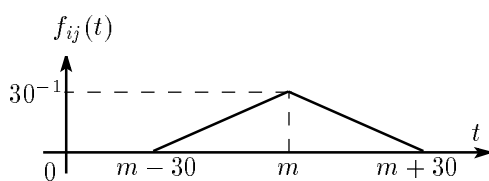


図 2:  $f_{ij}(t)$  .

プの約 10 個のデータアイテムの中の 1 つに対してアクセス要求を発行し, 0.2 の確率で, 異なるグループの約 160 個のデータアイテムの中の 1 つに対してアクセス要求を発行する環境を表している .

さらに, プロセス発生時の先頭アイテムがデータアイテム  $i$  である確率  $q_i$  は, グループ 1 に属するデータアイテムのみ高いものとし, 他のグループに属するデータアイテムの 16 倍とした . タイムスロット毎のプロセス発生確率は 0.01, あるアクセスを終了した時点でプロセスを終了する確率も 0.01 とした .

各アイテム間  $i-j$  において, 要求発生の時間間隔 (データ利用時間) の確率密度関数を, 次式で与えた .

$$f_{ij}(t) = \begin{cases} 30^{-2}(t-m) + 30^{-1} & (m-30 \leq t < m) \\ -30^{-2}(t-m) + 30^{-1} & (m \leq t \leq m+30) \\ 0 & (t < m-30, t > m+30) \end{cases} \quad (6)$$

これは,  $t-f_{ij}(t)$  平面上では, 図 2 に示すように,  $(m-30, 0), (m, 30^{-1}), (m+30, 0)$  を結ぶ突起型となる . 評価では,  $i-j$  間のデータ利用時間の平均値を表す  $m$  を,  $(\mu-10, \mu+10)$  の区間の一様分布で与えた .

サーバは, 全てのアイテムを一度ずつランダムに配置したプログラムを, 周期的に放送するものとした . 評価では, PT 法および CB-PT 法を比較対象とした . PT 法におけるアクセス確率  $p_i$  は,  $i$  が先頭アイテムである確率  $q_i$  で与えた . また, CB-PT 法における相関性  $p_{ij}$  は,  $p_{ij} = c_{ij}/M$  として与えた .

以上のようなシステム環境において, 3,000,000 タイムスロットのシミュレーション実験によって, 提案方式の性能評価を行った . 以下では, 各シミュレーション実験の結果を示し, 考察を行う .

### 5.1 データ利用時間の平均値の影響

キャッシュに格納可能なデータアイテム数 (キャッシュサイズ, 以下では  $|C|$  と表記する) を 100 とし, 平均データ利用時間  $m$  の平均値  $\mu$  を 50 から 400 まで

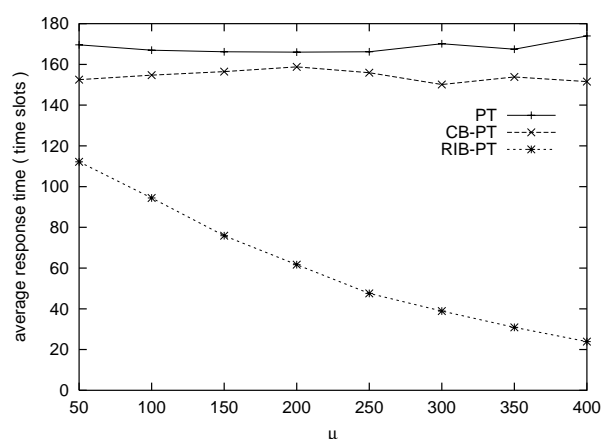


図 3:  $\mu$  と平均応答時間 ( $|C| = 100$ ) .

変化させた場合の平均応答時間を図 3 に示す . 図 3 より, 相関性およびデータ利用時間を考慮してキャッシュの置き換えを行う提案方式である RIB-PT 法が, PT 法および CB-PT 法と比較して良い性能を示していることがわかる . また,  $\mu$  が大きくなるほど, RIB-PT 法と他の 2 方式の差も大きくなっている . これは, 平均データ利用時間が長い方が, クライアントが最近のアクセス要求を発行してから, 次回のアクセス要求時刻までに放送されるアイテム数が多く, 効果的なデータアイテムをキャッシュする機会が多くなるためである .

### 5.2 キャッシュサイズの影響

$\mu = 50, 150, 400$  のそれぞれの場合について, キャッシュサイズ  $|C|$  を変化させたときの平均応答時間を図 4, 図 5, 図 6 に示す . 前節の結果と同様に, RIB-PT 法の性能は, データ利用時間の平均値が大きいほど良くなっており,  $\mu = 400$  である図 6 では常に三方式の中で最も良い性能を示している . 一方, データ利用時間が短い図 4, 図 5 においては,  $|C|$  の値が小さい場合にのみ, 三方式の中で最も良い性能を示している .

データ利用時間が短い場合, クライアントがキャッシュ内のデータアイテムにアクセスを行うと, 最近のアクセス要求の発行時刻と, 次回のアクセス要求の発行時刻の間隔が小さくなる .  $|C|$  が大きい場合は, キャッシュのヒット率が高いため, 次回にアクセス要求を発行する可能性の高いアイテムをあらかじめキャッシュしておく機会が少なく, RIB-PT 法によってあまり応答時間を短縮することができない . 一方,  $|C|$  が小さい場合, アクセス要求されているキャッシュ内にはデータアイテムの次回の放送時刻までに, 次回にアク

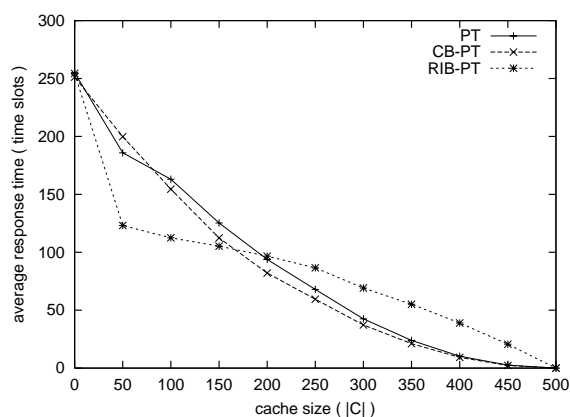


図 4: キャッシュ可能データアイテム数  $|C|$  と平均応答時間 ( $\mu = 50$ ) .

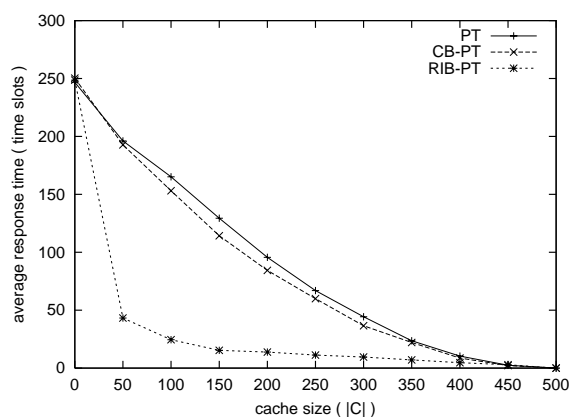


図 6: キャッシュ可能データアイテム数  $|C|$  と平均応答時間 ( $\mu = 400$ ) .

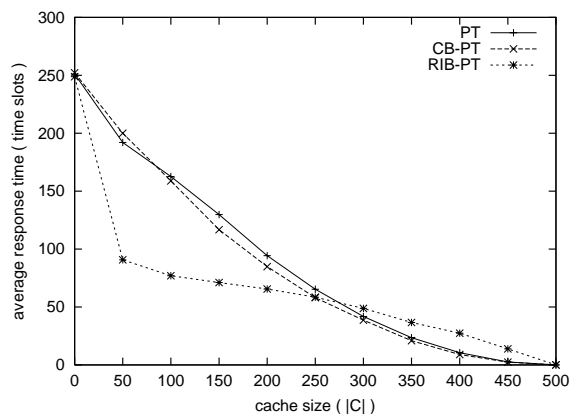


図 5: キャッシュ可能データアイテム数  $|C|$  と平均応答時間 ( $\mu = 150$ ) .

セス要求する確率の高いデータアイテムをキャッシュする機会が多いため、効率的にキャッシュを管理することができる。

## 6 おわりに

本稿では、クライアントが相関性のあるデータアイテムのアクセス要求を、データ利用時間において連続的に発行するプッシュ型情報システムを想定して、応答時間短縮のためのキャッシング方式を提案した。提案方式では、データ利用時間およびデータアイテム間の相関性に基づいて、次回に発行するアクセス要求に対する応答時間の利得を計算し、最も利得が大きいようにキャッシュを置き換える。

さらに本稿では、シミュレーション実験により、提案方式の性能評価を行った。性能評価の結果、データ利用時間が長い場合、およびキャッシュサイズが小さい場合には、データ利用時間を考慮しない従来方式と比較して、提案方式が応答時間を大幅に短縮することを確認した。提案方式は、キャッシュサイズやデータ利用時間などのシステム特性に応じて、異なる性能を示すため、実環境ではそれらを考慮して利用する必要がある。なお本稿では、特定のアクセス特性を想定して評価を行ったが、今後はデータ利用時間や相関性などについて、さまざまな環境を想定した評価を行う予定である。

また、提案方式は次回に発行されるアクセス要求に対する応答時間のみを考慮するため、データ利用時間が短い、キャッシュサイズが大きいなどの理由により、アクセス要求の発行間隔が小さい場合には性能が劣化する。今後は次回に発行されるアクセス要求のみでなく、次々回以降に発行されるアクセス要求に対する応答時間も考慮するように、提案方式を拡張する予定である。

## 謝辞

本研究は、文部科学省特定領域研究(14019063)および文部科学省科学技術振興調整費「モバイル環境向P2P型情報共有基盤の確立」の研究助成によるものである。ここに記して謝意を表す。

## 参考文献

- [1] Acharya, S., Alonso, R., Franklin, M., and Zdonik, S.: Broadcast Disks: Data Management for Asymmetric Communication Environments, Proc. ACM SIGMOD'95, pp. 199–210 (1995) .
- [2] Acharya, S., Franklin, M., and Zdonik, S.: Prefetching from a Broadcast Disk, Proc. Int'l Conf. on Database Engineering, pp. 276–285 (1996) .
- [3] Acharya, S., Franklin, M., and Zdonik, S.: Balancing Push and Pull for Data Broadcast, Proc. ACM SIGMOD'97, pp. 183–194 (1997) .
- [4] 青野正宏, 田窪昭夫, 渡辺尚, 水野忠則: データ放送におけるスケジュール決定法「二重循環法」の提案と評価, 情報処理学会論文誌, Vol. 40, No. 3, pp. 1267–1275 (1999) .
- [5] 青野正宏, 田窪昭夫, 渡辺尚, 水野忠則: RAID 型放送システムの提案, 情報処理学会論文誌, Vol. 40, No. 7, pp. 3043–3050 (1999) .
- [6] Erçetin, Ö. and Tassiulas, L.: Push-Based Information Delivery in Two Stage Satellite-Terrestrial Wireless Systems, IEEE Transactions on Computers, Vol. 50, No. 5, pp. 506–518 (2001) .
- [7] 箱守聰, 田辺雅則, 石川裕治, 井上潮: 放送型通信 / オンデマンド型通信を統合した情報提供システム, 情処研報, Vol. 34, No. 8, pp. 55–60 (1997) .
- [8] Hameed, S. and Vaidya, N. H.: Efficient Algorithms for Scheduling Data Broadcast, ACM-Baltzer Journal of Wireless Networks, Vol. 5, No. 3, pp. 183–193 (1999) .
- [9] 石川裕治, 田辺雅則, 箱守聰, 井上潮: HTML 文書間のデータ共有を考慮した放送型情報提供方式, 情報処理学会論文誌, Vol. 40, No. 7, pp. 3051–3062 (1999) .
- [10] カン ギョウビ, 浅田 一繁, 飯沢 篤志, 古瀬 一隆: 2 次元的な放送モデルにおける配信間隔と配信スケジュールリング, 情報処理学会論文誌: データベース, Vol. 42, No. SIG 10(TOD 11), pp. 54–63 (2001) .
- [11] Lin, L. and Xingming, Z.: Heuristic MultiDisk Scheduling for Data Broadcasting, Proc. Int'l Workshop on Satellite-Based Information Services (WOSBIS'97), pp. 1–5 (1997) .
- [12] Su, C. J., Tassiulas, L., and Tsotras, V. J.: Broadcast Scheduling for Information Distribution, ACM-Baltzer Journal of Wireless Networks, Vol. 5, No. 2, pp. 137–147 (1999) .
- [13] Su, C. J. and Tassiulas, L.: Joint Broadcast Scheduling and User's Cache Management for Efficient Information Delivery, ACM-Baltzer Journal of Wireless Networks, Vol. 6, No. 4, pp. 279–288 (2000) .
- [14] Tassiulas, L. and Su, C. J.: Optimal Memory Management Strategies For a Mobile User in a Broadcast Data Delivery System, IEEE Journal on Selected Areas in Communications, Vol. 15, No. 7, pp. 1226–1238 (1997) .
- [15] 内田渉, 原隆浩, 塚本昌彦, 矢島悦子, 西尾章治郎: データ間の相関性とアクセス頻度を考慮した放送スケジュールリング, 情報処理学会論文誌: データベース, Vol. 43, No. SIG 2 (TOD 13), pp. 146–157 (2002) .
- [16] 内田渉, 原隆浩, 西尾章治郎: アクセス要求発生頻度の時間的変化を考慮した相関データの放送スケジュールリング, 情報処理学会論文誌: データベース, Vol. 43, No. SIG 9 (TOD 15), pp. 28–38 (2002) .
- [17] Vaidya, N. H. and Hameed, S.: Scheduling Data Broadcast in Asymmetric Communication Environments, ACM-Baltzer Journal of Wireless Networks, Vol. 5, No. 3, pp. 171–182 (1999) .
- [18] 矢島悦子, 原隆浩, 塚本昌彦, 西尾章治郎: 相関性をもつデータ間の放送時間間隔について, 情報処理学会論文誌, Vol. 40, No. 1, pp. 188–196 (1999) .
- [19] 矢島悦子, 原隆浩, 塚本昌彦, 西尾章治郎: アクセス要求発生時の時間間隔を考慮した放送データのキャッシング方式, 情報処理学会第 119 回データベースシステム研究会報告 ( 99-DBS-119 ), pp. 1–6 (1999) .
- [20] 矢島悦子, 原隆浩, 塚本昌彦, 西尾章治郎: データ間の相関性を考慮した放送データのスケジュールリング法およびキャッシング法, 情報処理学会論文誌, Vol. 40, No. 9, pp. 3577–3585 (1999) .