

プッシュ型放送における Peer-to-Peer ネットワークを用いた プリフェッチ型協調キャッシング方式について

前田 和彦[†] 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 〒565-0871 吹田市山田丘 2-1

E-mail: †{maeda,wataru,hara,nishio}@ist.osaka-u.ac.jp

あらまし 近年、有線および無線の通信技術の発展に伴い、放送型通信を用いてデータを配送するプッシュ型放送に関する研究の関心が高まっている。本稿では、プッシュ型放送における複数のクライアントが接続することで構成される Peer-to-Peer ネットワークを利用して、協調的にキャッシングを行う方式を提案する。提案方式では、各ピアが、自ピアのアクセス確率と、他ピアが発行するクエリ情報に基づいて、Peer-to-Peer ネットワーク全体の応答時間の利得が大きいデータアイテムを各時点でプリフェッチする。さらに、本稿では、シミュレーション評価によって、提案方式の有効性を検証する。

キーワード プッシュ型放送, Peer-to-Peer ネットワーク, 協調キャッシング

On a Prefetch-type Collaborative Caching Strategy in a Peer-to-Peer Network for Push-based Broadcast

Kazuhiko MAEDA[†], Wataru UCHIDA[†], Takahiro HARA[†], and Shojiro NISHIO[†]

[†] Dept. of Multimedia Eng., Graduate School of Information Science and Technology, Osaka University
2-1 Yamadaoka, Suita, Osaka 565-0871, Japan

E-mail: †{maeda,wataru,hara,nishio}@ist.osaka-u.ac.jp

Abstract Recently, there has been an increasing interest in research of push-based broadcast that delivers data using broadcast in both wired and wireless environments. In this paper, we propose a new collaborative caching strategy by clients in a push-based broadcast environment where clients construct a peer-to-peer network by connecting with other clients. In the proposed strategy, a peer takes into account its own access probabilities and information on queries issued by other peers, and prefetches data items with large benefit of the response time at each time. Moreover, we verify the effectiveness of the proposed strategy by simulation experiments.

Key words push-based broadcast, peer-to-peer network, collaborative caching

1. はじめに

近年、放送型通信を用いてデータを配送するプッシュ型放送に関する研究の関心が高まっている。プッシュ型放送では、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へ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々のデータを放送するプル型情報システムとは異なり、プッシュ型放送では、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わらない。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。そのためプッシュ型放送は、動画や音声、ニュースやヒットチャートなどの公共性の高い情報を、衛星や地上波の放送帯域を用いて多数のユーザに配信するサービスなどに適用される。

しかし、プッシュ型放送では所望のデータが放送されるまで待たなければならないため、一般にデータアクセスの応答時間が長いという問題がある。そこで、プッシュ型放送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[17],[18],[20]や、クライアント側でのデータのキャッシング戦略[1],[2],[10],[19]などがこれまでに提案されている。

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、Vaidyaら[20], Suら[17]が提案した方式がある。この方式では、クライアントが頻繁にアクセスするデータの放送頻度を高くすることで、平均応答時間を短縮している。キャッ

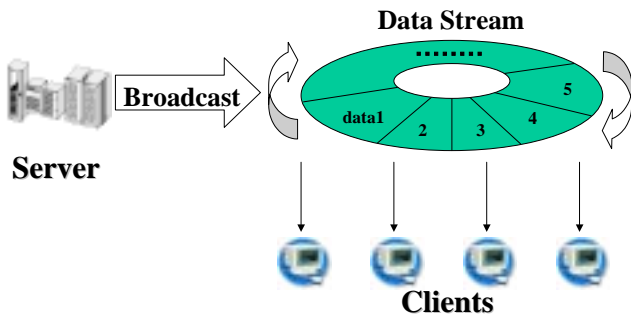


図 1 プッシュ型放送 .

Fig.1 Push-based broadcast.

シング戦略では, Acharya らが提案した PIX 法 [1] や PT 法 [2] などが著名である. これらの方式はクライアントの各データに対するアクセス確率と, 放送周期や次回の放送時刻までの時間から, キャッシュした場合の応答時間の利得が大きくなるデータをキャッシュする. これらのプッシュ型放送に関する研究は, クライアントが放送帯域のみからデータを取得する環境を想定している.

一方, インターネットにおける情報共有のための新たなネットワークモデルとして Peer-to-Peer (P2P) 型情報共有サービスが開始されており [8], [9], [15], これに関する研究も盛んに行われている [4], [6], [12], [13], [16]. P2P 型情報共有では, 各端末をピアと呼び, 各ピアは複数の別のピアに接続することで, 物理ネットワークとは独立した, ピア同士で接続された論理ネットワーク (P2P ネットワーク) を構成する. ピアはデータアクセスの際, 論理ネットワーク内の隣接ピアへアクセス要求に関する情報 (クエリ) を送信する. データを保持しない隣接ピアは, 再びそのピアの隣接ピアへクエリを伝播する. このようにして伝播されたクエリが要求データを保持するピアへ到達すると, 要求を行ったピアとデータを保持するピアの間でデータの受け渡しが行われ, データアクセスが完了する. このように, P2P ネットワーク上の各ピアが, 自律分散的にデータの送受信を行うことによって, スケーラビリティの高いデータ共有が行われる.

プッシュ型放送において, 各クライアントがキャッシュするデータを, P2P ネットワークを用いて共有することで, データアクセスの平均応答時間をさらに短縮できるものと考えられる. 筆者らの研究グループでは, プッシュ型放送のサービスを受ける複数のピアが P2P ネットワークを構成する環境を想定し, 協調的にキャッシングを行う方式を提案している [14]. 以下では, この方式を C-PIX (Collaborative PIX) 法と呼ぶ. C-PIX 法は, 自ピアのアクセス確率に加えて, 他ピアから発行されたクエリも考慮して P2P ネットワーク全体の応答時間の利得が大きくなるようにキャッシュの置き換えを行うことで, ネットワーク全体の平均応答時間の短縮を図っている.

プッシュ型放送では, サーバが 1 つのデータアイテムを放送するたびに, 各データアイテムの次回の放送時刻が変化する. C-PIX 法はキャッシュの置き換えの際に, データアイテムごとの放送間隔という静的なパラメータを用いて応答時間の利得を

計算しているため, 必ずしも各時点において最適なデータアイテムをキャッシュしているとは限らない. そこで本稿では, 応答時間の利得の計算において, 各データアイテムが次回に放送されるまでの時間を用いることで, 各時点において応答時間の利得が大きいアイテムをプリフェッチするキャッシング方式を提案する.

以下では, 2. でプッシュ型放送における従来のキャッシング方式について説明し, 3. で本稿の想定環境について述べ, 4. で本稿で提案するプリフェッチ型協調キャッシング方式について説明する. 5. で提案方式の性能評価を行い, 6. で関連研究について述べる. 最後に 7. で本稿のまとめを述べる.

2. 従来のキャッシング方式

プッシュ型放送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. 本章ではこれらの内, 平均応答時間の短縮を考慮したキャッシング方式である, PIX 法 [1] および PT 法 [2] について説明する. これらの方式は, 各クライアントにおいて重要度 (応答時間の利得) が高いと判断したデータをあらかじめキャッシュに格納しておく.

PIX 法や PT 法では, クライアントは自身のキャッシュ, もしくはサーバの放送帯域のみからデータを取得する環境を想定している. ここで, データはデータアイテムと呼ばれる単位にまとめられ, 各データアイテムはサーバから周期的に放送されているものとする. この環境では, クライアントはあるデータアイテムに対してアクセス要求を発行する際, まず自らのキャッシュに蓄積されたデータアイテムを確認する. もしキャッシュ内にそのデータアイテムが存在すれば, クライアントは即座にそのアイテムにアクセスすることが可能であり, それ以外の場合はそのアイテムの次の放送時刻までの時間が応答時間となる.

2.1 PIX 法

PIX 法の手順を下記に示す.

(1) 受信開始時に, 全てのデータアイテム $j (1 \leq j \leq M)$ に次式で表される評価関数 $K(j)$ を与える.

$$K(j) = p_j \cdot y_j \quad (1)$$

ただし, M はサーバが放送している総データアイテム数, p_j はデータアイテム j に対するアクセス確率, y_j はデータアイテム j の放送周期である.

(2) すべてのデータアイテムの中で, $K(j)$ の値が高いものをキャッシュ可能な個数だけキャッシュする.

PIX 法は, クライアントのアクセス確率が高く, 放送周期が大きいデータアイテムをキャッシュすることで, 頻繁にアクセスするが, あまり放送されないデータアイテムの応答時間の短縮を図る. また, クライアントのアクセス確率や各アイテムの放送周期が変化しない限り, 一旦必要なデータアイテムをキャッシュしてしまえば, キャッシュの置き換えは起こらない. キャッシュアイテムの置き換えによる負荷が他手法に比べて少ないため, モバイル端末のように限られた電力しか保持しない端末をクライアントと想定した環境においては, 電力消費を少なくする有効なキャッシング方式である.

2.2 PT 法

PT 法の手順を下記に示す。

(1) 各データアイテムの放送開始時に、キャッシュ内のデータアイテムおよび放送されるデータアイテムに PT 値と呼ばれる値を与える。データアイテム j ($1 \leq j \leq M$) に与える PT 値は次式で表される。

$$L_j = p_j \cdot (u_j(Q) - Q) \quad (2)$$

ただし、 $u_j(Q)$ は時刻 Q におけるデータアイテム j の次回の放送時刻とする。

(2) 放送されるデータアイテム k の PT 値 L_k が、キャッシュ内で PT 値が最小となるデータアイテム m の PT 値 L_m より大きい場合、データアイテム k と m を置き換える。キャッシュ内で PT 値が最小となるデータ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 値はあるデータアイテムに対してアクセス要求を発行したとき、そのデータ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の期待値を示している。すなわち、PT 法では、各データアイテムの放送開始時に、各々のデータ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データアイテムをプリフェッチする。PT 法は、各時点での放送予定を見て、あらかじめ応答時間の大きいデータアイテムをキャッシュするため、放送間隔だけを見て固定のデータアイテムをキャッシュする PIX 法よりもデータアクセスの平均応答時間を短縮できる。

3. 想定環境

本稿における想定環境の概念図を図 2 に示す。想定環境において、ピア同士は P2P ネットワークを利用して互いにデータアイテムの受け渡しが可能とする。各ピアはキャッシュに有限個数のデータアイテムを蓄えることが可能であり、自ピアのキャッシュへのアクセス、放送帯域からの受信、他ピアのキャッシュからの P2P ネットワークを利用した受信のうち、要求するデータアイテムへのアクセスが可能であるものの中で最も応答時間が短い方法でデータアクセスを行うものとする。

各ピアはデータアクセスの際、アクセス要求を行うデータアイテムを自ピアのキャッシュに保持していれば、キャッシュ内のデータアイテムにアクセスを行う。ここで、キャッシュに保持しているデータアイテムに対するアクセス要求の応答時間は 0 とする。アクセス要求を行うデータアイテムを、自ピアのキャッシュに保持していない場合は、そのアイテムの次回の放送時刻までの時間と P2P ネットワーク内の他ピアからデータアイテムの受信を完了するまでの時間を比較する。もしも、次回の放送時刻までの時間の方が短ければ、そのアイテムが放送されるまで待ち、放送帯域からデータアイテムの取得を行う。

他ピアからデータアイテムの受信を完了するまでの時間の方が短い場合は、フラッディング [13] により、周囲のピアに要求するデータアイテムを保持するものが存在するか否かを問い合わせる。フラッディングでは、ある TTL (Time to Live) 値を

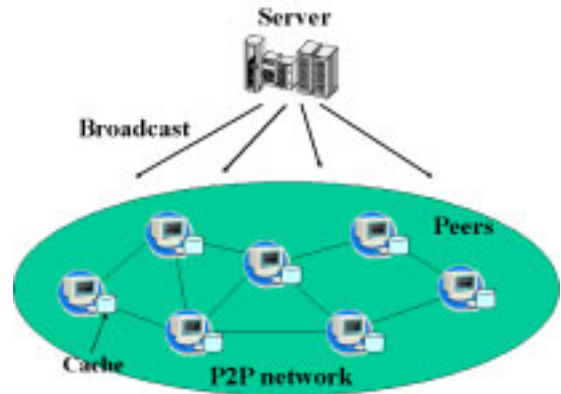


図 2 想定環境。

Fig. 2 Assumed environment.

定めたクエリを発行し、P2P ネットワーク内で隣接する全てのピアにブロードキャストする。クエリを受け取った各ピアは、自ピアのキャッシュ内に要求されたアイテムがなければ、クエリを発行したピアからの論理ホップ数が設定された TTL を超えない限り、そのクエリを再び隣接する全てのピアにブロードキャストする。自ピアのキャッシュ内にクエリで要求されたアイテムがあれば、その情報をクエリを発行したピアに返信する。このときの返信情報は、そのクエリが中継されたピアを逆向きにたどって転送される。クエリを発行したピアは、そのような返信情報を 1 つ以上受け取った（以下では、クエリがヒットしたと表記する）なら、アクセス要求を行うデータアイテムを保持するピアのうち、自ピアからの論理ホップ数が最も少ないピアから受信を行う。この受信は、クエリを発行したピアとデータアイテムを保持するピアの間で、直接通信で行われる。同じ論理ホップ数のピアが複数存在する場合は、等確率でその一つのピアを選択する。クエリがヒットしなかった場合は、そのアイテムが次回に放送されるまで待ち、放送帯域からデータアイテムを受信する。放送帯域からアイテムを受信するときの応答時間は、アクセス要求が発行されてから、そのアイテムの次回の放送時刻までの時間とする。

想定環境のその他の詳細を以下に示す。

- 放送サーバは 1 つとし、各ピアはサーバへアクセス要求を送信しない。
- 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。
- ピア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。
- データの更新は発生しない。
- クエリの伝播遅延および各ピアにおけるクエリの処理時間は無視する。すなわち、あるクエリを発行してから、そのクエリがヒットするか否かを検証するまでの時間は、十分小さく無視できるものとする。
- P2P ネットワーク内のピア間の帯域幅は全て等しく、1 つのデータアイテムの他ピアからの受信に必要な時間は、どの

ピア間も等しいものとする。

4. プリフェッチ型協調キャッシング方式

2. で説明した従来方式の PIX 法および PT 法は、各クライアントが放送帯域からのみデータを受信することを想定しており、各クライアントは単独でキャッシュの置き換えを決定する。しかし本稿の想定環境では、P2P ネットワーク上の他のクライアント（ピア）がキャッシュしているデータアイテムを考慮してキャッシュ内のデータアイテムの置き換えを行うことで、応答時間をさらに短縮できる。本章では、P2P ネットワークを用いたプリフェッチ型協調キャッシング方式を提案する。

4.1 他ピアからのクエリ情報

P2P ネットワークでは、各ピアに到達したクエリの到達頻度やそれぞれの検索結果などのクエリ情報から、周囲における各データアイテムの要求度を計算することができる。本稿の提案方式では、この要求度と、自ピアの各データアイテムのアクセス確率を用いてキャッシングを行うことで、P2P ネットワーク全体の平均応答時間を短縮する。本節では、このクエリ情報の分類方法、およびその分類によって分かる周囲のピアの状況について説明する。

あるクエリが各ピアに到達した場合、そのクエリは、隣接ピアに伝播したクエリの検索結果によって、次の3種類のクエリのうちの1つとしてカウントされる。

- F(Failure) クエリ: 自ピアに要求データアイテムを発見することができず、隣接ピアに伝播したクエリのうち、いずれも要求データアイテムを保持するピアに到達しなかったクエリ。
- S(Success) クエリ: 自ピアに要求データアイテムを発見した、もしくは隣接ピアに伝播したクエリのうち、1つ以上が要求データアイテムを保持するクエリに到達したクエリ。
- C(Connected) クエリ: S クエリのうちで、要求データアイテムを保持するピアから、実際にデータアイテムのダウンロードが行われたもの。

それぞれの種類のクエリは、各ピアにおいて、データアイテム毎にカウントされる。

例として、ピア b, c, d の3つの隣接ピアをもつピア a について考える(図3)。ピア a, b, c, d のうち、ピア d のみがキャッシュにデータアイテム i を保持しており、ピア a の隣接ピア b からデータアイテム i に対するクエリが伝播されたとする。ピア a のキャッシュ内にはアイテム i が存在しないため、クエリは隣接ピア c, d にブロードキャストされる。隣接ピア c はアイテム i を持っていないが、隣接ピア d が要求データアイテム i を保持しており、ピア a はデータアイテム i を発見したクエリの伝播経路上にいるため、ピア a ではピア b から到達したそのクエリを、S クエリとしてカウントする。もし、そのクエリを発行したピアが、ピア d からアイテム i を取得した場合、S クエリの代わりに C クエリとしてカウントする。ここで、C クエリをカウントすることをルート上のピアに伝えるために、ピア d は取得が行われたことを知らせるメッセージを伝播経路上のピアに送信する。

また、同じピアから発行されたクエリが複数の経路から到達

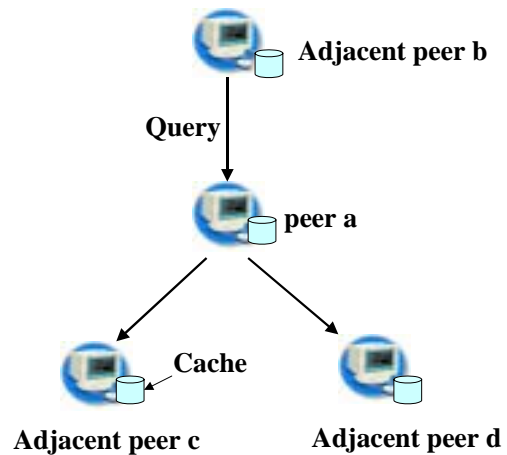


図3 クエリの伝播。

Fig. 3 Query distribution.

したとしても、そのクエリは1つしかカウントしない。その際の優先順位は、C クエリ、S クエリ、F クエリの順とする。例えば上記の例で説明したクエリとは別に、同じピアから発行された、データアイテム i に対するクエリが異なる経路を通じてピア a 到達したとする。このクエリの TTL が 0 になったとすると、そのクエリは隣接ピア d まで到達できず、要求されるデータアイテムを発見することができないため、同じピアから発行されたクエリが、S クエリと F クエリの2つのクエリとしてピア a に到達することになる。しかし、上記の優先順位によりピア a においては S クエリの1つしかカウントされない。

このようにクエリを分類してカウントすることにより、次のことが分かる。

- あるアイテムの F クエリが多くカウントされるピアの周辺は、そのアイテムを P2P ネットワークを用いて取得したいピアが多く存在するが、それを保持するピアが存在しない。
- あるアイテムの S クエリがカウントされるピアの周辺(そのピアも含む)には、そのアイテムを保持するピアが存在する。
- あるアイテムの C クエリが多くカウントされるピアの周辺(そのピアも含む)には、そのアイテムを保持するピアが存在し、実際に周辺のピアから必要とされている。

以上のようなクエリ情報は、各ピアがキャッシュしているアイテムが入れ替わったときに変化する。例えば、あるピアの周辺においてアイテム i が存在しないとし、周辺ピアのアイテム i に対する要求度が高いとする。このときアイテム i の F クエリは多くカウントされることになる。しかし、周辺ピアのいずれかがアイテム i をキャッシュしたとすると、今まで F クエリが多くカウントされていたピアには、S クエリおよび C クエリがカウントされるようになる。その後、再びそのアイテム i が周辺ピアに存在しなくなった場合は、F クエリが多くカウントされるようになる。

4.2 これまでの協調キャッシング方式

筆者らの研究グループではこれまでに、他ピアから送られるクエリの情報をもとに、他ピアの P2P ネットワークからのデー

タの取得も考慮して、P2P ネットワーク全体の平均応答時間の短縮を図る協調キャッシング方式である、C-PIX 法 [14] を提案している。本節では、C-PIX 法の概要を説明する。

C-PIX 法

C-PIX 法では、各ピアは 4.1 節で説明した方法で到達したクエリを分類し、以下のような手順でキャッシュの置き換えを行う。

各アイテムの放送開始時に、キャッシュ内のデータアイテムおよび放送されるデータアイテムに評価値を与える。キャッシュ内のデータアイテム j の評価値 U_j は次式を用いて求める。

$$U_j = P_j \cdot y_j + C_j \cdot (y_j - l) \quad (3)$$

ここで、P2P ネットワーク内のピア間でデータを受け渡しするのに必要な時間を l とする。また、アイテム j に対する単位時間当たりのアクセス確率を P_j 、単位時間当たりの F クエリ数を F_j 、S クエリ数を S_j 、C クエリ数を C_j とする。また、放送データアイテム k の評価値 U_k は次式を用いて求める。

$$U_k = \begin{cases} P_k \cdot y_k + F_k \cdot \alpha_k \cdot (y_k - l) & (S_k = 0) \\ P_k \cdot l + F_k \cdot \alpha_k \cdot (y_k - l) & (S_k > 0) \end{cases} \quad (4)$$

α_k はデータアイテム k をキャッシュしていない時点で、単位時間当たりに F_k 個到達している F クエリのうち、アイテム k をキャッシュしたときに C クエリに変わる割合の予測値を示す。

キャッシュしているアイテムを破棄するときは、次式を用いて α_i の値を変更する。ただし、式中では $\acute{\alpha}_i$ は変更前の α_i を表している。

$$\alpha_i = x \cdot \acute{\alpha}_i + (1 - x) \cdot C_i / \{F_i + C_i\} \quad (5)$$

ここで、 x ($0 \leq x \leq 1$) は以前の α_i の値をどの程度参考にするかを表すパラメータであり、1 に近いほど以前の α_i の情報を重視して、 α_i を決定することになる。以前の α_i の情報を重視しすぎると、ネットワークの状況の変化に応じて、 α_i の値を敏感に変化させることができなくなる。そのため、ネットワークの特性を考慮して、適当な x の値を決定する必要がある。

C-PIX 法では、キャッシュ内で最も評価値が小さくなるデータアイテム m の評価値 U_m と放送データアイテムの評価値 U_k を比較し、 $U_k > U_m$ の場合、キャッシュ内のアイテム m を k に置き換える。C-PIX 法は、自ピアのアクセス確率に加えて、他ピアから発行されたクエリも考慮してキャッシュの置き換えを行うことで、ネットワーク全体の平均応答時間の短縮を図る。しかし、プッシュ型放送においてはサーバが 1 つのデータアイテムを放送するたびに、各データアイテムの次回の放送時刻が変化したが、このキャッシング方式は応答時間の期待値の計算においてデータアイテムごとの放送間隔という静的なパラメータを用いている。そのため、必ずしも各時点において最適なデータアイテムをキャッシュするとは限らない。

4.3 提案方式

本稿で提案するプリフェッチ型協調キャッシング方式は、PT 法を P2P ネットワーク上の他ピアからのデータアクセスを考慮するように拡張したものである。キャッシュ内のデータアイテム

ム、および放送データアイテムそれぞれについて、P2P ネットワーク全体での応答時間の期待値を評価値として計算し、それをもとにキャッシュすべきアイテムを決定する。以下では、この方式を C-PT (Collaborative PT) 法と呼ぶ。C-PIX 法と同様に、C-PT 法も 4.1 節で説明した方法で各ピアに到達したクエリを分類し、それをもとにキャッシングを行う。

データアイテム i の評価値を G_i で表す。キャッシュ内のデータアイテム j については、それを失うことによる P2P ネットワーク全体での応答時間の増加分の期待値を評価値 G_j とし、次式を用いてこれを求める。

$$G_j = P_j \cdot (u_j(Q) - Q) + C_j \cdot \{(u_j(Q) - Q) - l\} \quad (6)$$

キャッシュしていない放送データアイテム k については、それを放送帯域から受信し、キャッシュすることで減少する P2P ネットワーク全体での応答時間の期待値を評価値とする。 $S_k = 0$ の場合、P2P ネットワークにおける近隣のピアの中に、データアイテム k をキャッシュに保持するピアが存在しない可能性が高いことを意味し、自ピアの発行するアクセス要求の応答時間は $(u_k(Q) - Q)$ となる。 $S_k > 0$ の場合、近隣のピアの中にデータアイテム k をキャッシュに保持するピアが存在することを意味し、現在はそのピアからデータアイテムの取得が可能のため、自ピアの発行するアクセス要求の応答時間は l となる。また、周辺のピアが発行するアクセス要求の応答時間は $(u_k(Q) - Q)$ から l になる。したがって、放送データアイテム k の評価値 G_k は次式で与えられる。

$$G_k = \begin{cases} P_k \cdot (u_k(Q) - Q) \\ + F_k \cdot \alpha_k \cdot \{(u_j(Q) - Q) - l\} & (S_k = 0) \\ P_k \cdot l + F_k \cdot \alpha_k \cdot \{(u_j(Q) - Q) - l\} & (S_k > 0) \end{cases} \quad (7)$$

C-PIX 法と同様に、 α_k は F クエリのうち、アイテム k をキャッシュしたときに C クエリに変わる割合の予測値を表す。また、 α_k の計算方法も式 (5) と同様である。

C-PT 法では、キャッシュ内で最も評価値が小さくなるデータアイテム m の評価値 G_m と放送データアイテムの評価値 G_k を比較し、 $G_k > G_m$ の場合、キャッシュ内のアイテム m を k に置き換える。キャッシュの置き換えが発生したとき、破棄したアイテムとキャッシュしたアイテムのクエリ情報 (F クエリ、S クエリ、C クエリの数) は破棄して、0 からカウントを再開する。

また、あるデータアイテムがキャッシュされてから、ある程度の数のクエリを受け取るまでに必要なタイムスロット数を、ウォームアップ時間 T とする。データアイテム i がキャッシュされてから T タイムスロットが経過するまでは、評価値 G_i は式 (6) を用いず、次式を用いて計算する。

$$G_i = P_i \cdot (u_i(Q) - Q) + \acute{F}_i \cdot \{(u_j(Q) - Q) - l\} \quad (8)$$

ここで、 \acute{F}_i は、データアイテム i がキャッシュされる前の単位時間当たりの F クエリ数を表す。

C-PIX 法が静的なパラメータである各データアイテムの放送間隔を利用するのに対し、C-PT 法は各時点で変化する値であ

る各データアイテムの次の放送時刻までの時間を利用する。そのため、各時点において、次に発行するアクセス要求に対するネットワーク全体での利得を最大にすることができる。

5. 性能評価

本章では、提案した手法の性能評価のために行ったシミュレーション実験の結果を示す。

5.1 シミュレーション環境

これまでの多くの研究により、インターネット上で構成される P2P ネットワークは、べき法則 (power-law) に従うことが示されている。ここで、ネットワーク上のあるピア j における隣接ピアの数をピア j の次数と呼び、 d_j で表す。 d_j の分布が以下で示す関数に従うとき、このネットワークはべき法則に従うという。

$$f(d_j) = [j^\beta] \quad (9)$$

この法則に従って、各ピア同士をランダムに接続したネットワークは PLRG(Power-law Random Graph) と呼ばれる。本稿の性能評価では、文献 [3] にならって、ピア数 500、 $\beta = 0.63$ の PLRG 型 P2P ネットワークを用いた。

各ピアのそれぞれのデータアイテムに対するアクセス確率は、Zipf 分布 [21] に従うものとした。ここで各ピアのそれぞれのデータアイテムに対するアクセス確率の分布は、次の 2 種類を用いた。

アクセス分布 1

全ピアが各データアイテムを、データアイテムの識別子の降順に高い確率でアクセスする分布。ただし、ピア間で各データアイテムに対するアクセス確率の大きさは等しくならず、多少のばらつきがあるものとする。これは、ネットワーク上の全てのユーザが同じような嗜好をもっている環境を想定している。具体的には、ピア j のアイテム i に対するアクセス要求確率 p_{ji} を、次式のような Zipf 分布を用いて与えた。

$$p_{ji} = \frac{i^{-\theta_j}}{\sum_{k=1}^M k^{-\theta_j}} \quad (10)$$

θ_j はピア j におけるアクセス確率を決定するための Zipf 係数と呼ばれるものであり、この値が大きいほど一部のデータアイテムが頻繁にアクセスされることを示している。この θ_j もまた、次式のように Zipf 分布を用いて求めた。

$$\theta_j = \frac{j^{-0.8}}{\sum_{k=1}^{MAX_PEER} k^{-0.8}} \quad (11)$$

ここで、 MAX_PEER は、総ピア数を表す。

アクセス分布 2

データアイテムに対するアクセス確率の高さの順序が、全ピアで異なる分布。これは、ネットワーク上の各ユーザが異なった嗜好をもっている環境を想定して

表 1 パラメータ設定.

Table 1 Parameter setting.

パラメータ	値
直前の α_i を参考にする度合い x	0.1
データアイテム数	1,000
クエリの TTL	3
ウォームアップ時間 T	500
ダウンロード時間 l	10

いる。具体的には、ピア j のアイテム i に対するアクセス確率 p_{ji} を次式のような Zipf 分布を用いて与えた。

$$p_{ji} = \frac{\{(i - r_j + 1) \bmod M\}^{-0.5}}{\sum_{k=1}^M k^{-0.5}} \quad (12)$$

ここで、 r_j はピア j が最も頻繁にアクセスするアイテムとする。あるピア j が、データアイテム i を最も頻繁にアクセスする ($r_j = i$ とする) 確率 q_i は、次式のような Zipf 分布を用いて求めた。

$$q_i = \frac{i^{-0.8}}{\sum_{k=1}^M k^{-0.8}} \quad (13)$$

各ピアがあるタイムスロットにおいてアクセスを発行する確率を 0.1 とした。よって、ピア j のデータアイテム i に対する単位時間当たりのアクセス頻度 P_{ji} は、次式のようになる。

$$P_{ji} = p_{ji} \times 0.1 \quad (14)$$

サーバは、すべてのデータアイテムが 1 つずつランダムな位置に配置されたプログラムに従って、周期的に放送を行うものとした。さらに各ピアは、初期状態において PT 法に基づいてデータアイテムをキャッシュしているものとし、各ピアの全てのデータアイテム i に対する α_i の初期値は 1 とした。その他のパラメータは、総データアイテム数を 1,000、クエリの TTL を 3 とし、また、各アイテムのデータサイズは 100Mbit (=12.5MByte)、各ピアの P2P ネットワークにおける受信帯域は 10Mbps と想定した。3. で示したとおり、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ので、P2P からの 1 アイテムのダウンロード時間 l を 10 タイムスロットと設定した。

以上のようなシステム環境において、300,000 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。比較対象として、C-PIX 法、および PIX 法、PT 法を用いてキャッシュの置き換えを行う場合の性能を評価した。ただし、PIX 法、PT 法では、P2P ネットワークからのデータアイテムの取得が可能であるとした。

5.2 シミュレーション結果

各ピアのキャッシュに格納可能な最大データアイテム数 (以下では、キャッシュサイズと表記する) を 0 から 1,000 まで変化させた場合の P2P ネットワーク全体での平均応答時間を図 4、図 5 に示す。図 4 は、各ピアのそれぞれのデータアイテムに対するアクセス確率をアクセス分布 1 に従って与えた場合の結果、

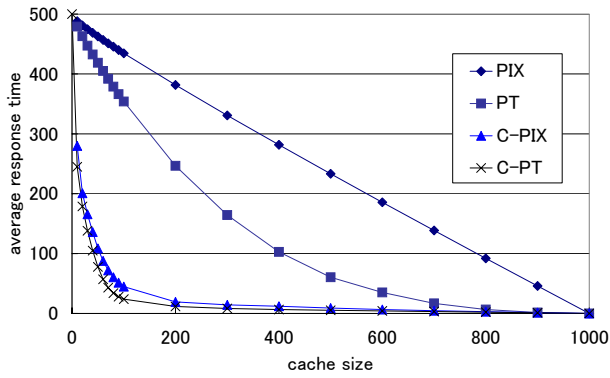


図 4 キャッシュサイズと平均応答時間 (アクセス分布 1) .

Fig. 4 Cache size vs. average response time (access distribution 1).

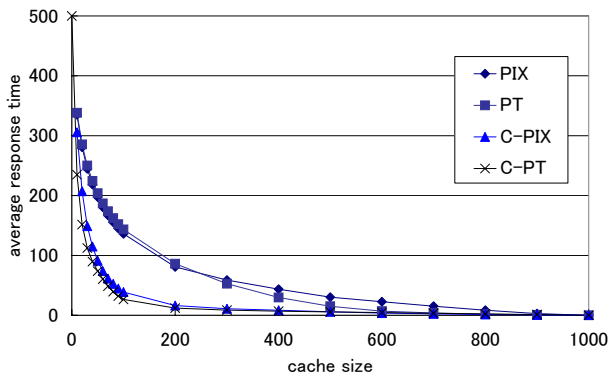


図 5 キャッシュサイズと平均応答時間 (アクセス分布 2) .

Fig. 5 Cache size vs. average response time (access distribution 2).

図 5 は、アクセス分布 2 に従って与えた場合の結果を表す。

図 4, 図 5 共に、提案方式である C-PT 法が最も良い結果を示している。ここで、図 4 と図 5 を比較すると、図 4 の方が周囲のピアのクエリ情報を考慮しない方式である PIX 法や PT 法と、周囲のピアのクエリ情報を考慮する C-PIX 法や C-PT 法の性能差が大きい。これは、図 4 では各ピアのデータアイテムに対するアクセス確率が類似したアクセス分布 1 を用いているためである。アクセス分布 1 に従ってアクセスを行うピアが PIX 法や PT 法を用いた場合、自ピアのアクセス確率しか考慮せずにキャッシングを行うため、各ピアが同じアイテムをキャッシュしてしまう。この場合、自ピアのキャッシュに存在しないアイテムを、P2P ネットワーク上の他ピアから取得可能な場合がほとんどない。しかし、C-PIX 法や C-PT 法を用いれば、近くのピア同士が協調してキャッシングを行うため、それぞれのキャッシュを有効に活用できる。よって各ピアのキャッシュサイズが小さい場合でも、クエリのヒット率を向上することが可能になり、ネットワーク全体での平均応答時間を短縮できる。

また、図 4, 図 5 のいずれにおいても、キャッシュサイズが 200 以下であるとき、C-PT 法の C-PIX 法に対する性能の優位性は高い。キャッシュサイズが小さい場合、C-PT 法の方が各

時点でもより正確に平均応答時間の期待値を計算しているため、少ないキャッシュ領域を有効活用できていることがわかる。

性能評価では、一部の中心となるピアがほとんどのピアに接続している PLRG 型 P2P ネットワークを用いているため、クエリの TTL が 3 とした少ない場合でも、多くのピアにクエリが伝播する。そのため、キャッシュサイズが小さい場合でもほとんどのクエリがヒットし、平均応答時間を短縮することができている。

6. 関連研究

Web キャッシングの分野では、複数のプロキシサーバによる Web コンテンツの協調キャッシングに関する研究が行われてきた [7]。この研究は、プロキシが階層的に管理され、データ要求経路に明確な方向づけが行われているのに対し、本研究は各ノードが特定の方針に従わずに接続しているネットワークを用いており、想定する環境が異なっている。また、この研究は負荷分散を目的とし、積極的に複数のプロキシにコンテンツをキャッシュする。それに対し、本研究では自ピアおよび P2P ネットワークからの取得確率の向上による各ピアの平均応答時間短縮が目的となっているため、各ピアのキャッシュデータができるだけ重複しないように、他ピアのキャッシュデータの推測を行っている点で、この研究と異なる。

また、分散ファイルシステムやアドホックネットワークの分野では、アクセス要求に対する応答時間を短縮するための協調キャッシングに関する研究が行われている。

文献 [5] では、分散ファイルシステムにおいて、各クライアントが自身の記憶領域へのアクセス、他のクライアントの記憶領域へのアクセス、そしてサーバから受信といった 3 つの方法から選択可能なもの内、応答時間が最短となる方法を選択するような環境を想定し、各クライアントのデータ配置方法についても議論している点で、本研究と類似する。文献 [5] ではクライアント数は数十程度と少なく、データの配信方式として一対一のプル型の通信を行う環境を想定しており、各データを保持するクライアントの情報をサーバが管理する。そのため、各クライアントはどのクライアントにデータがキャッシュされているか、およびどのデータをキャッシュするべきかをサーバに問い合わせることで知ることができる。

アドホックネットワークの分野では、プッシュ型放送とのハイブリッドなシステムを想定し、各移動体が協調的にキャッシングを行う、本研究と類似した研究がある [11]。この研究では、発掘調査を行う研究チームなど同じ目的を持った複数の移動体がある範囲に散らばって協調的に作業するような環境を想定しており、このような近接した移動体グループ内でデータを最適に配置することで、アクセス要求に対する応答時間の短縮を図っている。文献 [11] では比較的少数の移動体グループによる協調を想定している。そのため、ネットワーク全体で最適な配置を決定することができる。

本研究では上記の 2 つの研究とは違い、数百万規模のユーザが存在する大規模ネットワークを想定している。そのため、特定のサーバや移動体に、配置情報の管理や配置決定を委ねる

ことは困難である．そこで，各ピアが周辺の情報のみを送受信されるメッセージをもとに推測し，自律的にキャッシュすべきデータを決定している．

7. ま と め

本稿では，プッシュ型放送のサービスを受ける複数のピアが P2P ネットワークを構成しているような環境を想定し，ピアが協調してキャッシングを行う方式を提案した．提案方式では，自ピアのアクセス確率に加えて，他ピアから発行されたクエリも考慮し，各時点での応答時間の利得が大きいデータアイテムをプリフェッチすることで P2P ネットワーク全体の平均応答時間の短縮を図る．

シミュレーション実験の結果から，提案方式である C-PT 法がキャッシュサイズに関わらず，既存のキャッシング方式よりも良い結果を示すことを確認した．

本稿の想定環境では，各ピアは P2P ネットワーク上に必要とするデータアイテムを発見するためにクエリをフラッディングしているが，この手法ではネットワークトラフィックが大きくなってしまふ．この問題を解決するために，k-walker random walk [13] などの手法が提案されており，本稿で提案した協調キャッシング方式は，このような検索手法を用いた場合にも適用可能である．今後は，これらの検索手法を用いた場合のシミュレーション評価も行う予定である．

また本稿の性能評価では，TTL や各ピアのアクセス確率などのパラメータが一定である環境で評価を行ったが，データアイテムやユーザの設定によってこれらの値は変化する．また，本シミュレーションで用いた PLRG 型ネットワークはパラメータの設定により各ピアの繋がりが密になっているが，実環境において必ずしもこのような密なネットワークになるとは限らない．そこで今後は，それらを考慮した，様々な環境におけるシミュレーション評価を行う予定である．

謝 辞

本研究は，文部科学省 21 世紀 COE プログラム「ネットワーク共生環境を築く情報技術の創出」，文部科学省特定領域研究(15017262) 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．

文 献

- [1] S. Acharya, R. Alonso, M. Franklin, and S. Zdonik, "Broadcast Disks: Data Management for Asymmetric Communication Environments," Proc. ACM SIGMOD'95, pp. 199–210, 1995.
- [2] S. Acharya, M. Franklin, and S. Zdonik, "Prefetching from a Broadcast Disk," Proc. Int'l Conf. on Database Engineering, pp. 276–285, 1996.
- [3] T. Bu and D. Towsley, "On Distinguishing Between Internet Power Law Topology Generators," Proc. INFOCOM'02, pp.638–647, 2002.
- [4] E. Cohen and S. Shenker, "Replication Strategies in Unstructured Peer-to-Peer Networks," Proc. ACM SIGCOMM'02, pp.177–190, 2002.
- [5] M. Dahlin, R. Wang, T. Anderson, and D. Patterson, "Cooperative Caching: Using Remote Client Memory to Improve File System Performance," Proc. 1st Symp. on Operating Systems Design and Implementation, pp. 267–280, 1994.
- [6] E. Damiani, S.D.C.D. Vimericati, and S. Paraboschi, "A Reputation-Based Approach for Choosing Reliable Resources in Peer-to-Peer Networks," Proc. ACM Conf. on Computer and Communications Security, pp.207–216, 2002.
- [7] P. Danzig, R. Hall, and M. Schwartz, "A Case for Caching File Objects Inside Internetworks," Proc. ACM Conf. on Communications Architectures, Protocols and Applications (SIGCOMM), pp. 239–248, 1993.
- [8] FreeNet, <URL:http://freenet.sourceforge.net>.
- [9] Gnutella, <URL:http://gnutella.wego.com>.
- [10] V. Grassi, "Prefetching Policies for Energy Saving and Latency Reduction in a Wireless Broadcast Data Delivery System," Proc. MSWIM 2000, pp. 77–84, 2000.
- [11] 原 隆浩: アドホックネットワークにおけるプッシュ型配送データのための協調キャッシング方式, 日本データベース学会 Letters, vol. 1, no. 1, pp. 67–70, 2002.
- [12] I. Stoica, R. Morris, D. Karger, F. Kaashoek, and H. Balakrishnan, "Chord: A Scalable Peer-to-Peer Lookup Service for Internet Applications," Proc. ACM SIGCOMM'01, pp.149–160, 2001.
- [13] Q. Lv, P. Cao, E. Cohen, K. Li, and S. Shenker, "Search and Replication in Unstructured Peer-to-Peer Networks," Proc. the 16th Int'l Conf. on Supercomputing, pp.84–95, 2002.
- [14] 前田和彦, 内田 渉, 原隆浩, 西尾章治郎, "Peer-to-Peer ネットワークにおけるプッシュ型情報システムのための協調キャッシング方式について," 情報処理学会マルチメディア通信と分散処理ワークショップ論文集, pp.67-72, 2003.
- [15] Napster, <URL:http://www.napster.com>.
- [16] S. Ratnasamy, P. Francis, M. Handley, and R. Karp, "A Scalable Content-Addressable Network," Proc. ACM SIGCOMM'01, pp.161–172, 2001.
- [17] C. J. Su, L. Tassiulas, and V. J. Tsotras, "Broadcast Scheduling for Information Distribution, ACM-Baltzer Journal of Wireless Networks", vol. 5, no. 2, pp. 137–147, 1999.
- [18] 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な変化を考慮した相関データの放送スケジューリング," 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.
- [19] 内田 渉, 原隆浩, 西尾章治郎, "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," 情報処理学会第 129 回データベースシステム研究報告 (DBS-129/BCC-4), pp. 67–74, 2003.
- [20] N. H. Vaidya, and S. Hameed, "Scheduling Data Broadcast in Asymmetric Communication Environments," ACM-Baltzer Journal of Wireless Networks, vol. 5, no. 3, pp. 171–182, 1999.
- [21] G. K. Zipf, "Human Behavior and the Principle of Least Effort," Addison-Wesley, 1949.