

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

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 大阪府吹田市山田丘2-1
E-mail: {k.maeda,wataru,hara,nishio}@ist.osaka-u.ac.jp

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情報システムに関する研究の関心が高まってきている。本稿では、プッシュ型情報システム内の複数のクライアントが 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の情報をもとに、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キャッシュする。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。

On a Collaborative Caching Strategy for Push-based Information Systems in Peer-to-Peer Network

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

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

Recently, there has been an increasing interest in research of push-based information systems that deliver data using broadcast in both wired and wireless environments. In this paper, we propose a new collaborative caching strategy in a push-based information system in which clients construct a peer-to-peer network and collaboratively cache broadcast data. In the proposed strategy, a peer takes into account its own access probabilities and informations on queries issued by other peers and caches data items with long expected response times. Moreover, we verify the effectiveness of the proposed strategy by simulation experiments.

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情報システム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情報システムは、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、プッシュ型情報システム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、クライアント側でのデータのキャッシュ戦略 [1, 2] がこれまでにいくつか提案されている。この代表的なものとして、Acharyaらによる PIX 法 [1] がある。この方式は各クライアントにおけるデータのアクセス確率や放送周期から、応答時間の利得が大きくなるように効果的にキャッシュの置き換えを行っている。PIX 法を含む、プッシュ型情報システムにおける従来のキャッシュ戦略に関する研究の大半は、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) 型情報共有に関する研究が盛んに行われている [4, 5, 6]。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、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

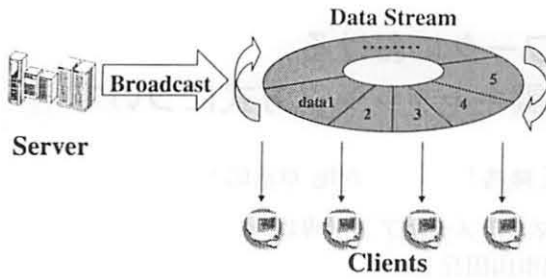


図 1: プッシュ型情報システム

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、データの共有が行われる。

プッシュ型情報システムにおいて、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短縮できるものと考えられる。しかし、筆者らの知る限り、そのような研究は、これまでには行われていない。そこで本稿では、プッシュ型情報システム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、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、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。

以下では、2 章で想定環境について述べる。3 章では従来のキャッシング方式について説明する。4 章では本稿の提案方式について説明する。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述べる。

2 想定環境

本稿における想定環境を図 2 に示す。想定環境において、データはデータアイテムと呼ばれる単位でサーバから放送されるものとする。各ピアはキャッシュをもち、自ピアのキャッシュへのアクセス、放送帯域からの取得、他ピアのキャッシュからの P2P ネットワークを利用した取得のうち、データ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。

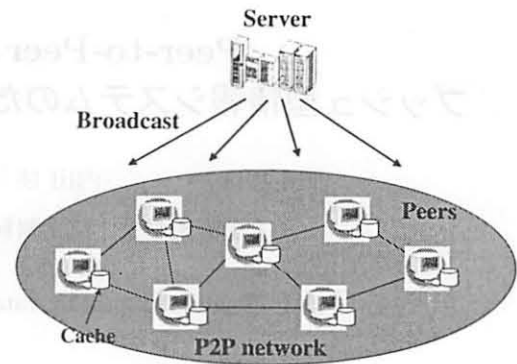


図 2: 想定環境

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る場合は、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ダウンロードを完了するまでの時間を比較する。もしも、次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ダウンロードを完了するまでの時間の方が短い場合は、周囲のピアにアクセス要求を行う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ために、ある TTL(Time to Live) 値を定めたクエリを発行し、隣接する全てのピアにそれを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ダウンロードを行う。このダウンロードは、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、P2P 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、等確率でその中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取得する。放送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次に示す。

- 放送サーバは1つとし、プル型配送は行わない。
- 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。
- 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取得に必要な時間は、どのピア間も等しいものとする。

3 従来のキャッシング方式

本章では、プッシュ型情報システムにおける平均応答時間短縮のためのキャッシング方式である、PIX 法 [1] について説明する。PIX 法は、クライアントのアクセス確率が高く、放送間隔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を短縮できる。PIX 法の手順を次に示す。

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

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

ただし、 p_j はデータアイテム j に対するアクセス確率、 x_j はデータアイテム j の放送間隔である。

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サイズだけキャッシュする。

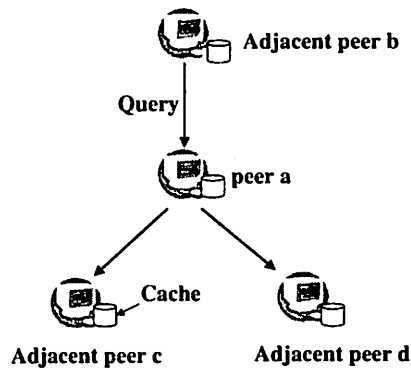


図 3: クエリの伝播

PIX 法は、クライアントのアクセス確率が変化しない場合、一度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。

PIX 法を本稿の想定環境に適用した場合、各ピアは自ピアのアクセス確率のみを考慮して、キャッシングを行う。

4 協調キャッシング方式

本章では、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を提案する。

4.1 他ピアからのクエリ情報

提案方式では、P2P ネットワーク上の各ピアに到達するクエリを、3つのカテゴリに分ける。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、クエリによって発見された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、3つの隣接ピアをもつピア a について考える (図 3)。そのピア a の隣接ピア b からクエリが伝播されたとする。ピア a のキャッシュ内に要求されたアイテムが存在しない場合、クエリは隣接ピア c, d にブロードキャストされる。もし、隣接ピア c がそのアイテムを持っていなかったとしても、隣接ピア d が要求データアイテムを保持していれば、ピア a は要求データアイテムを発見したクエリの伝播経路上にいるため、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 から要求データアイテムをダウンロードした場合、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、その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。例えば上記の例において、ピア a に同じピアから発行されたクエリが別の経路を通して到達したとする。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には、そのピアも含め、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には、そのピアを含め、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。

4.2 評価値

アイテム i の放送周期を L_i とし、P2P ネットワーク内のピア間でデータを受け渡すのに必要な時間を l とする。また、アイテム i に対する単位時間当たりのアクセス確率を P_i 、単位時間当たりの F ク

エリ数を F_i 、S クエリ数を S_i 、C クエリ数を C_i とする。

本稿で提案する協調キャッシング方式は、キャッシュ内のデータアイテム、および放送データアイテムそれぞれについて、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。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とする。データアイテム j の評価値 G_j は次式のように表される。

$$G_j = P_j \cdot L_j / 2 + C_j \cdot (L_j / 2 - l) \quad (2)$$

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、自ピアを含めた周辺のピアの発行するアクセス要求の応答時間は $L_k / 2$ となる。また、 $S_k > 0$ の場合は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。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。

$$G_k = \begin{cases} (P_k + F_k \cdot \alpha_k) \cdot L_k / 2 & (S_k = 0) \\ P_k \cdot l + F_k \cdot \alpha_k \cdot L_k / 2 & (S_k > 0) \end{cases} \quad (3)$$

ここで、 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。この詳細については後述する。

提案方式では、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

トが経過するまでは、評価値 G_i は式 (2) を用いず、次式を用いて計算する。

$$G_i = P_i \cdot L_i/2 + \hat{F}_i \cdot (L_i/2 - l) \quad (4)$$

ここで、 \hat{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。

キャッシュしているアイテムを破棄するときは、 α_i の値を、変更前の α_i を $\hat{\alpha}_i$ の値として、次式を用いて計算する。

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

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。ここで、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、 x の値を決定する必要がある。

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。このとき、 d_j の分布が以下で示す関数 $f(d_j)$ に従うとき、このネットワークはべき法則に従うという。

$$f(d_j) = j^\beta \quad (6)$$

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を PLRG (Power-law Random Graph) と呼ぶ。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。また、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、データアイテム 1 を頂点とした Zipf 分布 [7] に従うものとし、データアイテム i に対するアクセス確率 p_i を次式を用いて決定した。

$$p_i = \frac{i^\theta}{\sum_{j=1}^M j^\theta} \quad (7)$$

表 1: パラメータ設定

パラメータ	値
直前の α_i を参考にする度合い x	0.02
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	50
ダウンロード時間 l	10

ここで θ は Zipf 係数と呼ばれる値であり、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環境を表す。また、各ピアにおける、アクセス要求確率を決定するためのパラメータとなる θ も、Zipf 分布を用いて決定した。そのときの Zipf 係数の値を 0.8 とした。ここで、あるピアのデータアイテム i に対する単位時間当たりのアクセス頻度 P_i は次式のようになる。

$$P_i = p_i \times 0.1 \quad (8)$$

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 PIX 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を表 1 に示す。ここで、各アイテムのデータサイズは 100 Mbit (= 12.5 MByte)、各ピアの P2P ネットワークにおける受信帯域は 10 Mbps と想定した。そのため、ダウンロード時間 l は 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、PIX 法を用いた。以下ではシミュレーション評価の結果を示し、考察を行う。

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 に示す。図 4 において、“proposed-strategy” は提案方式を用いた場合の平均応答時間を示している。図 4 の結果から、提案方式がキャッシュサイズに関わらず PIX 法よりも平均応答時間を大幅に短縮していることがわかる。特に、キャッシュサイズが 100 の付近で、PIX 法と提案方式の性能の差が最も大きくなっている。既存の方式はキャッシュサイズが増加するにつれて平均応答時間が線形的に減少するが、提案方式では協調的に

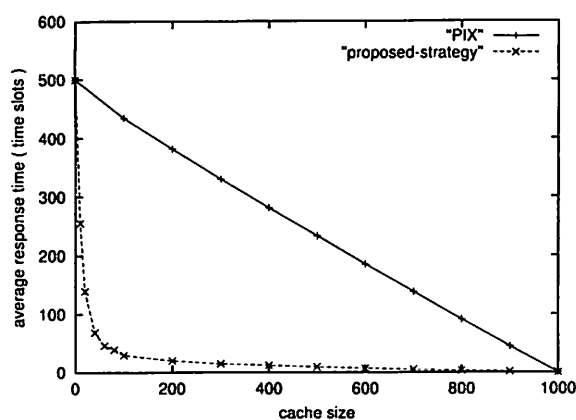


図 4: キャッシュサイズと平均応答時間

キャッシングを行っているので、キャッシュサイズが 0 から 100 の間で急激に減少している。

PIX 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、各ピアはアクセス確率が高ければ、周辺のピアが保持しているデータアイテムでもキャッシュしてしまい、他ピアのキャッシュを有効活用することができない。提案方式を用いれば、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である。特に、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 といった少ない場合でも、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、平均応答時間を大幅に短縮している。

6 まとめ

本稿では、プッシュ型情報システム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て 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、提案方式がキャッシュサイズに関わらず、既存の方式よりも平均応答時間を大幅に短縮することを確認した。

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ができるとは限らない。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 (研究拠点形成費補助)、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。

参考文献

- [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, D. Towsley, "On Distinguishing between Internet Power Law Topology Generators," Proc. INFOCOM'02, pp.638-647, 2002.
- [4] E. Cohen, S. Shenker, "Replication Strategies in Unstructured Peer-to-Peer Networks," Proc. ACM SIGCOMM'02, pp.177-190, 2002.
- [5] E. Damiani, S.D.C.D. Vimericati, 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.
- [6] Q. Lv, P. Cao, E. Cohen, K. Li, S. Shenker, "Search and Replication in Unstructured Peer-to-Peer Networks," Proc. the 16th Int'l Conf. on Supercomputing, pp.84-95, 2002.
- [7] G. K. Zipf, "Human Behavior and the Principle of Least Effort," Addison-Wesley, 1949.