

コンテンツ流通制御を考慮したウェブコンテンツ共有システムの実現

中村 聡史[†] 塚本 昌彦^{††} 西尾 章治郎^{††}

本稿では、インターネットで公開されているウェブコンテンツをローカルディスクに蓄積し、P2Pネットワーク上で相互共有するウェブコンテンツ共有システムを提案・実装する。本システムでは、コンテンツを検索する際、URLだけでなく、更新日時やサイズなどのメタ情報を付加することを可能とする。本システムを利用することにより、すでに存在しない過去に公開されていたコンテンツも、他のユーザが蓄積していればアクセスできるようになる。また、コンテンツ作成者向けに、再配布禁止、配布通知など、コンテンツの流通を制御する仕組みも実現する。

On Realizing a Web Contents Sharing System with Contents Delivery Control Functions

SATOSHI NAKAMURA,[†] MASAHIKO TSUKAMOTO^{††}
and SHOJIRO NISHIO^{††}

In this paper, we propose a web contents sharing system which allows a user to save web contents presented on internet and share them on P2P (peer to peer) network. In this system, a user can use the URL, the last update date and time, the size or some other information in order to search web pages. In addition, a user can get web pages that have existed and do not currently exist if they are stored in other systems. Furthermore, we realized a mechanism for authors to control contents delivery, such as the prohibited of forwarding the contents and notification to the authors on forwarding contents.

1. はじめに

近年、コンピュータの低価格化・高性能化、ネットワークインフラの整備などにより、インターネット上のウェブサイトの数が驚くべき速度で増加しつつある。これらのウェブサイトの中には良質のコンテンツを提供しているところも多数あり、ユーザは検索サービスなどを利用することで、必要な情報を効率良く探すことができる。しかし、インターネット上で公開されているウェブコンテンツは日々変化するものであり、恒久的なものではない。そのため、検索して望むコンテンツを見つけても、アクセスしたときには新しいコンテンツに置き換えられてしまっていることがある。また、サーバメンテナンスやネットワークの断線、過負荷状態によるサーバの動作不良などにより、コンテンツにアクセスできないこともある。

このようにウェブコンテンツは常時利用可能なもの

ではないため、ローカルのディスクにコンテンツをダウンロードし、必要なときにローカルのコンテンツを開いて利用できるようにシステムが開発されている^{13),14)}。また、Googleのキャッシュサービス⁵⁾や、Internet Archive⁷⁾などのサービスでは、過去のスナップショットを記録し、すでに削除されているような過去のコンテンツへのアクセスを可能としている。しかし、こうしたサービスは万能ではなく、インターネット上のすべてのコンテンツが保存されているわけではない。また、新しい情報への対応が遅いという問題がある。

一方、Gnutella⁴⁾やNapster¹¹⁾、Freenet⁶⁾に代表されるP2P(peer to peer)型データ共有システムが近年、注目を浴びている¹⁵⁾。こうしたシステムでは、ユーザは見ず知らずのユーザとネットワークを介して相互に接続し、各自が保持するドキュメントやマルチメディアデータなどを共有利用することを可能とする。ユーザは、ファイル名やファイルサイズなどを用いて検索を行い、探しているコンテンツを見つけた場合は、HTTPリクエストを利用してファイルの要求メッセージを送り、コンテンツを取得する。多くのユーザがネットワークに参加することにより、共有さ

[†] 大阪大学大学院工学研究科
Graduate School of Engineering, Osaka University

^{††} 大阪大学大学院情報科学研究科
Graduate School of Information Science and Technology, Osaka University

れるデータも莫大なものとなり、巨大なストレージが形成される^{1),3),8)}。このようなサービスは、新しいインターネットサービスとして期待されているが、著作権を無視したコンテンツ共有などにより違法性も指摘され、問題になっている。

そこで本研究では、各ユーザが興味のあるウェブコンテンツをローカルディスクに蓄積し、そのコンテンツを P2P 型のネットワークで相互に共有し、利用するための仕組みを実装する。蓄積時に URL や更新日時などの情報を保存しておくことにより、他のユーザは、URL や更新日、タイトルなどを指定した検索問合せを可能とする。また、ユーザによる好みの違いなどにより、共有されるコンテンツの種類は多岐にわたると予想される。さらに、保存のタイミングによる異なるバージョンの共有も期待でき、すでに公開されていない過去のコンテンツへのアクセスも可能になると考えられる。しかし、コンテンツ共有環境の実現により、ウェブコンテンツが作成者の意図に沿わず、著作権を無視して違法にコピーされ、ばら撒かれる可能性があるため、システム全体でこうした違法複製を防止し、コンテンツ作成者がコンテンツの流通を制御できる仕組みもあわせて提案する。

以下、2 章では P2P 型ウェブコンテンツ共有システムについて述べ、3 章で流通制御について提案を行う。4 章でシステムの実装について述べ、5 章で実現したシステムに関して考察を行う。最後に 6 章でまとめと今後の課題について述べる。

2. ウェブコンテンツ共有システム

ウェブコンテンツ共有システムでは、ユーザはウェブコンテンツを、URL をベースとしたローカルディスクのディレクトリに保存する。このとき、取得元の URL や、ウェブページのタイトル、ファイルサイズ、更新日時、取得日時などもあわせて保存する。また、ユーザのコンピュータ上で動作するサーバントシステム(ピア)は、近くに存在するピアを探し、相互に接続することにより、Gnutella のような PureP2P 型のコンテンツ共有ネットワークを構成する。ローカルディスクに保存したコンテンツは、このネットワーク上で共有される。

ユーザは情報(コンテンツ)を探す際、オリジナルの URL に加え、タイトル、ファイルサイズ、更新日時、取得日時などのメタ情報を用いた問合せを行うことができる。たとえば、「2002 年 12 月 31 日の NIKKEI WEB のトップページ」といった問合せを作成し、共有ネットワークに対して送信すると、接続しているピ

アは問合せを順次隣のピアにリレーする。問合せの条件に一致するコンテンツを持つピアは自動的に応答し、自分の IP Address や問合せに合致するコンテンツの URL、サイズ、更新日時、取得日時などの情報を問合せ元に戻す。この応答により、ユーザは指定したコンテンツを持つユーザ(ピア)を調べることができ、問い合わせした結果作成されるリストの中から好みの情報を選び分け、必要な情報にアクセスすることができる。コンテンツを保存する際に、「サッカー」や「ワールドカップ」など、意味のあるフォルダ名をつけ、コンテンツをその中に整理しておくことにより、保存フォルダをターゲットとした検索も可能となる。また、同じ趣味を持つユーザ同士が集まってプライベートなウェブコンテンツ共有ネットワークを構成することにより、より良い情報収集が可能になると考えられる。この、趣味の似通ったユーザ同士での情報の共有においては、フォルダ名を用いた問合せが特に有効になると考えられる。さらに、各ユーザは好みの情報を発見したときにウェブコンテンツをローカルディスクに保存するため、共有ネットワーク上でつねに新しい情報に対して、検索・問合せを行うこともできる。

ウェブコンテンツ共有システムの実現と多くのユーザの参加により、増え続ける膨大なウェブコンテンツ群を広くカバーし、共有することができるようになる。結果として、ユーザはこれまで以上に多くの情報を得ることができるようになる。しかし、こうしたウェブコンテンツの共有を無制限にすると、コンテンツ作成者の意図に沿わない可能性が高くなるため、コンテンツの流通を効果的に制御する機構が必要となる。

3. コンテンツの流通制御

本章では、ウェブコンテンツ作成者がウェブコンテンツ共有環境において、コンテンツ流通を制御するために最低限必要であると考えられる保存制限と転送制限の機能について説明する。

3.1 保存制限

保存制限としては以下の 6 つが考えられる。

- ローカルディスクへの保存(許可・不許可)
- 完全保存(許可・不許可)
- 部分保存(許可・不許可)
- 相対パス変換(許可・不許可)
- 改変(許可・不許可)
- 保存通知

一番厳しい制限は、ローカルディスクへの保存を許可しないというものである。この制限を加えることにより、ユーザはそのウェブコンテンツをローカルディ

スクに保存できなくなり、ネットワーク上での共有もできなくなる。当然、ネットワークの断線などによりサーバに接続できないときは、そのウェブページを閲覧することができない。コンテンツの完全保存制限では、コンテンツを完璧にローカルディスクに保存する際の許可・不許可を指定し、コンテンツの部分保存制限では、コンテンツの一部をローカルディスクに保存する際の許可・不許可を指定する。部分取り込みでは、取り込むコンテンツにタイトルと取得 URL、取得日時が付加される。広告や著作権表示などをカットされることがコンテンツ作成者の意図に反する場合、コンテンツ作成者はコンテンツの部分保存を不許可とすることで、広告がカットされることを未然に防ぐことができる。

ウェブコンテンツを取り込む際は、ネットワークから切断された状態でもコンテンツを閲覧可能とするため、画像やスタイルシートなど、ページを表示する際に利用するコンテンツへのパスを相対パスへと変換し、一緒に取り込んでいる。相対パス変換制限は、このパス変換を不許可とし、取り込み元 URL の絶対パスに変換することを強制するものである。この制限により、ユーザはコンテンツを取り込んでいても、取り込み元サイトと通信可能でない場合は、画像を含んだような完全なコンテンツを閲覧することができない。特に、スタイルシートで表示位置を指定している場合は、スタイルシートが存在しないとページをまともに見ることができないため、制限としてかなり有効であると考えられる。

コンテンツの改変制限では、そのコンテンツの編集に関する許可・不許可を指定する。たとえば、取り込んだコンテンツを編集し、コンテンツ中にメモを書き込むことや、アンダーラインを引いて部分的に強調することが利用方法として考えられるが、改変制限ではこうした利用を制限する。

保存通知制限は、保存する際にウェブコンテンツ作成者などに E-Mail など通知・連絡することを義務付けるものである。連絡をした場合はウェブコンテンツを保存することを許可する。

3.2 転送制限

転送制限としては以下の 7 つが考えられる。

- 転送（許可・不許可）
- 完全保存時転送（許可・不許可）
- 部分保存時転送（許可・不許可）
- 改変時転送（許可・不許可）
- 転送回数制限
- 切断時のみ許可

● 転送通知

転送制限では、ローカルディスクに保存したウェブコンテンツを他ユーザへ転送する際の許可・不許可を指定する。転送不許可を指定されている場合は、このコンテンツをウェブコンテンツ共有環境で共有することができない。ローカルディスクに保存されたウェブコンテンツの転送を禁止することにより、簡単に流通を制限することができる。完全保存時転送制限では、ローカルディスクに保存したウェブコンテンツが完全保存されているときの転送に関する制限を設定する。また、部分保存時転送制限では、ローカルディスクに保存されたコンテンツがウェブコンテンツの一部であり、完全でないときの転送に関する制限を設定し、改変時転送制限では、ローカルディスクに保存したウェブコンテンツを編集した場合のコンテンツの転送に関する制限についてそれぞれ設定する。たとえば、著作権表示をカットするなど、部分的にしかコンテンツを保存していない場合、他ユーザへの共有を禁止するというのも可能である。

転送回数制限では、ウェブサーバからローカルディスクに保存したコンテンツを、他ユーザへと転送できる回数を制限する。この回数制限が 1 以上であるときは、その回数だけ他ユーザへとコンテンツを転送することができ、転送するたびにこの回数を減らす。0 になったときは転送不可となる。

切断時のみ許可という制限は、そのウェブコンテンツを公開しているサーバが運用を停止しているときや、アクセスが集中しすぎてサーバにアクセスできないときのみ、ユーザ間でコンテンツをやりとりすることを許可するというものである。この機能により、ウェブサーバに問題があるときにだけ他のユーザからコンテンツを手に入れるよう制限することができる。

転送通知制限は、コンテンツを他ユーザへ転送するときに E-Mail などコンテンツ作成者に通知することを義務付けるものである。これらの指定により、コンテンツ作成者はコンテンツの流通具合を調べることができる。

4. 実 装

本章では、本研究で実現したプロトタイプシステムの設計および実装について述べる。

4.1 ウェブコンテンツ共有システムの実装

ウェブコンテンツ共有システムは、図 1 のように、ユーザ間を P2P ネットワークで相互に接続し、問合せや問合せ結果を送信・リレーするサーバントシステムと、ウェブコンテンツを保存、管理するクライアント

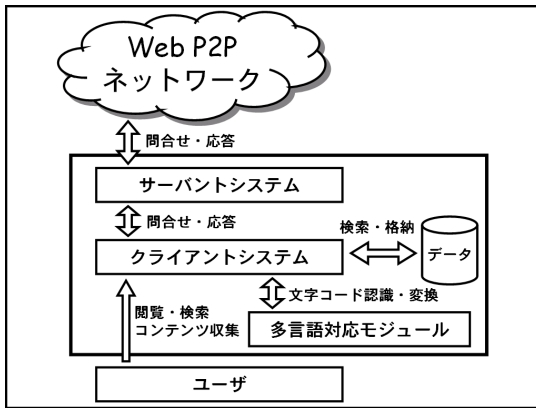


図 1 システム構成

Fig. 1 Design of a system.

トシステムで構成する。

クライアントシステムには、ウェブコンテンツを閲覧する機能、ウェブコンテンツを保存する機能、ウェブコンテンツを管理する機能、ウェブコンテンツを検索する機能が必要となる。

本プロトタイプシステムでは、ウェブコンテンツ閲覧機能を提供するため、ブラウザコントロールを利用する。また、保存については、ウェブコンテンツをローカルのみで閲覧可能とするために、すべてのコンテンツを、URL をベースとしたフォルダに保存し、ページ中の画像やスタイルシートへのリンクを、相対パス変換して保存する。なお、URL はフォルダやファイル名として利用できない文字を含むことがあるため、利用可能なファイル名に変換する仕組みも導入する。ウェブコンテンツの管理については、ツリービューとリストビューを組み合わせたインターフェースを利用し、整理できるようにする。階層管理を可能とすることにより、ローカルに保存しているコンテンツの整理が容易になると考えられる。また、ツリー・リストインターフェースから、ネットワークで共有するコンテンツの入ったフォルダを指定できるようにする。このインターフェースを利用することにより、共有したくないコンテンツも容易に指定できると考えられる。ウェブコンテンツの検索については、URL 検索、タイトル検索、サイズを指定した検索、更新日時を指定した検索、コンテンツ内部に含むキーワードでの検索をサポートする。これらを組み合わせて利用することにより、意図する情報を効率良く見つけることができるようになると考えられる。

P2P ネットワークを構成するピアどうしは、TCP/IP を利用して相互接続する。P2P のシステムとして、jnutella⁹⁾ や Jxta¹⁰⁾ などの利用も考えられるが、実

装環境の統一と拡張性を考え、独自のシステムを実装する。

本システムでは、サーバントシステムは接続可能ピアリストを持ち、他のピアと相互接続することにより P2P ネットワークを構成する。また、サーバントシステムは、現在直接接続しているピアとソケットをセットにした接続ピアテーブルと、問合せ ID と問合せ内容、問合せ元のソケットをセットにした問合せテーブルを持ち、問合せのリレーなどに利用する。問合せを受信した場合には、問合せを接続ピアにリレーすると同時に、クライアントシステムに問合せを送信し、検索の要求などを行う。ピアどうしの詳しいやりとりについては 4.2 節で述べる。

流通制御の実現においては、このサーバントシステムと、クライアントシステムに加え、流通制御システムが必要となる。本プロトタイプシステムでは、ピアシステム、クライアントシステムが内部に流通制御部を持ち、「共有可能か?」「転送可能か?」「保存可能か?」などを判断し、動作に反映させる。流通制御の実装については 4.3 節に述べる。

4.2 プロトコル

ピア間でやりとりするプロトコルは以下のとおりである。

- P2P ネットワークへの参加
- P2P ネットワークへの参加承認
- P2P ネットワークからの退出
- 別ピアの紹介
- P2P ネットワークへの問合せ
- 問合せ結果の返信
- データの取得

それぞれについて簡単に説明する。

P2P ネットワークへの参加する際には、P2P ネットワークに対して以下のような参加要求を送る。

```
WeBoXP2P Connect\n\n
```

このメッセージを受信したピアは、送信してきたピアに対して参加許可を出すかどうか判断する。参加許可を出す場合は、

```
WeBoXP2P ok\n\n
```

というメッセージを返信し、接続ピアテーブルに要求を送ってきたピアを追加する。このメッセージを受信したピアは、P2P ネットワークへの参加が認められたことを確認し、接続ピアテーブルに接続したピアを追

加する。もし、OKを受信しなかった場合は接続に失敗したことを意味するため、別のピアへ接続要求を出す必要がある。また、P2Pネットワークからの退出においては

```
WeBoXP2P quit\n\n
```

というメッセージを送ることで、P2Pネットワークから抜けることを明示的に示すことができる。メッセージを受信したピアは、メッセージを送信してきたピアを接続ピアテーブルより削除する。また、ソケットが破棄された場合、ピアは、破棄されたソケットに割り当てられていたピアがP2Pネットワークから退出したと見なし、接続ピアテーブルからこのピアを自動的に削除する。他ピアからの接続があまりに多いピアは、接続要求をしてきたピアに対して、別のピアを紹介することもできる。

```
WeBoXP2P introduce\n
ピアアドレス 1:ポート番号 1\n
ピアアドレス 2:ポート番号 2\n\n
:
```

このメッセージを受信したピアは、紹介されたリストの中から適当なピアを選び、接続要求を行う。これらの紹介機能を使用することにより、ピアは負荷分散を行うことができる。なお、接続可能ピア数は設定ファイルで指定することができ、その数を超えて接続要求が発生したときは、自動でこのメッセージを返すようになっている。

問合せにおいては、

```
Query\n
Query-ID: 2003010203040533@192.168.1.13\n
URL: www-nishio.ist.osaka-u.ac.jp\n
Contents: Satoshi Nakamura\n
HopLimit: 5\n\n
```

というようなフォーマットをとる。問合せにおいてキーとして利用できるのは表1のとおりである。問合せにおいてすべてのキーを指定する必要はないが、これらのキーを組み合わせることで、柔軟な検索をすることができる。なお、問合せ時に作成するQuery-IDは問合せ結果を受信するときのために問合せテーブルに格納しておく。問合せを受信したピアは問合せテーブルをチェックし、格納されていないQuery-IDを持つ問合せを受信したときは、問合せテーブルにQuery-

表1 Queryのキーリスト

Table 1 Key list of query.

キー	意味
Query-ID	一意な値を設定する。必ず設定しなければならない。
URL	URLを指定した問合せに利用。部分URLでもよい。
Title	タイトルを指定した問合せに利用。ただし、クライアントによってはタイトルが変更されていることがある。
Size	コンテンツのサイズを指定した問合せに利用。バイト単位で指定。範囲指定の場合はハイフンで区切る。
LastUpdate	コンテンツの更新日時を指定した問合せに利用。YYYY/MM/DD HH:MM:SSのフォーマットで指定。範囲指定の場合はハイフンで区切る。
Folder	ウェブコンテンツが保存されているフォルダ名を指定した問合せに利用。
Contents	保存されているウェブコンテンツが含む文字列を指定した問合せに利用。
HopLimit	何ホップ先まで問合せを行うかを指定する。デフォルトでは5。リレーされるたびに減らす。
LifeTime	問合せの有効期限を世界標準時刻で指定する。期限切れになった問合せは自動的に破棄される。

IDと問合せ内容、送信元ピア用ソケットを格納する。このとき、ローカルディスクに対して検索を行い、HopLimitを1減らした後、接続ピアテーブルの中から問合せを送信してきたピアを除いたすべてのピアに対して同様の問合せを送信する。ただし、HopLimitが0のときや、LifeTimeを過ぎているときは問合せを送信しない。また、受信したことのあるQuery-IDである場合は、受信した問合せを破棄することによりループを回避する。

問合せ結果は、

```
Queryhit\n
Query-ID: 2003010203040533@192.168.1.13\n
Host: 192.168.1.100\n
Port: 20000\n
URL: http://www-nishio.ist.osaka-u.ac.jp/~nakamura/\n
Title: Satoshi Nakamura's Web Page\n
Size: 8023\n
LastUpdate: 2002/12/31 22:35:34\n
Folder: 西尾研究室\n\n
```

というようなフォーマットをとる。この問合せ結果を収集し、リスト表示してユーザに提示する。なお、QueryhitにはQueryと同じIDがついているため、問合せテーブルをチェックすることで問合せを送信してきたピアを順次たどることができる。つまり、問合せの送信元ソケットに対してQueryhitを送信する処理を繰り返すことにより、Queryを作成したピアに対

して、Queryhit が届けられることになる。

データの取得においては、現バージョンではデータを持つピアに直接接続し、HTTP プロトコルの GET リクエストを利用することでコンテンツの取得を行う。

```
GET http://www-nishio.ist.osaka-u.ac.jp/~nakamura/
```

GET リクエストを受信したピアは応答し、HTTP のヘッダ情報とあわせてコンテンツを転送する。

なお、ピアがネットワークから退出した場合には、問合せテーブルに格納されているそのピアに関連する情報もあわせて削除する。

4.3 流通制御

コンテンツ流通を制御する機能については 3 章で述べたとおりである。本プロトタイプシステムではウェブコンテンツ内部に P2P 関連の保存や転送の制御を行う P2P タグを埋め込むことで流通制御機能を実現する。以下は P2P タグをウェブコンテンツに埋め込んでいる様子である。コンテンツの保存および転送を許可し、改変は禁止している。また、転送回数制限が 5 に指定されている。

タグ内で指定できる設定の一覧を表 2 に示す。

これらのタグはウェブコンテンツを受信する際に読み取り、関連する制御を行う。たとえば、保存不許可のタグが指定してある場合は、エラーメッセージを出力し、コンテンツを保存しない。また、ユーザの保存操作（完全保存・部分保存）がタグで禁止されている場合は、エラーメッセージを出力し、保存しない。さらに、保存通知が設定されている場合は、指定された方法で通知を行う。たとえば、mailto が指定してある場合は、メールにより通知を行う必要がある。なお、コンテンツを保存する際には、ユーザのコンテンツ保存操作に加え、コンテンツのタイトルや URL、ファイルサイズ、コンテンツの MD5 (Message Digest 5) 値もあわせて保存する。

コンテンツの改変に関しては、コンテンツの最終更新日、サイズ、あらかじめ保存しておいたコンテンツの MD5 値などを利用してチェックを行う。本システムはブラウザ部にエディタ切替え機能を実装しており、ユーザは簡単にコンテンツを編集できるが、改変禁止と指定されているコンテンツについては編集できないようになっている。なお、本システムでは改変が行われたコンテンツの転送に関しては簡単に制限することができるが、保存されたコンテンツの改変チェックは困難なものとなっている。たとえば、他のエディタでローカルディスクに保存されたコンテンツを開き、編

表 2 P2P 設定一覧

Table 2 Options for the P2P tag.

キー	意味
AuthorID	コンテンツ作成者の ID。指定する必要はない。自作のコンテンツの追尾などに利用。
Contact	通知先アドレスを指定する。
Save	保存制限に関する設定。allow または deny で指定。これを allow にすると保存に関する設定はすべて allow になる。以下 allow/deny を a/d と表記。
SaveFull	完全保存制限。a/d
SavePart	部分保存制限。a/d
Convert RelativePath	相対パス変換に関する設定。 デフォルトは 1。
Change	改変に関する設定。a/d
SaveRequire Contact	保存時に指定されたアドレスに通知する必要あり。デフォルトは 0 で 1 を設定すると要通知となる。
Forwarding	転送制限に関する設定。a/d。これを allow にすると転送に関する設定はすべて allow になる。
ForwardingFull	完全保存時の転送制限。a/d
ForwardingPart	部分保存時の転送制限。a/d
ForwardingChange	改変時の転送制限。a/d
LimitOf Forwarding	転送回数制限。回数を数字で指定。
Forwarding OnlyDisconnection	サーバとの接続が確立できないときの転送設定。a/d
Forwarding RequireContact	転送時に指定されたアドレスに通知する必要あり。

集した場合はチェックすることができないため、コンテンツの改変を許してしまう。今後は、本システムが保存改変制限されたコンテンツが改変されたことを検知した場合には、コンテンツを自動的に削除するような機能などを考えていく必要がある。

転送制限においては、他のピアから転送要求が発生したときに初めてチェックを行い、コンテンツが完全保存されているのか、部分保存されているのか、また改変されているかを保存されたユーザの操作や、改変チェックを利用してチェックし、転送が禁止されている場合は転送をとりやめ、許可されている場合は応答する。また、このとき転送回数もチェックし、転送回数が 1 以上のときは転送し、その回数を減らす。切断時のみ許可の場合、サーバへのコネクション確立を試み、コネクション確立に失敗したときのみ転送を許可する。転送通知については、保存通知と同じように動作する。

なお、P2P に関するタグが設定されていない場合、デフォルトの設定では、部分・完全ともに保存許可、ローカルでの使用については改変も許可、転送は不許可となっている。



図 2 スクリーンショット

Fig. 2 A screenshot of a system.



図 3 P2P 問合せ用ダイアログ

Fig. 3 A dialogue for query creation.

名前	URL
国見 4発で王手	http://192.168.1.176:80/www.yomiuri.co.jp/whochi#so...
小森王国だ！ 既に3連覇視野	http://192.168.1.176:80/www.yomiuri.co.jp/whochi#so...
国見 GK・徳重が通和入りへ	http://192.168.1.176:80/www.yomiuri.co.jp/whochi#so...
2002年ボウリングのページ	http://192.168.1.176:80/www.nishio.ise.eng.osaka-u...
1978掲示板	http://192.168.1.176:80/www.nishio.ise.eng.osaka-u...
松本ユリテイル memo	http://192.168.1.176:80/www.stryokokuac.jp/~kjm#...
ASAHI	http://192.168.1.176:80/www.asahi.com/index.html
K. Moriyama's diary	http://192.168.1.176:80/www.moriyama.com#diary#42...

図 4 問合せ結果

Fig. 4 Results of a query.

4.4 実装環境

ウェブコンテンツ共有システムを、Microsoft Windows 2000 Professional, Microsoft Visual C++ 6.0 上で実装した¹⁶⁾。クライアントシステムには、Microsoft Internet Explorer のブラウザコンポーネントを利用した。現段階ではウェブコンテンツ共有システムのプロトコルは日本語のみの対応となっている。今後、多言語への対応に向け、テキストのエンコードシステムなどを実現する予定である。

図 2 は、クライアントシステムを利用してコンテンツを閲覧している様子を表している。また、図 3 は P2P での問合せ用インタフェースを、図 4 は検索結果のリストをクライアントシステムで表示している様

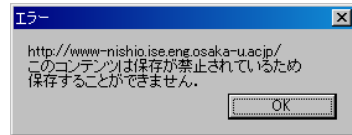


図 5 エラー表示

Fig. 5 Show an error message.

子である。さらに、図 5 は保存不可のコンテンツを保存しようとして、警告を受けている様子である。

5. 考 察

ウェブコンテンツ共有システムの実現により、ユーザ間で保持する情報が共有され、すでにリンク切れとなった情報や、昔掲載されていた情報などにアクセスすることができる。また、部分保存の許可により、ユーザは不要な部分をカットして、必要な部分だけウェブページを保存することができるため、ハードディスクを圧迫しないなど、ユーザにとってメリットは大きいと考えられる。一方、広告バナーがついているようなウェブページで、完全保存のみを許可している場合、転送されるコンテンツにはつねに広告が付加される。そのため、広告が含まれるようなページの場合、ウェブコンテンツ共有ネットワークを介して多くのユーザに広められることにより広告効果が増すことが考えられる。

今回実現したシステムでは、コンテンツを取得する際、コンテンツを保持しているピアに対して直接通信を行う必要があるため、外部ネットワークからコネクションを確立できないようなプライベートなネットワークにコンテンツを保持しているピアが属している場合は、コンテンツを要求することができないという問題があった。今後は、コンテンツ要求が発生したとき、コンテンツ保持ピアからコンテンツ要求ピアに対して順次パケットリレー式にコンテンツを転送する仕組みの実装を行う予定である。パケットリレー式のコンテンツ転送により、どのようなネットワークに属していても、コンテンツを確実に取得することができる。

現在のプロトタイプシステムでは、コンテンツ検索の高速化についてはまったく考慮していない。実際にウェブコンテンツ共有環境を巨大な P2P ネットワーク上で運用する場合には、このコンテンツ検索にかかる時間が重要になると考えられる。P2P ネットワークにおいて、高速にコンテンツを発見する手法として、広域コンテンツ発見に適している分散ハッシュテーブル²⁾ や、局所的コンテンツ発見に適している Attenuated Bloom Filter¹²⁾ などがある。今後、こうした手

法をウェブコンテンツ共有システムに応用した検索システムについても考えていく必要がある。なお、一般的な P2P システムとは異なり、ウェブコンテンツ共有システムではコンテンツ内部に含まれる語句を利用した検索が多く発生すると考えられる。このコンテンツ内部に対する検索は負荷が大きく、各ピアでの検索にかかる時間が、ウェブコンテンツ共有システム全体の検索性能のボトルネックになる可能性がある。コンテンツの問合せや問合せ結果を中継ピアが随時キャッシュしておき、同一の問合せが発生したときにはキャッシュされた問合せ結果を即座に返信するように改良すると、こうした検索における負荷を軽減することができると考えられる。

また、一般的なウェブサービス同様、人気のあるコンテンツを保持しているピアにはアクセスが集中し、ピアがサービス不能状態に陥るという問題がある。この問題は、先に述べたバケツリレー式のコンテンツ転送を利用し、コンテンツをリレーする際に中継ピアでコンテンツをキャッシュしておき、同一のコンテンツに対する問合せが発生したときには、中継ピアでそのキャッシュを要求ピアに転送することで解決できると考えられる。中継ピアにおけるコンテンツのキャッシュを効率良く利用することで、一部ピアに負荷が集中することを避けることができると考えられる。しかし、キャッシュをどこまで保存するか、キャッシュをいつタイムアウトするのかなど難しい点が多い。

本研究では、コンテンツ作成者が自作のコンテンツの流通を制御できる方法を考え、保存および転送に関する制限機能をいくつか実現した。これらの機能により、コンテンツ作成者は最低限のコンテンツ流通制御が可能となるが、これらの機能では不十分である。たとえば、設定を途中で変えることができないため、すでに出回ってしまったコンテンツに対して転送不可や保存不可などの条件を加えることができず、いったん流通させてしまうと、後で制御方法を変えることができない。また、必要なタグを付与せずに流通させてしまった場合、現時点でどの程度コンテンツが広がっているかどうかを調べることができない。さらに、自作のコンテンツが不当に改ざんされ、広められてしまったときに対処できない。こうした流通制御を実現するため、保存や転送などの基本的な機能に加え、コンテンツ更新機能、コンテンツ監視機能、コンテンツ追尾機能、コンテンツ削除機能について考える必要がある。

コンテンツ更新機能は、すでに各ユーザのローカルディスクに保存されているコンテンツに修正を加えるために利用する。コンテンツの更新が可能となると、

コンテンツに誤りがあり、コンテンツが流通することにより間違った情報が広まる危険性があるような場合に、コンテンツを保持しているピアに対してコンテンツの更新を要求することで、誤った情報が広まってしまふことを防ぐことができると考えられる。また、コンテンツの拡散を防止したいときに転送制限の設定を更新することにも利用できる。

コンテンツ削除機能は、コンテンツ作成者が望んでいないコンテンツが他ユーザに保存され、共有されているときなどに、ユーザのローカルディスクから削除できるような機能である。この機能を使うことにより、著作権を無視して改変されたようなコンテンツを削除できる。なお、削除については、権限の問題もあるため、認証などの別の機能が必要となる。

コンテンツ追尾機能は、自分が作成したコンテンツが現在どこに存在しているのかを調べるために利用する。この機能により発見したコンテンツに対して、コンテンツの更新機能や削除機能を適用することにより、効率良くコンテンツの流通を制御することができる。

コンテンツ監視機能は、他のピアに自分のコンテンツを検出したときに通知をするよう依頼する機能である。この機能により、コンテンツ作成者は、自分のコンテンツがどのように使用されているのかをチェックできる。また、この機能は自作のコンテンツが改ざんされ広められていないかをチェックすることにも利用できる。改ざんされているコンテンツを発見したときには、削除機能を利用して削除するなどの対処ができる。

しかし、これら機能を実現しても、コンテンツ自体を操作し、コンテンツ情報テーブルを変更されると対処できない。今後、コンテンツの暗号化による保護についても考えていく必要がある。

6. ま と め

本研究では、P2P のネットワークを利用したウェブコンテンツ共有環境について述べ、クライアントシステム、サーバントシステムを実装した。これらのシステムを利用することにより、ネットワークの断線やサーバの停止状態でウェブサーバにアクセスできないときや、コンテンツが改変されてしまい以前のコンテンツにアクセスできないときなどでも、ユーザは P2P ネットワークからコンテンツ（情報）を手に入れることができる。また、コンテンツ作成者が自作のコンテンツの流通を制御するため、いくつかの仕組みを提案、実装した。保存や、転送に制限を与えることにより、コンテンツの流通をある程度制御することができる。

現在、P2P コンテンツ共有機能を除いたクライア

ントシステムをインターネット上に公開している。現時点でどの程度のユーザがいるかを把握してはいるが、1カ月に1万回以上クライアントシステムがダウンロードされていることから、かなりの数のユーザがローカルディスクにコンテンツを収集してオフライン状態で閲覧利用していると考えられる。ユーザ数の増加により、ウェブ共有環境はより良いものになると考えられるが、P2P対応版を公開するためにはP2Pネットワークの特性などについて十分に調査し、また流通制御機能についても強化していく必要がある。そのため、今後は、コンテンツを追跡する機能や、削除する機能などに加え、より多くの流通制御の仕組みも考案し、実装していく予定である。また、CGIやServletなど動的に変化するようなページに関しても考えていく予定である。

謝辞 本研究の一部は、文部科学省21世紀COEプログラム(研究拠点形成費補助金)、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」によるものである。ここに記して謝意を表す。

参 考 文 献

- 1) Rowston, A. and Druschel, P.: Storage Management and caching in PAST, a large scale, persistent peer-to-peer storage utility, *Proc. ACM SOSP 2001* (2001).
- 2) Zhao, B., Kubiatowicz, J. and Joseph, A.: Tapestry: an infrastructure for fault-tolerant wide-area location and routing, Technical Report UCB/CSD-01-1141, Computer Science Division (2000).
- 3) Dabek, F., Kaahoeck, M., Kaeger, D., Morris, R. and Stoica, I.: Wide-area cooperative storage with CFS, *Proc. ACM SOSP '01* (2001).
- 4) Gnutella. <http://www.gnutella.com/>
- 5) Google Technology Inc.: Google. <http://www.google.com/>
- 6) Clarke, I., Wiley, B. and Hong, T.W.: Freenet: a distributed anonymous information storage and retrieval system, *Proc. ICSI Workshop on Design Issues in Anonymity and Unobservability* (2000).
- 7) Internet Archive. <http://www.archive.org/>
- 8) Kubiatowicz, J., Bindel, D., Chen, Y., Czerwinski, S., Eaton, P., Geels, D., Gummadi, R., Rhea, S., Weatherspoon, H., Weimer, W., Wells, C. and Zhao, B.: OceanStore: An architecture for global-scale persistent storage, *Proc. ASPLOS 2000* (2000).
- 9) Jnutella. <http://www.jnutella.org/>
- 10) Sun Microsystems, Inc.: JxtaProject. <http://www.jxta.org/>
- 11) Roxio, Inc.: Napster. <http://www.napster.com/>
- 12) Rhea, S. and Kubiatowicz, J.: Probabilistic location and routing, *Proc. IEEE INFOCOM '02* (2002).
- 13) MediaPolice., Co. Ltd: ThinkBook. <http://www.mediapolice.com/>
- 14) Nakamura, S.: WeBoX. <http://www-nishio.ist.osaka-u.ac.jp/~nakamura/webox/>
- 15) 伊藤直樹: P2P コンピューティング—技術解説とアプリケーション, ソフトリサーチセンター.
- 16) 中村聡史, 塚本昌彦, 西尾章治郎: コンテンツ流通制御機能をもつウェブコンテンツ共有システム, 電子情報通信学会データ工学ワークショップ論文集 (2003).

(平成 15 年 5 月 6 日受付)

(平成 15 年 10 月 16 日採録)



中村 聡史

1999年大阪大学工学部情報システム工学科卒業。現在、同大学院工学研究科情報システム工学専攻博士後期課程在籍。インタフェース、ウェアラブルコンピューティング、P2Pに興味を持つ。ヒューマンインタフェース学会会員。



塚本 昌彦(正会員)

1987年京都大学工学部数理工学科卒業。1989年同大学院工学研究科修士課程修了。同年シャープ(株)入社。1995年大阪大学大学院工学研究科情報システム工学専攻講師、1996年同専攻助教授、2002年より同大学院情報科学研究科マルチメディア工学専攻助教授となり、現在に至る。工学博士。ウェアラブルコンピューティングとユビキタスコンピューティングの研究に従事。ACM, IEEE等8学会の会員。



西尾章治郎(フェロー)

1975年京都大学工学部数理工学科卒業．1980年同大学院工学研究科博士後期課程修了．工学博士．京都大学工学部助手，大阪大学基礎工学部および情報処理教育センター助

教授，大阪大学大学院工学研究科情報システム工学専攻教授を経て，2002年より同大学院情報科学研究科マルチメディア工学専攻教授となり，現在に至る．2000年より大阪大学サイバーメディアセンター長，2003年より大阪大学大学院情報科学研究科長を併任．この間，カナダ・ウォータールー大学，ビクトリア大学客員．データベース，マルチメディアシステムの研究に従事．現在，ACM Trans. on Internet Technology，Data & Knowledge Engineering，Data Mining and Knowledge Discovery等の論文誌編集委員．情報処理学会フェロー含め，ACM，IEEE等9学会の会員．
