

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

中村 聡史[†] 塚本 昌彦[‡] 西尾 章治郎[‡]

[†] 大阪大学大学院工学研究科 〒565-0871 大阪府吹田市山田丘 2-1

[‡] 大阪大学大学院情報科学研究科 〒565-0871 大阪府吹田市山田丘 2-1

E-mail: [†] nakamura@ise.eng.osaka-u.ac.jp, [‡] {tuka, nishio}@ist.osaka-u.ac.jp

あらまし 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、インターネット上で相互接続することによりコンテンツを共有する P2P 型のシステムを提案する。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。すでに存在しない過去に公開されていたコンテンツも、他のシステムが保存している場合には、アクセスできるようにする。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。

キーワード P2P, ウェブ, データ共有, コンテンツ配送, 流通制御

A web contents sharing system with contents delivery control functions

Satoshi NAKAMURA[†] Masahiko TSUKAMOTO[‡] and Shojiro NISHIO[‡]

[†] Graduate School of Engineering, Osaka University

2-1 Yamadaoka, Suita-shi, Osaka, 565-0871 Japan

[‡] Graduate School of Information Science and Technology, Osaka University

2-1 Yamadaoka, Suita-shi, Osaka, 565-0871 Japan

E-mail: [†] nakamura@ise.eng.osaka-u.ac.jp, [‡] {tuka, nishio}@ist.osaka-u.ac.jp

Abstract 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.

Keyword Peer-to Peer, Web, Date Sharing, Contents Delivery, Delivery Control

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、アクセスした時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なシステムが開発されている[12,14]。また、Google のキャッシュサービス[4]や、Internet Archive[6]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、Gnutella[3]や Napster[10], Freenet[5]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[12]。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[1,2,5]。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可能とす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 章にて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、取得日時などもまとめて保存する。また、ユーザのコンピュータ上で動作するサーバントは、近くに存在するサーバントを探し、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）を探し、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。例えば、「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。接続サーバントは順次問合せをリレーする。問合せの条件に一致するコンテンツをもつサーバントは自動的に応答し、自分の IP Address や問合せに合致するコンテンツの URL、コンテンツのサイズ、更新日時、取得日時などの情報を問合せ元に返す。この応答により、ユーザは指定したコンテンツをもつユーザ（サーバント）を調べることができ、問合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、同じ趣味をもつユーザ同士が集まりプライベートなコンテンツ共有ネットワークを構成し、各自が保持しているウェブコンテンツを共有することで、よりよ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時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常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等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、部分保存を不許可とすることで広告などがカットされることを未然に防ぐことができる。

通常のウェブコンテンツ取り込みでは、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を相対パスへと変換し、一緒に取り込む。相対パス変換制限は、ローカルでも閲覧可能とするための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閲覧できないため、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。例えば、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こむことや、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 などで通知・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時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。例えば、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ことも可能である。

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、転送する度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アクセスが困難である場合のみ、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、ウェブサーバが問題なく運用されている時は、ユーザはウェブサーバからコンテンツを取得し、問題があるときにだけ他ユーザからコンテンツを手に入れるといったように使いわけることができる。

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 などで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システムで構成される。

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

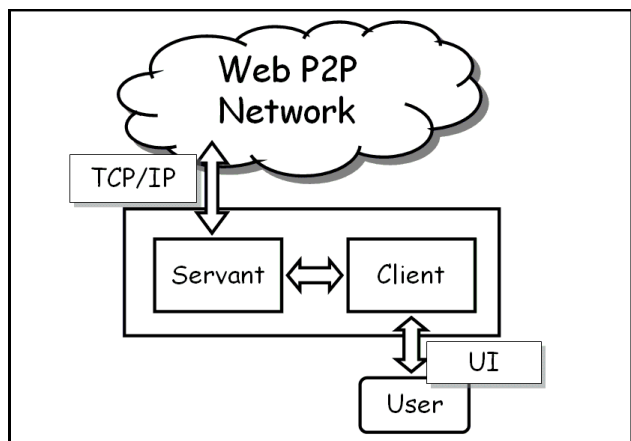


図 1. システム構成

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合せたインタフェースを利用し，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，コンテンツ内部に含むキーワードでの検索などをサポートする。これらを利用することにより，意図する情報を見つけるまでの速度が向上すると考えられる。

P2P ネットワークを構成するサーバント同士の接続には，TCP/IP を利用して相互に接続する。P2P のシステムとして，jnutella[6]や Jxta[7]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，各サーバントは，接続可能サーバントリストをもち，他のサーバントに接続することにより P2P ネットワークへと参加する。また，サーバントは現在直接接続しているサーバントとソケットをセットにした接続サーバントテーブルと，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

テーブルをもち，問合せのリレーなどに利用する。問合せがきた場合は，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。サーバント同士の詳しいやり取りについては 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，サーバントシステム，クライアントシステムが内部に流通制御部をもち，「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

現状で実装されているサーバント間でやりとりするプロトコルは以下の通りである。

- 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
- 別サーバントの紹介
- 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サーバントは，送信してきたサーバント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，接続サーバントテーブルに要求を送ってきたサーバントを追加する。このメッセージを受信したサーバントは，P2P ネットワークへの参加が認められたことを確認し，接続サーバントテーブルに接続したサーバントを追加する。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，別のサーバント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サーバントは，メッセージを送信してきたサーバントを接続サーバントテーブルより削除する。また，ソケットが破棄された場合，サーバントは，破棄されたソケットに割り当てられていたサーバントが P2P ネットワークから抜けたとみなし，接続サーバント

トテーブルからこのサーバントを自動的に削除する。
 他サーバントからの接続があまりに多いサーバントは、接続要求をしてきたサーバントに対して、別のサーバントを紹介することもできる。

```
WeBoXP2P introduce¥n
サーバントアドレス 1:ポート番号 1¥n
サーバントアドレス 2:ポート番号 2¥n
:
```

このメッセージを受信したサーバントは、適当なサーバントを選び、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、サーバントは負荷分散を行うことができる。

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: 中村聡史¥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、これらのキーを組合せることにより、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

表 1. Query のキーリスト

キー	意味
Query-ID	一意な値を設定する。必ず指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。

合せ結果を受信する時のために問合せテーブルに格納しておく。問合せを受信したサーバントは問合せテーブルをチェックし、格納されていない Query-ID をもつ問合せを受信したときは、問合せテーブルに Query-ID と問合せ内容、送信元サーバント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、接続サーバントテーブルの中から送信元を除いたすべてのサーバントに対して同様の問合せを送信する。ただし、HopLimit が 0 の時や、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/
File: ¥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サーバント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 を作成したサーバントに対して、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もつサーバントに直接接続し、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/
```

この手法では、コンテンツをもつサーバントに対してコネクションをはれないと情報が取得できない。例えば、グローバルなネットワークに所属するサーバントから企業などのプライベートネットワークに所属するサーバントに対しては、コネクションをはることができないため、折角コンテンツを保持するサーバントを発見しても、コンテンツ自体にたどりつけないという問題がある。GET のリレーを行い、コンテンツを問合せ元まで配送する手法を取るによりこうした問題に対応できると考えられる。このコンテンツリレー手

法の実装は今後の課題である。

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章で述べた通りである。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。

```
<html>
<head>
<title>Satoshi Nakamura's Web Page</title>
<p2p
  Save="allow" Forwarding="allow" Change="deny"
  LimitOfForwarding=5
  Contact="mailto:nakamura@ise.eng.osaka-u.ac.jp"/>
</head>
:
```

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、コンテンツを保存する際にタイトルやURL、ファイルサイズなどと一緒に記録する。また、ユーザのコンテンツ保存操作が完全保存だったか、部分保存だったかについてもフラグとしてあわせて記録する。保存不許可のタグが指定してある場合は、エラーメッセージが出力され、コンテンツは保存しない。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。例えば、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、デフォルト設定では自動的にシステムが通知するようになっている。

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、コンテンツのMD5ダイジェスト値などでチェックを行い、変更されている場合は改変に関するフラグをたてる。本プロトタイプシステムでは、改変不許可のコンテンツについて、改変が行われた場合の転送に関する制限は与えているが、保存に関しては制限を加えていない。保存改変制限に関して、コンテンツが改変された場合は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。また、この改変チェックは完璧ではないため、新たな改変チェックに関する手法も考えていく必要がある。

転送制限においては、他のサーバント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フラグでチェックし、応答する。また、このとき転送回数もチェックし、転送回数

表 2. P2P 設定一覧

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.
Email	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
LimitOfForwarding	転送回数制限. 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.

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に関する設定がしてない場合、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。

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。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。

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。



図 2. クライアントシステムでのブラウジング



図 3. 問合せインタフェース

名前	URL
国見 4発で王手	http://192.168.1.176-80/www.yomiuri.co.jp
小嶺王国だ！既に3連覇視野	http://192.168.1.176-80/www.yomiuri.co.jp
国見 Gk・徳重が浦和入りへ	http://192.168.1.176-80/www.yomiuri.co.jp
2002年ボウリングのページ	http://192.168.1.176-80/www-nishio.ise.eng
1976掲示板	http://192.168.1.176-80/www-nishio.ise.eng
セキュリティホール memo	http://192.168.1.176-80/www.str.yukok.u.ac
ASAHI	http://192.168.1.176-80/www.asahi.com
K.Moriyama's diary	http://192.168.1.176-80/www.moriyama.com

図 4. 検索結果の表示

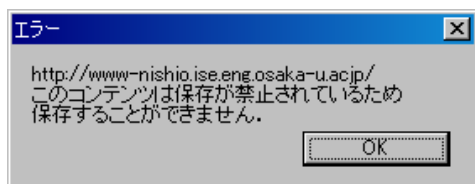


図 5. 保存制限による警告

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、ハードディスクを逼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常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。

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。しかし、これらの機能では不十分である。例えば、設定を途中で変更す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、一旦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。例えば、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ことを防ぐため、クライアント（サーバント）に対してコンテンツの更新を要求する際に利用できる。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、効率よ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、他のサーバント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確認することにも利用でき、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。

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全く考慮していない。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 Filter[12]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。

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ないが、1ヶ月に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よ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] A. Rowston and P. Druschel, "Storage Management and caching in PAST, a large scale, persistent peer-to-peer storage utility," Proc. ACM SOSP 2001, Oct. 2001.
- [2] B. Zhao, J. Kubiatowicz, and A. Joseph, "Tapestry: an infrastructure for fault-tolerant wide-area location and routing," Technical Report UCB/CSD-01-1141, Computer Science Division, April 2000.
- [3] F. Dabek, M. Kaahoe, D. Kaeger, R. Morris, and I. Stoica, "Wide-area cooperative storage with CFS," Proc. ACM SOSP '01, Oct. 2001.
- [4] Gnutella, <http://www.gnutella.com/>.
- [5] Google, <http://www.google.com/>.
- [6] I. Clarke, Sandberg, B. Wiley, and T. W. Hong, "Freenet: a distributed anonymous information storage and retrieval system," Proc. ICSI Workshop on Design Issues in Anonymity and Unobservability, June 2000.
- [7] Internet Archive, <http://www.archive.org/>.
- [8] J. Kubiatowicz, D. Bindel, Y. Chen, S. Czerwinski, P. Eaton, D. Geels, R. Gummadi, S. Rhea, H. Weatherspoon, W. Weimer, C. Wells, and B. Zhao, "OceanStore: A architecture for global-scale persistent storage," Proc. ASPLOS 2000, Nov. 2000.
- [9] Jnutella, <http://www.jnutella.org/>.
- [10] JxtaProject, <http://www.jxta.org/>.
- [11] Napster, <http://www.napster.com/>.
- [12] S. Rhea and J. Kubiatowicz, "Probabilistic location and routing," Proc. IEEE INFOCOM '02, June 2002.
- [13] ThinkBook, <http://www.mediapolice.com/>.
- [14] 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.
- [15] 紙, <http://www.ki.rim.or.jp/~kami/>.