2008－EVA－24 (6)
2008／3／19

非構造化P2Pファイル共有ネットワークにおけるポイゾニングによるファイル流通制御方式の提案と評価

吉田雅裕[†1] 大坐畠智[†2] 川島幸之助[†2]

非構造化 P2P ファイル共有ネットワークは，プロトコルにファイル流通制御機能を備えていない場合がある．そのようなネットワークでは，既にネットワーク上を流通しているファイルを制御することが難しい．そのような場合は，ポイゾニングを用いることが有効だと言われている．そこで本稿では，Winny ネットワークを対象に，ポイゾニングによるファイル流通制御を行うための制御ノードを実装し，評価実験によりポイゾニングの効果を確認する．

An evaluation of a file distibution control method by the poisoning technology for the unstructured P2P file sharing network

Masahiro Yoshida,[†1] Satoshi Ohzahata[†2] and Konosuke Kawashima[†2]

Some unstructured P2P file sharing applications do not have a file eliminate in their protocol. It is difficult to control a file distribution in such networks. In those case, the poisoning is effective. In this report, we implement poisoning system for Winny network, and the results of evaluation experiment show effectiveness of poisoning methods.

1. はじめに

近年，個人が所有する計算機の高性能化やネットワークのブロードバンド化に伴い，コンピュータネットワークにおける P2P 技術が注目されている．この P2P 技術を応用したアプリケーションには P2P 音声通信ソフトや P2P ファイル共有ソフトなどが存在する．特に P2P ファイル共有ソフトは国内外で大きく広まっているが，この P2P ファイル共有ソフトを介した著作権侵害や個人情報流出が大きな問題となっている．国内ではピュア P2P ファイル共有ソフトの 1 つである Winny が広く利用されているが，Winny プロトコルはファイルの流通制御機能を備えていない．そのため，既に Winny ネットワーク上を流通しているファイルを第三者の手でネットワーク上から完全に削除することは不可能である[1]．参考文献 2) によると，2006 年 10 月 6 日から 10 月 15 日の Winny ノード数の調査では，24 時間当たり約 35 万ノードが観測されたと報告している．東京農工大学 川島研究室の調査でも，2007 年 7 月 21 日から 7 月 28 日の調査で，24 時間当たり約 30 万ノードが観測されている．このように，Winny ネットワークは数十万規模のネットワークを形成するが，このネットワーク上を流通するファイルのほとんどが著作権を侵害したファイルである．さらに Winny には，Winny ネットワークに個人情報を流出させることを目的としたウィルスが存在する．このウィルスによる個人情報流出の問題は，近年さらに深刻化してきており，Winny ネットワークに対するファイル流通制御の要望が高まっている．

そこでこの問題を技術的に解決するために，ポイゾニングと呼ばれる手法が提案されている[3)4)]．本稿では，ポイゾニング手法を用いたファイル流通制御の効果を明らかにするために，Winny ネットワークに対してファイル流通制御を行うための制御ノードを実装し，ポイゾニング手法の評価を行う．

†1 東京農工大学 工学部
Faculty of Engineering, Tokyo University of Agriculture and Technology
†2 東京農工大学 大学院 共生科学技術研究院 先端情報科学部門
Institute of Symbiotic Science and Technology, Tokyo University of Agriculture and Technology

2. Winny のファイル流通方式

2.1 ファイルキーを用いたファイル検索

Winny におけるファイル流通には，ファイルの要約情報が格納されたファイルキーを利用する．Winny ではファイル流通におけるスケーラビリティの実現と，ファイル流通に要するトラヒックを削減するために，ファイル本体の代わりにファイルキーを流通させる．

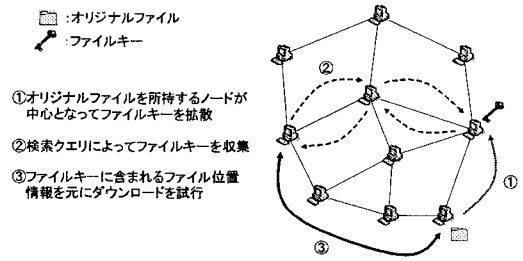

図 1 ファイルキーの拡散と検索
Fig. 1 File diffusion and search by the file key.

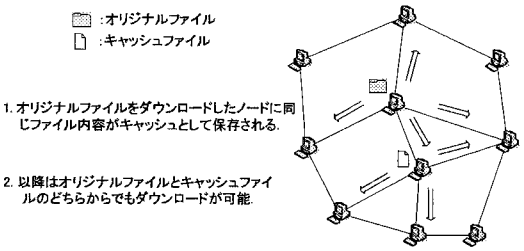

図 2 Winny ネットワークにおけるキャッシュシステム
Fig. 2 A cache system on Winny network.

ファイルキーには，アップロードファイルのファイル名，ファイルサイズなどの情報が含まれる．また，ファイルのネットワーク上での位置情報として，そのファイルをアップロードしているノードの IP アドレスとポート番号が含まれている．さらに，キータイマーと呼ばれる，ファイルキーのネットワーク上での生存時間を示すデータが含まれている．キータイマーが存在する理由は，ファイルをアップロードしているノードがネットワーク上に存在しなくなった場合に，ネットワーク上からファイルキーを消滅させるためである．

隣接ノードにファイル情報の問い合わせを行うことを検索クエリという．ファイルのダウンロードを行うノードは，まず，検索クエリによってネットワーク上からファイルキーを収集する．その後，ファイルキーに記載されたファイルの位置情報を元に，ノード間でファイルを転送する（**図 1**）．

2.2 キャッシュシステム

Winny では，P2P ファイル共有ネットワークにおける匿名性を実現するために，キャッシュシステムを利用している．キャッシュシステムとは，ノード間通信によって取得したデータをキャッシュとして保持し，その後の通信において同じデータが必要になった場合にはキャッシュに保持しておいたデータを使用する技術である．

ノードがファイルをダウンロードした場合は，そのノードにオリジナルファイルと同じデータ内容のキャッシュファイルが作成される．他ノードから同じファイルのダウンロード要求があった場合は，オリジナルファイルを所持するノードからだけでなく，そのオリジナルファイルと同じ内容のキャッシュファイルを保持するノードからもダウンロードが可能である（**図 2**）．キャッシュシステムを利用すれば，人気のあるファイルは頻繁にキャッシングされるため，P2P ファイル共有ネットワークは全体的に効率がよくなる．また，キャッシュシステムを使用することにより匿名性を高めることが可能である．これは，ファイルをダウンロードしているユーザは，そのファイルがオリジナルファイルなのかキャッシュファイルなのかを区別できないからである．

3. Winny におけるポイゾニング

3.1 インデックスポイゾニング

インデックスポイゾニング[3] とは，制御用に加工したファイル情報をネットワーク上に拡散させることで，ファイル流通制御を行う手法である．Winny ネットワークに対するインデックスポイゾニングは，ファイル流通制御用に加工したファイルキー（ダミーファイルキー）をネットワーク上に拡散する．拡散するダミーファイルキーは，ファイル名を制御対象ファイルと同じにし，ファイルの持ち主をネットワーク上に存在しない架空のノードにする．ノードが検索クエリによってこのダミーファイルキーを入手した場合は，ネットワーク上に存在しない架空のノードにダウンロードを試行することになるため，ダウンロードを失敗させることが可能である（**図 3**）．

3.2 アイテムポイゾニング

アイテムポイゾニング[4] とは，制御用に加工したファイル本体をネットワーク上に拡散することで，ファイル流通制御を行う手法である．Winny ネットワークに対するアイテムポイゾニングでは，ファイル名を制御対象ファイルと同じにし，ファイルの持ち主に制御用ノードの位置情報を格納したダミーファイルキーをネットワーク上に拡散する．ノードが検索クエリによってこのダミーファイルキーを入手した場合は，制

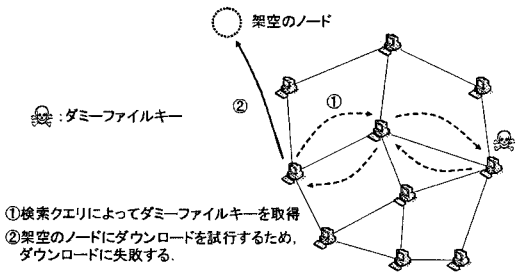

図 3 インデックスポイゾニング
Fig. 3 Index poisoning for Winny network.

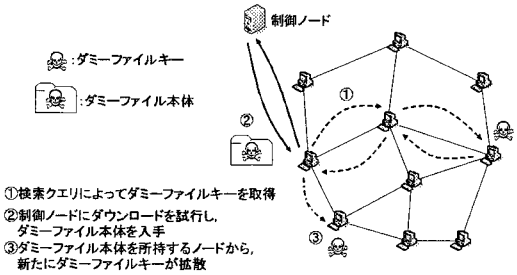

図 4 アイテムポイゾニング
Fig. 4 Item poisoning for Winny network.

御ノードにダウンロードを試行することになる．制御ノードは，他ノードから要求があった場合に制御対象ファイルとファイル内容が異なるが，ファイル名が同じファイル（ダミーファイル本体）をアップロードする．このダミーファイル本体は，制御対象ファイルとファイル内容が全く異なっており，制御対象ファイルを要求するユーザにとっては無益のファイルとなる．Winny では，ファイル本体を所持するノードからファイルキーが拡散されるため，ダミーファイル本体を所持するノードはダミーファイルキーを拡散する．さらに，Winny のキャッシュシステムによってダミーファイル本体も拡散される．このような理由から，ダミーファイル本体を所持するノードが，制御ノードと同等の働きをすることが期待できる（図 4）．

4. 実　　装

Winny ネットワークに対する制御ノードの実装は，Visual Studio2005 環境で開発言語に C#2.0 を用いて行った．制御ノードには，

- ネットワーク上をクローリング（巡回）し，各ノードが持つファイルキーを収集する機能
- ダミーファイルキーを配布する機能
- ダミーファイル本体をアップロードする機能

が実装されている．1 つの Winny ノードが所持できるファイルキーの最大数は 35,000 個であるが，実装した制御ノードが，1 つのノードに 35,000 個のファイルキーを配布するために要する時間は数秒である．また，マルチスレッド化により同時に 100 ノードにダミーファイルキーの配布が可能である．

5. 評　　価

5.1 評価環境

評価環境は，仮想 PC エミュレータの VMWare[5] を用いて，合計 61 ノードの Winny ネットワークを構築した．Winny には回線速度の速いノードを上流，回線速度の遅いノードを下流とする概念がある．構築した Winny ネットワークでは，上流に 15 ノード，中流に 35 ノード，下流に 5 ノードを割り当てた．これらのノードは，Winny プロトコルに従ってファイルキーの中継のみを行う．また，制御対象ファイルのダウンロードを試行するノードを 5 ノード参加させる．最後に制御対象ファイルをアップロードするノードを 1 ノード参加させた．ダウンロードを試行するノード，ならびに制御対象ファイルのアップロードを行うノードは中流に割り当てた（**図 5**）．また，あらかじめネットワーク上には制御に関係ない内容のファイルキーを 10,000 個流通させておく．これは，実ネットワークでは制御対象ファイル以外のファイルキーも流通していることを考慮に入れるためである．

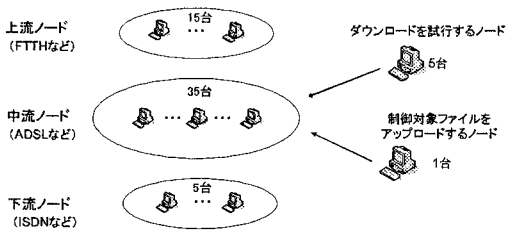

図 5 評価環境
Fig. 5 Evaluation setups.

5.2 評価シナリオ

評価シナリオを**図 6** に示す．1 回の評価時間は 2 時間とした．制御対象ファイルのアップロードを行うノードは，評価開始後すぐにネットワークに参加する．評価開始から 10 分後に，ダウンロードを試行するノードがダウンロードを開始する．制御開始から 1 時間経過した段階で，制御を終了させた．

5.3 評価項目

評価項目として

(1) ダウンロードを試行するノードが，制御対象ファイルのダウンロードを完了するまでに要した時間
(2) 制御対象ファイルのファイルキーの発見率
(3) 制御を行う際に制御ノードに発生したトラヒックの 3 つを設定した．

ダウンロードに要した時間の測定は，制御対象ファイルのダウンロードが終了した時刻と実験開始時刻との差分を取ることにより求めた．また，ダウンロード

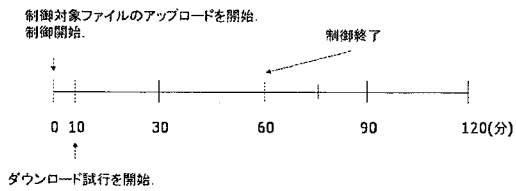

図 6 評価シナリオ
Fig. 6 Evaluation scenario.

に要した時間はダウンロードを試行するノード5台の平均を求めて算出した．

制御対象ファイルのファイルキーの発見率を求めるために，ポイゾニングによって制御されたWinnyネットワーク上を流通するファイルキーを収集した．制御対象ファイルのファイルキーの発見率は，式(1)のように定義した．

R:制御対象ファイルのファイルキーの発見率
A:制御対象ファイルのファイルキーの発見回数
B:ダミーファイルキーの発見回数

$$R = \frac{A}{A+B} \quad (1)$$

制御に要したトラヒックは，Windows標準のパフォーマンスモニタ機能を使用して計測した．パフォーマンスモニタを使用することで，制御ノードのNICに発生した送受信トラヒックを記録することができる．トラヒックは2時間当たりの送受信トラヒックの総計を算出した．

5.4 インデックスポイゾニングの評価

5.4.1 ダミーファイルキーの内容

インデックスポイゾニングでは，**表1**に示す内容のダミーファイルキーをネットワーク上に拡散する．制御ノードは制御開始後から3分間，ネットワーク上の各ノードを巡回しながらダミーファイルキーを配布する．制御ノードは一つのノードに数秒でダミーファイルキーを配布することが可能であり，なおかつ，同時に100ノードに配布することができるため，3分間だけの制御であっても十分であると考え，このような時間設定を行った．

5.4.2 ダウンロードに要した時間

インデックスポイゾニングの評価は，ネットワーク上に拡散するダミーファイルキーのユニーク数を変えながら評価を行った．**図7**に，インデックスポイゾニングによる制御において，ダウンロードを試行するノードが制御対象ファイルのダウンロードに要した時間の平均値を示す．図7には，制御対象ファイルのダウンロードに要した時間の最大値と最小値も記してある．

全く制御を行わない場合は，制御対象ファイルのダウンロードが平均18分で完了することがわかる．同様に，拡散するダミーファイルキーのユニーク数が少ないときは，全く制御を行わないときとほぼ同様の時間で，制御対象ファイルのダウンロードが完了することもわかる．しかし，ユニークなダミーファイルキーを1,000個拡散した場合は，制御開始から1時間以上経過しなければ，5台のダウンロードを試行するノードが制御対象ファイルのダウンロードを完了することができていない．このことから，ダミーファイルキーをネットワーク上に一定数だけ拡散すれば，ファイル流通制御効果を出すことができることがわかる．

表1 インデックスポイゾニングにおけるダミーファイルキーの内容
Table 1 Contents of a dummy file key for index poisoning.

項目	内容
ファイル名	制御対象ファイルと同じ
ファイルサイズ	制御対象ファイルと同じ
ファイル位置情報	架空ノード
キータイマー	3,600秒

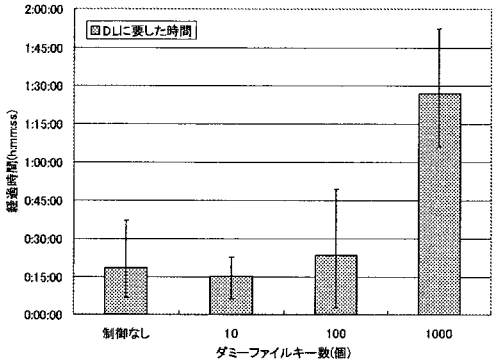

図7 インデックスポイゾニングによる制御でダウンロードに要した時間
Fig. 7 Time to download completion under index poisoning control.

5.4.3 制御対象ファイルのファイルキーの発見率

インデックスポイゾニングによる制御中の，制御対象ファイルのファイルキーの発見率を**図8**に示す．図8は，制御対象ファイルのファイルキーの発見率を1分間ごとにプロットしている．拡散するダミーファイルキーのユニーク数が少ないと，制御対象ファイルのファイルキーの発見率があまり下がっていない．しかし，ダミーファイルキーを1,000個配布した場合は，制御対象ファイルのファイルキーの発見率を1%未満にすることができている．5.4.2項において，配布したダミーファイルキーのユニーク数が1,000個の時は，ネットワーク上にダミーファイルキーが残っている間，DLノードによる制御対象ファイルのダウンロードを阻止することができていた．これは，ダミーファイルキーによって制御対象ファイルのファイルキーの発見率を低くすることができたからである．

また，ダミーファイルキーのキータイマは1時間に設定したにも関わらず，実際にネットワーク上からダミーファイルキーが消えたのは1時間を少し過ぎてからである．これはノード間でファイルキーの交換を行った際に，ファイルキーを受け取ってその中のキータイマの値を減らし始めるまでに多少の時間差があるからである．そのため，実際の時間経過とキータイマの値の間に差が生じてしまい，ネットワーク上からダ

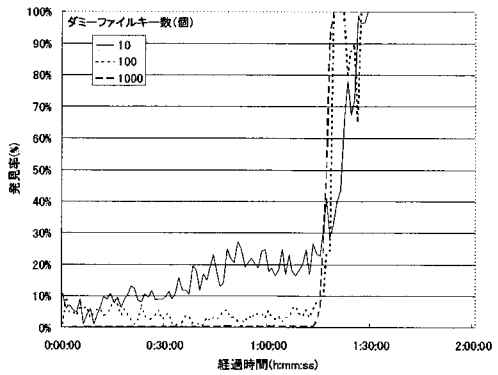

図 8 制御対象ファイルのファイルキーの発見率
Fig. 8 Discovery rate of target file key.

ミーファイルキーが消えるまでの時間が延びている.

5.5 アイテムポイゾニングの評価

5.5.1 ファイルキーの内容

アイテムポイゾニングでは,**表 2** に示す内容のダミーファイルキーをネットワーク上に拡散する.制御ノードはインデックスポイゾニングと同様に,制御開始後から 3 分間,ネットワーク上の各ノードを巡回しながらダミーファイルキーを配布する.その後,ダミーファイルキーの配布は終了するが,ダウンロードを試行するノードからダミーファイル本体のアップロードを要求された場合は,ダミーファイル本体をアップロードする.

制御対象ファイルとダミーファイル本体のファイルサイズを同じにすれば,ユーザがそれらを見分けることがさらに困難になり,ユーザがダミーファイル本体をダウンロードする確率が高くなることが期待できる.しかし,制御対象ファイルのファイルサイズが大きい場合,ダミーファイル本体のファイルサイズも大きくなってしまう.ダミーファイル本体のファイルサイズが大きければ,ダミーファイル本体のアップロードに要する時間が長くなり,短時間で大量のダミーファイル本体を拡散することが困難になる.その上,ダミーファイル本体の拡散のために大量のトラヒックが必要となる.Winny では自動ダウンロード機構が備わっており,ユーザが設定したキーワードをファイル名に持つファイルは自動的にダウンロードする仕組みになっている.よって,ダミーファイル本体のファイルサイズが小さい場合でも,ダミーファイル本体をダウンロードするノードは十分存在することが考えられる.以上の理由から,アップロードするダミーファイル本体は 151〜152Byte の小さなサイズとした.

5.5.2 ダウンロードに要した時間

アイテムポイゾニングの評価は,ネットワーク上に拡散するダミーファイルキーのユニーク数を変えながら評価を行った.拡散するダミーファイルキーのユニーク数が増加すれば,制御ノードに対するダミーファイル本体のダウンロード要求も増加することになる.

図 9 に,アイテムポイゾニングによる制御でダウンロードを試行するノードが,制御対象ファイルのダウンロードに要した時間の平均値を示す.図 9 には,制御対象ファイルのダウンロードに要した時間の最大値と最小値も記してある.アイテムポイゾニングにおいても,拡散するダミーファイルキーのユニーク数が 1,000 個の時は,ネットワーク上にダミーファイルキーが残っている間,ダウンロードを試行するノードによる制御対象ファイルのダウンロードを阻止することができている.

また,インデックスポイゾニングでは拡散するダミーファイルキーのユニーク数が少ない時は,ダウンロードを試行するノードがダウンロードに要した時間が,全く制御を行わない場合とほとんど変わらなかった.しかし,アイテムポイゾニングでは拡散するダミーファイル本体の数が増えるにつれ,制御対象ファイルのダウンロードに要する時間が徐々に増加していくこともわかる.インデックスポイゾニングでは,ダミーファイルキーに記載されたファイルの位置情報が架空のノードとなっているため,ダウンロードを試行するノードがダミーファイルキーの情報を元にダウンロードを試行した場合,すぐにダウンロードに失敗してしまう.しかし,アイテムポイゾニングでは,ダミーファイルキーに記載されたファイルの位置情報が制御ノードとなっており,制御ノードによってダミーファイル本体のアップロードが行われる.そのため,拡散したダミーファイルキーのユニーク数が少なくても,インデックスポイゾニングに比べてファイル流通制御効果を出すことができている.

5.5.3 制御対象ファイルのファイルキーの発見率

アイテムポイゾニングによる制御中の制御対象ファイルのファイルキーの発見率を**図 10** に示す.アイテムポイゾニングの場合もインデックスポイゾニングと同様に,ダミーファイルキー数が多ければ,制御対象ファイルのファイルキーの発見率を低い値に保つことができている.

また,インデックスポイゾニングではダミーファイルキーのキータイマが尽きると,ネットワーク上からダミーファイルキーが消滅するため,制御対象ファイルのファイルキーの発見率が 100%になる.しかしアイテムポイゾニングでは,制御ノードが拡散したダミーファイルキーがネットワーク上から消滅しても,

表 2 アイテムポイゾニングにおけるダミーファイルキーの内容
Table 2 Contents of a dummy file key for item poisoning.

項目	内容
ファイル名	制御対象ファイルと同じ
ファイルサイズ	152〜153Byte
ファイル位置情報	制御ノード
キータイマー	3,600 秒

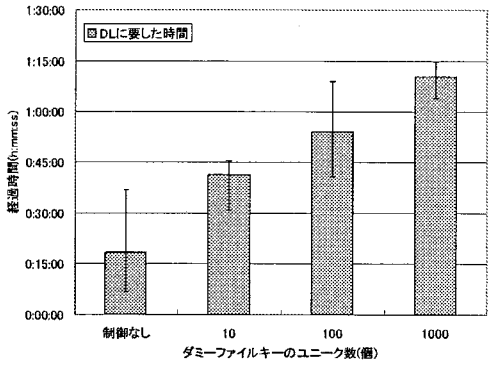

図 9 アイテムポイゾニングによる制御でダウンロードに要した時間
Fig. 9 Time to download completion under item poisoning control.

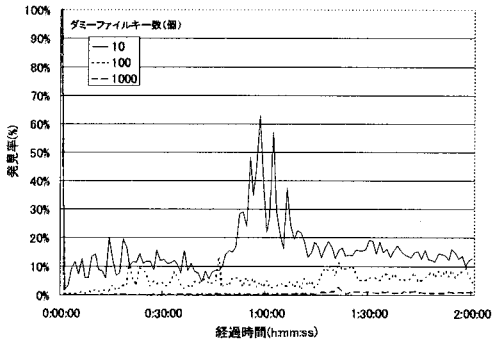

図 10 制御対象ファイルのファイルキーの発見率
Fig. 10 Discovery rate of target file key.

制御対象ファイルのファイルキーの発見率を低い値に保つことができている．これは，ダミーファイル本体を所持するノードによってダミーファイルキーの拡散が継続されるからである．つまり，アイテムポイゾニングにより制御を行った場合は，ダミーファイル本体を所持する全てのノードがファイルのアップロードをやめない限り，ダミーファイルキーによるファイル流通制御効果が持続することになる．以上から，長期的な制御ではアイテムポイゾニングによる制御が効果的であると言える．

5.6 制御に要したトラヒック

インデックスポイゾニング，ならびにアイテムポイゾニングにおいて，制御ノードが拡散するダミーファイルキーのユニーク数が 1,000 個の時に，制御ノードに発生したトラヒックを示す．インデックスポイゾニングによる制御では，評価時間の 2 時間で制御ノードに発生した送受信トラヒックの合計は 27.14MByte であった．アイテムポイゾニングによる制御では，評価時間の 2 時間で制御ノードに発生した送受信トラヒックの合計は 108.14MByte であった．アイテムポイゾニングは，インデックスポイゾニングの約 4 倍のトラヒックを制御に必要とする．しかしアイテムポイゾニングにおいては，ネットワーク上にダミーファイル本体を拡散すれば，その後は制御ノードが全く制御を行わなくてもファイル流通制御効果が持続する．インデックスポイゾニングによって長期的な制御を行う場合，定期的にダミーファイルキーを拡散する必要があるため，アイテムポイゾニングよりもトラヒックを要することが考えられる．長期的な制御が必要ない場合はインデックスポイゾニング，必要である場合はアイテムポイゾニングという使い分けをすることで，それぞれの長所を生かすことができる．

6. おわりに

本稿ではポイゾニングを用いて Winny ネットワークにおけるファイル流通制御を行い，その効果を確認した．インデックスポイゾニングによる制御では，制御対象ファイルのファイルキーの発見率を低下させることができるため，制御対象ファイルの流通制御が可能であることを確認した．また，アイテムポイゾニングによる制御では，ダミーファイル本体を所持するノードによってダミーファイルキーの拡散が継続される．この効果により，制御ノードによる制御を終了した後でもファイル流通制御効果を持続させることが可能であることも分かった．今後は，より大きな Winny ネットワークでのポイゾニングが可能となるように，ポイゾニング手法と制御ノードの改良を行い，実ネットワークでの評価を行う予定である．

謝辞 本研究の一部は科研費基盤 C(No. 18500047)，および，電気通信普及財団の研究助成の支援を受けており，ここに記して感謝する．

参考文献

1) 金子 勇, “Winny の技術,” アスキー, 2005.
2) 寺田真敏, 鵜飼裕司, 金居良治, 畑田充弘, 松木隆宏, 宮川雄一, “クローリング手法を用いた P2P ネットワークの観測,” 情報処理学会研究報告 *CSEC*, vol. 2007, no. 48, pp. 51–56, 2007.
3) J. Liang, N. Naoumov, and K. W. Ross, “The Index Poisoning Attack in P2P File-Sharing Systems,” *Proceedings of IEEE Infocom'06*, pp. 1–12, April 2006.
4) J. Liang, R. Kumar, Y. Xi, and K. W. Ross, “Pollution in P2P File Sharing Systems,” *Proceedings of IEEE Infocom'05*, pp. 1174–1185, March 2005.
5) VMWare, http://www.vmware.com.