

非構造化 P2P ファイル共有ネットワークにおける 非協力ピアの参加によるファイル流通制御方式の提案と評価

上野 真弘^{†1} 大坐 島 智^{†2} 川島 幸之助^{†2}

近年、P2P によりファイル共有を行うアプリケーションが増加しており、それらが発生するトラフィックは IP ネットワークに大きな影響を及ぼしている。そこで本稿では、トラフィックを減少させるために、ファイル検索に関して非協力的な制御ピアをネットワーク上に参加させることにより、ピアの検索効率を低下させ、共有されるファイル数を制御する方式を提案する。また、Gnutella と Winny を想定したシミュレーション実験を行い、本方式の有効性を評価した結果、いずれのネットワークに対しても検索効率の低下が見られることを確認した。特に、Winny ネットワークに対する効果は大きく、少数の制御ピアでも十分な効果が得られることを示した。

An Evaluation of File Control Method with Uncooperative Peers in Unstructured P2P File Sharing Networks

MASAHIRO UENO,^{†1} SATOSHI OHZAHATA^{†2}
and KONOSUKE KAWASHIMA^{†2}

P2P file sharing applications are increasing recently, and the traffic affects to IP networks. In this paper, we propose a file sharing control method to reduce file sharing efficiency. In our method, control peers join the file search network and we can control number of shared files of networks. We evaluated our method by computer simulations for Gnutella and Winny networks. From these results, we confirmed decrease of the efficiency of file search for both networks. Especially, the effect on Winny networks is more effective than Gnutella and few control peers can give enough effect on the network.

1. はじめに

近年、Peer-to-Peer (P2P) ネットワークを構成することによりファイル共有を行うアプリケーションが増加しており、特に欧米では Gnutella¹⁾ プロトコルを応用した LimeWire が、日本では Winny²⁾ が人気である。しかし、これらのアプリケーションによるファイル転送で発生するトラフィック量は大きく、IP ネットワーク全体に大きな負荷を掛けており、他のネットワークアプリケーションにも影響がでている。また、P2P ネットワーク全体のトラフィックは集中的に管理することができないため、悪意ある者により一度コンピュータウィルスを拡散されてしまえば、ネットワーク上のピア全体に危険が及ぶ。このような問題を解決するためにトラフィックを管理するための手法として、

ファイル転送の前段階であるファイル検索の効率を低下させることにより、共有されるファイル数を制御する方法が考えられる。

そこで本稿では、現在最も普及している非構造化の P2P ネットワークを対象に、ネットワーク上にファイル検索に対して非協力的な制御ピアを参加させることにより、検索効率を低下させる方法を提案する。提案方式では、本来ならばファイル検索を行うためにピア間でパケツリレー式に転送される検索要求 (クエリ) や、拡散されるファイル情報 (キー) を転送しないことによりファイル検索を妨害する。また、代表的な非構造化ファイル共有アプリケーションである Gnutella と Winny を想定したシミュレーション実験を行い、各ネットワークにおける提案方式の有効性を評価する。

2. ピア P2P ファイル共有における検索方式

2.1 Gnutella

Gnutella は、プロトコルのバージョンにより 0.4 と 0.6 の 2 種類が存在している。バージョン 0.4 では、すべてのピアが対等に接続し合ってネットワークを構

^{†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

Institute of Symbiotic Science and Technology,
Tokyo University of Agriculture and Technology

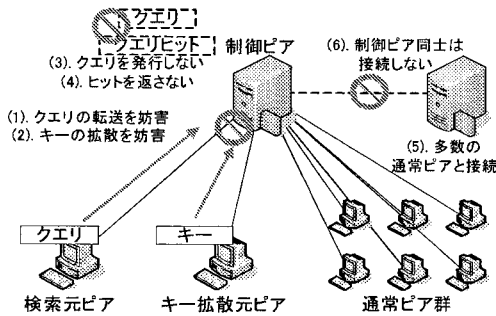


図 3 制御ピアの動作
Fig. 3 Controls of the control peer.

(1) と (2) は、前述したファイル検索に用いるメッセージの転送を妨害する非協力的な動作である。(3) と (4) は、制御目的以外の動作を行わないことを示している。(5) は、制御ピア 1 台あたりの妨害の効率を向上させるためである。(6) と (7) は、無駄な接続を無くし、よりクエリが集中しやすい接続を確立することにより、制御ピアの接続 1 本あたりの妨害の効率を向上させるためである。

3.2 Gnutella ネットワークに対する影響

Gnutella では、クエリは転送を繰り返すたびにフラディングにより増殖するため、検索元ピアと目的ファイルの保有ピア間を結ぶ複数の経路のうち、1 つでも制御ピアによりクエリが妨害されない経路が存在すれば、ファイル保有ピアへクエリが到達する。

制御ピアの接続数が少数である場合、検索元ピアとファイル保有ピア間を結ぶ経路のうち、いくつかは制御ピアにより塞がれるが、残りの経路によってクエリが到達する可能性は十分に残される。ただし、塞がれた経路が 2 ピア間を結ぶ唯一の最短経路であった場合、塞がれていない別の経路が新たな最短経路になるため、到達するまでのクエリの最短ホップ数は増加する。

しかし、制御ピアの接続数の増加に伴い、2 ピア間を結ぶ経路はすべて塞がれてクエリは到達できなくなる可能性が増加する。クエリが転送される経路に 1 台でも制御ピアがある場合、その転送は妨害されるため、クエリはホップ数が増加するほど妨害される確率は高くなる。従って、制御ピアの接続数が少数である場合とは逆で、クエリはネットワーク上で検索元ピアから近いピアにしか到達できない。

3.3 Winny ネットワークに対する影響

Winny では、検索クエリは転送のたびに複製されないため、制御ピアにより 1 つの検索クエリの転送を妨害できれば、ファイル検索を完全に妨害することができる。同時に、検索条件に当てはまったキーは検索クエリと共に転送されるので、検索元ピアが検索結果を得ることができなくなる。

また、Winny 上の制御ピアは拡散クエリによるキー

の拡散も妨害するため、ネットワーク全体へのキーの拡散速度が遅くなる。キーには寿命が存在するため、拡散速度が遅くなるとネットワーク上の遠くのピアまでキーを拡散できなくなる。ただし、キーは拡散するたびに複製されるため、検索クエリの妨害とは異なり、1 つの経路を塞いだとしても別の迂回路から拡散する可能性があるため、1 つのキーの拡散を完全に妨害することは困難である。

4. 実験評価

4.1 実験環境

本稿では、ネットワークシミュレータとして ns2³⁾ を用いた。その上で、Gnutella を想定したシミュレーションには、ns 上で動作する Gnutella のシミュレータである GnutellaSim⁴⁾ を用いた。Winny を想定したものについては、さらに GnutellaSim を基にモデル化した。ただし、実験にあまり関係がない計算処理の省略のために、実際のアプリケーションの実装とは異なる部分がある。変更点を以下に示す。

ピア情報取得の簡易化 Gnutella では Ping および Pong メッセージにより、Winny では接続を拒否したピアから他のピアの情報 (IP アドレス等) を入手する必要があるが、本稿ではブートストラップサーバからのみ入手する。

QRP の省略 二階層型のネットワークを構成する Gnutella では Leaf が保有ファイル情報を隣接の Ultrapeer へ通知するが、本稿では省略する。代わりに、Ultrapeer は隣接の Leaf へクエリを転送する。

クラスタリングの省略 Winny では嗜好性が近いピア同士を近くに配置してネットワークのクラスタリングを行うが、本稿では省略する。

P2P ネットワークは、通常ピアの台数を 100 台として途中での参加や離脱はないものとする。1 回の実験では、通常ピアの最大接続数 C_{norm} を固定し、次の式 (1) の α を 0~10 まで 1 刻みで変化させることにより、制御ピアの最大接続数 C_{cont} を変化させて測定する。つまり、 α は C_{norm} を基準とした C_{cont} の倍率であり、 $\alpha = 0$ のとき $C_{cont} = 0$ なので、制御ピアはネットワークに参加しない。

$$C_{cont} = C_{norm} \times \alpha \quad (1)$$

1 回の実験は 1000 秒行い、前半の 500 秒をピア間の接続時間とし後半を測定時間とした。これを 1 つの条件につき 10 回行い、それらの平均を測定値とした。

4.2 Gnutella

評価は非階層型と二階層型について、それぞれ表 1 に示すとおり、タイプ A とタイプ B の 2 種類の C_{norm} を用意し、計 4 種類の C_{norm} で行った。タイプ B はタイプ A の約 2 倍である。Leaf 受入とは、Ultrapeer が接続を受け入れる Leaf の台数である。制御ピアは 10 台で固定した。また、二階層型においては、Ultrapeer

表 1 想定した通常ピアの接続形態と接続数 (Gnutella)
Table 1 Network type and degree patterns of normal peers (Gnutella).

接続形態	接続名	接続数	Leaf 受入
非階層型	タイプ A	3	-
	タイプ B	6	-
二階層型	タイプ A	Ultrapeer:3, Leaf:2	5
	タイプ B	Ultrapeer:6, Leaf:3	10

表 2 想定した通常ピアの接続数 (Winny)
Table 2 Degree of normal peer (Winny).

接続名	上流への接続数	下流への接続数
タイプ A	2	5
タイプ B	4	10
タイプ C	5	5

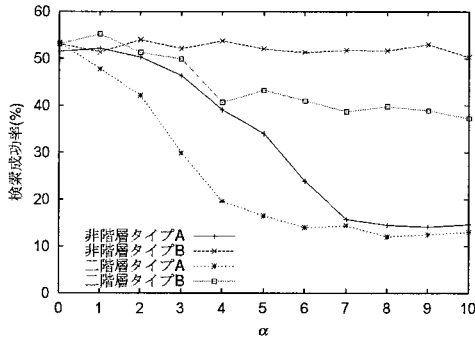


図 4 制御ピアの参加による検索成功率 (Gnutella)
Fig. 4 Success rates with control peers (Gnutella).

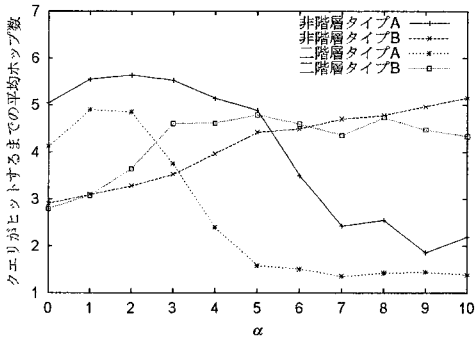


図 5 制御ピアの参加によるクエリがヒットするまでの
ホップ数 (Gnutella)
Fig. 5 Average number of hops for query hit
with control peers (Gnutella).

と Leaf の割合を 1:1 とした。評価項目は以下に示す 2 項目とした。

検索成功率 1 回の検索に対し、検索元ピアが 1 つでもクエリヒットを受信した場合を検索成功と定義し、すべてのピアの検索回数に対する成功回数を測定した。

クエリがヒットするまでのホップ数 3.2 節で述べた現象が発生していることを確認するために、クエリがヒットするまでの平均ホップ数を測定した。

制御ピアの参加による検索成功率の変化を図 4 に示す。まず、非階層と二階層のどちらにおいても C_{norm} が大きいタイプ B よりも小さいタイプ A の方が検索成功率の低下が大きいことが分かる。これは、接続数

が多い方がフラッディングによるクエリの増殖数が多くなるため、制御ピアの接続数を増加させても容易にはクエリの増殖を阻止することができないためであると考えられる。ただし、タイプ A も 10 数% までは低下したが、それ以上の低下を確認できなかった。このことより、制御ピアの最大接続数を増加させても、検索元ピアの近隣のピアへのクエリの転送を妨害することは困難であると考えられる。また、それぞれのタイプにおいて、非階層型よりも二階層型の方が検索成功率の低下が大きい。これは、二階層型ではクエリが Ultrapeer へ集中するため、制御ピアが効果的にクエリの転送を妨害できるためであると考えられる。

次に、制御ピアの参加によるクエリがヒットするまでの平均ホップ数の変化を図 5 に示す。非階層と二階層のどちらにおいても、タイプ A は α の増加に伴い一時的にホップ数が上昇しているため、3.2 節で述べた現象が発生していると考えられる。しかし、タイプ B はホップ数が徐々に上昇するのみで下降しない。これは、前述した実験と同様にクエリの増殖を阻止できなかったことが考えられる。また、タイプ A の非階層型と二階層型を比較して後者の方がホップ数が大きく減少する α が小さいのも同様に、クエリが Ultrapeer へ集中するためと考えられる。

4.3 Winny

評価は表 2 に示すとおり、タイプ A、タイプ B、タイプ C の 3 種類の C_{norm} で行った。タイプ A は通常の Winny と同じ接続数である。通常ピアの申告速度は、東京農工大学 川島研究室が開発した Winny クローラにより実測したデータを基に、上流から 1000KB/s、120KB/s、50KB/s の 3 種類を 3:6:1 の割合に設定した。制御ピアは 3 台で固定し、申告速度は上流と同じく 1000KB/s とした。また、制御ピアは検索クエリの妨害を目的とするため、通常ピアとは異なり上流よりも下流への接続を優先的に試行する。シミュレーションは、キーを十分に拡散させるために、接続時間と測定時間の間にさらに 500 秒設け、すべての時間においてキーを拡散させた。さらに、実際のキーの寿命は 1500 秒であるが、今回はネットワークの規模が小さいため 300 秒に変更した。評価項目は以下に示す 2 項目とした。

検索成功率 1 回の検索に対し、検索クエリが検索元ピアへ戻り、かつ 1 つ以上のキーがヒットした場合を検索成功と定義し、すべてのピアの検索回数に対する成功回数を測定した。

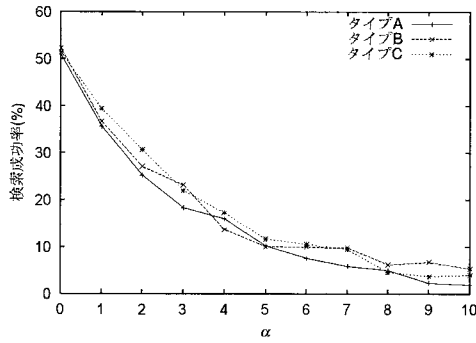


図 6 制御ピアの参加による検索成功率 (Winny)
Fig. 6 Success rates with control peers (Winny).

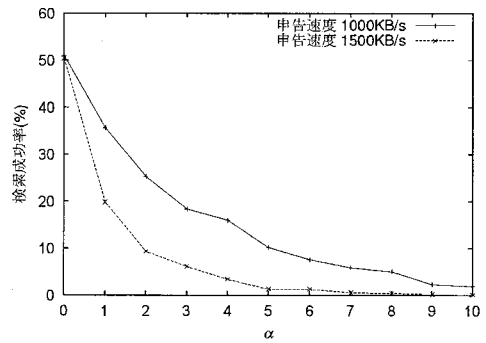


図 8 制御ピアの申告速度による検索成功率 (Winny)
Fig. 8 Success rates with different link speed (Winny).

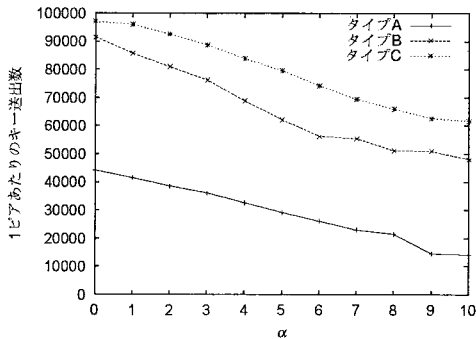


図 7 制御ピアの参加による 1 ピアあたりのキー送出数 (Winny)
Fig. 7 Number of transferred keys per peer with control peers (Winny).

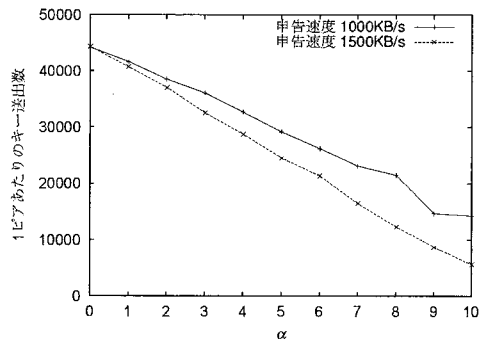


図 9 制御ピアの申告速度による 1 ピアあたりのキー送出数 (Winny)
Fig. 9 Number of transferred keys per peer with different link speed (Winny).

1 ピアあたりのキー送出数 Winny において検索のために要するトラフィックの大半はキー情報であるため、トラフィック量を把握するために、1 ピアあたりのシミュレーション中に送出したキー数を測定した。

制御ピアの参加による検索成功率の変化を図 6 に示す。いずれの接続数においても、 α の増加に伴い検索成功率が指数関数的に減少しているが、接続数の違いによる効果の違いはあまり見られなかった。これは、Gnutella ではピアの接続数によってクエリの増殖数が増えるのに対し、Winny ではクエリを単一のピアへのみ転送するため、接続数による変化がないものと考えられる。

制御ピアの参加による 1 ピアあたりのキー送出数の変化を図 7 に示す。いずれの接続数においても、 α の増加に伴いほぼ比例して減少している。これは、前述したようにキー交換はプル型の通信を行っており、これに非協力的である制御ピアは隣接ピアへ拡散要求を送信しないため減少したと考えられる。また、接続ごとにキー拡散を行い、特に上流への拡散は最低転送個数が規定されているため、上流への接続数が多い方がキーの送出数が多くなる。

次に、制御ピアの申告速度を通常ピアの上流より速い 1500KB/s に設定して同様の実験を行い、前述した申告速度 1000KB/s のものと比較した。接続数はいずれもタイプ A とした。

制御ピアの申告速度の違いによる検索成功率の比較を図 8 に示す。申告速度 1500KB/s は 1000KB/s と比較して、 α の増加に伴う成功率の減少が大きい。いずれの α においても 1000KB/s の半分以上の成功率まで減少しており、 α が 5 のときの成功率はほぼ 0% である。これは、Winny の検索クエリがより上流へ移動するように規定されているため、通常ピアの上流よりも申告速度が速い制御ピアへ集中しやすいためであると考えられる。

最後に、制御ピアの申告速度の違いによる 1 ピアあたりのキー送出数を図 9 に示す。どちらも α にほぼ比例して減少しているものの、申告速度 1500KB/s の方が減少が大きいことが確認できる。これは、今回の実験の中でキーの転送個数が一番大きい接続である、上流の通常ピアの上流への接続の大半が制御ピアに接続したため、その接続によるキーの拡散が行われな

ことが原因として考えられる。

4.4 Gnutella と Winny の比較

本稿では、Gnutella と Winny の両方のネットワークにおいて検索成功率を測定しており、クエリの転送方式の違いはあるものの、いずれも 1 回の検索に対して 1 つ以上の保有ファイル情報が返ってくる場合を検索成功と定義した部分は共通である。そこで、図 4 と図 6 および図 8 を比較すると、Gnutella に対する制御では α の増加に伴い成功率が徐々に減少したのに対し、Winny に対する制御では指数関数的に大きく減少した。さらに、本稿で用いた制御ピアの台数も Winny に対する制御の方が少数であることから、本制御方式は Gnutella よりも Winny に対してより効果が大きいことが分かる。

この理由として 2 点考えられる。まず、クエリの転送方式の違いによるものである。Gnutella では、目的のファイルを保有するピアはクエリを受信した時点でクエリヒットを検索元ピアへ返すのに対し、Winny では、条件に当てはまるキーを検索クエリに詰め込んでクエリが TTL に到達するまで共に転送されるので、その経路上に 1 台でも制御ピアがある場合に検索元ピアへ結果が返らないという点である。もう一点は、最大でも 2 階層しかない Gnutella ネットワークに対して、Winny ネットワークは速度を基準としてより階層が深いため、一部の上流ピアへよりクエリが集中しやすく、制御ピアが効果的にクエリを妨害できるという点である。

5. おわりに

本稿では、P2P ファイル共有ネットワーク上に検索に対して非協力的な制御ピアを参加させることにより、ファイルの検索効率を減少させる方式を提案し、Gnutella と Winny ネットワークを想定したシミュレーション実験によって、その有効性を評価した。実験の結果、いずれのネットワークに対しても本制御方式により検索効率をある程度減少できることを確認した。特に、Gnutella に対する制御では、制御の効果がネットワークの接続形態や各ピアの接続数に大きく依存しており、Winny に対する制御では、制御ピアの接続数の増加に伴い検索効率が指数関数的に大きく減少することが分かった。また、両ネットワークを比較すると、Gnutella のように多数のクエリを拡散させるような検索方式に対する効果は弱いのにに対し、Winny のようなクエリを増殖させない方式に対する効果は強いことが判明した。

今後は、より実環境で稼働しているネットワークへ近づけるために、大規模なネットワークに対する実験やピアの参加や離脱を考慮した実験を行い、本方式の有効性を評価していく予定である。また、CAN(Content-Addressable Network)⁵⁾ や Chord⁶⁾ のような構造化

ネットワークや文献 7) のようなピア間の評価機能を備えたネットワークに対して、本方式を適用する方法を考えていく必要がある。

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) Gnutella: <http://www.gnutella.com/>.
- 2) 金子 勇: Winny の技術, アスキー (2005).
- 3) ns2: <http://www.isi.edu/nsnam/ns/>.
- 4) Q.He, M.Ammar, G.Riley, H.Raj and R.Fujimoto: Mapping Peer Behavior to Packet-level Details: A Framework for Packet-level Simulation of Peer-to-Peer Systems, *Proc. IEEE/ACM MAS-COTS 2003*, pp.71-78 (2003).
- 5) S.Ramasamy, P.Francis, M.Handley, R.Karp and S.Shenker: A Scalable Content-Addressable Network, *Proc. ACM SIGCOMM'01*, pp.161-172 (2001).
- 6) I.Stoica, R.Morris, D.Karger, M.F.Kaashoek and H.Balakrishnan: Chord: A Scalable Peer-to-Peer Lookup Service for Internet Applications, *Proc. IFIP/ACM Middleware 2001*, pp.149-160 (2001).
- 7) Q.Sun and H.Garcia-Molina: SLIC: A Selfish Link-based Incentive Mechanism for Unstructured Peer-to-Peer Networks, *Proc. IEEE ICDCS'04*, pp.506-515 (2004).