

複数レイヤでの部分測定を用いた P2P ファイル共有アプリケーションの規模推定

亀井 聡^{†,††a)} 内田 真人[†] 森 達哉[†] 高橋 豊^{††}Estimating Scale of Peer-to-Peer File Sharing Applications
Using Multi-Layer Partial MeasurementSatoshi KAMEI^{†,††a)}, Masato UCHIDA[†], Tatsuya MORI[†], and Yutaka TAKAHASHI^{††}

あらまし ウェブの次代を担う技術として、P2P や分散コンピューティングを代表とするインターネット上でオーバーレイネットワークを構成する自律分散型システムの普及が目覚ましい。このような大規模システム上ではその性質上、アプリケーションの内部プロトコルを変更することなしに、その全体の特性を測定することは困難である。本論文では、P2P アプリケーションに汎用的に適用可能な測定法と、一部にその測定法を適用し得られた測定結果から、P2P ネットワーク全体の挙動や規模の推定を行うための汎用的な手法を提案する。更に実 P2P ネットワークに本手法を適用し、その有効性を具体例で示す。

キーワード P2P, トラヒック, ファイル共有, ネットワーク測定, アプリケーションレイヤ測定

1. ま え が き

端末機器の高性能化とともに、アクセスネットワークのブロードバンド化、常時接続化の進行が目覚ましい。これに伴い、これまでネットワークのコア側に配備されていたストレージやデータベース、計算資源といった機能の一部やすべてをエッジやアクセス側の端末機器側へと委譲する Peer-to-Peer (P2P) と呼ばれる技術が普及を始めている。

このような自律分散型のシステムでは情報の集約がほとんど行われなため、ネットワーク上の一点でその挙動を把握することは困難である。また、代表的な P2P アプリケーションにより構築される P2P ネットワークである、FastTrack [1] や GnutellaNet [2] 等ではそのノード規模は常時 P2P ネットワークに接続しているものだけで数万～数百万ノードにも達する [3]

とされるが、ネットワークの規模も把握は困難である。

一方で、このようなオーバーレイネットワークは、その独特の挙動から下位ネットワークに与える影響が大きく、その規模拡大に従い下位レイヤに対するインパクトも大きくなる可能性が指摘されている [4]。このため、IP ネットワークを設計・制御する上でも、P2P ネットワークの規模や挙動の把握は重要な課題である。また、オーバーレイネットワーク上に構築される Skype [5] のようなサービスや、Distributed Hash Table (DHT) [6], [7] 等に見られる大規模な自律分散システムでは、各参加端末のもつリソース上でインデックスの共有を行うため、各ノードの生存状況や全体の端末数がシステム運用上大きな意味をもつ。

本論文では、全体の情報を測定により得ることが困難であるような大規模自律分散ネットワークにおいて、部分的な測定からその全体の挙動、特に規模を推定するための汎用的手法を提案する。また、その適用例として実 P2P ネットワークの部分測定を行い、得られた結果を用いてノード規模の推定を行う。

以下、2. で P2P アプリケーションによって構築される P2P ネットワークについて概説を行い、3. で既存の測定法とその問題点について簡単に説明した後、4. で P2P ネットワークのトラヒック測定法、5. で規

[†] 日本電信電話株式会社 NTT サービスインテグレーション基盤研究所, 武蔵野市

NTT Service Integration Laboratories, NTT Corporation, 3-9-11 Midori-cho, Musashino-shi, 180-8585 Japan

^{††} 京都大学情報学研究科システム科学専攻, 京都市
Department of Systems Science Graduate School of Informatics, Kyoto University, Yoshidahonmachi, Sakyo-ku, Kyoto-shi, 606-8501 Japan

a) E-mail: kamei.satoshi@lab.ntt.co.jp

模推定法を示す．6. で提案手法の実際の P2P ネットワークに対する適用例を示した後，7. でまとめる．

2. P2P ネットワーク

2.1 P2P 技術と P2P ネットワーク

P2P ファイル共有と呼ばれる分野は Napster [8] をその起源とし，参加ユーザのディスク上にあるファイルを検索するメカニズムと，検索問合せにより各ユーザが発見したファイルを直接，あるいはキャッシュや他ノードを介して間接的に自分のノードへと転送する機能をその特徴とする．Napster においてはファイル検索のためのサーバを必要とするため，検索問合せのトラヒックはウェブと似たパターンを示すが，トラヒックの大部分を占める音楽ファイルのファイル転送トラヒックは，各ノード間で直接発生する．Napster の後に登場した Gnutella においてはファイル転送のみならず，検索問合せも各ピア間で自律分散的に交換され，サーバを必要としない．このため，全トラヒックはピア間で発生し，そのトラヒックパターンはウェブとは大きく異なる．

P2P の定義については様々な提案があるが，本論文では Internet Research Task Force (IRTF) の Peer-to-Peer Research Group (P2PRG) [9] の定義に従う．P2PRG では，P2P とは個々のノードが対称的な役割をもつような分散アプリケーションを構築する技術と定義しており，P2P アプリケーションではサーバとクライアントの役割が分離されておらず，一つのノードがクライアントとサーバの両方の役割を果たしている．

P2P ネットワークは，ピアと呼ばれる P2P アプリケーションの動作ノードと，ピア間を相互に接続するアプリケーションレイヤでの論理リンクから構成される．これらにより構築される論理ネットワークは IP ネットワークのトポロジー上にオーバーレイする形で構築されるため，オーバーレイネットワークとも呼ばれる．このネットワークは情報の集約個所や制御の中心をもつ必要がない．このため，ネットワークは自律分散的に構築，制御される．これにより，高いスケーラビリティとフィージビリティを実現し，更に下位レイヤの IP ネットワーク層に標準で備わっていない様々な機能を追加することができる．一方で情報が集約されないために，アプリケーション側で情報集約を行わない限り，集中的な制御や管理は困難となる．また，情報集約はスケーラビリティとのトレードオフとなるため，

大規模なネットワーク構築には向かない．

2.2 P2P ネットワークの規模と挙動

P2P アプリケーションの普及に伴い，オーバーレイの P2P ネットワーク上を流れるトラヒックは下位レイヤに大きな影響を及ぼす規模 [10] に達しており，IP ネットワークの設計や制御，運用に際してこのようなネットワークの挙動規模把握が重要な課題となっている．規模把握が有用な局面は大きく二つあり，一つはインフラ管理の立場からそのネットワーク上を流れるトラヒックの増加傾向やこれまでとは異なる振舞いを早期に検知するため．もう一つは，このような分散アプリケーションを開発する立場から利用者数の変化を検知し，その振舞いを把握するためである．

インフラ管理の観点から見た場合，特に P2P ファイル共有に関しては，その参加ユーザ規模の 2^n に比例してコンテンツ量が增大するとする，Group Forming Networks (GFN) [4] の考え方も提唱されており，その規模拡大を早期に検知することが肝要となる．また，文献 [11] で提案したような P2P トラヒックの分離法においても，全体の何割程度を分離できたかを知るための目安として，対象となる P2P ネットワークの規模把握は必要である．更に，実際に P2P トラヒックに対して制御を行うにあたり，どの程度の利用者に影響を与えるのかを知る上でも，P2P ユーザ規模の把握は IP ネットワークの運用上有用である．

一方で，P2P 技術を用いたアプリケーションの開発や，このようなオーバーレイのインフラを用いた分散コンピューティングやストレージといったサービスの開発も盛んに行われるようになってきている．中でも，Chord [6] や CAN [7] といった，大規模分散環境での一貫性を保ったデータ保持技術は近年盛んに研究されている．こういった技術は，P2P ネットワーク上に存在するノードのもつディスク領域等のリソース上にデータを格納する形をとる．このため，ノードの数や，各ノードの生存状況，平均生存時間といった情報を把握することが，データ格納の冗長性，ノード消滅時の堅牢性等の評価にあたって必要となる．

3. 既存の測定法と課題

3.1 ポート番号を用いた測定法

P2P アプリケーションの挙動解析については，いくつかの検討がある．文献 [12], [13] では，ネットワークレイヤでの測定結果から P2P トラヒックを分離し，特性解析を行っている．また，文献 [14] では国内ネット

ワークでウェブの割合はわずか 20%にとどまり、P2P の割合がウェブの約 3 倍に達している状況が示されている。これらの研究では、P2P アプリケーションの識別のために IP パケットヘッダの TCP の着ポート番号を用いている。この手法は電子メールのようにポート番号が固定であったり、ウェブのように省略時に固定ポートが利用される等、サーバ側の運用が厳密に定義されている場合には有効だが、現行の P2P アプリケーションではポート番号が可変であることが多く、アプリケーション起動時にランダムに割り当てられる場合さえある。したがって、ポート番号による識別ではその一部しかとらえられない。

3.2 シグネチャを用いた測定法

文献 [15], [16] では、ペイロードに記録された特定パターンを用いた P2P トラフィック測定法が提案されている。こうした手法は、レイヤ 7 スイッチ等での実装 [17], [18] や、Intrusion Detection System (IDS) 技術、Passive Fingerprinting 技術等 [19] で用いられているものと同等である。本手法では、パターンマッチングのためにペイロードを監視する必要があるため、大容量回線での利用についてはスケラビリティに限界がある。また、プロトコル仕様が不明であったり暗号化されているアプリケーションへの対応は困難であり、誤認識の問題もある。

3.3 トラフィックパターンを用いた測定法

P2P アプリケーションの特徴的トラフィックパターンを用いるものもある。文献 [20] ではセッション追跡による Napster トラフィックの識別法が示されている。このようなトラフィックパターンマッチングでは、ヘッダ情報とパケットサイズの組合せをホスト間の通信ごとに管理するため、大量のリソースを消費する。また、適用範囲が一定のトラフィックパターンや中央サーバをもつ P2P アプリケーションに限られる。

より汎用的なトラフィックパターンを用いた測定法として、文献 [21] にあるようなヒューリスティックな方法、文献 [22] にある、フローのサイズを用いた方法も提案されている。しかし、これらの手法は P2P 的な振舞いをするトラフィック一般を測定することはできるが、個別のアプリケーションに対して解析を行うことはできない。

3.4 オーバレイネットワーク特有の課題

測定技術だけにとどまらない問題もある。それは、P2P の特徴的な機能であるアプリケーションレイヤでのオーバレイネットワークがネットワークレイヤと

独立して存在する点に由来する。ネットワークレイヤの測定ではその測定物理リンク上に載ったオーバレイネットワークのリンクは測定できるが、それはオーバレイネットワーク全体に対しては限られた一部分にすぎないため、ネットワーク全体の情報を取得するためにはバックボーンの多数のポイントに物理的に測定機器を設置する必要がある、非現実的である。

バックボーンへの機器設置を必要としないアプリケーションレイヤでの測定に関しては、文献 [23] 等に見られる個別アプリケーションの解析による検討のほかにも、著作権団体と組んで測定を実施するモニタリング会社が数社存在する。これらについては、技術詳細が公開されていないためその手法の有効性は検証できないが、国内でも [24] 等が同様のサービスを行っている。しかし、これらはアプリケーションレイヤでの測定は一般に特定のアプリケーションに依存した解析が必要となるために汎用性は低い。また、内部情報が利用できる場合でも、一般的には局所的な情報しか集約できない。これは、例えば Gnutella がその TTL 範囲内しか検索ができないことや、Winny [25] のファイル被参照量の値が不完全であることからもうかがえる。

以上のように、P2P に対する既存測定法は、スケラビリティ、アプリケーションへの依存、局所的情報への限定という点に課題を抱えていることが分かる。

4. 複数レイヤでのトラフィック測定法

前章で述べたような既存測定法の問題点を解決するためには、アプリケーションレイヤでの測定とネットワークレイヤでの測定を組み合わせる測定を行う必要がある。本章ではその具体的手法を提案する。

4.1 ピアの情報

P2P ネットワークにおける一般的な新規参加ピアの接続動作は以下のような流れとなる。

(1) 既に P2P ネットワークに参加している端末 (ピアと呼ばれる) の IP アドレスとサービスポートの組を何らかの方法で入手。

(2) 取得したピア情報 (一般的には複数) をもとに、ピアへ接続。

(3) ピア間で論理リンクを確立。

(4) 検索パケットや、定期的に各ピアから発行される生存確認パケットから更なるピア情報を入手。

(5) 論理リンクがピアの消滅や、ネットワークトラブルにより切断されたときには (4) で得たピアへ論理リンクが一定数になるまで再接続を試みる。

P2P ネットワークにおいて汎用的に収集可能な情報は、ここで挙げたピアの動作アドレスとサービスポートを組とした、ピア情報になる。アプリケーションレイヤで情報を収集する場合には、アプリケーション内部で保持する情報から (4) で流通するピア情報を直接収集することも可能だが、このためにはアプリケーションの改造かパケットのデコードが必要になる。しかし、多数のユーザを獲得している P2P アプリケーションの中には、ソースコードやプロトコルが非開示であるものも少なく、改造が不可能であることが多い。また、ネットワークレイヤでの測定から直接ピア情報を取得することはアプリケーションへの依存、スケラビリティの両観点から容易とは言いがたい。

4.2 論理ネットワークでのピア情報収集

以上見てきたとおり、ネットワークレイヤでの測定だけによるピアの情報取得には様々な課題がある。また、論理レイヤでの測定だけでは特定アプリケーションへの依存を抑えることはできない。

これらの課題を解決するため、本論文では論理レイヤとネットワークレイヤでの測定を組み合わせるピアの動作ノードの情報取得を試みた。本手法は、アプリケーションを通常の手続きで動作させることができれば、tcpdump 等の通常のネットワークレイヤ測定との組合せにより、他ピアのノード情報を得ることができるという特徴をもっており、特定アプリケーションへの依存を低く抑えることができる。

測定手順は図 1 に示した環境において、以下に示したとおりとなる。

(1) 任意の地点において P2P アプリケーションを 1 ユーザとして動作、P2P ネットワークに接続。

(2) P2P 動作ノードと同一ノード、あるいはダムハブ等によって分岐させたノードにおいてレイヤ 3 でのパケットダンプを実施。

(3) ダンプデータから P2P 動作ノードの通信相手となったピアが動作しているノードの IP アドレスと利用ポートを収集。

以上の手順により、測定対象となった P2P ネットワーク上に存在するピアの情報が得られる。

4.3 P2P トラフィック識別法

以上提案した論理ネットワークでのピア情報収集方式では、内部構造が不明な P2P アプリケーションを対象とした場合でも通常のユーザとしての動作手続きに従いアプリケーションを動作させ、P2P 動作ノードでの IP レイヤでの通信を測定することによりピア情

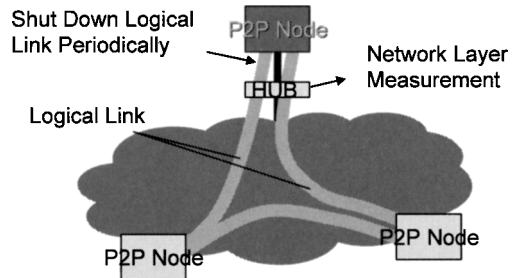


図 1 オーバレイネットワークのトラフィック測定
Fig. 1 Traffic measurement for overlay network.

報が収集可能であり、適用範囲は広い。

本技術の応用例として [11] でも提案したトラフィック識別法がある。識別においては、本手法によって得られたノード情報と特定リンクを流れるトラフィックのヘッダ情報の照合による特定リンク上での P2P トラフィックが識別できる。

P2P トラフィック識別法の手順は以下のとおりになる。

(1) 任意の測定点で 4.2 で述べた手法を用いて論理ネットワークでのピア情報を収集。

(2) P2P トラフィックを含むバックボーンネットワークでレイヤ 3 の測定を行い、IP アドレスとポート番号を含んだヘッダをダンプ。

(3) (1) で得られたピア情報とヘッダ部分がマッチしたパケットを P2P トラフィックと識別する。

本手法では、ペイロード監視が不要であるため、バックボーン等の高集約高帯域回線でもスケラビリティを保ったままで適用できる。

4.4 提案測定法の課題

これは他の測定方法にも共通していえることであるが、ある特定のネットワークリンクやオーバレイでの論理リンクでの測定では、P2P ネットワークの特質上全体の情報を得ることは困難である。もっとも、提案手法は論理レイヤでの測定を併用しているために、ネットワークレイヤのみでの測定で見られるような地理的条件には束縛されず、十分な測定サンプルを取得できればこの問題を解決できる可能性が高い。

しかし、P2P ネットワークにおいてはそのノード規模は膨大であり、ノードアドレスの頻繁な変更や出入りが発生するため、十分な測定サンプルを得ることは容易ではない。

図 2 は Winny において論理ネットワークのトラフィック測定から、2 拠点からのそれぞれの収集結果と、両測定点での測定結果を合わせてユニークなアドレス

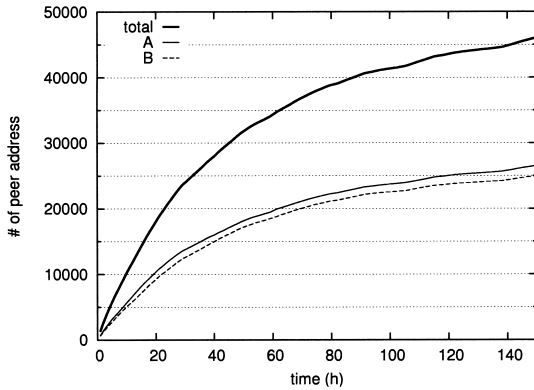


図 2 Winny における複数測定点からのピア情報収集状況
Fig. 2 Peer information from multi-point measurement.

を抽出したものと比較である。測定を行うノードを増やしても収集状況が容易には向上しないのが見取れる。また、両地点からの測定結果に重複してピア情報が含まれるため、2 点で測定しても収集能力が単純に 2 倍にはなっていないことが示されている。

このため、実ネットワーク上での運用において P2P トラフィックの識別や、ノード状況の精査に本手法を適用するためには、全体のどの程度の部分が測定できているのかを把握することが重要となる。これは、測定対象となっている P2P ネットワークの全体規模を推定することとほぼ同意である。

5. P2P ネットワークの規模推定

5.1 既存推定法とその問題点

全網把握が困難な大規模システムに対して、その部分測定から全体を推定する方法として、ウェブページの規模推定法 [26] が提案されている。ここでは、求めるウェブの総ページ数を N としたときにあるクエリー x に対して検索エンジン a が返すページ数を N_a 、検索エンジン b が返すページ数を N_b 、両者の重複ページ数を N_{ab} としたとき、収集の偏りがないとすると、検索エンジン a の収集率 $P_a = N_a/N$ は、 N_{ab}/N_b と等しいと推定される。これより、求めるウェブの総ページ数は $N = N_a N_b / N_{ab}$ と推定される。

以上を P2P ネットワークに適用する場合、推定対象はある P2P アプリケーションによって構成される P2P ネットワークにおける参加ピア動作ノードのアドレス数となり、論理ネットワークでのトラフィック測定によって得られたピア情報からの推定が可能となる。モデルとしては、その規模の大きさゆえに一部分しか

収集できない情報を用いて全体の規模を推定するというウェブの場合に近い。

しかし、単純に収集したピア動作ノード数を用いる場合、収集ピア数そのものは図 2 のように単調増加となるために推定全体ピア数も単調増加となり、一定数には収束しない。推定結果が一定数に収束するためには、全ピア情報を収集し、収集ピア数が定常状態になるまで測定点を増やす必要がある。これでは全ピア情報の収集を行うことになり、推定の意味がない。

このような現象が起きるのは、測定期間中に P2P ネットワークにリアルタイムで各ピアが参加、退出を行うためと考えられる。したがって、大規模 P2P アプリケーションのように部分測定しかできないシステムで、なおかつノード数が測定期間中に変動するようなシステムでは推定法の改良が必要と考えられる。

5.2 提案手法

5.2.1 既存手法の改良

ノード数変動の影響を減らすために、アクティブピア情報の取得と、測定範囲の限定の二つの方法による既存手法の改良を試みる。

前者は、刻一刻と移り変わるピアの情報に追従するために、P2P ネットワークにおいてアクティブなピア情報に限定して取得するという方法であり、手順は以下のとおりとなる。

- (1) 論理ネットワークのトラフィック測定によって得られたピア情報をもとに、該当 IP アドレスのサーバサポートに対して SYN パケットを送信。
- (2) SYN パケットに対して戻ってきたパケットが、SYN ACK の場合はノードがアクティブであると判断し、アクティブピアとして扱う。
- (3) その他の場合は無効とする。
- (4) 一定時間ごとに各ピアの状態をポーリングする。

後者の測定範囲の限定は必須の修正ではないが、ピア数の変動率を抑え、収束速度を早めるための修正である。すべての P2P アプリケーションに実装されているわけではないが、P2P アプリケーションの中には論理リンクの設定先を一定のルールに基づいて限定するものがある。例えば、Winny では探索ファイルを意味するクラスタワードや、回線帯域によってピアの優先度を決定する。これによって、興味や回線速度の近いピア同士が近隣に固められる。このような性質の利用により、限定されたピア集合への測定が実施可能となり、収束時間を短縮できる。

以上をまとめると、P2P システムの規模推定手順は以下ようになる。

(1) 測定点 A, B を設け、P2P アプリケーションを動作。

(2) A, B にそれぞれ異なるクラスタを測定するように設定し、測定範囲を限定。

(3) 論理ネットワークの測定を実施。得られたピア集合をそれぞれ、 N_a , N_b とする。また、両測定点で重複して得られたピア集合を N_{ab} とする。

(4) 得られたピアの IP アドレスとポート番号に対して一定時間ごとに SYN パケットを送信。

(5) SYN ACK が得られたピアをアクティブピアとし、A, B, $A \cap B$ のアクティブピアをそれぞれ、 A_a , A_b , A_{ab} とする。

このとき、P2P ネットワークに参加しているアクティブなピア集合を A とすると、全数 $|A|$ は、

$$|A| = \frac{|A_a| + |A_b|}{|A_{ab}|}$$

と推定される。

本手法ではピア集合の収集がそれぞれ独立で、かつ偏りがないことを前提としている点が誤差の原因となる可能性がある。具体的には独立性が低い場合、 $|A_{ab}|$ の値が小さくなるため、推定全ピア数 $|A|$ は小さめに出ると予測される。

また、SYN パケットによる生存確認においては NAT 内のノードが離脱ノードとして扱われるため、こちらにも全ピア数を小さめに見積もる要因となる。

5.2.2 ポートの使用比率を利用した規模推定

デフォルトのサービスポート利用が一定割合を占めるアプリケーションについては、バックボーンにおけるネットワークレイヤの測定と組み合わせた上で、ポートの使用比率を用いて規模が推定できる。

手順は以下のとおりになる。

(1) P2P トラフィックを多数含むバックボーンリンクで、4.3 で述べた手法で P2P 通信を行っているノード情報を抽出、その集合を R_p とする。

(2) バックボーントラフィックにおいて、デフォルトポート利用のノード集合を抽出、 R_d とする。

(3) R_d と R_p の重複を R_{dp} 、P2P ノードの集合全体を N とする。

これら集合に偏りがないとすると、 $|R_{dp}|/|R_d| = |R_p|/|N|$ が成り立つ。このとき、全数 $|N|$ は、

$$|N| = \frac{|R_p| \cdot |R_d|}{|R_{dp}|}$$

と推定される。

本方式にはアクティブピアと非アクティブピアの存在は考慮されていないため、5.2.1 の手法とは逆に、推定結果が大きくなる。もちろん 5.2.1 の手法と同様に各ノードへのパケット送信により生存状態を管理することにより回避可能だが、ある程度容量が大きいバックボーンにおいて収集したアドレスに対してアクティブ測定を行うのは現実的ではない。

また、本手法では、物理バックボーン網での測定を行うため、測定バックボーンに収容されている利用者の特性によりバイアスがかかるという問題もある。

本手法はもともとデフォルトのポート番号をもたないアプリケーションに対してはそのまま適用できないが、現状のファイル共有ソフトにおいては利用ポート番号は偏る傾向があり、実際にデフォルトポートが存在しない Share においても 10% 程度の利用者が上位 3 種類のポートを利用している。更に、ポート番号に限らずネットワークレイヤで取得できる情報を併用できれば、本手法は適用可能であり、例えば一部のトラフィックが簡易なシグネチャで識別できる場合等にも応用できる。

5.2.3 複数地点測定による規模推定

論理網上でピア情報の収集は比較的容易に多数の測定を行うことができる。この特性を利用して、論理網上で複数測定点からのピア情報収集を実施し、測定点同士が互いの存在を発見するまでの時間分布を用いた推定法が考えられる。

手法としては、

(1) P2P ネットワーク上に n 個所の測定点を設け、ピア情報の収集を開始。

(2) 全ピア数を N とし、各ノードが同じ速度でピアの収集をランダムに行っていると仮定する。

(3) ある一つの測定点に注目した場合、このノードが初めて残りの $n-1$ 個のノードのうちの一つを発見するまでの収集試行回数を X とする。

(4) このとき、 X は幾何分布に従うので、 X 回目に初めて残りの $n-1$ 個のノードのうちの一つを収集する確率は、

$$Pr\{X = x\} = (1-p)^{x-1}p$$

となり、期待値 $E[X] = 1/p$ となる。

(5) $p = n/N$ より、 $E[X] = N/n$ 。

本方式においては、収集対象となるノードをアクティブなピアに限定することにより、ノードの出入り

も考慮できる。

5.2.4 推定結果の誤差

したがって、これら三つの手法による規模の推定結果は、定性的には、

$$\begin{aligned} (\text{ウェブ推定法}) &\leq (\text{複数地点測定}) \\ &\leq (\text{ポート使用比率}) \end{aligned}$$

となることが予測される。

なお、このようにして得られた結果についてその絶対的推定誤差を示すことは、ウェブの規模推定を行った文献 [26] ~ [30] 等にもあるように困難である。したがって、実システムに対する適用においては、複数の手法を同時に適用し、それらの結果及び定性的な予測との比較からその妥当性を評価する必要がある。また、実際に 2.2 で述べたようなネットワーク運用や、分散アプリケーション開発時のパラメータ決定に用いる場合は継続的な測定によるノード数の変化傾向をとらえることが重要であり、提案するような測定及び推定を継続的に行うことは有効だと考える。

6. 実システムへの適用と推定

提案測定法と推定法の適用対象として、国内で最も広く用いられている P2P アプリケーションの一つである Winny (Version2)[25] を用い、論理ネットワーク上の複数地点において測定を実施した。論理ネットワーク上のトラフィックの測定手法に関しては、4.2 で述べたように図 1 のような環境下において論理ネットワークに接続されたアプリケーションノードを複数台用意し、IP レイヤでの測定を実施。ピア情報として通信相手の IP アドレスと利用ポートを収集した。

この際収集したピア情報について通信を確認した後、1 時間ごとに SYN パケットを送ることにより、生存確認を行った。測定結果を図 3 に示す。

また、Winny では同じ単語を登録したピア間での接続を優先的に行うクラスタワードという仕組みがあり、これによって同じ興味をもつピアを近づける。今回の測定では「映画」と「音楽」というクラスタワードを設定し、収集するピア空間を限定した。これらのキーワード選定にあたっては、24 時間 Winny ネットワークに接続してファイル名空間を得、その中からそれぞれの空間が独立であり、なおかつ多数のサンプルを収集可能であるキーワードを選んだ。表 1 に「映画」、「音楽」それぞれのクラスタワードを指定したときと、指定しなかったときのファイル名空間を示す。ここで、

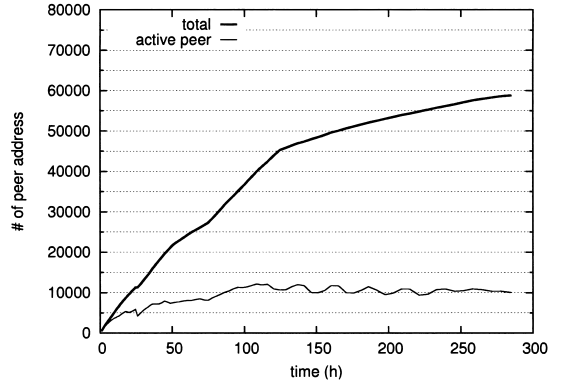


図 3 Winny のピア情報収集結果
Fig. 3 Peer information for Winny.

表 1 Winny のファイル名空間
Table 1 Names of files on Winny network.

指定クラスタ	なし	映画	音楽
ファイル数			
映画	661	2011	535
音楽	5084	5260	7656
重複	43	76	116
全体	39881	43984	37901

音楽ファイルは mp3 を名前に含むファイル、映画は「映画/邦画/洋画」のどれかを同じく名前に含むファイルとしているため、映画のファイル数は少なめに見積もられている。

生存確認の SYN パケットに対し SYN ACK を返したノードをアクティブピアと定義し、映画と音楽それぞれのクラスタごとに収集した全ピア数と、各クラスタごとのアクティブピア数の変遷を図 4 に示す。

アクティブピア数の変遷については、ほぼ 24 時間での周期性が認められる。しかし、ウェブ等の通信状態について同様の測定をした場合と異なり、全アクティブピアに占める変動率は 10% 前後と極めて小さい。ここから、Winny の利用者の多数は PC の電源を常時 ON にした状態で長時間サービスを起動していることが分かる。

6.1 ウェブページの規模推定法の適用

5.2.1 で述べたように、映画、音楽の各クラスタで収集したアクティブピアをそれぞれ、 A_a 、 A_b 。両クラスタで共通に収集されたアクティブピアを A_{ab} とすると、それぞれの収集率に偏りがないとすれば Winny の測定時点における全参加ピア数 A は $A_a A_b / A_{ab}$ と推定できる。図 5 に推定参加ピア規模の変遷を示す。

以上より、Winny のアクティブピアの規模は本手法

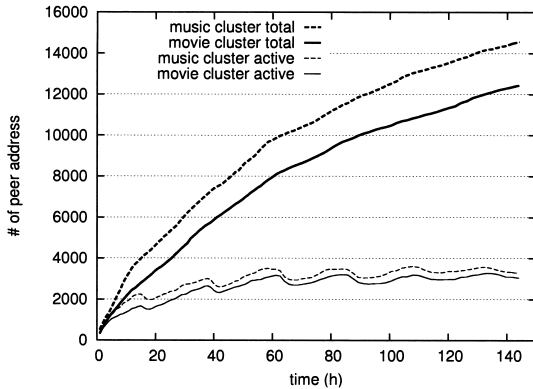


図4 Winny のクラスタごとのピア情報収集結果
Fig. 4 Peer information for Winny with cluster.

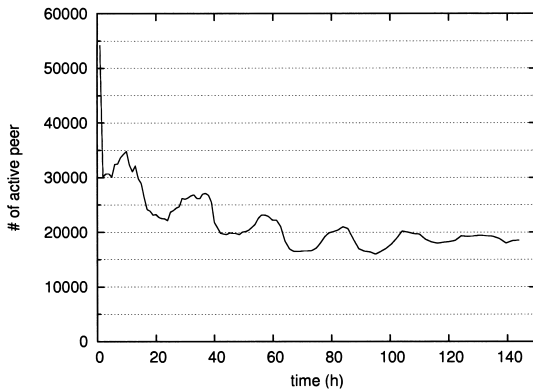


図5 Winny の推定参加ピア規模
Fig. 5 Scale of Winny network.

によるとおよそ2万弱と推定される。また、結果の収束時間から、あるクラスタ内のピア情報を偏りなく収集するためには80~100時間程度の連続測定が必要であると考えられる。なお、Winnyはその特性上、クラスタは他クラスタと完全に遮断されていないため、ここでの推定結果はWinny全空間のノード数になる。

6.2 ポートの使用比率を利用した規模推定

ピア情報収集と時間帯を併せて、ある組織のバックボーントラヒックの送受信アドレスとポート番号の組の測定を24時間実施した。その結果、バックボーントラヒックのうち、7743ポートを利用しているノード数 $|R_d| = 12,013$ 、収集ピア情報を用いて抽出されたノード数 $|R_p| = 6,698$ となった。また、これらの重なりである $|R_{pd}| = 2,041$ である。

したがって、5.2.2で述べた手法によるピアの全数 N' の推定結果は、 $|N'| = |R_p||R_d|/|R_{pd}| = 39,423$

となる。

図3を参照して1時間当りの N からノードの離脱数 v_N をおよそ350ホスト/時間と見積り補正を行った結果、 $N = N' + 24 \times v_N = 30,843$ となる。

6.3 複数地点測定による規模推定

六つのピア情報収集ノードを設置し、それぞれのノードの遭遇までに要する時間を測定した。物理的な測定点は1個所から行ったが、P2Pネットワークにおいては論理系エントリーポイントを変更することにより、独立なアドレス収集が可能である。測定を5回行ったところ、遭遇に要する平均値は26.2時間となった。

また図3より、アクティブピアの収集速度 v_a は原点付近ではおよそ250ホスト/時間になることから、5.2.3で述べた手法を用いることにより、

$$\frac{N}{5} \times \frac{1}{v_a} = 26.2$$

となり、 N は32,750と推定できる。

6.4 考察

以上、三つの手法を用いたWinnyのノード規模の推定を実測結果を用いて行った。ウェブページの規模推定法の改良による推定結果は5.2.1でも述べたように、試行の独立性が薄いために結果が少なめに出ている可能性があるため、2万弱というこの結果がおおよその下限であると考えられる。

一方、ポート番号の使用比率を用いた推定では、5.2.2でも述べたとおり、ノードの生存状況を考慮していないために補正前では結果が大きめに出ると予測される。また、バックボーン測定を行った個所での収容ユーザの特性が強く出ている可能性もある。

複数地点測定による推定結果とポート番号の使用比率を用いた推定結果にアクティブピアの補正を加えた結果は、これら二つの推定結果の間に位置する結果となっており、定性的な予測からも結果の信頼性は高いと考えられる。

以上の結果から、WinnyのP2Pネットワークにおいてはおよそ3万強程度のピアがアクティブ状態で存在していることが示された。

7. むすび

本論文では全体測定が困難であるような大規模自律分散ネットワークにおける全体推定法の提案を行い、実際の大規模P2PネットワークであるWinnyのネットワークについて提案手法の適用を試みた。その結果、3万強程度のアクティブピアの存在を明らかにし、本

手法の有効性を示すことができた。

文 献

- [1] “KaZaA.” <http://www.kazaa.com>, 2001.
- [2] “Gnutella.” <http://www.gnutella.wego.com>, 2000.
- [3] “SLYCK — File sharing news and info.” <http://www.slyck.com>, 2001.
- [4] D.P. Reed, “That sneaky exponential — Beyond metcalfe’s law to the power of community,” Context magazine Spring, 1999. <http://www.reed.com/Papers/GFN/reedslaw.html>
- [5] “skype.” <http://skype.com>, 2003.
- [6] I. Stoica, R. Morris, D. Karger, M.F. Kaashoek, and H. Balakrishnan, “Chord: A scalable peer-to-peer lookup service for internet applications,” ACM SIGCOMM, pp.149–160, Aug. 2001.
- [7] S. Ratnasamy, P. Francis, M. Handley, R. Karp, and S. Schenker, “A scalable content-addressable network,” ACM SIGCOMM, pp.161–172, Aug. 2001.
- [8] Napster Inc. <http://www.napster.com>, 1999.
- [9] IRTF Research Groups, “Peer-to-peer research group.” <http://www.irtf.org/charters/p2prg.html>, 2003.
- [10] 亀井 聡, 森 達哉, 大井恵太, “P2P ファイル共有の実態と課題,” 信学技報, CQ2003-40, July 2003.
- [11] 亀井 聡, 森 達哉, 木村卓巳, “P2P トラヒック分離法とその評価,” 2003 信学ソ大, B-7-19, Sept. 2003.
- [12] S. Sen and J. Wang, “Analyzing peer-to-peer traffic across large networks,” IEEE/ACM Trans. Netw., vol.12, pp.219–232, April 2004.
- [13] C. Fraleigh, S. Moon, C. Diot, B. Lyles, and F. Tobagi, “Packet-level traffic measurements from the sprint IP backbone,” IEEE Netw. Mag., vol.17, no.6, pp.6–16, Nov. 2003.
- [14] 岡田昭宏, 川原亮一, “IP 網におけるトラヒック特性分析の一考察,” 信学技報, NS2003-5, April 2003.
- [15] S. Sen, O. Spatscheck, and D. Wang, “Accurate, scalable in-network identification of P2P traffic using application signatures,” 13th International Conference on World Wide Web, pp.512–521, May 2004.
- [16] “Hack the KaZaA participation level.” <http://www.davesplanet.net/kazaa/>
- [17] “Ellacoya networks.” <http://www.ellacoya.com>
- [18] “Sandvine.” <http://www.sandvine.com>
- [19] “Snort.org signature search.” <http://www.snort.org/cgi-bin/sigs-search.cgi?sid=P2P>
- [20] D. Plonka, “Uw-madison napster traffic measurement.” <http://net.doit.wisc.edu/data/Napster/>
- [21] T. Karagiannis, A. Broido, M. Faloutsos, and K. Claffy, “Transport layer identification of P2P traffic,” ACM SIGCOMM/USENIX Internet Measurement Conference (IMC’04), pp.121–134, Oct. 2004.
- [22] Caspian Networks Inc., “P2P control.” <http://www.caspiannetworks.com/page.asp?intNodeID=12&intPageID=193>
- [23] S. Saroiu, P.K. Gummadi, and S.D. Gribble, “A measurement study of peer-to-peer file sharing systems,” Multimedia Computing and Networking (MMCN), pp.156–170, Jan. 2002.
- [24] NETARC, “P2PFinder.” <http://www.netarc.jp/p2pfinder.html>
- [25] aki, “winny.info.” <http://winny.info>
- [26] S. Lawrence and C.L. Giles, “Searching the World Wide Web,” Science, vol.280, no.5360, pp.98–100, 1998.
- [27] S. Lawrence and C.L. Giles, “Accessibility of information on the Web,” Nature, vol.400, no.6740, pp.107–109, 1999.
- [28] K. Bharat and A. Broder, “A technique for measuring the relative size and overlap of public web search engines,” Proc. 7th International World Wide Web Conference, pp.107–117, April 1998.
- [29] 内田 斉, “メディアとしての web の成長を測る,” マス・コミュニケーション研究, no.62, March 2003.
- [30] 水谷正大, 大森貴博, 来住伸子, 小川貴英, “検索エンジンを利用した日本語 web ページ数の統計的推定の研究,” 東京情報大学論文集, vol.5, no.1, pp.33–45, 2001.
(平成 17 年 3 月 2 日受付, 6 月 14 日再受付)



亀井 聡 (学生員)

平 9 京大・工・数理卒。平 11 同大大学院・工・応シス修士課程了。同年日本電信電話(株)入社, サービスインテグレーション基盤研究所所属。次世代ネットワーク技術の品質及びトラヒックにかかわる性能評価, 制御技術の研究に従事。平 16 から京大大学院情報学研究所システム科学専攻博士後期課程在学。情報処理学会会員。平 17 本会学術奨励賞受賞。



内田 真人 (正員)

平 11 北大・工・情報卒。平 13 同大大学院修士課程了。同年日本電信電話(株)入社。平 17 九工大・ネットワークデザイン研究センター・助教授。博士(工学)。通信トラヒックデータ解析, 統計的学習理論の研究に従事。平 15 本会コミュニケーションクオリティ研究会奨励賞受賞。平 16 本会学術奨励賞受賞。



森 達哉 (正員)

平 9 早大・理工・応用物理卒, 平 11 同大大学院修士課程了。同年日本電信電話(株)入社。博士(情報科学)。コンピュータネットワークの計測と分析に関する研究に従事。



高橋 豊 (正員)

昭 50 京大・工・数理卒．昭 55 同大大学院博士課程単位取得退学．同年同大・工・助手．平元同助教授．平 8 奈良先端大情報科学研究科教授．平 11 より京大大学院情報科学研究科システム科学専攻教授．工博．待ち行列理論・トラヒック理論，情報システム・情報ネットワーク等のモデル化と性能評価に関する研究に従事．Telecommunication Systems などの Associate Editor．IFIP WG6.3 Former Cochairman．Elected Full Member of IFIP WG6.3 and WG7.3．日本オペレーションズ・リサーチ学会，情報処理学会，応用数学会，システム制御情報学会各会員．