5　機械学習理論の応用

5 Applied Machine Learning Theory

5-1　P2P 環境におけるネットワークトラフィックのモニタリングと解析

5-1 Monitoring and Analysis of Network Traffic in P2P Environment

班　　涛　　安藤類央　　門林雄基
BAN Tao, ANDO Ruo, and KADOBAYASHI Youki

要旨

通信ネットワークに対する最近の統計的研究によれば、ピアツーピア（P2P）ファイル共有が増加の一途をたどり、現在では全インターネットトラフィックの約 50 ～ 80 %を占めることが分かっている[1]。また、ストリーミング、インターネット電話、インスタントメッセージといったネットワークアプリケーションは、P2P 通信の形態を取るものがますます増えている。P2P アプリケーションは本質的に多くの帯域を占有するため、P2P トラフィックは利用ネットワークに大きな影響を与える可能性がある。そのため、この種のトラフィックを解析し、その特性を明らかにすることは、作業負荷モデルを作成し、かつネットワークトラフィックのエンジニアリング及び容量計画を改善する上で不可欠な作業である。本稿は、有用な P2P トレースの捕捉と解析を実施するための、適応システムについて紹介する。そのシステムは、限られたリソースを効率的に体系化することによって、信頼性とトレース可能性を共に備えたネットワークを構築することができる。システムが捕捉したトレースの解析を通じて Winnyの挙動特性を明らかにした上で、非常に興味深い結果を報告する。

Recent statistical studies on telecommunication networks outline that peer-to-peer (P2P) file-sharing is keeping increasing and it now contributes about 50–80 % of the overall Internet traffic [1]. Moreover, more and more network applications such as streaming media, internet telephony, and instant messaging are taking a form of P2P telecommunication. The bandwidth intensive nature of P2P applications suggests that P2P traffic can have significant impact on the underlying network. Therefore, analyzing and characterizing this kind of traffic is an essential step to develop workload models and possible amelioration in network traffic engineering and capacity planning? In this paper, we introduce an adaptive system for handy P2P trace capture and analysis. The system can efficiently organize limited resources to build a network both reliable and tractable. Traces captured by the system are analyzed for characterization of Winny behavior with very interesting results reported.

[キーワード]
トラフィックのモニタリングと解析，仮想機械，P2P ネットワーク，ファイル共有
Traffic monitoring and analysis, Virtual machine, Peer to peer network, File sharing

1 はじめに

インターネットの信頼性、可用性及び安定性を高めるため、研究とネットワーク管理の両面において、P2P ネットワーク解析という研究分野が注目を集めている。その理由は次の点にある。(1) 現在、インターネットに占める P2P トラフィックの比率が高く、しかも増え続けている。(2) 多数の P2P アプリケーションは多くの帯域を占有するために、過度のネットワーク輻輳を招くほか、ユーザの不満やチャーン(短期間で他社に乗り換えること)の原因となる。(3) P2P ファイル共有は主に著作権侵害などの法的問題によって、常に多くの論争を巻き起こしている。(4) 大半の P2P クライアントはマルウェアの攻撃に対して脆弱であり、対策を怠ると深刻な情報漏えいやその他の危機的問題に発展する。

ところが、複雑な特性を示す、現在発展途上の P2P ネットワークアプリケーションを解析するには、従来のモニタリング及び解析方法では対応できない。第 1 の理由は、現在の P2P ネットワークが従来のものより高度なインフラとなり、トラフィックパターンが従来のアプリケーションに比べて複雑化している点である。第 2 の理由は、現在の大半の P2P ネットワークが、カスタマイズされた、または動的に割り当てられたポート番号を使用していることである。 P2P クライアントは、HTTP のポート 80 上においても容易に動作できる。そのため、IANA[2] によって割り当てられた周知のポート番号を基に動作する従来の解析方法が、P2P トラフィックの解析には適用できない。三つ目として、フィルタ方式のファイアウォールと法的問題の両方を回避するため、最近の P2P アプリケーションはカスタム設計による非標準の独自プロトコルで動作することに加え、生成されるトラフィックを通常のトラフィックに偽装する。四つ目に、P2P プロトコルはペイロード暗号化をサポートする傾向が強まっている。

P2P ネットワークのモニタリングと解析をするための使いやすいソリューションの実現に向け、本稿では二つの問題について取り上げる。最初に、限られた量のネットワークリソース及びコンピュータリソースを効率的に使用して、比較的大きな P2P ネットワークを実現するシステム構成を提案する。そのシステムにはアピールポイントが幾つか存在する。一つ目は、システムが収集するトラフィックトレースが実際のネットワークトレースと同じ特性を持つことである。二つ目は、複数の P2P ネットワークに容易に適応できることが保証される点である。最後のポイントは、大規模ネットワークにアクセスする必要がないことである。本稿で論じる二つ目の問題は、Winny ネットワーク挙動の特性評価である。Winny は日本で作成された代表的な匿名 P2P ネットワークである。Winny ネットワークの特性評価は、ネットワークエンジニアリングとセキュリティの面で極めて重要であると考えられる。私たちは主にフローレベルの情報を用いて解析する。

本稿の以下の構成は次のようになっている。**2** では、P2P ネットワークのトレーシングと特性評価に関するこれまでの研究を概観する。**3** では、ネットワークアプリケーションのトレース捕捉に対して私たちが提唱するシステム構成について述べる。**4** では、前記の提案システムを用いて実施した、P2P トラフィック解析の予備的実験結果を幾つか報告する。最後に、**5** でまとめをする。

2 P2P ネットワークトラフィックのモニタリングと解析に関するこれまでの研究

インターネットトラフィックのモニタリングと解析はこれまで常に便利なツールであり続け、ネットワークに悪影響を及ぼす脅威の防衛、重要なインターネットリソースの悪用もしくは不正利用の防止、悪意ある者もしくはソフトウェアによって生じる危害及び損害の最小化に役立ってきた。しかしながら、P2P ネットワークは通常前述のように、従来のインターネットアプリケーションに比べて複雑な特性を示すため、従来型のトラフィックモニタリング・解析システムを P2P トラフィックの解析に応用するには様々な処置が必要になる。本節では、P2P トラフィックのモニタリングと解析に関する研究について概観する。

2.1 ネットワークレベルのトレーシング

ネットワークレベルのトレーシングとは通常、ネットワーク設備の適当な地点において IP レベ

ルのパケットモニタリングを実施することを言う。ネットワークレベルのトレーシングはP2Pネットワークに対して透過的であり、しかも同時に複数のP2Pシステムを別分野のアプリケーションと分析・比較できるため、これまでこの方法が研究の主流であった。ネットワークレベルのトレーシングにおける欠点の一つは、ネットワークのアクセスポイントや識別精度によっては、大きな局所的偏りの発生が考えられることである。そのため、十分なトラフィックの標本を採取するにはネットワーク設備のキーポイント（例えば学術ネットワークにつながるゲートウェイ）にモニタリングプログラムを仕掛ける必要がある。

2.1.1　トランスポート層による解析

従来のアプリケーション、例えばウェブ、FTP、Telnetなどによって発生するネットワークトラフィックは、IANAポートリスト[2]に登録された周知のポートを用いて識別できる。P2P以前の時代であれば、1024未満のポート番号もしくはIANAポートリストに登録されたポート番号を用いることで、大半のインターネットトラフィックを十分に識別できた。しかし現在では、P2Pやストリーミングなどの新たなアプリケーションによるトラフィックを判別するのに、この方法は使用できない。ポート番号を用いた方法が使用できないのは、以下の状況が発生するためである。一つ目は、多くのアプリケーション（例えばMS Windows Media Server/Player）が動的なポート番号を使用すること。二つ目は、異なるアプリケーションが同じポート番号を同時に使用する場合があること。もう一つは、独自のプロトコルが未登録のポート番号を使用する可能性があることである。とはいえ、トラフィック解析においてトランスポート層の情報は極めて重要である。特にフローベース方式では、トランスポート層の情報に基づいてパケットのシーケンスを定義し、その統計情報を用いてその後の解析をする。文献[3]によれば、トランスポート層の情報は完璧ではないものの、依然としてP2Pトラフィックのかなりの部分の識別に役立てることができるという。

2.1.2　ペイロード検査による方法

ストリーミングやP2Pのアプリケーションで動的に割り当てられるポート番号の検出においては、ペイロード検査が最も強力であると同時に、最も多くのリソースを消費する方法であると推定される。メディアストリーミングの場合、クライアント・サーバ間で制御セッションとデータセッションが確立される。データセッションのポート番号は、制御セッションにおけるクライアント・サーバ間のネゴシエーションによって動的に決定される。したがって、制御セッションを詳しく調べればデータセッションのポート番号を見つけることができる。しかし、パケットペイロードの捕捉と解析は通常、法律、プライバシー及び財務上の壁にぶつかる上に、技術的な欠陥もある。一方、規定文書が不十分であるにもかかわらず増加の一途をたどるP2Pプロトコルをリバースエンジニアリングすることは、一般に退屈で、気が滅入る作業だと考えられる。他方、ペイロードの暗号化のためのP2Pプロトコルの場合、ユーザペイロードの復号作業は技術的に無理である可能性がある。ストリーミングトラフィックと他のインターネットトラフィックとをペイロード検査によって区別する作業には、mmdump[4]及びSMONのツール[5]が使用される。

2.1.3　シグネチャマッピングによる方法

幾つかの研究では、インターネットトラフィックをシグネチャ方式で識別する方法が一部の状況において有望であることが示されている[6]。ある種のアプリケーションにおいてシグネチャを抽出するには、他のパケットと区別可能な情報を含んだペイロードの一部をすべての関連アプリケーションについて検査する。それらのペイロードは通常、プライバシーの観点からIPヘッダとその直後の少数のバイトしか含まない。限られたペイロード情報では増加の一途をたどるアプリケーションに対応できないことがある。例えばP2Pアプリケーションがその存在を偽装する状況を考えてみよう。シグネチャ方式が有するもう欠点の一つは、各アプリケーションのシグネチャを見つけるために膨大なオフライン作業が必要になる点である。したがって、解析担当者の作業量を減らすには、自動シグネチャ生成が有望な方法である[7]。

2.1.4　フローレベルの特性評価

P2Pトラフィックの特性評価に関する研究には、パケットフローの統計量やパターンを用いるものがある[9][10]。パケット交換型ネットワークの場合、パケットフローないしトラフィックフ

ローとは、ある特定の発信元から一つのあて先に至る一連のパケットであると定義される。IP ネットワークの場合、トラフィックは発信元 IP、あて先 IP、プロトコル、発信元ポート及びあて先ポートの 5 要素によって、複数のフローに分割することができる。一般に普及しているフロータイムアウトは、文献[8]に提唱されている 64 秒である。言い換えると、ある特定のフローにおいて 64 秒の間にパケットが一つも到達しない場合、そのフローはタイムアウトする。トラフィックの特性評価にはホスト分布やトラフィックボリュームなどの特徴が使用される。フローベースの解析をすれば P2P トラフィックの特性について貴重な洞察を得られる一方で、アプリケーションレベルの詳細情報が得られないという制限がある。

2.1.5 ハイブリッドシステム

文献[4]には、P2P トラフィックの識別に関してペイロード方式による方法と非ペイロード方式による方法が提唱されている。ペイロード方式では、周知のポート番号、16 バイトペイロードに多く見られるシグネチャ、発信元及びあて先 IP アドレスといった発見的方法が用いられる。それに対し、非ペイロード方式ではユーザペイロード情報に関する知識は用いられず、ユーザを特定する TCP/UDP のペアやポートのペアに関する統計量が同時に多くの独特なコネクションを保持している。彼らの方法は P2P フローの 95 % が識別でき、偽陽性率（陰性の標本集団のうち、誤って陽性であると判定された標本の割合）は 10 % 前後であると報告された。文献[11]では、P2P トラフィックの解析においてパケットレベルの情報とフローレベルの情報が共に用いられている。実験では、異なる P2P ネットワークでは通信トラフィックの特性がかなり異なる場合があることが示されており、深い理解を得るために様々な P2P 環境における通信トラフィックを詳しく調査することが提唱されている。

2.2 アプリケーションレベルのトレーシング

発見したい特性によっては、ある特定アプリケーションのトラフィックに対するトレース及び解析に対して、アプリケーションレベルのトレーシングツールを使用するという方法もある。アプリケーションレベルのトレーシング方法は、モニタリングプログラムの動作モードに応じて以下の 2 種類に大別される。

2.2.1 受動的なアプリケーションレベルトレーシング

受動的なアプリケーションレベルトレーシングは、他のピアとの通信中に P2P ノードが送受信する、アプリケーションレベルのリソース発見メッセージ及びネットワークメンテナンスメッセージをモニタリングすることによって実施される。このとき使用されるのは通常、ルーティングを要求されるメッセージを受動的にロギングし、それ以外のインタラクションには参加しないように改変されたクライアントである。受動的なアプリケーションレベルトレーシングは、ネットワーク設備のキーポイントにアクセスする必要がなく、容易に実施できる。ただし、対象となる P2P ネットワークに対して透過的なだけで、P2P ネットワークの重要なサブセットをトレースすることは期待できない。

2.2.2 能動的なアプリケーションレベルトレーシング

能動的なアプリケーションレベルトレーシングは、ネットワーク設備にアクセスできないときにグローバルなネットワーク情報を発見するという問題に対して有効である。この方法はクエリングと接続に関して積極的な方針を採用し、モニタリング側のピアはできるだけ多くの P2P ネットワークに接続し、調査することを試行する。P2P ネットワークでクロールをするピアをトレースデータの大きさと典型度が最大になるように導くことができる。なお、このピアは再接続やリソース発見メッセージのより多い通常のピアとは挙動がかなり異なるため、収集されるトレースを鵜呑みにはできない場合がある点に注意する必要がある。

本稿では上述したすべての方式をバランスよく取り扱う予定である。一方において、ある特定のネットワークの状態を知るためにプロトコルを一つに限定したネットワークを構築し、そのネットワークのトレースを収集して解析する。このネットワークから収集されたデータは、当ネットワークのアプリケーションのみに属すると考えることができる。そのような排他的ネットワークには二つのメリットがある。一つは、アプリケーション

に特有の特性がトレースデータから抽出できる点である。二つ目は、データのラベル付けが高い信頼度で自動的に実施でき、ほとんど労力を要しない上、教師あり学習によって詳しい解析が可能である点である。他方、トレースの収集がネットワークレベルで実施されることから、ある P2P ネットワークに使用される方法を別の P2P ネットワークに対して容易に一般化することができる。その際、インストールした様々な P2P アプリケーションに対して同じ実験をすることにより、必要とされる P2P アプリケーションを蓄積することが可能である。このように、扱いやすいネットワークを一つ用意するだけで、十分なばらつきをもったトレースを得ることができる。大規模ネットワークのネットワーク設備にアクセスする必要はない。

3 P2P トラフィックトレーシングシステムの設計と実装

この節では、P2P トラフィックトレーシングシステムの設計と実装について述べる。図 1 はシステムの全体構成を示したものである。これは、ネットワークレイヤ、サーバレイヤ及び仮想機械レイヤの 3 層のレイヤから成る。

3.1 ネットワークレイヤ

ネットワークレイヤには、外部ネットワークへのアクセスと高速ストレージサービスの二つの機能がある。WAN インタフェースは、インターネットへの直接的なアクセスによってブロードバンドインターネットに接続される。ファイアウォールの検査ルールを細かく調節することにより、ローカルマシンにインストールされたクライアントから特定の P2P ネットワークにアクセスすることができる。他方、LAN インタフェースは信頼性の高い高速イーサネットに接続され、ローカルマシンはトレースファイルをストレージサーバに送ることができる。性能の観点から、私たちのシステムではローカルマシンとリモートストレージサーバとの接続に iSCSI プロトコルを採用している。ストレージサーバを別個に設ける目的は、ローカルマシンの作業負荷を軽減することである。なお、仮想機械によってはゲスト OS 及びトラフィック捕捉ツールが複数、同時に動くものがある。その場合、トラフィック解析を同時にはできないことがある。また、複数の仮想機械から収集したデータは統計的に、信頼できるネットワーク情報を提供してくれる可能性がある。

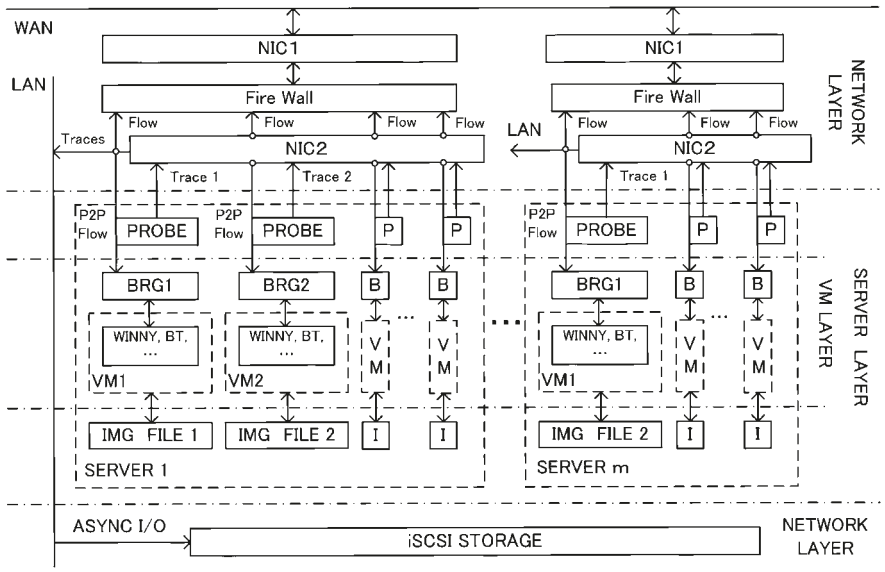

図1 システムの全体構成

3.2 サーバレイヤ

サーバレイヤにはシステムの仮想機械（VM）モニタ（別名ハイパーバイザ）[12]が実装され、実際の物理機械を異なる仮想機械間で多重化できる。物理機械のシステムサービスをシミュレートするハイパーバイザ上に複数のゲスト OS が実装され、それぞれが独立した機械として機能する。仮想機械を使えば物理機械のネットワークに比べて以下のメリットが得られる。(1) 1 台のコンピュータに複数の OS 環境が共存できる。しかも互いの独立性が強い。比較的大規模なネットワーク環境を多くの機械を使わずに構築する場合に役立つ。(2) システムリソースが効率的に使用できる。これはグリーンコンピューティングの主要課題の一つである。ある特定の P2P アプリケーションに関するトラフィックを収集する場合、一つの OS 上に一つの P2P クライアントが実装及び実行されるだけである。これは一般に、機械の処理能力の観点で大きな無駄につながる。しかし、仮想機械を使えば、システム仕様に従って適切な数のタスク、すなわち仮想機械の数を各サーバに割り当てることができる。同一量のリソースでより大きなネットワーク環境が実現できることは明らかである。(3) 三つ目の理由は、P2P ネットワークはセキュアでないため、システムが何らかの攻撃を受ける可能性が存在する。仮想機械はサンドボックスによって OS を保護するのに役立ち、想定されるリスクに対してハイパーバイザを安全に維持することに寄与する。(4) もう一つ重要な点は、仮想機械ではシステムの復旧及びリブートが高速で実施できるため、実験の再実行や他の P2P プロトコルの解析のためのシステム変更が、同数の物理機械を扱う場合に比べてはるかに容易なことである。

サーバレイヤのもう一つの機能は、各ゲスト OS に対して別々のトレースを捕捉し、そのデータをイーサネット経由でストレージサーバに送ることである。

3.3 仮想機械レイヤ

仮想機械レイヤでは、ハイパーバイザによってシミュレートされる仮想 NIC インタフェースによってゲスト OS とサーバレイヤが接続される。トラフィックはそこからインターネットに送出される。毎回、各ゲスト OS 上に一つの P2P クライアントのみを実装し、それを外部の P2P ネットワークに接続する。このとき、KVM、VMware、Xen といったハイパーバイザの最新のインプリメンテーションはトラフィック制御に対応していないため、帯域はゲスト OS が均等に共有する。異なるネットワーク条件がシミュレートできる多目的システムを構築するには、各ゲスト OS にトラフィック制御ソフトを実装すればよい。都合の良いことに、多くの P2P アプリケーションはファイル共有のために割り当てた帯域を制御するオプションを備えている。

私たちの実験では、Dell PE 2950 サーバにおいてそれぞれ 20 以上の仮想機械を運用した。また、手元の 6 台のサーバを使ってノード数が 100 を超える P2P ネットワークを構築することができた。私たちの目的が P2P ネットワークの単独シミュレーション環境を作ることでない点をここで再確認したい。帯域などのオプションを 100 個のノードで個別に設定することにより、信頼性の高いネットワーク特性を反映した P2P トレースを得ることができる。

4 実験

今回提案するシステム構成の実現可能性とパフォーマンスを調べるため、Winny、BitTorrent 及びその他のウェブアプリケーションという 3 種類のアプリケーションのトラフィック解析にこのシステム構成を用いる。6 台の PowerEdge サーバ各々に対し、Windows 2000 を実装した仮想機械を 16 台運用する。パフォーマンス上の理由から、ハイパーバイザソフトには Qemu を使用する。

4.1 実験 I

多くの第 2 世代 P2P ネットワークと同様、Winny もネットワークのスケーラビリティ改善のためにスーパーノードを使用する。二つのピア間で通信が確立されると、Winny は両者の上り帯域設定値を比較し、値が大きいほうのピアを格上のノードと見なす。その後、サーチは主に格上の方向に送出される。この仕組みによって三つのノードグループが生まれる。スーパーノードは他より広い帯域と高い計算能力を持ち、主に他のピアに

対するプロキシやIndexing Serverとして機能する。同時に、必要なファイルをダウンロードする機会もそれだけ多い。中間レベルノードは通常のネットワークリソース及び計算リソースを持ち、スーパーノードからファイルを取得して下位ノードにサービスを提供する。下位ノードは限定的なサービスを他のノードに提供するだけである。

Winnyネットワークでのノードの役割を決めるにあたり、上り帯域は明らかに最も重要なパラメータである。ノードの挙動に関係するもう一つの要因はWinnyの動作モードである。Winnyには、拡散クエリモードとダウンロード(アップロード)モードという二つの動作モードがある。拡散クエリとは、Winnyが近隣ノードに検索メッセージを送出して応答を受信するプロセスをいう。拡散クエリの処理中は、ピア間で主として制御メッセージが交換される。一方、ダウンロードの処理中は主にコンテンツメッセージが転送される。そのため、動作モードが異なる場合、捕捉されるトラフィックトレースに多少の違いがあると考えられる。第1の実験では、上り帯域と動作モードがWinnyの挙動にどのように影響するのかを確認することを目指している。Winnyクライアントの上り帯域を819.2、409.6、204.8、102.4、51.2、25.6、12.8、6.4 kbpsの8段階に設定した。またWinnyクライアントの半数を拡散クエリモードに、残りをダウンロードモードとした。以下に報告する結果は、収集したトレースを1時間ごとに平均した値である。

図2は、あるWinnyクライアントとその近隣ノードとの間で開始された会話の数を示す。ダウンロードモードのWinnyのほうが拡散クエリモードのものより一般に発生会話数が多いことが分かる。詳しい解析によると、帯域の小さいWinnyが生成した会話は通常、その接続速度と持続時間が小さい。拡散クエリモードの広帯域Winnyも、他ピアとのクエリ転送に積極的に参加しており、それによって比較的多くの会話が開始されている。

図3の結果は明瞭である。ダウンロードモードのWinnyでは、帯域が大きいほど単位時間に転送されるパケット数が多い。拡散クエリモードのWinnyでも状況は似ているが、転送されるパケット数はダウンロードモードほど多くない。図4は、

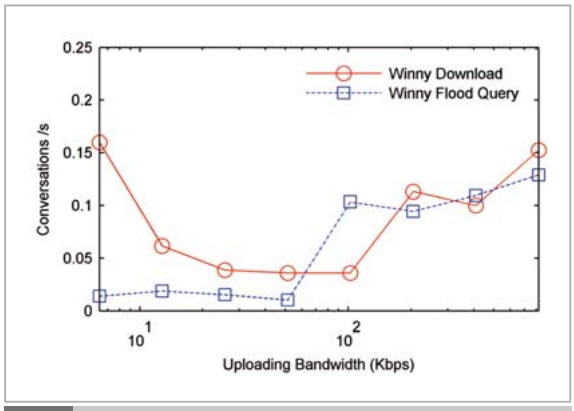

図2 1秒間の会話数

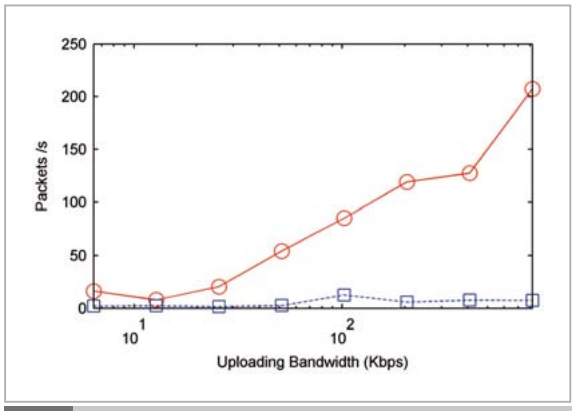

図3 1秒間のパケット数

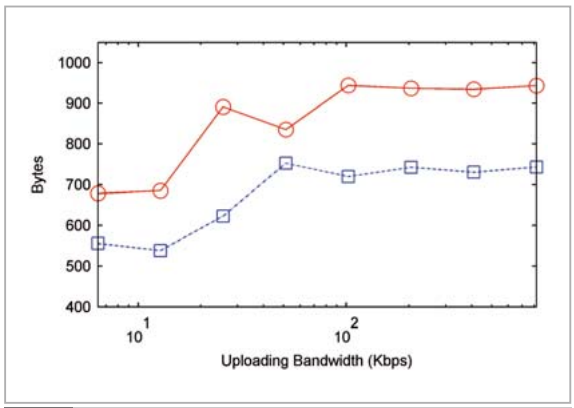

図4 平均トラフィックサイズ

制御メッセージとコンテンツメッセージはサイズが異なるとする私たちの予測が正しいことを示している。そのため、異なる動作モードにあるWinnyクライアントを特定するための基本的な発見的方法では、平均パケットサイズの違いを使用することができる。

4.2 実験Ⅱ

第 2 の実験では、各種ネットワークアプリケーションのパケットサイズ分布に対する判別能力について調査をする。Winny、BitTorrent 及びその他のウェブアプリケーション（ウェブブラウザ、FTP クライアント、SSH クライアントを含む）に対して同じ仮想機械ネットワークを構築する。

図 5 に、各記載アプリケーションのパケットサイズ分布を示す。Winny のトラフィックは両動作モードで分布が非常に似通っているように見える。このことは、Winny のトラフィック特性を知る上でパケットサイズ分布が信頼できる特徴となる可能性があることを意味する。BitTorrent とその他のウェブアプリケーションは Winny のトラフィックとかなり異なる。しかし、4 種類のトレースすべてで、40 ～ 79 と 1280 ～ 2559 のパケット長が多くなっている。このことは、パケット長分布がこれらのアプリケーションの大きな判別要素にならないことを示唆している。各アプリケーションを区別するには、パケット長の区間幅をもっと狭めた詳しい分布情報を調べるか、更なる知識や発見的方法を取り入れる必要がある。

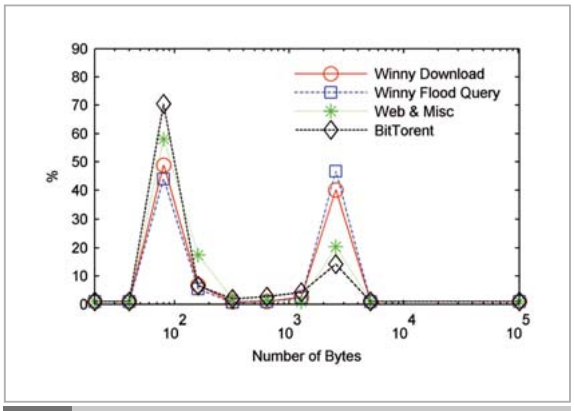

図5 パケットサイズ分布

5 まとめ

本稿では、仮想機械を用いたトラフィックモニタリングシステムの枠組みを提案した。仮想機械を用いれば、このシステムは使用可能なネットワークリソース及び計算用リソースを有効に活用することによって、比較的大規模なネットワーク環境を構築することができる。大規模なネットワークにアクセスしなくても、ある特定の P2P プロトコルのみを使用する排他的ネットワークを設定することにより、明解で信頼性の高いネットワークレベルのトレースが収集可能である。また、このシステムは、少しの労力で様々な P2P ネットワーク及びネットワークアプリケーションに適合するように修正することができる。

実験の節では、ネットワークによって収集したトレースが統計的に妥当な特性を示した。パケットレベル、フローレベル及びトランスポートレベルの情報を更に考慮すれば、このシステムは、発展しつつある P2P ネットワーク及び各種の新しいネットワークアプリケーションの挙動について知見を与えてくれる、有望なツールになると考えられる。

参考文献

1 http://www.ipoque.com/news_&_events/internet_studies/internet_study_2007
2 IANA. Internet Assigned Numbers Authority (IANA), "http://www.iana.org/assignments/port-numbers".
3 T. Karagiannis, A. Broido, M. Faloutsos, and K. Klaffy, "Transport Layer Identification of P2P Traffic", Proceedings of the 4th ACM SIGCOMM Conference on Internet Measurement (IMC 2004), pp.121-134, Italy, Oct. 2004.
4 J. van der Merwe, R. Caceres, Y. Chu, and C. Sreenan, "mmdump-A Tool for Monitoring Internet Multimedia Traffic", ACM Computer Communication Review, Vol.30, No.5, 2000.
5 H. Kang, H. Ju, M. Kim, and J. W. Hong, "Towards Streaming Media Traffic Monitoring and Analysis", APNOMS 2002, Sep. 2002, Jeju, Korea.
6 S. Sen, O. Spatscheck, and D. Wang. "Accurate, Scalable In-Network Identification of P2P Traffic Using Application Signatures". In Proceeding of the 13th International Conference on World Wide Web (WWW 2004), New York, NY, USA, 2004, pp.512-521.

7 P. Haffner, S. Sen, O. Spatscheck, and D. Wang, "ACAS: Automated Construction of Application Signatures", ACM SIGCOMM Workshop on Mining Network Data (MineNet 2005), pp.107-202, Philadelphia, PA, USA, Aug. 2005.
8 K. Claffy, H. W. Braun, and G. Polyzos. "A Parametrizable methodology for Internet traffic flow profiling", In IEEE JSAC, 1995.
9 S. Sen and J. Wang, "Analyzing peer-to-peer traffic across large networks", Proceedings of the 2nd ACM SIGCOMM Workshop on Internet measurement, pp.137-150, 2002.
10 M. Kim, H. Kang, and J. W. Hong, "Towards Peer-to-Peer Traffic Analysis Using Flows", DSOM 2003: 55-67.
11 R. Bolla, R. Rapuzzi, and M. Sciuto, "Monitoring and Classification of Teletraffic in P2P Environment", Proc. of the 2006 Australian Telecommunication Networks and application Conference (ATNAC 2006), Melbourne, Australia, Dec. 2006, pp.313-318.
12 J. E. Smith and R. Nair, "The Architecture of Virtual Machines". Computer 38 (5): 32-38. 2005.

班　涛(Ban Tao)
情報通信セキュリティ研究センター
レーサブルネットワークグループ専攻研究員　博士(情報工学)
ネットワークセキュリティ、機械学習

安藤類央(あんどうるお)
情報通信セキュリティ研究センター
レーサブルネットワークグループ研究員　博士(政策・メディア)
ネットワークセキュリティ、ソフトウェアセキュリティ

門林雄基(かどばやしゆうき)
情報通信セキュリティ研究センター
レーサブルネットワークグループ客員研究員　博士(工学)
ネットワークセキュリティ