

平成21年 5月20日現在

研究種目：基盤研究(C)

研究期間：2007～2008

課題番号：19500057

研究課題名（和文）パッシブ計測に基づく大規模広域ネットワークトモグラフィ

研究課題名（英文）Passive Measurement-Based Network Tomography  
For Large-Scale Wide-Area Networks

研究代表者

鶴 正人 (TSURU MASATO)

九州工業大学・大学院情報工学研究院・教授

研究者番号：40231443

研究成果の概要：大規模広域なネットワークにおける通信品質の劣化検知とその原因箇所（ネットワーク区間）の特定のために、パケットのサンプリング観測から通過フローの性能特性を調べ、性能特性間の相関に基づいて劣化区間を推定する手法（Network Tomography の一種）を検討した。計測や推定に必要な技術を考案または改良し、実用的な手法としての可能性を日本を横断する商用インターネット上での実験によって評価した。

交付額

(金額単位：円)

	直接経費	間接経費	合計
2007年度	2,200,000	660,000	2,860,000
2008年度	1,200,000	360,000	1,560,000
年度			
年度			
年度			
総計	3,400,000	1,020,000	4,420,000

研究分野：情報ネットワーク工学

科研費の分科・細目：計算機システム・ネットワーク

キーワード：インターネット計測

## 1. 研究開始当初の背景

インターネットは、既に重要な社会インフラであり、電話もインターネット利用の形態への移行が急速に進み、また、放送とインターネットの融合も始まりつつある。そこで、次世代インターネットにおいては、来るべきICT社会の基盤として、経済性・拡張性と信頼性・適正な（必要な）品質維持を両立することが求められている。そのためには、ネットワークの内部状態・特性を常時かつ実時間で把握し、品質を考慮して動的・適応的にネットワーク資源を割り当て、また利用していくことが必須である。

一般に、計測方式は、対象ネットワークの端に置かれた計測ノード（PC等）間で試験ト

ラヒックを送受信し、それを計測することで通過ネットワーク経路（パス）の状態を調べるアクティブ計測と、対象ネットワーク内部に置かれた計測ノードで局所状態やそこを通過する実トラヒックを計測するパッシブ計測に大別され、ネットワーク内部を調べるには後者の方法が直接的である。しかし、大規模・高速ネットワークの内部の多地点で常時かつ直接的に通過トラヒックをパッシブ計測・監視することはコスト的にも困難であるので、間接的または部分的な計測からネットワーク状態・特性を推定する技術が重要であり、活発に研究されている。

Network Tomography とは、空間的・時間的に異なる複数の複合的現象（要素的現象の異

なる組み合わせ)を観測することによって、元になる要素的現象を統計・情報理論的に推定する枠組みである。従来の Network Tomography では、通常、ネットワークの端(周辺部)に複数の試験トラフィック送受信用 PC を置き、アクティブ計測によって複数のパスの状態を調べ、その組み合わせから内部区間の局所状態を推定する。しかしその方式では、大規模ネットワーク全域をカバーするには、多数のアクティブ計測用 PC を配置し、すべてのペア間で試験トラフィックを交換する必要があり、コストが大きくスケールしない。そこで、パッシブ計測に基づく Network Tomography の研究も始まり、例えば、大規模 web サーバにおいて、そこへアクセスする多数のクライアントとの web トラフィックの packet ロス率を計測し、サーバを頂点とする木構造ネットワークの内部(各枝)区間の状態を推定する研究が有名である。しかし、packet ロスはそのほど頻繁に起こるものではなく、また、web トラフィックを計測する手法は、ひとまとまりの通信(TCP フロー)が短く、方向も限定的で、一般に大規模ネットワーク全域にわたる内部状態把握には適さない。そこで別のアプローチが必要である。

## 2. 研究の目的

大規模ネットワーク全域での内部状態の実時間把握において、内部の多地点でのパッシブ計測による直接計測は困難であり、他方、周辺部の多地点にアクティブ計測用の専用計測端末を多数設置することも同様に困難であるので、それらを克服するために、比較的少ない地点に設置した計測端末とインターネット上の多数の P2P ネットワーク参加端末との間の TCP フローのパッシブ計測に基づく新しい Network Tomography 推定手法を検討し、具体的には以下の課題に取り組む。図 1 に全体構想を示す。

まず、広域 P2P トラフィックの TCP フロー単位の計測技術の開発を行なう(i)。高速ネットワークにおいて多数のフローを同時にパッシブ計測するために、PC 上で動作可能な汎用的なソフトウェアベースの packet フィルタの高速化を行なう。さらに packet サンプリングに基づき多数の TCP フローの特性を比較的短い(ネットワークの状況の変化に追従できる程度の)時間スケールで把握する方式を確立する。

次に、TCP フロー特性を利用した Network Tomography 手法に基づくネットワーク内部の状態推定技術の開発を行なう(ii)。(i)で得た、多数の TCP フローの特性とそれらが通過するパスのトポロジー情報から、部分区間の状態変化(輻輳、障害等)を推定する方式を確立する。簡単な例として、図 1 の 2 本のピンク色のフロー性能が同時に悪化し、同時

に回復した場合、それらに共通する緑の(右向き)ネットワーク区間が性能劣化の原因になっている可能性が高い。このような推定を TCP フローの特性を考慮して効率よく行う。

さらに、大規模ネットワーク内部状態推定技術の開発とその評価実験を行う(iii)。前項(i),(ii)を連携し、PlanetLab 等の広域ネットワークテストベッドを利用して、インターネット上での実験により評価する。

本研究によって、従来実現できなかった、大規模ネットワーク全域にわたる内部状態の実時間での把握が実用化可能になり、インターネットの適応的品質制御・管理につながると期待できる。

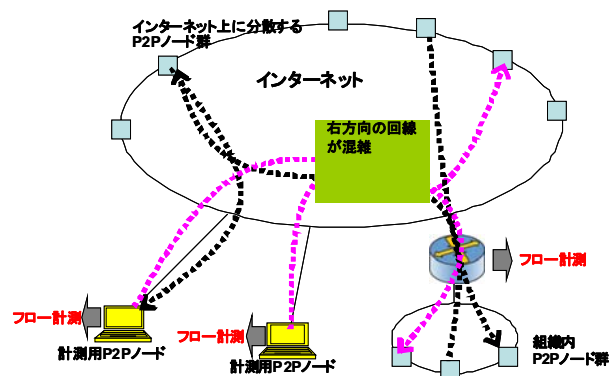


図 1 : 計測方式の全体構想

## 3. 研究の方法

(1) 広域 TCP のフロー単位の計測技術の開発 :

ソフトウェアベースの packet フィルタに関しては、ソフトウェアパイプライニングに基づくマシン語レベルの並列処理による高速化を調査・検討し、事前に実トラフィックから取得したサンプル packet 列を使い、Linux PC 上での模擬実験によって性能を評価する。

個々の TCP フローの特性を packet サンプリングから推定する方法に関しては、TCPDUMP による packet 内の TCP シーケンス番号の観測に基づく TCP フロースループットの推定を調査・検討する。さらに、packet サンプリングからフローの RTT や packet ロス率を推定する方法も調査・検討する。

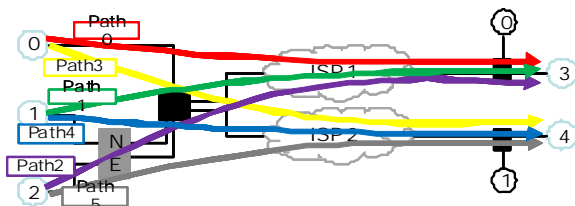


図 2 : 日本のインターネット上での基礎実験環境

なお、それらを実インターネット環境で評価するために、学内 LAN 経由の SINET 接続に加え、商用 ISP 3 社から高速光インターネットサービスを購入し、そこに計測用 Linux PC を接続する。これにより、日本中を横断する多様なパスに沿った TCP フローを計測できる実験環境を用意する。図 2 はその一例であるが、5 台のアクティブ計測（試験）端末、2 台のパッシブ計測端末（図の右上と右下の 0 と 1）、擬似的なロスや遅延を発生させるネットワークエミュレータから構成される。

(2) TCP フロー特性を利用したネットワーク内部の状態推定技術の開発：

前項(1)で計測・推定した各 TCP フローのスループットの時間変化の相関から、同じ原因で品質劣化（スループット低減）を被っているフロー群を推定（クラスタリング）する手法を検討する。TCP フローのスループット自身は RTT をはじめとする様々な環境パラメータによってその平均レートが異なり、またネットワーク区間の品質劣化以外の要因でも変動するために、それらを考慮したクラスタリングが必要である。

また、異なるフロー群の TCP レート変化の相関の分析にはそれらが通過する複数パスの共有区間の最小リンク帯域幅の情報が有用であるので、まずはアクティブ計測によるその推定の手法を検討する。将来的にはこれもパッシブ計測にできることが望ましい。

(3) 大規模ネットワーク内部状態推定技術の開発とその評価実験：

前項(1)、(2)を合わせて、さらに P2P トラヒックの誘導や計測、パスの経路トポロジーの推定などを行う計測・推定システム全体としての方式を検討・評価する。そのために推定結果の検証手法に関しても検討する。

これらの評価実験は、前項(1)で述べた、日本のインターネットを横断する実験環境をまず利用し、次に Planetlab に参加して、世界規模のインターネット環境において実施する。

#### 4. 研究成果

(1) 広域 P2P トラヒックのフロー単位の計測技術の開発：

高速ネットワーク上の多数フローを同時に計測するために必要なソフトウェアベース・パケットフィルタの高速化に関しては、条件分岐を含むループのためのソフトウェア・パイプライン化手法を適用し、インテル IA-64 Itanium 2 プロセッサを対象に、まず TCPDUMP の場合に多数のパケットの並列処理による高速化（雑誌論文[3]）を行い、さらに長いフィルタ規則の場合に多数のルールパターンの並列処理による高速化（雑誌論文

[1]) を達成した。図 3 は TCPDUMP の場合で、右の 3 つ（提案方式）はインタプリタ実装の約 20 倍、既存研究の最適コンパイル実装の約 2 倍の性能向上を示している。

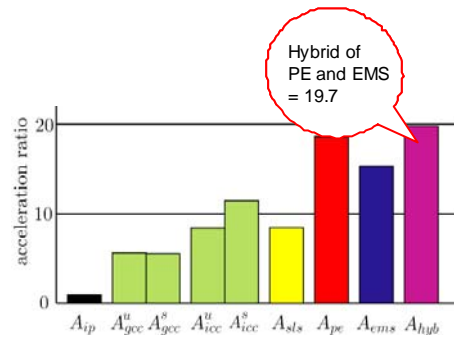


図 3：インタプリタに対する性能比

個々の TCP フローのスループットのパケットサンプリングからの推定に関しては、様々な環境でパッシブ計測とアクティブ計測の結果を比較して実験した結果、100 分の 1 のサンプリングレートでもある程度の特徴把握ができる目処を得た。

(2) TCP フロー特性を利用したネットワーク内部の状態推定技術の開発：

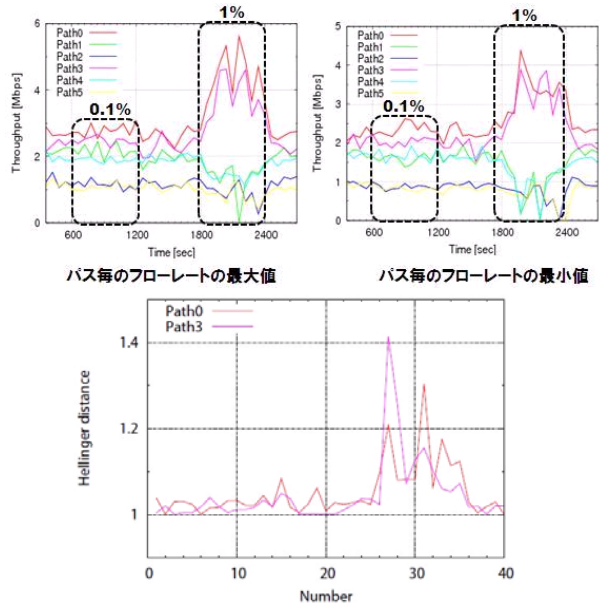


図 4：異なるパスを通る TCP フロー群の特性の時間変化の相関～スループットの最大値（上左）と最小値（上右）、スループット分布の変化のヘルンジャー距離（下）

前項(1)のパッシブ計測による TCP スループット推定を使い、同じパスを通る複数の TCP フロー（フロー群）の特徴量を抽出し、それらの相関から品質劣化区間を推定する手法を、実インターネットを用いた実験によって検討した。図 4 は、共通の輻輳区間（パケットロス率が 0.1% の場合と 1% の場合）を

通過するパスを類別し輻輳箇所を推定するためのフロー群の特徴量として、図上はある単位時間区間内のスループットの最大値と最小値を見た場合、図下は単位時間区間内スループットの分布そのものの時間変化（隣り合う分布間の Hellinger 距離）を見た場合の例である。前者の手法では短い時間の輻輳発生を検出できない場合があるので、後者を検討し、基本的な目処を得た。しかしスループットが低下する（輻輳区間を通過する）フロー群と向上する（前者のフロー群と他の場所で競合している）フロー群の区別が必要であり、最終目標の実現には引き続き研究が必要である。それに関連して、スループット以外の特性（ロス率や RTT）を推定し、それらを組み合わせてパスの類別を行う手法の検討も始めた。

一方、異なるフロー群の TCP レート変化の関連の分析にはそれらが通過する複数パスの共有区間の最小リンク帯域幅の情報が有用であるが、アクティブ計測によるその推定のための疑似パケットペアを用いた従来手法に対して、遅延時間と到着時刻差の両方考慮することで少ないパケット数での推定を可能する改良手法を考案した（雑誌論文[2]）。図5に示すようにネットワーク全体の利用率が50%という混雑した状況において、推定に必要な計測時間と計測パケット数が大幅に減っている。

	共有区間 利用率	20%	30%	40%	50%
現行手法	推定成功率	100%	100%	100%	80%
	推定時間	66sec	69sec	174sec	248sec
	パケット量	6000	6000	15000	20000
提案手法	推定成功率	100%	100%	100%	80%
	推定時間	22sec	22sec	22sec	22sec
	パケット量	2000	2000	2000	2000

図5：共有区間帯域幅推定手法に関する必要計測時間と計測パケット量の改善

### (3) 大規模ネットワーク内部状態推定技術の開発とその評価実験

前項(1)や(2)を組み合わせ、日本のインターネット上での実験環境を用いた手法の分析は行うに留まり、世界規模のインターネット環境を用いた評価実験には至らなかった。主たる理由は、フロー群のクラスタリングに関する手法が不十分であり、まだ検討の必要性がある点と、Planetlab への参加やP2P誘導方式などを含む評価実験環境の準備が遅れたためである。なお、品質劣化箇所特定の検証のためにアクティブ計測を用いて指定した区間のロス率や遅延分散を推定する

ツールに関しては、実験を通じて目処を得た（学会発表[1]）。今回の研究開発で得られた知見を生かし、計測・推定システム全体としての統合を行ない、大規模環境での評価を行うことは今後の課題である。

## 5. 主な発表論文等

（研究代表者、研究分担者及び連携研究者には下線）

〔雑誌論文〕（計3件）

- [1] 山下義行, 鶴 正人, “高速パケットフィルタの実装と評価”, 情報処理学会論文誌コンピューティングシステム (IPSJ ACS), 1(1):1-11, 2008年、査読有
- [2] 辻村直也, 鶴 正人, 尾家祐二, “疑似パケットペアを用いた共有区間最小帯域幅推定手法の改良と評価”, 電子情報通信学会技術報告, 108(457):21-26, 2008年、査読無
- [3] Y. Yamashita, M. Tsuru, “Software Pipelining for Packet Filters”, Proc. the 9-th International Conference on High Performance and Communications (HPCC07), LNCS4782:446-459, 2007年、査読有

〔学会発表〕（計1件）

- [1] Y. Sakai, M. Uchida, M. Tsuru, Y. Oie, “Locating Causal Segments Along A Performance-Degraded End-to-End Path,” the 10-th Passive and Active Measurement Conf. (PAM 2009) Poster Session, 2009年4月1日、ソウル大（韓国）

## 6. 研究組織

### (1) 研究代表者

鶴 正人 (TSURU MASATO)

九州工業大学・大学院情報工学研究院・教授

研究者番号：40231443

### (2) 研究分担者

なし

### (3) 連携研究者

なし

### (4) 研究協力者

山下義行 (YAMASHITA YOSHIYUKI)

佐賀大学・理工学部・教授

研究者番号：20220353