

令和 2 年 6 月 5 日現在

機関番号：32689

研究種目：基盤研究(C)（一般）

研究期間：2017～2019

課題番号：17K00135

研究課題名（和文）観測空間上の情報量を用いた確率モデルの探索・生成・統合に基づくトラヒック行列推定

研究課題名（英文）Traffic Matrix Estimation Based on Estimation, Generation and Integration of Probabilistic Models Using Information in Observable Space

研究代表者

内田 真人 (Uchida, Masato)

早稲田大学・理工学術院・教授

研究者番号：20419617

交付決定額（研究期間全体）：（直接経費） 3,600,000円

研究成果の概要（和文）：トラヒック行列推定においては、対地間フロー流量を直接観測できないという制約により、その推定精度も直接評価できないという問題がある。すなわち、トラヒック行列推定の精度を非観測空間（対地間フロー流量が物理的に取り得る状態の集合）上において直接評価することは原理的に不可能である。この問題を解決するために、本研究では、トラヒック行列推定の精度を、非観測空間上ではなく観測空間（集約フロー流量が物理的に取り得る状態の集合）上において間接評価する手法を提案した。また、実トポロジー上に定義されたトラヒック行列推定が、疑似トポロジー上に定義されたトラヒック行列推定に等価に変換できることに着目した手法も提案した。

研究成果の学術的意義や社会的意義

情報化社会の進展により、人々の社会・経済活動の基盤としての通信インフラの重要性が高まっている。これに伴い、通信障害や通信品質の低下が与える影響も高まっている。こうした事象を未然に防止し、迅速に復旧させるためには、ネットワークの内部状態を把握することが必要である。しかし、通信インフラの複雑化・大規模化により、ネットワークの内部状態を直接把握することは容易ではない。本研究では、確率モデルの探索・生成・統合というコンセプトに基づき、観測可能な情報からネットワークの内部状態を間接的に推定する手法を提案したものであり、学術的・社会的意義が大きい。

研究成果の概要（英文）：In traffic matrix estimation, there is a problem that the accuracy of the estimation cannot be directly evaluated because of the constraint that the direct observation of the individual flow rate is not possible. That is, it is in principle impossible to directly evaluate the accuracy of the traffic matrix estimation in a non-observational space (the set of states in which the individual flow rate can be physically taken). To solve this problem, we propose a method to evaluate the accuracy of the traffic matrix estimation indirectly on the observational space (the set of states in which the aggregate flow rate can be physically taken) rather than on the non-observational space. We also proposed a method focusing on the fact that the traffic matrix estimation defined on the real topology can be transformed into an equivalent traffic matrix estimation defined on the pseudo-topology.

研究分野：機械学習、情報ネットワーク

キーワード：ネットワークトモグラフィ

様式 C-19、F-19-1、Z-19 (共通)

1. 研究開始当初の背景

あらゆる社会・経済活動の基盤である通信インフラは、国民生活を支えるライフラインとして、その役割がますます増大している。これに伴い、通信障害が国民生活に与える影響や、その発生頻度も増大している。こうした背景の中、総務省では「多様化・複雑化する電気通信事故の防止の在り方に関する検討会(平成25年4月~10月)」や、法令上の重大事故の第三者検証を目的とした「電気通信事故検証会議(平成27年5月~)」を開催し、再発防止のための提言を行っている。この提言の中では、ネットワークの利用状態を的確に把握する技術の整備が、通信障害の長期化・大規模化を抑制する上で重要であると主張されている。

しかしながら、広域に分散管理されたインターネットにおいては、ネットワークの利用状態を直接的に把握することは容易ではない。そのため、直接観測することが困難(場合によっては不可能)なネットワークの利用状態を間接的に推定する技術が必要となる。このような推定技術はネットワークトモグラフィと呼ばれる。本研究では、直接観測することが困難なネットワーク上の大域的な対地間トラフィックフローの流量(対地間フロー流量)を、それらが通過するルータにおいて容易に観測できる集約されたフロー流量(集約フロー流量)から推定するためのネットワークトモグラフィ技術(トラフィック行列推定)について検討する(図1参照)。

この方法は、ルータを通過する個々のパケットの始点/終点IPアドレスを調べることで対地間フロー流量を直接観測する方法よりも運用管理コストが低いため、平常時はもちろんのこと、障害発生の検知や障害箇所の特定制といった非常時における広域ネットワークの運用管理において極めて有用である。トラフィック行列推定においては、観測が困難な対地間フロー流量の状態を表現する数理モデルに関する仮定の適否が、その推定精度に大きく影響する。例えば、代表的な先行研究としては、対地間フロー流量が互いに独立な正規分布に従うことを仮定したモデルや、集約フロー流量と対地間フロー流量との間にある種の比例関係を仮定したモデルが知られているが、これらの仮定が成り立たない場合には推定精度が著しく低下することが知られている。

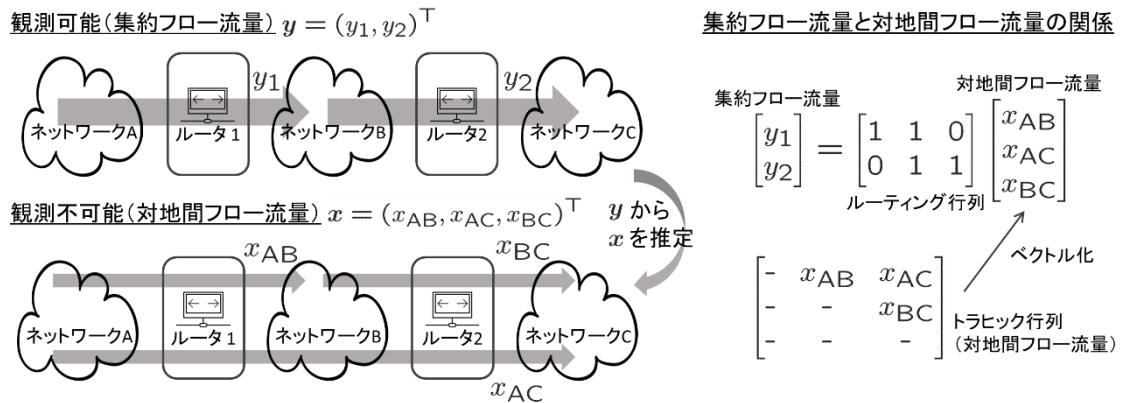


図1: 最も基本的なネットワークトポロジーにおけるトラフィック行列推定

2. 研究の目的

上述の背景を踏まえ、本研究ではまず、直接観測することが困難な対地間フロー流量の統計的な状態を表現する数理モデル(確率モデル)を、容易に観測できる集約フロー流量の観測結果との比較に基づいて探索する手法について検討する。本研究ではこの手法を、対地間フロー流量の確率モデルに基づく推定結果から集約フロー流量を再構成すると共に、その再構成された集約フロー流量と実際に観測された集約フロー流量との整合性を評価することで実現する。この着想は、対地間フロー流量に関する確率モデルを先験的に仮定するという先行研究のアプローチとは根本的に異なるものであり、集約フロー流量の観測結果と矛盾しない確率モデルを経験的に探索するという考え方に基づいている。本研究では次に、与えられた観測データのリサンプリングを繰り返すことで多様な確率モデルを生成し、それらを統合することで、推定精度の変動を抑制する。以上のことから、対地間フロー流量の統計的な状態に関する仮定が不要となるだけでなく、その仮定の適否に伴う推定結果の変動を抑制することができ、トラフィック行列推定の精度向上が可能になると考えられる。

3. 研究の方法

トラフィック行列推定においては、対地間フロー流量を直接観測できないという制約により、その推定精度も直接評価できないという本質的な問題がある。すなわち、トラフィック行列推定の精度を非観測空間(対地間フロー流量が物理的に取り得る状態の集合)上において直接評価することは原理的に不可能である。そこで本研究では、トラフィック行列推定の精度を、非観測空間上ではなく、観測空間(集約フロー流量が物理的に取り得る状態の集合)上において間接評価する手法について検討する。具体的には、対地間フロー流量の確率モデルに基づく推定結果から集約フロー流量を再構成し、その再構成された集約フロー流量と、実際に観測された集約フロー流量との整合性を評価する手法について検討する。これにより、対地間フロー流量の統計的な状態を的確に表現する確率モデルを、先行研究のような先験的な仮定を前提とせず、経験的に探索することが可能になると考えられる。

上述の手法は、容易に観測できる集約フロー流量の観測データを用いることで、直接観測できない対地間フロー流量の確率モデルの探索を可能とするものである。しかしながら、この探索に用いた集約フロー流量の観測データ（図3の☆）は、背後にある未知なる真の確率構造（図3の★）から生み出された1つのサンプルに過ぎない。すなわち、仮にこの確率構造から何度もサンプルを得られるならば、それらのサンプルには偏りやバラつきが生じ（図3の灰色の領域）、サンプル毎に探索される確率モデルも変化することになる。したがって、ある確率モデルが上述の手法により探索されたとしても、それは偶然に過ぎないものであるという可能性がある。

しかしながら、実際にはこの未知なる真の確率構造（図3の★）からは1つのサンプル（図3の☆）しか得ることができない（図3の①）。そこで本研究では、集約フロー流量の観測データに対してリサンプリングを繰り返すことで多数の複製データ（図3の○）を作成し（図3の②）、これらの複製データに対して上述の手法を適用することで多様な確率モデル（図3の●）を生成する（図3の③）。あるいは、ある観測空間において定義されたトラヒック行列推定を、それと等価となる疑似的な観測空間におけるトラヒック行列推定に変換することで、多様な確率モデルを生成する。これらの確率モデル（図3の●）は、トラヒック行列推定における弱い意味での解の集合を与え、真の解（図3の★）の周辺に分布するものと期待される。そこで本研究では、これらの弱い解を統合する（図3の④）。この統合解（図3の◎）を用いることにより、トラヒック行列推定における推定結果の偏りやバラつきを低減することが可能となり、上述の手法における推定精度のさらなる向上を達成できるものと考えられる。

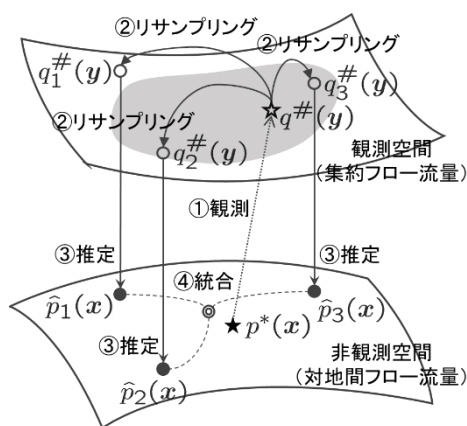


図3: 研究の方法(2)

4. 研究成果

トラヒック行列推定の解法として、自由度の高い離散分布モデルを用いた逆関数法が提案されている。しかし、逆関数法には、以下に示すように、大きく三つの問題がある。

- 対地間フロー流量が0となる確率が正であるという仮定が成り立たない場合には実行することができない
- 計算手続が再帰的であることに起因し、集約フロー流量の計測データの変動に、推定精度が大きく影響されてしまう
- 対地間フローの数よりも観測地点であるルータの数が少ないという不良設定問題であるため、解の候補が一つに定まらない

本研究では、逆関数法が抱えるこれらの欠点を解消するために、以下に示す三つの観点から研究を行った。

(1) 一般化逆関数法の提案 [1]

逆関数法は対地間フロー流量の分布が互いに独立であり、かつ、それぞれの対地間フロー流量が0となる確率が正であるという仮定の下で実行可能な手法である。この仮定が成り立つとき、逆関数法により集約フロー流量の確率分布から対地間フロー流量の確率分布が一意に定まる。しかし、この仮定は現実のネットワーク環境においては必ずしも成立しない。そこで本研究では、従来の逆関数法における上記のような仮定が成立しない場合にも適用可能となる一般化された手法を提案し、その有効性を数値実験により検証した。具体的には、対地間フロー流量の最小値（以下、最小実効流量と呼ぶ）が既知であるという仮定のもとで実行可能となるように逆関数法を一般化した。従来の逆関数法では、対地間フロー流量の最小実効流量が既知であり、なおかつそれが0であるという場合に限定されていたが、本研究で提案した一般化逆関数法ではこれが0に限定されない。これにより、逆関数法の適用可能範囲を拡大することができた。

(2) 観測データのリサンプリングによる解候補の生成と選別による推定精度向上 [2, 3]

逆関数法における上述の仮定（対地間フロー流量の分布が互いに独立であり、かつ、それぞれの対地間フロー流量が0となる確率が正であるという仮定）が成り立っている場合であっても、逆関数法で用いられる離散分布モデルの自由度が高いことや、逆関数法の計算手続が再帰的であることに起因し、逆関数法における推定精度は、推定に用いる集約フロー流量の計測データの変動に大きく影響される。そこで本研究では、集約フロー流量の計測データの変動が逆関数法における推定結果に与える影響を抑制するのではなく、その影響の様子を把握することによって推定精度の向上を可能とする手法を提案した。提案手法においては、集約フロー流量の計測データに対するリサンプリングを繰り返しながら対地間フロー流量が従う確率分布を推定する。さらに、リサンプリングを繰り返すことで得られる多数の確率分布（解の候補）のばらつきを傾向を加味した上で、推定精度が低いと思われる解を除外する。実トラヒックデータを用いた評価の結果、多くの場合において、提案手法は従来手法よりも高い推定精度を達成することを確認した。

(3) 擬似ネットワークトポロジーを用いた推定精度向上 [4]

トラヒック行列推定は、対地間フローの数よりも観測地点であるルータの数が少ない、すなわち未知変数に対して情報が不足している不良設定問題である。不良設定問題では、未知変数間で誤差が相殺される場合がある。そのため、上述の推定結果の単純な統合や集約フロー流量の期待値による選別では、解の不定性への対策が不十分である場合があり、異なる観点から解を絞り込む必要がある。そこで本研究では、実トポロジー上に定義されたトラヒック行列推定が、それと等価な疑似トポロジー上に定義されたトラヒック行列推定に変換できることに着目した。そして、実トポロジーと疑似トポロジーのそれぞれに対して逆関数法を適用して得られる複数の推定結果を統合することにより解の不定性を緩和し、対地間フロー流量分布の推定精度の向上を実現した。

参考文献

- [1] 土井 裕貴, 鶴 正人, 内田 真人, “OD トラヒック行列推定における逆関数法の一般化,” 電子情報通信学会 ネットワークシステム研究会, 信学技報 Vol. 117, No. 459, pp. 275-280 (NS2017-216), 2018 年 3 月.
- [2] 加瀬 史門, 宮田 高道, 鶴 正人, 内田 真人, “観測データのリサンプリングによる解候補の生成と選別に基づいた OD トラヒック行列推定の精度向上,” 電子情報通信学会 ネットワークシステム研究会, 信学技報 Vol. 117, No. 459, pp. 281-286 (NS2017-217), 2018 年 3 月 1-2 日.
- [3] Shimon Kase, Masato Tsuru, and Masato Uchida, “Accurate OD Traffic Matrix Estimation Based on Resampling of Observed Flow Data,” Asia-Pacific Signal and Information Processing Association Annual Summit and Conference 2018 (APSIPA-ASC 2018), pp. 1574-1579, November, 2018.
- [4] 長谷川 陽平, 内田 真人, “擬似ネットワークトポロジーを用いた OD トラヒック行列推定の精度向上,” 電子情報通信学会 コミュニケーションクオリティ研究会, 信学技報 Vol. 119, No. 455, pp. 81-86 (CQ2019-150), 2020 年 3 月.

5. 主な発表論文等

〔雑誌論文〕 計1件（うち査読付論文 1件 / うち国際共著 0件 / うちオープンアクセス 0件）

1. 著者名 Naoki Fukushi, Daiki Chiba, Mitsuaki Akiyama, and Masato Uchida	4. 巻 E103.B
2. 論文標題 Exploration into Gray Area: Toward Efficient Labeling for Detecting Malicious Domain Names	5. 発行年 2020年
3. 雑誌名 IEICE Transactions on Communications	6. 最初と最後の頁 375 ~ 388
掲載論文のDOI（デジタルオブジェクト識別子） https://doi.org/10.1587/transcom.2019NRP0005	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

〔学会発表〕 計19件（うち招待講演 0件 / うち国際学会 6件）

1. 発表者名 Naoki Fukushi, Daiki Chiba, Mitsuaki Akiyama, and Masato Uchida
2. 発表標題 Exploration into Gray Area: Efficient Labeling for Malicious Domain Name Detection
3. 学会等名 IEEE COMPSAC 2019（国際学会）
4. 発表年 2019年

1. 発表者名 Tokio Takahashi and Masato Uchida
2. 発表標題 Optimal Hand Sign Selection Using Information Theory for Custom Sign-Based Communication
3. 学会等名 IEEE COMPSAC 2019（国際学会）
4. 発表年 2019年

1. 発表者名 Yohei Hasegawa and Masato Uchida
2. 発表標題 Predicting Network Outages Based on Q-drop in Optical Network
3. 学会等名 IEEE COMPSAC 2019（国際学会）
4. 発表年 2019年

1. 発表者名 野村 和也, 小川 翔大, 中野 翔, 座間味 卓臣, 内田 真人
2. 発表標題 ユーザーの行動パターンに注目したECサイト上の不正取引検知
3. 学会等名 第18回情報科学技術フォーラム
4. 発表年 2019年

1. 発表者名 長谷川 陽平, 内田 真人
2. 発表標題 擬似ネットワークボロジエを用いたODトラヒック行列推定の精度向上
3. 学会等名 電子情報通信学会 コミュニケーションクオリティ研究会
4. 発表年 2020年

1. 発表者名 藤田 孝志, 福士 直翼, 内田 真人
2. 発表標題 スキャン活動のON/OFFパターンに着目したブラックリストの掲載継続ポリシーの有効性調査
3. 学会等名 電子情報通信学会 コミュニケーションクオリティ研究会
4. 発表年 2020年

1. 発表者名 中野 翔, 内田 真人
2. 発表標題 不均衡データクラスタリングを用いた機械学習の判断根拠説明における安定化手法の提案
3. 学会等名 電子情報通信学会 総合大会
4. 発表年 2020年

1. 発表者名 福土 直翼, 千葉 大紀, 秋山 満昭, 内田 真人
2. 発表標題 クラウドサービス悪用攻撃の大規模実態調査
3. 学会等名 コンピュータセキュリティシンポジウム2018
4. 発表年 2020年

1. 発表者名 戸祭隆行, 千葉 大紀, 秋山 満昭, 内田 真人
2. 発表標題 ドメインパーキングを利用するドメイン名の大規模実態調査
3. 学会等名 コンピュータセキュリティシンポジウム2018
4. 発表年 2020年

1. 発表者名 Suguru Goto, Masahiro Shibata, and Masato Tsuru
2. 発表標題 Dynamic optimization of multicast active probing path to locate lossy links for OpenFlow networks
3. 学会等名 ICOIN 2020 (国際学会)
4. 発表年 2020年

1. 発表者名 Takumi Nakamura, Masahiro Shibata, and Masato Tsuru
2. 発表標題 On Retrieval Order of Statistics Information from OpenFlow Switches to Locate Lossy Links by Network Tomographic Refinement
3. 学会等名 INCoS 2019 (国際学会)
4. 発表年 2019年

1. 発表者名 中村 勇海, 後藤 卓, 木村 亮太, 柴田 将弘, 鶴 正人
2. 発表標題 ネットワークトモグラフィーを用いた統計情報取得順序決定による障害リンク検知の効率化
3. 学会等名 電子情報通信学会 情報通信マネジメント研究会
4. 発表年 2019年

1. 発表者名 Shimon Kase, Masato Tsuru, and Masato Uchida
2. 発表標題 Accurate OD Traffic Matrix Estimation Based on Resampling of Observed Flow Data
3. 学会等名 APSIPA-ASC 2018 (国際学会)
4. 発表年 2018年

1. 発表者名 福土 直翼, 齊藤 秀, 内田 真人
2. 発表標題 発現量分布の類似性に着目したマイクロアレイ測定データの統合手法の提案
3. 学会等名 情報処理学会 バイオ情報学研究会
4. 発表年 2018年

1. 発表者名 高橋 時生, 内田 真人
2. 発表標題 情報理論に基づいた高信頼かつ高効率なハンドサインの設計
3. 学会等名 第17回情報科学技術フォーラム
4. 発表年 2018年

1. 発表者名 長谷川 陽平, 内田 真人
2. 発表標題 光学レイヤ特性に基づく広域光ネットワークの停止予測
3. 学会等名 第17回情報科学技術フォーラム
4. 発表年 2018年

1. 発表者名 福土 直翼, 千葉 大紀, 高田 雄太, 秋山 満昭, 内田 真人
2. 発表標題 Active LearningとEnsemble Learningを用いた悪性ドメイン名検知システム
3. 学会等名 コンピュータセキュリティシンポジウム2018
4. 発表年 2018年

1. 発表者名 土井 裕貴, 鶴 正人, 内田 真人
2. 発表標題 ODトラヒック行列推定における逆関数法の一般化
3. 学会等名 電子情報通信学会 ネットワークシステム研究会
4. 発表年 2018年

1. 発表者名 加瀬 史門, 宮田 高道, 鶴 正人, 内田 真人
2. 発表標題 観測データのリサンプリングによる解候補の生成と選別に基づいたODトラヒック行列推定の精度向上
3. 学会等名 電子情報通信学会 ネットワークシステム研究会
4. 発表年 2018年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究 分担 者	鶴 正人 (Tsuru Masato) (40231443)	九州工業大学・大学院情報工学研究院・教授 (17104)	