

遺伝的プログラミングと遅延トモグラフィーを用いたリンク遅延 分布推定とその応用

01014343 信州大学 池田 欽一 IKEDA Yoshikazu
01304556 九州大学 *時永 祥三 TOKINAGA Shozo
九州大学 呂 建軍 LU Jianjung

1 まえがき

遺伝的プログラミングによるトポロジー推定および遅延トモグラフィーを用いて観測されたデータからネットワーク内部のノード配置とリンクにおける遅延を同時に推定する方法を提案する。

2 遅延トモグラフィーとネットワークシステム構成の推定問題

ネットワークの外部的な情報である受信側の情報 (end to end network measurement) だけから中間のリンクにおける遅延情報を推定する方法が提案され、医療分野における Computer Tomography にならって Network Delay Tomography (以下では、単に遅延トモグラフィーと呼ぶ) として定式化されている [1][2]。

これまでの遅延トモグラフィーの解析では、あらかじめノードとこれを接続するリンクの配置が与えられていることが前提とされ、ネットワークのトポロジーとルーティングテーブルが固定されていることを仮定している。しかし、一般的にはネットワークのトポロジーとルーティングテーブルは固定されていないケースが多く存在する。本報告ではこれを同時に推定する方法を提案する。

以下では、ネットワークを構成するエンドノードを GP における関数表現の終端記号に対応させ、ノードにおけるリンクの結合を GP における演算記号に対応させることにより、ネットワークトポロジーを GP の個体として算術式と同様に表現する方法が適用できる。GP における個体のそれぞれは、ネットワークトポロジーに対応しているため、この構造を前提として、pseudo likelihood の方法を用いて個別のリンクにおける遅延時間分布を推定し適合度を計算する [1]。

3 遅延トモグラフィー問題の概要

$X = (X_1, X_2, \dots, X_J)'$ は J 次元のランダムベクトルであり、ネットワークにおいて推定したい数量、すなわちリンクの遅延、トラヒックの量などである。 $Y = (Y_1, Y_2, \dots, Y_I)'$ は I 次元の観測ベクトルである。遅延トモグラフィーの最終目的は、 Y が与えられた場合に、これをもとにして X を推定することである。

$$Y = AX. \quad (1)$$

ここで、行列 A はネットワークトポロジーであるが、われわれの拡張モデルでは、行列 A は GP により同様に推定される。更に、変数 X_j は f_j により代表される確率密度関数に従う分布をしていると仮定する。

$$X_j \sim f_j(\theta_j). \quad (2)$$

ここで、 θ_j は推定されるべきパラメータである。 T を連続す

る観測時刻であるとする。

PLE によるサブモデルの解法

これまでのいくつかの計算の単純化が提案されている。以下では、これらの中の 1 つの方法である Pseudo Likelihood Estimation (以下では PLE とよぶ) を用いることにする [1]。この PLE においては、全体の問題をいくつかの簡単な構造をもつサブ問題 (S) に分解し、これらのゆう度関数の積として全体を表現している。ルーティングテーブルの列の対をとりだし解析することで実現できる。 X_j はいくつかの有限の数値を取るように離散化されており、次のようにさだめる。 $X_j \in (0, q, 2q, \dots, mq, \infty)$ 。ここで m は定数であり、 X_j の値が無大となるのは、パケットがネットワークで紛失された場合に相当している。パケットの遅延時間が lq である確率を θ_j^l により与える。

$$\theta_j^l = \text{Prob}(X_j = lq). \quad (3)$$

また観測されたデータの中で、長さが lq であるデータ数を n_{jl}^S とする。このような前提のもとで、次に示す対数尤度関数を求めておき、これを最大化する方向に式に含まれるパラメータを変更しながら探索する。

パラメータ推定には EM (Expectation-Maximization) 法による逐次近似の方法を用いている。すなわち、同時にパラメータを最適化するのではなく、1 つの時期には最初のグループのパラメータを固定し、これをもとにして別のグループのパラメータを尤度最大化の方向に更新する。次の時刻には、逆に、更新されたパラメータを固定し、別のグループのパラメータを更新する。

(E-step) 次を計算する

$$\hat{n}_{jl} = \sum_{s \in S} E_{\theta^{s(k)}} \left(\sum_{t=1}^T 1\{X_{tj}^s = lq\} | Y_t^s \right). \quad (4)$$

(M-step) 次のように更新する

$$\theta_{jl}^{(k+1)} = \frac{\hat{n}_{jl}}{\sum_{r \in R} \hat{n}_{jr}}, R = [0, 1, \dots, m, \infty]. \quad (5)$$

本報告でも、ここに示す方法と同様に逐次的にパラメータの更新により推定を進めていく。

4 GP によるトポロジーの推定

これまでの遅延トモグラフィーの議論ではネットワークトポロジー (以下では、NT と略称する) は既知で固定されると仮定したが、以下ではこれも不明であり、推定する必要があるケースへと拡張する。本論文では、NT の構成を木構造

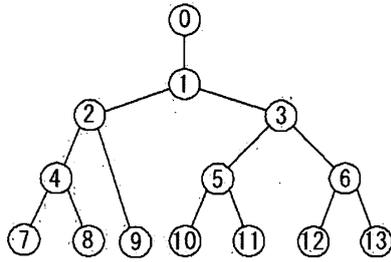


図 1: GP による推定に用いるネットワーク

で表現し、これを GP における個体とみなし最適なモデルの構成を得るために GP を適用する。

木構造との対比を考慮すると、ターミナルノードには終端ノードを、木構造における中間ノードをノードに対応させることができることが分かる。GP における個体のそれぞれは、1 つのトポロジーの実現形式に対応しているので、遺伝的操作を実施することにより構成を推定する。遅延時間分布の推定には、前に述べた PLE 法を用いる。このエンドノードにおける遅延時間分布の推定誤差の逆数を、個体の適合度とする [3]-[6]。

(ステップ 1): 個体の初期値生成

最初の個体の集合 (プール) を乱数をもとにして発生させる。この場合、トポロジーとして意味をなすものが得られるように中間、端末、それぞれのノード数をカウントする必要がある。(ステップ 2): 個体の適応度の計算

個体 i の与えるトポロジーを前提としてエンドノードでの累積遅延分布の予測値と観測値との 2 乗誤差の逆数を、個体 i の適合度 S_i として定義する。個体を適応度の大きい順に並びかえておく。

(ステップ 3): 交差処理

適応度に応じて 2 つの個体を取り出し、交差処理を実施する。交差処理では、乱数を 1 個発生させておいて、一方の個体の切断箇所とし、そのノード以下の中間、端末ノード数をカウントしておく。もう一方の個体のノード数を求め、同じ値となる場所を検出し、これらから任意に 1 個を選択して切断位置を決める。

5 シミュレーション

図 1 に示すネットワークの構成と遅延を本報告の手法により同時に推定する。図中の 0 はルートノードであり、ここからパケットが送信される。1~6 は中間ノードであり、7~13 は累積遅延時間が観測されるターミナルノードである。各遅延はノードをつなぐリンクにおいて発生すると仮定し、リンクの番号は木の下側のノード番号と同じもので表すこととする。GP による推定の結果がこれに一致するかを検証する。シミュレーションでは、 $q = 1, m = 20$, 各リンクでの遅延分布は平均 3~8 の指数分布、GP 個体は 20 とした。

シミュレーションの結果、GP の世代数 26 において推定対象の元のネットワーク構成と同じ構成を推定することができた。

図 2 にはリンク 7 での元の遅延時間分布と推定分布を示している。また、図 3 にはリンク 11 での元の累積遅延時間分布とネットワーク構成とともに推定された遅延分布を示している。各図において実線は元の分布、破線は推定された分布を表す。これらの結果から分かるように適切な分布の推定がなされていることが分かる。

6 むすび

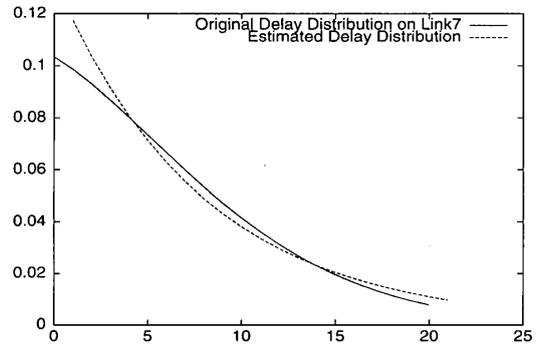


図 2: 個別遅延分布

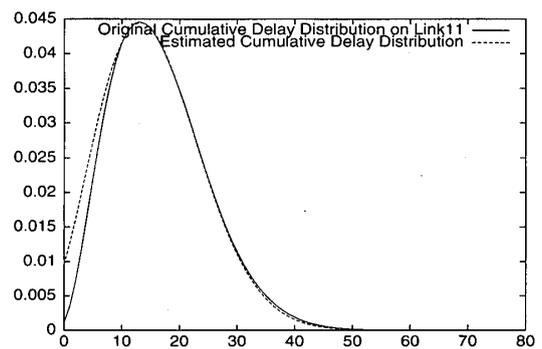


図 3: 累積遅延分布

GP および遅延トモグラフィーを用いて観測されたデータからネットワーク内部のノード配置とリンクにおける遅延とを同時に推定する方法を提案した。今後の課題として、実際的な問題適用する予定である。

参考文献

- [1] G.Liang and B.Yu, "Maximum pseudo likelihood estimation in networks tomography", IEEE Trans.Signal Processing, vol.51,no.8, pp.2043-2053, 2003.
- [2] Y.Tsang, M.Coates and R.D.Nowak, "Network delay tomography", IEEE Trans.Signal Processing, vol.51,no.8, pp.2125-2136, 2003.
- [3] Y. Ikeda and S.Tokinaga, "Approximation of chaotic dynamics by using smaller number of data based upon the genetic programming", Trans. IEICE, vol.E83-A,no.8, pp.1599-1607, 2000
- [4] Y. Ikeda and S.Tokinaga, "Controlling the chaotic dynamics by using approximated system equations obtained by the genetic programming Trans. IEICE, vol.E84-A, no.9, pp.2118-2127, 2001.
- [5] X. Chen and S.Tokinaga, "Multi-Agent-Based modeling of artificial stock markets by using the co-evolutionary GP approach", JORSJ to appear, 2004.