

# GRID における負荷分散のためのネットワーク スループット予測\*

3D-2

秋岡 明香<sup>†</sup> 村岡 洋一<sup>†</sup>  
早稲田大学理工学研究科<sup>§</sup>

## 1 はじめに

近年、高性能な計算機と高速ネットワークが広く普及している一方で、これらを使い切る利用者は極めて少なく、余剰計算機が溢れている。そこで、世界中に散在する余剰計算機群を大規模クラスタとして活用する GRID に関する研究が盛んに行われてきた。同時に GRID にタスクを最適に配置する負荷分散に関する研究も多く行われてきたが、GRID では通信遅延が大きく計算機やネットワークの負荷は刻々と変化するため、効率的な負荷分散を行うことは難しい。そこで本稿では、ネットワークスループットの予測アルゴリズムを提案する。提案アルゴリズムは、複数の予測アルゴリズムから隠れマルコフモデル (HMM) [1] により、最も良い予測値を出す確率が高いアルゴリズムを選択する。ネットワーク状況を予測することは負荷分散の助けになる他、これまで例が少ない並列プログラムを GRID で走らせるための大きな手助けとなることが期待できる。

## 2 従来研究

ネットワークの動向を探る研究は多く行われてきたが、負荷分散を目的としたものは少なく、同じ目的の研究には Network Weather Service (NWS) [2] がある。NWS は、サーバが計算機やネットワークの負荷情報を収集し、収集データから予測値を計算して負荷分散システム AppLeS [3] に提供する。NWS は常に複数手法による予測を行ない、前回最も誤差の小さい予測値を得た手法を次回用いる。

## 3 予測アルゴリズム

本稿で使用した予測アルゴリズムは 6 種である。以下、時刻  $t$  の測定値を  $v(t)$ 、予測値を  $p(t)$  とする。

**DV** 直前値を予測値として出力する手法。

$$p(t+1) = v(t)$$

**GD** 計測値の変化率を求め、前回の計測値に足したものを予測値とする手法。

$$p(t+1) = (v(t) - v(t-1)) + v(t)$$

**SG** 勾配法を利用して予測値を算出する手法。ただし  $0 < g < 1$  である。

$$p(t+1) = (1-g) * p(t) + g * v(t)$$

**SW** 過去の計測値の平均値を予測値として出力する手法。  $i$  は  $err_i(t)$  を最小にする値である。

$$p(t+1) = SW(t, K(t))$$

$$SW(t, K(t)) = \frac{1}{K(t)+1} \sum_{i=t-K(t)}^t v(i)$$

$$K(t+1) = K(t) + i$$

$$err_i(t) = (v(t) - SW(t, K(t) + i))^2$$

**MF** Median Filter を利用して予測値を求める手法。

$S_K$  は最近データ  $K$  個をソートした数列で、 $i$  は  $err_i(t)$  を最小にする値である。

$$p(t+1) = MF(t, K(t))$$

$$MF(t, K) = \begin{cases} S_K((K+1)/2) & (K : \text{odd}) \\ \frac{S_K(K/2) + S_K(K/2+1)}{2} & (K : \text{even}) \end{cases}$$

$$K(t+1) = K(t) + i$$

$$err_i(t) = (v(t) - MF(t, K(t) + i))^2$$

**AR** AR モデルを利用して予測値を算出する手法。時刻  $t$  の  $M$  次 AR モデルは次式により表現できる。

$$v(t) = \sum_{m=1}^M a(m)v(s-m) + err(t)$$

$$1 - a(z) - a(z^2) - \dots - a(M)z^M = 0$$

\*Network throughput prediction helping load distribution on GRID

<sup>†</sup>Sayaka Akioka

<sup>†</sup>Yoichi Muraoka

<sup>§</sup>Graduate School of Science and Engineering, Waseda University

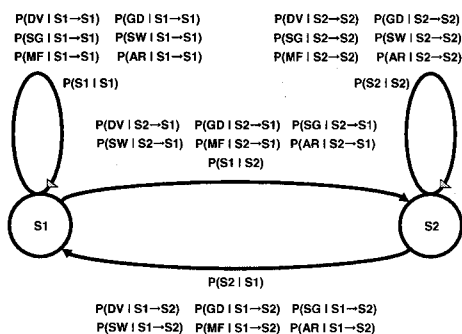


図 1: 隠れマルコフモデル

## 4 HMM

本稿で使用した HMM を図 1 に示す。S1 は前回と同じ予測手法を用いた状態を表し、S2 は前回とは別の予測手法を用いた状態を表す。本モデルではスループット計測の度に各状態遷移確率が再計算される。この時、最新の計測値と比較して最も誤差が小さかった手法を正解とする。毎回の予測時には、新たに計算し直された確率分布に基づき、前述の 6 種のアルゴリズムのうち次回最も有効である確率が高い手法で算出された値を予測値とする。スループット計測毎に確率分布の再計算を行なうことで、時間の経過と共にモデルが最適化されていくことが期待できる。

## 5 実験と考察

アルゴリズムの有用性を確認するために実データでの実験を行った。表 1 に示す環境に於いて、ある計算機からもう一方の計算機へ 64KB のデータ転送を行うのに要する時間を測定した。実験は 2001 年 6 月 7 日～13 日に行い、計測は約 30 秒に 1 回ずつ、各組み合わせについて 24 時間に渡って行った。

上記実験で得た 6 種のデータのうち 5 種を学習データとして用い、残り 1 種についてスループット予測実験を行った。なお勾配法では  $g = 0.5$  とした。実験全体、定常時、非定常時のそれぞれについて二乗平均により求めた予測誤差と、予測誤差率が 10 % 以下の場合が母集団に占める割合を表 2 にまとめる。

Schopf ら [4] によると、GRID で負荷分散を行う場合に負荷予測値が有効な判断材料として活用されるためには、予測誤差が 10 % 以内であることが必要であるとされているが、本手法では十分にその要件を満たしている。

## 6 まとめ

本稿では、GRID で負荷分散を行なう際に有効なネットワークスループットの予測アルゴリズムを提

表 1: ネットワークスループット測定時の実験環境

場所	環境
早大 (大久保)	100BASE-TX, SunOS 5.8, PentiumII 350MHz, Memory 64MB
早大 (喜久井町)	100BASE-TX, Linux 2.2.18, Athlon Thunderbird 1.0GHz, Memory 256MB
早大 (喜久井町)	100BASE-TX, SunOS 5.8, UltraSparcII-i 440MHz, Memory 512MB
千葉県市川市	64k ISDN, SunOS 5.8, PentiumIII 600EMHz, Memory 128MB

表 2: HMM を用いたスループット予測実験の結果

母集団	予測誤差平均値	誤差率 10 % 以下
全体	87.20	91.3 %
定常時	35.93	96.6 %
非定常時	545.68	43.3 %

案した。提案アルゴリズムの特徴は、複数の予測手法から HMM により状況に応じて最適である確率が高い手法を選択し、予測値を得る点である。また実データに基づいた実験では、定常時で 96.6 %、実験全体で 91.3 % の場合に予測誤差率が 10 % 以下となる良好な結果が得られ、本手法の有用性が確認された。今後は非定常時の予測精度の向上を目指してアルゴリズムの改良を重ねる予定である。

## 参考文献

- [1] 中川聖一. 確率モデルによる音声認識. 電子情報通信学会, 1998.
- [2] Rich Wolski, Neil T. Spring and Jim Hayes. The Network Weather Service: A Distributed Resource Performance Forecasting Service for Metacomputing. The Journal of Future Generation Computing Systems, 1999.
- [3] Fran Berman, Rich Wolski, Silvia Figueira, Jennifer Schopf and Gary Shao. Application-Level Scheduling on Distributed Heterogeneous Networks. In Proceedings of Supercomputing 1996, 1996.
- [4] Jennifer M. Schopf and Francine Berman. Performance Prediction in Production Environments. In Proceedings of IPPS/SPDP '98, 1998.