

統計的手法による計算機システムの評価

井須 芳美† 長嶋 雲兵† 関口 智嗣†† 佐藤 三久††

†お茶の水女子大学理学部 ††電子技術総合研究所

並列計算機システムの性能評価のために各種のベンチマークが提案されている。システムの様々な側面を理解するためには、複数のベンチマークによる性能評価が行なわれる。そこで本研究ではこれらのベンチマーク間の相関を統計的手法である主成分分析を用いて解析することを試みた。具体的にNAS Parallel Benchmarksの実行結果を用いて、NAS-PBのベンチマークセットに含まれるベンチマーク間の相関を解析した。さらに、並列システムの性能評価指標として、スケーラビリティという評価指標が用いられていることから、クラスタ分析を用いて、スケーラビリティを定量的に求めることを試みた。

The Principal Component Analysis of NAS Parallel Benchmarks Results and the Scalability on the Cluster Analysis

YOSHIMI ISU†, UMPEI NAGASHIMA†, SATOSHI SEKIGUCHI††
and MITSUHISA SATO††

†Ochanomizu University, ††Electrotechnical Laboratory

The correlation between the benchmark tests for performance evaluation of parallel computing systems was analyzed by the principal component analysis. NAS Parallel Benchmarks(NAS-PB) is to measure the performance of various highly parallel computer systems and consists of a set of eight benchmark problems. The correlation between four benchmarks performance results out of eight was analyzed. It was suggested that the set of NAS-PB has a good criterion for evaluation of the performance of various highly parallel computer systems.

Furthermore, we have proposed a new conventional technique to suggest the scalability of the parallel computing systems using the cluster analysis.

1. はじめに

単一CPU性能の限界が見えはじめたことを背景に、各種の並列アーキテクチャや並列計算機システムが開発され、その性能を評価するためのベンチマークプログラムも数多く提案されている。単一のベンチマークプログラムでシステムの様々な側面を評価することは困難であり、1つのシステムに関して複数のベンチマークによる評価が行なわれる。ここで問題となるのはシステムを正當に評価するために最適なベンチマークセットを選択することである。そのためにはまず選択された各ベンチマーク間の独立性およびそれらの相関が検証される必要がある。そこで本研究ではまずこれらの問題に対し主成分分析を用いて解析することを試みた。

主成分分析¹⁾とは1つのデータが複数の変数から成る場合にそれらの変数をそのまま扱うのではなく、変数間の相関関係を解析し、全体のもつ変動をなるべく小数の合成変数の変動で表すものである。1つのシ

テムに関して複数のベンチマークの結果が得られているときに、各ベンチマークの結果を1つの変数とみなし、主成分分析により各ベンチマーク間の相関関係の解析を行なうことにした。本研究で用いたデータはNAS Parallel Benchmarks²⁾の実行結果³⁾である。これは5個のカーネルと3個の応用プログラムからなる高並列計算機のためのベンチマークセットである。いくつかのマシンに対するこれら複数のベンチマークの結果が報告されているので本研究ではそれらを用いて各ベンチマーク間の相関の検証を行なった。

さらにベンチマークの方法論として性能評価指標として何を用いるべきか、得られた数値から何を読みとるべきかといった問題がある。最近並列システムの性能評価指標としてスケーラビリティという指標がよく用いられている。スケーラビリティとはプロセス数および計算サイズを変化させたときのシステム性能の振舞であるといえる。本研究ではスケーラビリティに対して定量的な評価指標を与えることを目的とし、統計的な手法であるクラスタ分析を用いて簡便にスケーラビリティを求める

ことを試みた。

クラスタ分析¹⁾とは複数個の特性によって決定されるデータをデータ間の類似性の指標をもとにいくつかのグループ(クラスタ)に分類する手法である。そこでプロセッサ数や計算サイズを変化させたときに得られるベンチマークの実行結果をプロセッサ数および計算サイズで正規化したデータに対してクラスタ分析を行なう。このとき同じクラスタに属するものは、プロセッサ数や計算サイズの変化に対してほぼ同程度の性能が得られると考えることができる。具体的には並列行列乗算プログラムのCM-5による実行結果を用いてクラスタリングを行ない、スケーラビリティを定量的に求めることを試みた。

2. 主成分分析によるベンチマークテストの統計処理

主成分分析法¹⁾とは1つのデータが複数の変数からなる場合に変数群内部の関連を明らかにしていくもので全体の変数の持つ変動をなるべく少数の合成変数の変動で表す手法である。ここで求められる合成変数のことを主成分と呼ぶ。そこで本研究ではNAS Parallel Benchmarksの実行結果を用いて主成分分析を行なった。NAS Parallel Benchmarksは5個のカーネルと3個の応用プログラムからなるベンチマークセットであるが、このうち4個のカーネルプログラムの結果を用いた。4個のベンチマークそれぞれを変数とみなし主成分を求めることにより、ベンチマーク間の関連を検証することを試みた。まず本研究で用いた4種類のカーネルプログラムの内容を簡単に述べる⁴⁾。

EP(Embarrassingly Parallel): 乗算合同法による一様乱数および正規乱数の生成。モンテカルロシミュレーションの典型。プロセッサ間通信を含まない。

MG(Multigrid): 簡略化されたマルチグリッド法。3次元ポアソン方程式の解法。長距離通信と短距離通信を含む。

CG(Conjugate Gradient): 正値対称な大規模疎行列の最小固有値を求めるための共役傾斜法。不規則な構造を持つ行列とベクトルの乗算を含む。不規則な長距離通信を含む。

FT(3-D FFT PDE): FFTを用いた3次元偏微分方程式の解法。規則的な長距離通信。

NAS-PBの特徴としてベンチマークがアルゴリズムで与えられていることがあげられる。そのため本研究で用いたデータは各並列計算機に対する最大限のチューニングを行なったプログラムによる実行結果である。

次に主成分分析の手順を簡単に述べる。 p 個の変数を (x_1, x_2, \dots, x_p) とすると主成分 f は p 個の変数の合成変数として以下のように表される。

$$f = x_1 w_1 + x_2 w_2 + \dots + x_p w_p \quad (1)$$

表1 相関係数行列—プロセッサ数で正規化した場合—

| | EP | MG | CG | FT |
|----|-------|-------|-------|-------|
| EP | 1.000 | 0.996 | 0.943 | 0.999 |
| MG | 0.996 | 1.000 | 0.964 | 0.997 |
| CG | 0.943 | 0.964 | 1.000 | 0.946 |
| FT | 0.999 | 0.997 | 0.946 | 1.000 |

表2 得られた主成分—プロセッサ数で正規化した場合—固有値

| | | | | |
|-----------------|-------|--------|-------|-------|
| | 3.922 | 0.075 | 0.002 | 0.001 |
| 寄与率 | | | | |
| | 0.981 | 0.019 | 0.000 | 0.000 |
| 累積寄与率 | | | | |
| | 0.981 | 0.999 | 1.000 | 1.000 |
| 主成分負荷行列(重みベクトル) | | | | |
| | 第1 | 第2 | | |
| EP | 0.994 | 0.102 | | |
| MF | 0.999 | 0.031 | | |
| CG | 0.972 | -0.233 | | |
| FT | 0.995 | 0.095 | | |

このとき f が p 個の変数の分散をなるべく説明できるように重みベクトル $w = (w_1, w_2, \dots, w_p)$ を定める。具体的には各変数の相関係数を成分とする相関係数行列(または分散共分散行列)の固有値及び固有ベクトルを求めることに帰着される。このとき求まった固有値のうち最大固有値に対応する固有ベクトルより求められた主成分は第1主成分と呼ばれ、全体の変数の持つ変動を最もよく説明する。また最大固有値は全ての変数の分散のうち第1主成分によって説明される分散を表す。従って最大固有値が全ての固有値の和に占める割合(以下寄与率と呼ぶ)が小さい場合には、2番目に大きい固有値の固有ベクトルより第2主成分を求める。以下同様に第3主成分、第4主成分を考慮することができる。

NAS-PBの4種類のベンチマーク(EP, MG, CG, FT)の結果を用いて主成分分析を行なった結果を示す。用いたデータはCRAY C90, CRAY J90, CRAY T3D, Fujitsu VPP500, IBM SP2-WN(Wide Nodes), IBM SP2-TN2(Thin Nodes 2), Convex Exemplar SPP1000, nCUBE-2S, CRAY T916, SGI Power Challenge XL(90MHz)のEP, MG, CG, FTの結果である。

まずデータをプロセッサ数で正規化して主成分分析を行なった。プロセッサ数を p 、実測値を $T(p)$ とすると $pT(p)$ によって正規化した。得られた相関係数行列、主成分をそれぞれ表1, 表2に示す。相関係数は1に近いほど(線形の)相関係数が強く、0に近いほど相関は弱くなる。表1を見ると全体的に4つのベンチマーク間の相関は高くなっている。これにより得られた主成分も第1主成分により4つの変数の持つ変動を98.1%表すことができる。しかしながらこれはマシンのクロック性能の違いによるデータの変動が大きいため、プロセッサ数の変化によるデータの変動が小さくなるためであると思われる。そこでさらにデータをクロック周波数により正規

表3 相関係数行列—クロック周波数で正規化した場合—

| | EP | MG | CG | FT |
|----|-------|-------|-------|-------|
| EP | 1.000 | 0.979 | 0.904 | 0.979 |
| MG | 0.979 | 1.000 | 0.903 | 0.973 |
| CG | 0.904 | 0.903 | 1.000 | 0.860 |
| FT | 0.979 | 0.973 | 0.860 | 1.000 |

表4 得られた主成分—クロック周波数で正規化した場合—
固有値

| | | | | |
|-------|-------|-------|-------|-------|
| 寄与率 | 3.801 | 0.160 | 0.023 | 0.017 |
| 累積寄与率 | 0.950 | 0.040 | 0.006 | 0.004 |
| | 0.950 | 0.990 | 0.996 | 1.000 |

主成分負荷行列(重みベクトル)

| | 第1 | 第2 |
|----|-------|--------|
| EP | 0.991 | 0.076 |
| MG | 0.989 | 0.070 |
| CG | 0.939 | -0.342 |
| FT | 0.979 | 0.180 |

化した。プロセッサ数を p 、実測値を $T(p)$ 、クロック周波数を f として $p \cdot T(p) \cdot f$ によりデータを正規化した。クロック周波数により正規化したデータを用いて主成分分析を行なって得られた相関係数行列を表3に示す。

表3を見ると4つのベンチマーク間の相関係数はプロセッサ数のみで正規化した場合に比較して若干小さくなっていることがわかる。これはクロック性能の違いによる変動がある程度除かれたためと考えられる。表3より、EP, MG, FT のベンチマーク間の相関はほぼ同じである。これに対し、CG と他の3つのベンチマーク間の相関係数の大きさは、小さくなっている。表4に主成分分析の結果を示す。

第1主成分は全体の変数の持つ変動を95.0%表している。第1主成分は4つのベンチマークがほぼ同じ比で合成され、4つのベンチマークによって総合的な性能を見ることができる。4つのベンチマークの独立性が高いならば、4つの主成分が表れるはずであるので、今回の結果は、4つのベンチマークの相関が高く、それぞれの一次従属性が高いことを示している。第1主成分は正の方向に大きくなるに従って、性能は低下する。これに対し、第2主成分は寄与率が4.0%で第1主成分に比べてかなり小さいが、EP, MG, FT に係る重みが正であるのに対し、CG に係る重みは負である。つまりCG以外のベンチマークの結果に対しCGの結果が悪くなる場合には第2主成分は負の方向に絶対値が大きくなる。また第2主成分の変化が大きいのものはEP, MG, FT に比べてプロセッサ数を増やすことによりCGの性能の低下が著しいことを示している。そこで図1に得られた第1主成分、第2主成分を軸とする平面に各マシンをプロットしたものを示す。図1を見ると負の方向に絶対値が大きいののはCRAY T3D, nCUBE-2Sの1024プロセッサのときである。ちなみにこれらは分散メモリ型の超並列計算機である。CGは不規則な長距離通信を含む

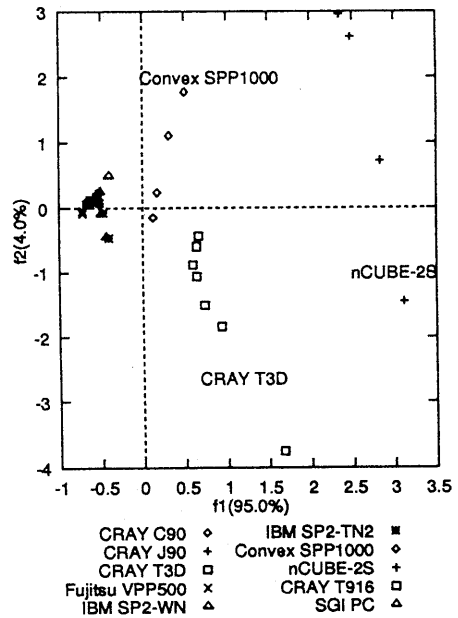


図1 主成分分析の結果—クロック周波数で正規化した場合—

表5 相関係数行列—TPPで正規化した場合—

| | EP | MG | CG | FT |
|----|-------|-------|-------|-------|
| EP | 1.000 | 0.687 | 0.552 | 0.621 |
| MG | 0.687 | 1.000 | 0.628 | 0.871 |
| CG | 0.552 | 0.628 | 1.000 | 0.622 |
| FT | 0.621 | 0.871 | 0.622 | 1.000 |

ため、分散メモリ型の場合、通信に必要なコストが増大し、性能が相対的に低下している。これに対し、Convex Exemplar SPP1000はプロセッサ数が増えるに従って第2主成分が正の方向に増えている。つまりCGの性能が良くなっている。第1主成分はプロセッサ数の増加にとまぬ、正の方向に大きくなっているため、全体的な性能は低下しているが、第2主成分によりCGの性能は向上していることがわかる。図より見ることはできないが、SGI Power Challengeもこのような傾向を示しており、これは他のマシンと比べて特異的である。

最後に Theoretical Peak Performance(TPP) によって正規化したデータを用いて主成分分析を行なった結果を示す。

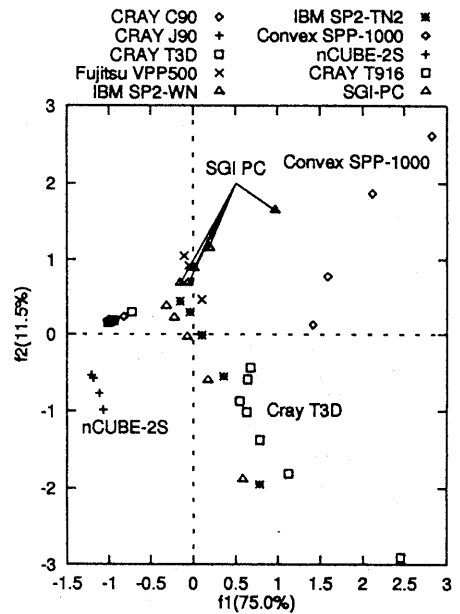
TPPは(1サイクルあたりに可能な浮動小数点演算回数)*(クロック周波数)*(プロセッサ数)で与えられ、単位時間あたりに可能な浮動小数点演算回数の理論値を与える。マシン M のプロセッサ数 p のときの TPP を $TPP_{M,p}$ 、そしてベンチマークの実測値を $T(M,p)$ とすると $TPP_{M,p} * T(M,p)$ によって正規化を行なった。TPPによって正規化されたデータをもとに、主成分分析を行なって得られた相関係数行列、主成分を表5、表6に示す。

表5を見ると、ベンチマーク間の相関はプロセッサ数

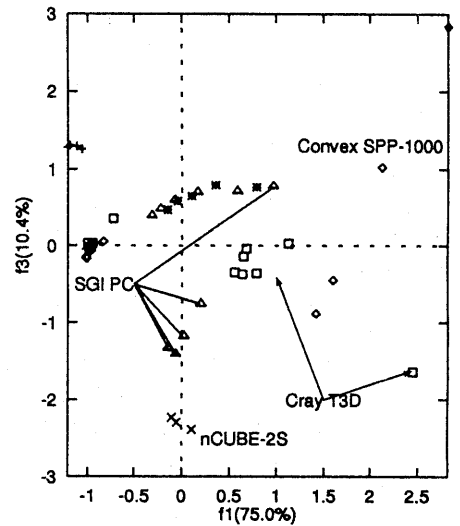
表6 得られた主成分-TPP で正規化した場合-固有値

| | | | | |
|-----------------|-------|--------|--------|--------|
| 寄与率 | 3.000 | 0.459 | 0.418 | 0.123 |
| 累積寄与率 | 0.750 | 0.115 | 0.104 | 0.031 |
| 主成分負荷行列(重みベクトル) | | | | |
| | 第1 | 第2 | 第3 | 第4 |
| EP | 0.821 | 0.263 | -0.504 | -0.043 |
| MG | 0.928 | 0.160 | 0.212 | 0.260 |
| CG | 0.800 | -0.591 | -0.101 | 0.005 |
| FT | 0.908 | 0.119 | 0.329 | -0.232 |

のみあるいはプロセッサ数とクロック周波数で正規化した場合に比べて、小さくなっている。これはマシンの違いによるデータの変動が小さくなり、プロセッサ数の変化に伴うデータの変動が顕著になったためである。表5より最も相関の強いベンチマークはFTとMGである。いずれも規則的な通信を含むベンチマークテストである。これに対し最も相関の弱いベンチマークはEPとCGである。EPはプロセッサ間通信を含まないためプロセッサ数に比例した効果が期待される。そのためEPとの相関を見ることにより、他の3つのベンチマークの結果のプロセッサ数に比例した性能が得られる度合を見ることができる。ちなみにCGは不規則な長距離通信を含むベンチマークである。TPPで正規化した場合も得られた主成分のうち第1主成分は、4つのベンチマークの全体的な傾向を示している。また、第2主成分はCG以外の3つのベンチマーク、特にEPとCGの差であるといえる。この傾向はクロック周波数で正規化した場合と変わらない。さらに第3主成分を見ると、主にMG、FTとEPの差であるといえる。つまりEPに比べてMG、FTの性能が悪くなる場合には、第3主成分は正の方向に大きくなる。そこで第1主成分、第2主成分を軸とする主成分平面上に各マシンをプロットしたときの図および第2主成分、第3主成分を軸とする主成分平面上に各マシンをプロットしたときの図をそれぞれ図2(a)、図2(b)に示す。図2よりTPPで正規化したことにより、マシン間のばらつきは小さくなっていることがわかる。また図2(a)と図1を比較すると、主成分平面上の各マシンの位置は正規化の方法により異なるため、異なるマシンを比較するには正規化の方法を検討する必要がある。しかしながらTPPで正規化したときも、Convex SPP-1000、SGI Power Challengeはプロセッサ数の増加に伴い全体的な性能は低下しているが、CGの性能は向上していることが読みとれる。またCRAY T3D、nCUBE-2S等はプロセッサ数の増加に伴いCGの性能の低下が著しいことを表している。次に図2(b)を見ると、Convex SPP-1000、SGI Power Challengeがプロセッサ数の増加に伴い第3主成分が正の方向に増えている。つまりEPに比べてMG、FTの性能が相対的に悪くなっている。これは他のマシンと



(a) 第1主成分、第2主成分を軸とする場合



(b) 第1主成分、第3主成分を軸とする場合

図2 主成分分析の結果-TPPで正規化した場合-

比べて特異的な傾向である。これに対しCRAY T3Dは、128プロセッサまではMG、FTの性能が相対的に良くなっている。

3. クラスタ分析によるスケーラビリティの考察

並列システムのスケーラビリティはプロセッサ数や計算サイズの変化に見合った性能が得られる度合を表す指標として、並列システムの性能評価の際に用いられてい

る。本研究ではスケーラビリティを定量的に与える試みとしてクラスタ分析による方法について考察する。

クラスタ分析¹⁾とは複数個の特性によって決定されるデータをデータ間の類似性の指標をもとにいくつかのグループ(以下クラスタと呼ぶ)に分類する手法である。クラスタ分析を適用するにあたり、まず分類の際のデータ間の類似性の指標となるデータ間の距離を定義する必要がある。一般的に用いられている距離としてはユークリッド距離、マハラノビスの汎距離などがある。さらにいくつかのデータをまとめて1つのクラスタとしたときにさらにクラスタ同士を併合して新たなクラスタを形成していくがこの際にクラスタ併合の基準となるクラスタ間の距離を定義する必要がある。クラスタ間の距離の定義の仕方によりいくつか方法があるが、ここではクラスタの重心間の距離をクラスタ間の距離の定義とする重心法を採用した。

具体的に関口、佐藤の報告⁵⁾にある Thinking Machine Corp. CM-5 の並行列乗算プログラムによる測定結果を用いてクラスタ分析を行なった。まずデータをプロセッサ数および問題サイズで正規化した。プロセッサ数を p 、行列のサイズを n として、行列サイズ n の問題を p 台のプロセッサで実行したときの実測値を $T(p, n)$ 、各プロセッサにおける単位時間あたりの計算量を $e(p, n)$ として $T(p, n) = n^3 / (e(p, n) * p)$ を仮定する⁶⁾。以下このように仮定した $e(p, n)$ に対してクラスタリングを行なうこととする。上述したようにクラスタ分析の適用にあたり、まずデータ間の類似性の指標となる距離を定義する必要がある。本研究においては距離を以下のように定義する⁷⁾。

$$t_i \equiv \log_2 e(p_i, n_i) \\ d_{ij} \equiv |t_i - t_j| \quad (2)$$

ここで式(2)によって定義された距離とマージナル効率との関連について、考察する。マージナル効率とはプロセッサ数を変化させたときに元のプロセッサ数での効率を100%と仮定したときの効率で、プロセッサ数 p_1 からプロセッサ数 p_2 へ変化させたときのマージナル効率は次式によって与えられる⁶⁾。

$$\text{Marginal Efficiency} \equiv \frac{p_1 \cdot T(p_1, n)}{p_2 \cdot T(p_2, n)} \quad (3)$$

マージナル効率が1に近いほどプロセッサ数を変化させても効率が低下しない、つまりスケーラブルな状態であるといえる。ここで式(3)を変形する。

$$|\log(\text{Marginal Eff.})| = \left| \log \frac{p_1 \cdot T(p_1, n)}{p_2 \cdot T(p_2, n)} \right| \\ = \left| \log \frac{e(p_1, n)}{e(p_2, n)} \right| \\ = |\log e(p_1, n) - \log e(p_2, n)| \quad (4)$$

式(4)より式(2)によって定義された距離はマージナル効率の拡張になっていることがわかる。式(2)によって定義された距離に基づいて構成されたクラスタにおい

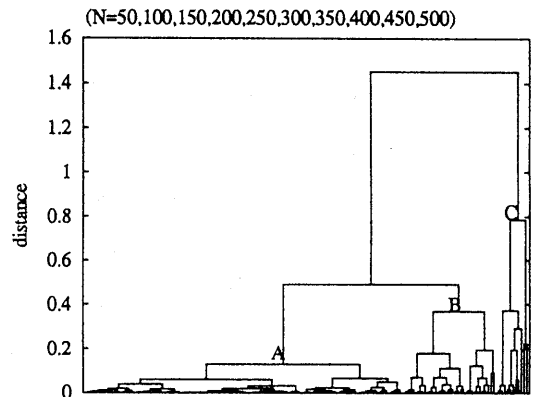


図3 クラスタ分析の結果—CM5 のスケーラビリティ—

て、同一クラスタに入る点はプロセッサ数および計算サイズの変化に対してマージナル効率がほぼ1、つまり効率が低下しないことを意味し、ほぼスケーラブルであるといえる。従って同一クラスタ内の点はほぼ同程度のスケーラビリティをもつ点であるといえる。

具体的にCM-5の並行列乗算プログラムの実行結果(行列の次元は50~500まで50刻に、プロセッサ数は2~32まで2刻に変化させたもの)を用いてクラスタリングを行なった結果の樹状図を図3に示す。

図3に示されているクラスタA、クラスタB、クラスタCに含まれるデータを表7に示す。表7に示されている値は1プロセッサに割り当てられる問題サイズ、ここでは行列の次元である。四角で囲まれているのがクラスタCに含まれるデータで、太字で示されているのがクラスタBに含まれるデータである。またそれ以外はすべてクラスタAに含まれる。図3を見るとクラスタAに含まれるデータは距離がほぼ0.1以下であり、つまりマージナル効率がほぼ1程度であるから最もスケーラブルな状態であるといえる。クラスタAに含まれるデータは表7より1プロセッサに割り当てられる行列サイズがほぼ10以上のものである。従って以上の結果はプロセッサに適度な演算量が与えられてくるとスケーラブルになるということを定量的に表している。図4に正規化したデータをプロットした。ここで示したCluster Aの領域が最もスケーラブルな領域である。クラスタ分析により、スケーラブルな領域を定量的に示すことができた。

4. まとめと今後の課題

本研究では統計的手法に基づく、並列計算機システムの評価手法についての検討としてベンチマーク間の相関性を検証する方法について検討した。具体的にNAS Parallel Benchmarksと呼ばれる高並列計算機のために開発されたベンチマークセットの中から4つのベンチマークを選び、統計的手法である主成分分析を用いて相関の検証を試みた。4つのベンチマーク間の相関係数の

表7 各クラスタに含まれるデータ

| 行列サイズ | 50 | 100 | 150 | 200 | 250 | 300 |
|--------|-------------------------|-----|-----|-----|-----|-----|
| プロセッサ数 | 問題サイズ (行列の次元)/processor | | | | | |
| 2 | 25 | 50 | 75 | 100 | 125 | 150 |
| 4 | 13 | 25 | 38 | 50 | 63 | 75 |
| 6 | 9 | 17 | 25 | 34 | 42 | 50 |
| 8 | 7 | 13 | 19 | 25 | 32 | 38 |
| 10 | 5 | 10 | 15 | 20 | 25 | 30 |
| 12 | 5 | 9 | 13 | 17 | 21 | 25 |
| 14 | 4 | 7 | 11 | 15 | 18 | 22 |
| 16 | 4 | 7 | 10 | 13 | 16 | 19 |
| 18 | 3 | 6 | 9 | 12 | 14 | 17 |
| 20 | 3 | 5 | 8 | 10 | 15 | 13 |
| 22 | 3 | 5 | 7 | 9 | 12 | 14 |
| 24 | 2 | 5 | 7 | 9 | 11 | 13 |
| 26 | 2 | 4 | 6 | 8 | 10 | 12 |
| 28 | 2 | 4 | 6 | 8 | 9 | 11 |
| 30 | 2 | 4 | 5 | 7 | 9 | 10 |
| 32 | 2 | 4 | 5 | 7 | 8 | 10 |

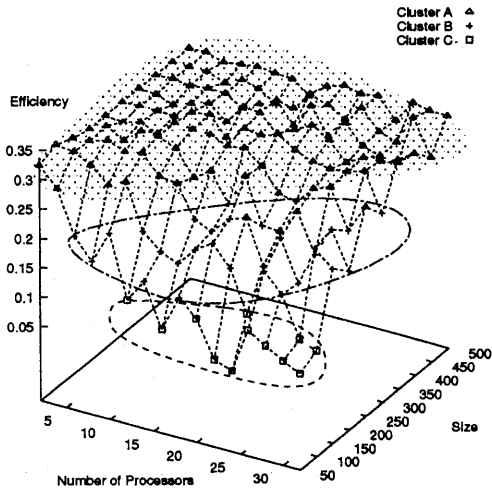


図4 正規化したデータの分類図

大きさはデータの正規化の方法により異なるが、傾向としては4つのうちMGとFTの相関が高く、相関が低いのはEPとCGであった。また得られた主成分のうち第1主成分は、4つのベンチマークをほぼ同じ比で合成し、総合的な性能を見ることができ、また第2主成分は主にEPとCGの差を表していた。さらにTPPで正規化した場合、第3主成分はMG、FTとEPの差を表していた。このことから4つのベンチマークは高い相関を持っているが、並列計算機の構成要素であるCPU性能や通信性能といった全体の性能をバランス良く評価していることが示唆された。また4つのうち1つはプロセッサ間通信を含まないベンチマークで演算性能を評価すること、また、1つは複雑な長距離通信を含むベンチマークで不規則な通信性能を評価することが意図された

ベンチマークであり、残りの2つのベンチマークは規則的な通信性能を評価することが意図されたベンチマークであるが、主成分分析の結果からもこれらのベンチマークを相対的に評価することにより通信性能を評価することができることが示唆された。

さらに性能評価においては性能評価指標として何をを用いるべきか、また得られた数値から何を読みとるべきかといった問題があるが、特に最近並列計算機システムの性能評価においてはスケラビリティという指標がよく用いられていることからクラスタ分析を用いてスケラビリティを定量的に扱うことを試みた。クラスタ分析においてはデータ間の類似性を距離によって定義するが距離をマージナル効率との関係に基づいて定義することにより同一クラスタ内に属する点は同程度のスケラビリティを持つということが出来る。具体的にCM-5の並列行列乗算プログラムの結果を用いてクラスタリングを行なったところ、各プロセッサに適当な演算が与えられてくるとスケラブルになるということを定量的に示すことができた。

今後は今回提案した評価手法をより多くのシステムに適用し検証を行なう。

謝辞 本研究を行なうにあたり、有益な御助言を頂きました筑波大学電子・情報工学系 朴泰祐助教授、中村宏助教授に感謝申し上げます。

参考文献

- 1) 柳井 晴夫, 高木 広文編著, 多変量解析解析ハンドブック, 現代数学社 (1986)
- 2) Bailey, D.; Barszcz, E.; Barton, J.; Browning, D.; Carter, R.; Dagum, L.; Fatoohi, R.; Fineberg, S.; Frederickson, P.; Lasinski, T.; Schreiber, R.; Simon, H.; Venkatakrishnan, V.; Weeratunga, S; The NAS Parallel Benchmarks, RNR Technical Report, RNR-94-007, NASA Ames Research Center, 1994
- 3) Saini, S.; Bailey, D.H. NAS Parallel Benchmarks Results 12-95, Report NAS-95-021, NASA Ames Research Center, 1995
- 4) 次世代スーパーコンピュータ性能評価技術調査研究事業報告書, 日本機械工業連合会 (1993)
- 5) 関口智嗣, 佐藤三久: 並列システムのスケラビリティによる評価, 研究報告 ARC-111, 情報処理学会 (1995)
- 6) 佐藤三久, 関口智嗣: 並列計算機システムのスケラビリティについて, 研究報告 HPC-49, 情報処理学会 (1993)
- 7) 関口智嗣, 佐藤三久, 井須芳美, 長嶋雲兵: 定量的な並列システムのスケラビリティ評価指標, JSPP'96, 投稿中