

機関番号：12601

研究種目：基盤研究（B）

研究期間：2009 ～ 2011

課題番号：20300007

研究課題名（和文） 階層的なアルゴリズム選択機構を有する自動チューニング方式の研究

研究課題名（英文） Auto-tuning with Hierarchical Algorithm Selection

研究代表者

佐藤 周行（SATO HIROYUKI）

東京大学・情報基盤センター・准教授

研究者番号：20225999

研究成果の概要（和文）：

本研究は、コンパイラの最適化や高速ライブラリの提供等で行われてきた従来型の性能チューニングが近年のコンピュータアーキテクチャの複雑化に対応しきれない現状の問題点を、コンピュータアーキテクチャに特化したライブラリ開発とその性能解析、アルゴリズム選択の自動化、コンパイラ最適化の拡張を主とした、それぞれの問題解決の抽象度に応じた選択機構の研究を行うことで、解決したものである。一部のチューニング技法や開発アルゴリズムが実用化を見据えている。

研究成果の概要（英文）：

This research aims at solving the complexity problems of conventional compiler-oriented or library-oriented performance tuning caused by explosive computer architecture complexities by doing architecture-specific library functions and their performance evaluation, automatic optimal algorithm selection, and enhanced compiler optimization framework that are classified into several abstraction levels in problem solving. A part of tuning methodologies and library functions are planned to be released.

交付決定額

（金額単位：円）

	直接経費	間接経費	合計
2009 年度	10,200,000	3,060,000	13,260,000
2010 年度	2,900,000	870,000	3,770,000
2011 年度	2,000,000	600,000	2,600,000
総計	15,100,000	4,530,000	19,630,000

研究分野：計算機科学

科研費の分科・細目：情報学（ソフトウェア）

キーワード：

並列処理 分散処理 自動チューニング
性能最適化

1. 研究開始当初の背景

高性能化に向けてコンピュータアーキテクチャの多様化が進展しつつある。さらに、近年の個々の分野における特定の問題を解決する数値計算アプリケーション、また Web サイエンス中での知識ネットワークの解析では、入力データの特徴を活かした特殊化ライブラリが性能をあげている。しかし、開発されたライブラリが他の分野に適用され得る

汎用性を持つかどうかについてはほとんど無関心であった。個々の計算機がハードウェアとして保持する性能、また個々のアルゴリズムの最適な適用場面の決定を、人手を介さずに最大限に利用する仕組みが求められていた。とりわけ、数値計算アプリケーションの共通基盤となるライブラリにおいては、様々な計算機上で、各種のデータに対して効率的に動作することが要求される。この状況

のもとで、ハードウェア的にはアプリケーションの性能パラメータを、対象計算機のキャッシュ、プロセッサ数などのアーキテクチャ特性に自動的にチューニングし、アルゴリズム的には対象入力データを解析して適用アルゴリズムを自動選択する**ソフトウェア自動チューニング技術**が注目を集めていた。

2. 研究の目的

計算機がハードウェアとして保持する性能を利用者が最大限に利用するためには、ソフトウェアによる性能の自動チューニングが必要である。これは、従来コンパイラの最適化、ライブラリの複数提供で提供されてきたが、これらは性能最適化の観点からはごく原始的なロジックで動作したり、人間の高度な知見を対話的に投入しないと動作しないなどの問題点を抱えていた。本研究のゴールとする自動チューニングにおいては、多様な計算機のアーキテクチャやリソース的な制約、また入力データの性質に適応してアルゴリズムを適切に選択する機能の提供が重要な役割をはたす。本研究では、同機能を実現する**＜アルゴリズム選択機構＞**に焦点をあてた自動チューニングの枠組みを提案し、それを実装・評価することを目的とする。特に需要の多い数値計算アプリケーションを重要なターゲットとする。

3. 研究の方法

ソフトウェア自動チューニング機能は、次の2つの機構で実現される。1. 性能パラメータの最適化を行う機構と2. 最適アルゴリズムを選択する機構である。アプリケーション実行のポリシーとしてユーザによって入力されるデータとして最大実行時間、使用する最大記憶量、計算結果の精度、許容する性能安定性など、また、計算機環境情報として、記憶容量、キャッシュサイズ、プロセッサ数、通信性能などを考える。出力は選択された最適アルゴリズムと最適化された性能パラメータである。

性能パラメータ最適化機構では、自動的にパラメータサーチを行う方法のほかにユーザからの関数指定に基づき最適化を行う機能がある。

最適アルゴリズム選択機構では、性能パラメータ最適化機構によってあたえられた最適性能パラメータを利用し、最適なアルゴリズムを決定する。最適アルゴリズム選択のため、統計手法を用いた自動分類方式、 λ 式を用いた実装導出方式を適用する。また、すべての可能性を全探索する方式をユーザが指定した

場合には、全探索により最適アルゴリズムを決定する。

これらの性能チューニングにおいて、プログラム実行中の計算機性能を、インストール時にモニタする必要がある。実行時間、キャッシュミス率、性能安定性などを測定し、チューニングの効果を定量的に予測する。予測結果に基づいて、チューニングの判定を行う。判定基準については、速度と精度、速度と記憶量、速度と記憶量と精度のトレードオフなどが存在する。対象が並列計算機の場合、プロセッサ数と速度のトレードオフがある。これらのトレードオフを勘案して、効果を予測し、性能パラメータおよび適用アルゴリズムを決定し、チューニングを実施する。

4. 研究成果

研究方法の2つの軸となる性能パラメータの最適化機構と最適アルゴリズムの選択機構について、前者については特に数値アプリケーションを題材に鳥、性能パラメータの網羅的な解析を行った。また後者については、連立一次方程式の求解法やコンパイラ最適化について新たな視点をもたらすことができた。具体的には以下のとおりである。

1. 数値アプリケーションを題材にした性能パラメータ最適化の研究

片桐、今村は具体的に固有値問題と疎行列問題において、アルゴリズム選択による性能向上の有効性を示した。さらに黒田はPCクラスタの計算性能・通信性能を網羅的に調査し、さらに既存のスーパーコンピュータとの比較を行い、パラメータチューニングでの基本データを蓄積することができた。片桐、黒田はOpenATLibを提案し、自動チューニングのAPIを提案した。特に並列・分散プラットフォーム上での優位性を明らかにした。さらに伊藤、今村は数値計算における反復法と固有値問題において、自動チューニング機構の基礎となる理論の展開と、関係するパラメータの渉猟を行い、それについての幅広い知見を得た。

2. アルゴリズム選択機構の研究

研究分担者のうち、伊藤は選択されるアルゴリズムのランドスケープを得るために、当初は経験的な視点から、アルゴリズムの分類を行い、性能との組み合わせを行い、大規模な表を得た。さらにアルゴリズム選択において、管理工学の手法を適用してアルゴリズムを網羅的に並べる手法を開発した。この副産物

として、連立一次方程式の求解法の分類について新たな知見を得た。佐藤はコンパイラ最適化において、イディオム認識法を拡張して、従来の中間表現の枠組みを本質的に拡張する手法を提案した。さらに、インライン展開を対象にして、アルゴリズム選択のための基準として新しい評価軸を与えた。また、片桐、黒田は、特にマルチコアシステムのシステム特性を研究することで、選択機構を構築する上で必要となるパラメタセットの導出に成果を残した。

研究期間内において、(自動)チューニングの問題点の所在と、その難しさが次第に明らかになった。この研究の成果が、その認識を深める大きなきっかけになったことを確信している。また、この研究と連携する形で多くの研究プロジェクトが立ち上がるのを見ることができた。これらとの連携は、毎年開催された合同ワークショップで強いものになった。

今後の展望としては、高速ライブラリ開発や方法論の展開といった自動チューニングの要素技術をさらに展開することが視野に入っている。さらにコンピュータアーキテクチャの性質を性能パラメタの網羅的なサーチによってさらに明らかにすることの重要性が強調される。性能に大きく寄与しながら、まだチューニングやコンパイラ最適化で認識されていない性能パラメタを探することができる環境を整えられた。また、この研究で得た知見を本格的なチューニングツールに実現することも視野に入れることができる。個々の要素技術の進展とともに、自動チューニングのフレームワークの提供を見据えた展開に注力する予定である。

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20件)

- ① Imamura, T., Yamada, S., and Machida, M. : Narrow-band reduction approach of a DRSM eigensolver on a multicore-based cluster system, Chapman, B., et al (Eds.), Parallel Computing: From Multicores and GPU's to Petascale, pp. 91-98 IOS Press, 2010. 査読有
- ② Imamura, T., Yamada, S., and Machida, M. : Development of a high performance eigensolver on the peta-scale next generation supercomputer system, Proceedings of Joint International

Conference on Supercomputing in Nuclear Applications and Monte Carlo 2010 (SNA+MC2010), 2010. [CD-ROM] 査読有

- ③ Pham, H. P., Imamura, T., Yamada, S., and Machida, M. : Novel approach in a divide and conquer algorithm for eigenvalue problems of real symmetric band matrices, Proceedings of Joint International Conference on Supercomputing in Nuclear Applications and Monte Carlo 2010 (SNA+MC2010), 2010. [CD-ROM] 査読有
- ④ Yamada, S., Imamura, T., Okumura, M., Igarashi, R., Onishi, H., and Machida, M. : High Performance Computing of Density Matrix Renormalization Group Method for 2-Dimensional Model: Parallelization Strategy toward Peta Computing, Proceedings of Joint International Conference on Supercomputing in Nuclear Applications and Monte Carlo 2010 (SNA+MC2010), 2010. [CD-ROM] 査読有
- ⑤ Imamura, T., Kano, T., Yamada, S., Okumura, M., and Machida, M. : High-Performance Quantum Simulation for Coupled Josephson Junctions on the Earth Simulator: A challenge to Schrodinger Equation on 256^4 Grids, International Journal of High Performance Computing, vol. 24 no. 3 319-334, Sage publication, 2010. 査読有
- ⑥ Yoshida, S., and Imamura, T. : Acceleration of two dimensional FMM using GPU, Proceedings of International Workshop on Modern Science and Technology 2010 (IWMST 2010) , at Kitami Institute of Technology, September 2010. 査読有
- ⑦ 今村俊幸 : 超並列環境向きの固有値計算アルゴリズムと自動チューニング, 応用数理, Vol.20, No. 3, 2010. 査読なし
- ⑧ 伊藤祥司, 杉原正顕, 姫野龍太郎, クリロフ部分空間法に対する前処理方式と収束判定について, 情報処理学会論文誌コンピューティングシステム(ACS), Vol.3, No.2, pp.9-19, 2010. 査読有
- ⑨ Shoji Itoh, Masaaki Sugihara, Systematic performance evaluation of linear solvers using quality control techniques, Software Automatic Tuning From Concepts to State-of-the-Art Results (eds. Ken Naono, Keita Teranishi, John Cavazos and Reiji Suda), pp.135-152, Springer, 2010. 査読有
- ⑩ Takahiro Katagiri and Shoji Itoh, A

Massively Parallel Dense Symmetric Eigensolver with Communication Splitting Multicasting Algorithm, LNCS (the post-conference book of VECPAR'10) (in press).査読有

- ⑪ Sato, H.: "Embedding Dataflow Information on Arrays into SSA and Extended Optimization Schema for Parallelization," Proc. Parallel and Distributed Computing and Networks 2009, 240-245, 2009.査読有
- ⑫ Sato, H.: "Analyzing Semantics of Documents by using a Program Analysis Method," Proc. COMPSAC 2009, 373-382, 2009.査読有
- ⑬ Nishimura, T., Sato, H.: "Analysis of a Security Incident of Opensource Middleware -- Case Analysis of 2008 Debian Incident of OpenSSL--," in MidArch 2009, Proc. 9th Int'l Symp. Applications and the Internet, 247-250, 2009.査読有
- ⑭ Sato, H.: "Idiom Recognition and Program Scheme Recognition based Program Transformation for Performance Tuning -- Beyond Compiler Optimizations--," Proc. Parallel and Distributed Computing, Applications and Technologies 2009, 272-279, 2009.査読有
- ⑮ 弓場, 片桐: "ソフトウェア自動チューニングの枠組み", 科学技術計算におけるソフトウェア自動チューニング特集号 (解説), 情報処理学会誌 Vol.50, No.6(2009年6月) 478-482.査読なし
- ⑯ 今村俊幸: <ソフトウェア自動チューニング技術の応用> MPI 通信ライブラリの自動チューニング, 情報処理 Vol.50, No.6, pp. 523-526, June 2009.査読なし
- ⑰ 伊藤祥司: ソフトウェア自動チューニングのための支援ツール, 情報処理, Vol.50, No.6, pp.499-504, (2009).査読なし
- ⑱ 黒田久泰, 直野健, 岩下武史: <ソフトウェア自動チューニング技術の応用> 6.自動チューニング機能付き数値計算ライブラリ, 情報処理学会会誌「情報処理」, 特集「科学技術計算におけるソフトウェア自動チューニング」, Vol.50 No.6, pp.505-511, 2009 査読なし
- ⑲ Sato, H.: "Delay Front and Skip Interval for Assessing Open Source Software Maintenance of System," Proc. 32nd IEEE Computer Software and Applications Conference (COMPSAC), 771-776, 2008.査読有

[学会発表] (計72件)

- ① Tajima, K. and Yuba, T.: "Speeding up a Sort Program by Algorithm Selection," First Data Intensive Workshop, Poster (Tokyo, March 9, 2011)
- ② 真子, 佐藤: インライン展開による最適化の加速の度合いの計測およびフィードバック手法, 情報処理学会プログラミング研究会 PRO 82, 宮古市中央公民館, 2011年1月21日
- ③ 今村俊幸: ペタスケール環境での高並列固有値ソルバの開発, 日本計算工学会 計算工学講演会論文集 Vol. 15, pp. 103-106, 九州大学, 2010年5月26日.
- ④ 黒田久泰, 片桐孝洋, 須田礼仁: 消費電力を抑えた基本演算ライブラリの実装と評価, 2010年並列/分散/協調処理に関する『金沢』サマー・ワークショップ (SWoPP2010), 金沢市文化ホール, 2010年8月4日
- ⑤ 今村俊幸, 山田進, 町田昌彦: ペタスケール環境における固有値ソルバについて, 日本応用数理学会 2009年度年会講演予稿集, pp. 1-2, 大阪大学 (豊中市) 2009.9月28日
- ⑥ 黒田久泰, 片桐孝洋, 須田礼仁: 省電力のための基本演算ライブラリの実装と評価, 2009年並列/分散/協調処理に関する『仙台』サマー・ワークショップ (SWoPP2009), フォレスト仙台, 日本応用数理学会「行列・固有値問題の解法とその応用」, 2009年8月4日
- ⑦ Yuba, T., Katagiri, T. and Suda, R.: "A Survey of Research Topics: Towards Challenges and Future Directions of Software Automatic Tuning Technology," Proc. 4th International Workshop on Automatic Performance Tuning (iWAPT2009), Poster, 73 (Tokyo, October 1, 2009)
- ⑧ 雪下, 佐藤: "シンボリックアサインメントグラフを用いたコンパイラ最適化器," 第26回ソフトウェア科学会大会, 3C-2, 松江, 2009年9月17日

[図書] (計0件)

[産業財産権] (計0件)

[その他]

ホームページ等

<http://www.atrg.jp/>

6. 研究組織

(1) 研究代表者

佐藤周行 (SATO HIROYUKI)
東京大学・情報基盤センター・准教授
研究者番号：20225999

(2) 研究分担者

片桐 孝洋 (KATAGIRI TAKAHIRO)
東京大学・情報基盤センター・特任准教授
研究者番号：40345434

今村 俊幸 (IMAMURA TOSHIYUKI)
電気通信大学・情報理工学部・准教授
研究者番号：60361838

伊藤 祥司 (ITO SHOJI)
東京大学・情報基盤センター・特任准教授
研究者番号：70333482

黒田 久泰 (KURODA HISAYASU)
愛媛大学・大学院情報理工学系研究科・准教授
研究者番号：60323507

(*) 研究協力者

弓場 敏嗣 (YUBA TOSHITSUGU)
電気通信大学・名誉教授