

## 科学研究費助成事業 研究成果報告書

平成 26 年 5 月 29 日現在

機関番号：10101

研究種目：基盤研究(C)

研究期間：2010～2013

課題番号：22500196

研究課題名(和文) ロバストかつスケーラブルな進化計算による超大規模最適化問題の解決

研究課題名(英文) Extremely large-scale optimization problem-solving using robust and scalable evolutionary computation

研究代表者

棟朝 雅晴 (Munetomo, Masaharu)

北海道大学・情報基盤センター・教授

研究者番号：00281783

交付決定額(研究期間全体)：(直接経費) 3,300,000円、(間接経費) 990,000円

研究成果の概要(和文)：本研究においては、進化計算アルゴリズムの改良として混合分布モデルへの拡張を行った BOA-MD、局所探索との融合を行った BHCS、MINLP (Mixed-Integer Non-Linear Programming) のための ARGA および BRGA を開発するとともに、進化計算の大規模並列化としてメニーコアアーキテクチャにおける大規模並列実装を行うとともに、クラウド環境への展開も図った。

現実の問題に対する適用例として、MINLP 問題としてのクラウド資源最適化問題や、薬剤の構造を自動的に最適化することを目指した de novo Ligand Docking 問題等へ適用し、その有効性を検証した。

研究成果の概要(英文)：We have developed a series of evolutionary algorithms such as BOA-MD, an extension of BOA introducing mixture distributions, BHCS by combining BOA with local search, ARGA and BRGA for MINLP (Mixed-Integer Non-Linear Programming). We also developed large-scale parallel evolutionary algorithms in a many core architecture and cloud computing environment.

As applications to real-world problems, we applied evolutionary algorithms to optimal resource allocation problem in cloud computing environment as a MINLP problem, de novo ligand docking problems to find promising structures of medicines automatically, and so on, to show the effectiveness of our approach.

研究分野：総合領域

科研費の分科・細目：情報学、感性情報学・ソフトコンピューティング

キーワード：進化計算 大規模並列化 メニーコア

## 1. 研究開始当初の背景

近年のテラフロップス級、ペタフロップス級のスーパーコンピュータの開発およびグリッドコンピューティングの発展により、大規模な最適化問題を一度に解くことを目指した大規模並列最適化アルゴリズムが現実性を増してきており、研究開発プロセスの大幅な効率化、スピードアップを図ることが期待されている。

現実問題の設計において大規模な最適化を行う場合、設計者の経験則に基づき部分問題(部品など)に分割した上で、それぞれの部分問題について、その問題にあわせてチューニングされた最適化アルゴリズムを用いることが多い。そのような人手による問題分割を行うことなく、大規模な問題をそのまま一度に解こうとする”Big-Bang”的なアプローチを可能とするスケーラブルなアルゴリズムを開発することで、人間の研究者・設計者は問題を定式化するだけで良く、解の探索を大規模計算資源にまかせることが期待される。従来は計算機的能力およびコストの問題から、このようなアプローチは非現実的であったが、計算機の性能向上、コスト低下により現実的になりつつある。

目的の実現のためには、問題の構造を動的に分析しつつ最適化を行うことで広範囲の問題を安定的に解ける「ロバスト」かつ「スケーラブル」なアルゴリズムの開発が必要条件となる。ロバストな進化計算アルゴリズムの開発に関しては、我々の研究グループにより、リンケージ同定など遺伝子間の相互依存関係を動的に検出しつつ最適化を進める手法を開発してきた。また、超大規模問題を解くためのスケーラブルな並列化の試みとして、進化アルゴリズムによる Billion-bit 級の超大規模最適化問題の解法に関する研究がなされている。

さらに、我々の研究グループでは、北海道大学の保有するスパコン(5.4TFlops) 上にて、複雑なリンケージ構造を有する 50 万ビットのテスト関数の最適化問題を、リンケージ同定を行いつつ解くことに成功している。ここで用いられたリンケージ同定アルゴリズム D<sup>5</sup> は、スケーラビリティにおいて BOA よりも優れているがそれでも  $O(l^2)$  の計算量を必要とし、スパコンでも 1 日以上計算を必要としている。

## 2. 研究の目的

以上の背景およびこれまでの研究成果に基づき、本研究においては、ロバストかつスケーラブルな進化計算アルゴリズムを開発するため、リンケージ同定や確率モデルの構築など依存関係の解析の対象となる遺伝子座群を適切な基準で選択することで、大幅に解析やモデル化のコストを削減し、スケーラビリティを向上させるための方法論につい

て研究を進める。具体的には、混合ペイジアンネットワークなど、複雑な確率モデル構築を可能とすることで、計算コストの増加を抑えつつ適切な遺伝子解析を行うことで、効率的かつ高精度なリンケージ同定やモデル構築を行うための理論的な基盤を確立する。

大規模並列化にあたっては、クラスタシステムにおける実装に加えて、マルチコア、メニーコアアーキテクチャへの実装を行い、近年の大規模計算環境に適合した大規模並列化を実現する。本研究では特に、1 チップあたり数百程度の多数のコアから構成される GPGPU アーキテクチャへの展開を図る。具体的には、GPGPU の開発フレームワークである CUDA を用いたアルゴリズムの大規模並列化を行うとともに、多数の GPGPU から構成され、高いコストパフォーマンスでテラフロップス超級の演算性能を達成できる GPGPU クラスタシステムを前提としてシステムの実装を行う。GPGPU アーキテクチャにおいては、多数の計算コアを並列に用いることが可能であるが、使用可能なローカルメモリなどに制約が多く、単純な並列化は困難であり、アーキテクチャの条件を十分考慮したアルゴリズムの開発および実装が求められる。

さらにグリッドコンピューティング・クラウドコンピューティングなど現実の大規模分散システム環境への展開を図る。グリッド環境下における問題解決支援システムについては、我々の研究グループで開発が進められ、小規模なプロトタイプを構築することができているが、本研究においてはその成果を踏まえて、大規模なグリッドコンピューティングシステム・クラウドコンピューティング環境への展開を図る。

## 3. 研究の方法

平成 22 年度においては、主にスケーラブルなリンケージ同定、および確率モデル構築による進化アルゴリズムの開発を中心に研究を進める。特に複雑な遺伝子間の依存関係を考慮しつつ現実的な計算時間で最適化を行うこと、さらに並列化に向いていることなどを判断基準として、モデル構築の高度化を実現する。

平成 23 年度以降においては、アルゴリズムの検証、改良を行うとともに、開発されたアルゴリズムの大規模並列化、メニーコアアーキテクチャへの対応を実現する。平成 23 年度においては、主にアルゴリズムの検証、改良を中心に行い、平成 24 年度においては、大規模並列化、メニーコアアーキテクチャへの対応、平成 25 年度においては、グリッド・クラウド環境への実装を中心に研究を進める。

アルゴリズムの検証については、平成 22 年度において開発されたアルゴリズムの、大規模ベンチマーク問題および、現実の大規模

組み合わせ最適化問題へ適用した実験を通して、スケーラビリティおよび解の精度に関する分析を行う。そこで得られた結果を基にして、アルゴリズムの改良およびそのパラメータ設定に関する理論的、実験的検証をさらにすすめる。

大規模並列化に関しては、クラスタ型の相互通信による並列化、マルチコア CPU への対応に加えて、比較的粒度の小さい演算装置を多数備えたメニーコアアーキテクチャの一つである GPGPU における大規模並列化に向けた並列化手法の開発をすすめる。

さらに、現実の複雑かつ大規模な最適化問題へ開発された大規模並列アルゴリズムを適用し、その有効性を検証するとともに、グリッドコンピューティングやクラウドコンピューティングなど現実の大規模並列分散システムへの展開を図る。

#### 4. 研究成果

##### (1) 進化計算アルゴリズムの改良

進化計算アルゴリズムの改良については、確率モデル構築による進化計算アルゴリズムである Bayesian Optimization Algorithm (BOA) において、その確率モデルにおいて、複数の Bayesian Network を混合した確率モデルへの拡張を行った手法、BOA-MD (BOA with Mixture Distribute) を開発し、その有効性を複雑な問題構造を有する最適化問題において検証した。さらに、確率モデル構築において、スキーマ上での曖昧性を考慮しつつ局所探索を融合させる手法についても検討をすすめ、BHCS (BOA with Hill-Climbing on Schemata) と呼ばれる手法を開発し、その有効性をテスト関数において検証した。

さらに、現実の設計問題等を解くために必要となる、整数および実数の変数が混在した非線形計画問題の解法として、進化計算手法を適用するための研究を行った。具体的には、広範囲の実数空間を部分空間に分割し、その探索の精度をバイナリ型の遺伝的アルゴリズムにより動的に制御しながら探索をする BRGA (Binary-Real coded GA)、適応的に精度を制御する ARGGA (Adaptive Resolution GA) を開発し、各種の代表的なベンチマーク問題において従来手法と比較した優位性を検証している。さらに制約付き最適化問題への対応を行うための動的制約制御アルゴリズムの導入を行った。その探索性能について評価するため、CEC2010 ベンチマーク問題での検証実験を行い、その有効性を確かめることができた。

##### (2) メニーコアにおける大規模並列化

大規模かつ複雑な問題のための進化計算手法の開発、およびメニーコアアーキテクチャにおける大規模並列実装の両面から研究

を行った。メニーコアアーキテクチャの実装に関しては、GPGPU のフレームワークである CUDA を用いた実装を行った。

対象問題としては、組み合わせ最適化問題として有名な問題の一つである巡回セールスマン問題において、百万都市レベルの超大規模問題の解決を目指した手法の開発を行った。具体的には Zoning Crossover と呼ばれる部分領域情報に基づく部分経路の生成手法を導入した交叉手法を開発することで、大規模な巡回セールスマン問題を進化計算による効率的に解くことを可能とした。

##### (3) 現実の大規模最適化問題への適用

具体例としては、現実の設計問題で多く見られる、整数と実数が混在した複雑な最適化問題である MINLP (Mixed-Integer Non-Linear Programming) への適用に加えて、創薬に必要な有望な薬剤構造を発見する探索問題である、De Novo Ligand Docking 問題への適用を行った。

MINLP への適用については、その GPGPU 向けの実装を行うことで、MINLP のベンチマーク問題において、CPU の数倍から数十倍の高速化を実現することができた。その応用として、クラウドコンピューティングにおける資源割当最適化問題への適用を行った。当該問題は、仮想化された計算資源をそのアプリケーションシステムの必要性に応じて物理資源上に最適配分する必要がある、多数の仮想マシンについてその割当先となる物理マシンの ID や割当資源量など、数多くの整数および実数のパラメータが混在する混合整数計画問題として定式化される。そのような多数の整数、実数が混在した困難な最適化問題を解決する手段として、BRGA が有効であることを検証した。

創薬への応用については、薬剤の構造を自動的に最適化する手法である de novo Ligand Docking 問題を対象に成果をあげた。具体的には、そのエネルギー計算に大量の計算量を要するため、CUDA による大規模並列実装を行い、GPU における各スレッドで並列に有望な薬剤構造に関する局所探索を行う手法を提案することで、並列化の効率を向上させることができた。結果として、従来の CPU による実装では現実的な時間で解を得られなかったところ、数日程度で必要な薬剤の構造を得ることに成功した。本研究の成果として、肝臓疾患に有望となる薬剤構造の候補を自動的に生成することができ、有望な薬剤構造のデータベースである ZINC データベースに正式登録された。

##### (4) クラウドコンピューティングへの展開

当初の計画では、進化計算の GPU における大規模並列実装が中心であったが、その部分についてはおおむね成果をあげたため、クラ

ウドコンピューティング環境下における大規模並列実装についても検討を進め、多数のバーチャルマシン、演算コアを利用できる Infrastructure as a Service (IaaS) 環境を前提とし、Hadoop による大規模並列データ処理環境における大規模並列化に加えて、CloudFoundry による先端的な Platform as a Service (PaaS) 環境への実装についても検討を進めることができ、当初の計画を上回る成果をあげた。

具体的にはクラウド環境における MapReduce による大規模並列化について、進化計算を用いた代表的な機械学習アルゴリズムである Classifier System の Hadoop による大規模並列実装を行い、ログデータから不正侵入パターンを学習する侵入検知問題へ適用することで、大規模なデータを高速に学習するためのフレームワークを実現した。

さらに、クラウドコンピューティングで近年注目を集めている PaaS 環境における大規模並列実装についても検討を行い、代表的な汎用 PaaS 環境である CloudFoundry 上において、多数のバーチャルマシンを用いた Interactive Evolutionary Computation (IEC) の大規模並列実行環境を実現し、多数のユーザからの入力に対応して自律的な進化を実現することのできるプロトタイプシステムを構築した。

#### 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22件)

1. Omar Arif Abdul-Rahman, Masaharu Munetomo, Kiyoshi Akama, An adaptive parameter binary-real coded genetic algorithm for constraint optimization problems: Performance analysis and estimation of optimal control parameters, Information Sciences, 査読有, Vol. 233, No. 1, pp. 54-86 (2013)
2. Masaharu Munetomo, Shintaro Bando: A Scalable Infrastructure of Interactive Evolutionary Computation to Evolve Services Online with Data, Proceedings of the 2013 IEEE International Conference on Big Data, 査読有 p.28, (2013)
3. Masataka Mizukoshi, Shitaro Bando, Martin Schlueter, Masaharu Munetomo: Implementation of Multiple Classifier System on MapReduce Framework for Intrusion Detection, Proceedings of the 2013 International Conference on Parallel and Distributed Processing Techniques and Applications, 査読有, pp.1-4 (2013)
4. 川勝崇史, 棟朝雅晴: 分散クラウド環境における SLA を考慮した WEB システムの多目的資源割当最適化, 情報処理学会研究報告, 査読無, Vol.2013-MPS-96, No.9, pp.1-6 (2013)
5. 堀伸哉, 棟朝雅晴, 赤間清: 混合ページアンネットワークを導入した分布推定アルゴリズム, 進化計算学会論文誌, 査読有, Vol.3, No.2, pp.63-72 (2012)
6. Omar Abdul-Rahman, Masaharu Munetomo, Kiyoshi Akama: An adaptive Parameters Binary-Real Coded Genetic Algorithm for Real Parameter Optimization: Performance Analysis and Estimation of Optimal Control Parameters, IJCSI International Journal of Computer Science Issues, 査読有, Vol.9, Issue 2, No.1, pp.37-53 (2012)
7. Mohamed Wahib, Asim Munawar, Masaharu Munetomo, Kiyoshi Akama: A GPU accelerated Fragment-Based de Novo Ligand Design by a Bayesian Optimization Algorithm, IPSJ Transactions on Bio Informatics (TBIO), 査読有, Vol.5, pp.7-17 (2012)
8. Omar Abdul-Rahman, Masaharu Munetomo, Kiyoshi Akama: Toward A Genetic Algorithm Based Flexible Approach for the Management of Virtualized Application Environments in Cloud Platforms, Proceedings of the 21th International Conference on Computer Communications and Networks (ICCCN2012), 査読有, pp.1-9 (2012)
9. 川勝崇史, 棟朝雅晴: インタークラウド環境における Web 三層モデルと CDN を組み合わせた多目的資源割当最適化に関する検討, 情報処理学会研究報告, 査読無, Vol.2012-MPS-90, No.13, pp.1-2 (2012)
10. Omar Abdul-Rahman, Masaharu Munetomo and Kiyoshi Akama: An Adaptive Resolution Hybrid Binary-Real Coded Genetic Algorithm, Journal of Artificial Life and Robotics, 査読有, vol.16, no.1, pp.121-124 (2011)
11. Omar Abdul-Rahman, Masaharu Munetomo and Kiyoshi Akama: An Adaptive Resolution Hybrid Binary-Real Coded Genetic Algorithm, Proceedings of the Sixteenth International Symposium on Artificial Life and Robotics 2011, 査読有, pp.359-362 (2011)
12. Asim Munawar, Mohamed Wahib, Masaharu Munetomo, Kiyoshi Akama: Solving extremely difficult MINLP problems using Adaptive Resolution micro-GA with Tabu Search, Learning and Intelligent Optimization (LION5), Lecture Notes in Computer Science

- 6683, 査読有, pp.203-217 (2011)
13. Mohamed Wahib, Asim Munawar, Masaharu Munetomo and Akama Kiyoshi: A GPU accelerated Fragment-Based De Novo Ligand Design by a Bayesian Optimization Algorithm, 情報処理学会研究報告, 査読無, Vol.2011-MPS-85, No.6, pp.1-6 (2011)
  14. 堀伸哉, 棟朝雅晴, 赤間清: 動的環境問題における BOA への混合ベイジアンネットワーク導入の効果について, 情報処理学会研究報告, 査読無, Vol.2011-MPS-85, No.16, pp.1-6 (2011)
  15. Omar Abdul-Rahman, Masaharu Munetomo and Kiyoshi Akama: An Improved Binary-Real Coded Genetic Algorithm for Real Parameter Optimization, Proceedings of the Third World Congress on Nature and Biologically Inspired Computing (NaBIC2011), 査読有, pp.156-163 (2011)
  16. Omar Abdul-Rahman, Masaharu Munetomo and Kiyoshi Akama: Multi-Level Autonomic Architecture for the Management of Virtualized Application Environments in Cloud Platforms, Proceedings of the IEEE 4th International Conference on Cloud Computing (IEEE CLOUD 2011), 査読有, pp.754-755 (2011)
  17. Mohamed Wahib, Asim Munawar, Masaharu Munetomo and Akama Kiyoshi: A Framework for Cloud Embedded Web Services Utilized by Cloud Applications, Proceedings of the IEEE 2011 World Congress on Services Computing (IEEE SERVICES 2011), 査読有, pp.265-271 (2011)
  18. Masaharu Munetomo: Realizing Robust and Scalable Evolutionary Algorithms toward Exascale Era, Proceedings of 2011 IEEE Congress on Evolutionary Computation, 査読有, pp.312-317 (2011)
  19. Mohamed Wahib, Asim Munawar, Masaharu Munetomo and Akama Kiyoshi: On the Optimization of Parallel Genetic Algorithms over Nvidia's GPU, Proceedings of 2011 IEEE Congress on Evolutionary Computation, 査読有, pp.803-811 (2011)
  20. Asim Munawar, Mohamed Wahib, Masaharu Munetomo and Kiyoshi Akama: Advanced Genetic Algorithm to solve MINLP problems over GPU, Proceedings of 2011 IEEE Congress on Evolutionary Computation, 査読有, pp.318-325 (2011)
  21. Omar Abdul-Rahman, Masaharu Munetomo, Kiyoshi Akama: Live Migration-based Resource Managers for Virtualized Environments: A Survey, Proceedings of The First International Conference on Cloud Computing, GRIDs, and Virtualization (CLOUD COMPUTING 2010), 査読有, pp. 32-40 (2010)
  22. Mohamed Wahib, Asim Munawar, Masaharu Munetomo, Akama Kiyoshi: A Bayesian Optimization Algorithm For De Novo Ligand Design Based Docking Running Over GPU, Proceedings of the 2010 IEEE Congress on Evolutionary Computation, 査読有, pp.17-24 (2010)
- [学会発表](計16件)
1. 平島慶典, 三浦克宜, 棟朝雅晴: Hadoop 環境上で動作する研究分野判定ツールの試作, 第75回情報処理学会全国大会, 2013.3.6, 東北大学
  2. 鈴木一史, 棟朝雅晴: リンケージツリー遺伝的アルゴリズムにおける計算量削減の検討, 進化計算シンポジウム 2012, 2012.12.15, ホテルマロウド軽井沢
  3. 堀伸哉, 棟朝雅晴, 赤間清: 混合ベイジアンネットワークを導入した分布推定アルゴリズム, 第2回進化計算学会研究会, 2012.3.10, 大阪大学
  4. 堀伸哉, 棟朝雅晴, 赤間清: BOA への混合ベイジアンネットワーク導入の効果について, 進化計算シンポジウム 2011, 2011.12.17, モンタナリゾート岩沼
  5. 鈴木一史, 棟朝雅晴, 赤間清: スキーマ上での局所探索によるベイジアン最適化アルゴリズムの評価, 進化計算シンポジウム 2011, 2011.12.18, モンタナリゾート岩沼
  6. 鈴木一史, 棟朝雅晴, 赤間清: スキーマを考慮した局所探索を導入したベイジアン最適化アルゴリズムの検討, 進化計算学会・人工知能学会第7回進化計算フロンティア研究会, 2011.9.9, 東京大学
  7. Omar Abdul-Rahman, Masaharu Munetomo, Kiyoshi Akama: An Adaptive Parameters Binary Real-Coded Genetic Algorithm with Dynamic Constraint Handling Technique for Constrained Problems Optimization, 進化計算学会・人工知能学会第7回進化計算フロンティア研究会, 2011.9.10, 東京大学
  8. 鈴木一史, 棟朝雅晴, 赤間清: スキーマを考慮した局所探索を導入した分布推定アルゴリズムの提案, 人工知能学会第6回進化計算フロンティア研究会, 2011.3.14, 名古屋大学
  9. 堀伸哉, 棟朝雅晴, 赤間清: BOA における

混合ベジアンネットワークの導入の効果について, 人工知能学会第 6 回進化計算フロンティア研究会, 2011.3.15, 名古屋大学

10. Asim Munawar, Mohamed Wahib, Masaharu Munetomo, Kiyoshi Akama: Adaptive Resolution Genetic Algorithm to solve MINLPs over nVidia Fermi GPU using CUDA, ACM/IEEE SC10 Poster, 2010.11.16, New Orleans Ernst N. Morial Convention Center
11. Mohamed Wahib, Asim Munawar, Masaharu Munetomo, Kiyoshi Akama: A novel Evolutionary Algorithm for Flexible Protein Docking Using Fragment-based de Novo Ligand Design Over GPU, 人工知能学会第 5 回進化計算フロンティア研究会, 2010.10.7, 北海道大学
12. 堀 伸哉, 棟朝 雅晴, 赤間 清: 混合ベジアンネットを使用した BOA の新手法についての検討, 人工知能学会第 5 回進化計算フロンティア研究会, 2010.10.7, 北海道大学
13. 仲田 善弘, 棟朝 雅晴, 赤間 清: Extended compact GA の Map Reduce による大規模並列化に関する検討, 情報処理北海道シンポジウム 2010, 2010.10.2, 北海道大学
14. Bang Minyoung, Omar Abdul-Rahman, Munetomo Masaharu, Akama Kiyoshi: A Mathematical Model for Optimal Resource Allocation in Cloud Systems, 情報処理北海道シンポジウム 2010, 2010.10.2, 北海道大学
15. Asim Munawar, Mohamed Wahib, Masaharu Munetomo, Kiyoshi Akama: Introducing Adaptive Resolution Technique to solve Mixed Integer Non-Linear Optimization Problems using Genetic Algorithms, 人工知能学会第 4 回進化計算フロンティア研究会, 2010.6.4, 東京工業大学
16. 棟朝雅晴: ロバストかつスケラブルな確率的最適化アルゴリズムの実現に向けて, 電気学会第 2 回確率的最適化アルゴリズムの設計技術調査専門委員会, 2010.7.27, 北海道大学

## 6. 研究組織

### (1) 研究代表者

棟朝 雅晴 (MUNETOMO, Masaharu)  
北海道大学・情報基盤センター・教授  
研究者番号: 00281783

### (2) 研究分担者

( )

研究者番号:

### (3) 連携研究者

( )

研究者番号: