

科学研究費助成事業 研究成果報告書

令和 2 年 6 月 22 日現在

機関番号：12201

研究種目：基盤研究(C) (一般)

研究期間：2016～2019

課題番号：16K00068

研究課題名(和文) 超大規模相互結合網のための高追従性分散制御方式の研究

研究課題名(英文) Study on highly adaptable distributed control of interconnection networks

研究代表者

横田 隆史 (Yokota, Takashi)

宇都宮大学・工学部・教授

研究者番号：90334078

交付決定額(研究期間全体)：(直接経費) 3,500,000円

研究成果の概要(和文)：相互結合網は並列計算機において通信の機能を司る構成要素であり、その成否はシステム全体の性能を大きく左右する。本課題は、並列プログラムの実行に伴い雪崩的に発生する通信に対して、(1) 個々の構成要素の自律動作により、(2) 通信状況に素早く追従して制御できる制御システムを構築する目的で遂行した。

本課題では、後追いの手法では対処が難しいことから、制御状態の目標となる理想状態を解明することに注力した。その結果、大パケット送出時刻の制御(時間的な最適化)および並列計算ノードの論理配置の制御(空間的な最適化)の2方向の検討に対し、人工知能での手法を導入することで論文誌への発表等多くの成果を得た。

研究成果の学術的意義や社会的意義

並列計算機の通信機構に関しては、従来、様々な通信状況に対応できる機構的な工夫が中心だった。本課題で理想状態を解明することにより、制御の可否次第で大きく通信性能を向上できることが明らかになった。この成果は、過重な機構・装備を備えていない簡素な構成であっても、使いこなし方の工夫によって遜色ない性能を達成できる可能性があることを示している。

また、本課題の研究の過程では、粒子群最適化(PSO)、遺伝的アルゴリズム(GA)、機械学習(ML)、深層学習(DL)などの人工知能の手法を積極的に導入し、顕著な成果を得ることができたことから、当該分野への人工知能技術の適用が進むものと期待される。

研究成果の概要(英文)：Interconnection network is an inevitable component in parallel computing systems. It largely affects the total performance of the parallel system. This program aims at a rapid control of the interconnection network by a highly distributed manner in autonomous routers. Since an initial study of the program suggested severe difficulty in appropriate reaction to the vast communication traffic, this program has employed some essential discussions on the ideal situations. The program has introduced some artificial intelligence methodologies, which are novel approaches in this area, and temporal and spatial optimizations are deeply discussed. Through the research period, this program has produced many results such as journal papers.

研究分野：計算機システム

キーワード：相互結合網 並列計算機 輻輳制御 時間的最適化 空間的最適化 大規模最適化

1. 研究開始当初の背景

相互結合網は並列計算機に特異な構成要素であり、並列処理において本質的に求められる通信を司る。その成否はシステム全体の並列処理性能を左右するのみならず、実現のコストや消費電力の大きさにも大きく関わることから、重要な研究対象と位置づけられている。

一般に相互結合網では、通信負荷が低く輻輳が生じない状況では負荷量に比例した性能が得られるが、輻輳が発生する負荷量に達すると極端に性能が低下する非線形性を示し、しかも解消までに多大な時間を要する問題がある。このため、最大性能を得られるポイントすなわち輻輳を生じない限界の条件で運用することが大きな研究課題となっている。さらに超大規模並列システムでの相互結合網では、実現性の観点から(集中的ではなく)分散的な制御手法を求められることも大きな障壁となっている。

こうした問題意識は既に広く共有されており多くの提案がなされているが未だ理想的な手法はない。実際の並列処理では、特定の通信パターンに従う一斉同時通信を繰り返すことが多く、相互結合網に対して突発的かつ極端な通信負荷変動が課せられる。しかし既往研究での提案手法ではこうした負荷変動に十分に追従できず、ブレイクスルーが求められていた。

2. 研究の目的

本研究課題では、相互結合網内での時間的・空間的挙動に着目し、突発的な負荷変動に即座に対応できる方式の創出を目指す。この目的のため、以下の体系的アプローチをとる。

(1). 制御オプションの整理

本研究課題では相互結合網の制御に際して以下の技術を適用する。いずれも排他的なものではなく、組み合わせで適用することが可能である。

- (a) パケット投入制御(スロットリング): 新たなパケットの生成を抑制する。
- (b) フロー制御: ルータ間でのパケット授受の制御を行う。なおパケットの滅失は許容しない。
- (c) ルーティング制御: パケットの転送方向を制御する。

(2). 代表的な通信パターン

種々の通信状況に対応できるようにするため、転置・シャッフル・ビット反転・ビット逆転・ビット回転・ランダム・トルネード・ホットスポットの通信パターンを想定する。

(3). パケットと相互結合網の詳細な動的挙動の把握

上記(1)、(2)の組合せによる広いバリエーションを前提にして、時間的・空間的にダイナミックに変動する挙動を効果的に表現・理解する手法を追求する。制御理論による数理的な表現手法の導入を探求しながら、これと並行して数理的表現では表せない現実挙動の表現手法を求める。この検討は従来型のシミュレーション実行だけでは遂行できないため、シミュレーション方式の革新も必要である。これらの成果により、制御オプションの選択の結果、相互結合網上のいつ・どこで・どのように影響するかを解明できる。

(4). 分散制御方式の創出

上記の動的挙動把握の検討の過程では、制御オプションとその適用方法について多くの試行がなされる。正確な挙動理解が進めば、それは即ち効果的な分散制御手法の創出に直結する。さらに本研究課題では、一時故障も含む様々な状況にも対応できる手法を追求する。

3. 研究の方法

(1). 相互結合網状態の時系列変化の把握と先行制御手法の初期検討

相互結合網での通信の効率が落ちるのは、結合網上にある複数の通信パケットの間で、相互に進行をブロックしあう状況が発生するためである。こうした状況を輻輳と呼び、たとえば、パケット間でのブロックが発生している度合いを以て輻輳度とする。

全ノードが一斉にパケットを送出すると、輻輳度はきわめて急峻な傾きにより上昇する。もともと本課題で前提にしている相互結合網には、輻輳状態を抑制する機構を備えているが、雪崩的に発生する輻輳に対しては制御が追い付かず、その結果、輻輳度は初期段階で大きなオーバーシュートのカーブを描き、そののち徐々に抑制が効いてくる遷移をする。

理想的には、結合網のすべての状態を一瞬のうちに把握し、即座に制御を効かせることであるが、現実的でない。実験的に、本課題より以前の研究で用いたシミュレータを改変し、少量の制御情報を通信パケットの数倍(2から8倍程度)の速度で伝播させる仕組みを仮想的に実装して評価実験を行った。具体的には、相互結合網を構成する各ルータでのバッファの使用状況を満杯か否かの2値で伝播させた。パケット群がこちらに押し寄せようとするのを前もって知ることができるので、制御に反映させようという発想である。しかしこの評価実験では、期待した効果は得られなかった。これは、理想的な制御方法を目指すには、輻輳の影響が顕在化してからの「事後」ではなく、その前に通信をナビゲートするような形で制御することが望ましいことを示唆している。

(2). 時間的最適化の粒子群最適化 (PSO) による求解

全ノードが一定の通信パターンに従った通信パケットを所定の個数だけ一斉に送出する一斉同期通信 (collective communication) を対象とし、全ノードが所定のパケットを受信し終えるまでと定義する通信時間を短縮する方法を検討した。一斉同期通信では、全ノードが「一斉に」パケットを送出するが、それは「同時」である必要はない。個々のノードでのパケット送出時刻をずらすことで、パケット間の干渉を抑えることができれば、通信時間は短縮するはずである。

個々のノードが、それぞれのパケットを送出する時刻を変数とし、通信時間が最短になる変数の値の組を求める。これが、時間的最適化の問題となる。本課題では、この求解のために、粒子群最適化 (PSO, Particle Swarm Optimization) を導入した。探索空間は、ノードの個数 \times 送出するパケットの個数の次元数になる。こうした広大な (高次元の) 探索空間に対して効率よく探索を行うための手法を工夫した。

(3). 時間的最適化の深層学習による求解

前項(2)の粒子群最適化により顕著な改善成果を得ることができた。しかし個々の結果から法則性を見出すことが難しかった。突出して最適なケースが唯一に決まることはむしろ少なく、複数の解がないし類似した性能の解が複数発見されるケースが多数である。もし、それぞれの解候補に対して、通信時間を短縮する傾向の有無により仕分けられれば、実用的に役立つはずである。

その作業は人間の目視等では到底かなわないため、個々の解候補をパターンとみなし、深層学習によるパターン分類の技術を適用することを考えた。このため、多数の解候補をランダムに生成し、それぞれの通信時間を測定することで、学習データを生成した。単純な神経回路網のモデルで評価実験を行った。

(4). シミュレータの高速化

本課題の先行研究で、セルオートマトンの原理を適用することで相互結合網の動作を高速に模擬できるシミュレーション手法を検討していた。本課題では、粒子群最適化や遺伝的アルゴリズムなど各種の大規模最適化技法を導入して研究を進めたため、膨大なシミュレーションをこなす必要が生じた。このため、セルオートマトンに基づく手法をさらに進め、高速シミュレータとして確立した。本課題の実施にあたって、相互結合網シミュレーションに関する部分は、すべてこのシミュレータをベースにして行った。

(5). 空間的最適化の遺伝的アルゴリズム (GA) による求解

並列処理の中で行われる通信のパターン (通信パターン) により有向グラフが形成される。また並列計算機の構成 (すなわち相互結合網) を別の有向グラフとして表すことができる。両者のグラフをできるだけ一致させることができれば、通信のオーバーヘッドを劇的に削減できるはずである。これは、トポロジ・マッピングとして並列処理分野で古くから検討されている内容である。通信パターンが1つのみの場合など手続き的に解決できるケースもあるが、一般的には困難である。

本課題では、2次元トーラスのトポロジを想定し、1つ以上のパターンにより通信させたとき、最適となる論理配置を求めた。この求解には遺伝的アルゴリズムを適用し、本問題に適した手法を開発し準最適解を求めた。

(6). 時間的最適化の機械学習による求解 (初期検討)

上記(2)の時間的最適化では、粒子群最適化により膨大な空間の中での探索を行った。この問題に対する求解の方法はひと通りに決まるわけではなく、他にさまざまな方法が考えられる。機械学習の適用もそのひとつである。本課題では、検討の初期段階として、探索空間を狭めて有効性を見極めをやすくする観点から、ひとつのノードに着目し、そこでのパケット送出時刻の最適値を求める問題として求解法を検討した。具体的には、既存手法であるQ学習を導入し、相互結合網全体を表す状態情報の遷移の様子を記録していきながら機械学習を試みた。

4. 研究成果

(1). 相互結合網状態の時系列変化の把握と先行制御手法の初期検討

研究の方法欄にも一部記載したとおり、全ノードが一斉にパケットを送出すると、輻輳度はきわめて急峻な傾きにより上昇する。輻輳度は通信の初期段階で大きなオーバーシュートのカーブを描き、そののち徐々に抑制が効き漸減する。少量の制御情報を通信パケットの数倍 (2から8倍程度) の速度で伝播させる仕組みを仮想的に実装して評価実験を行った結果、期待した効

果は得られなかった。この結果は、輻輳の影響が顕在化してからの制御ではなく、その前に通信をナビゲートするような形で制御することが望ましいことを示唆している。

(2). 時間的最適化の粒子群最適化 (PSO) による求解

i 番目のノードが j 番目に送出する通信パケットの送出時刻を x_{ij} とする。粒子群最適化により探索空間 $(x_{0,0}, x_{0,1}, \dots, x_{1,0}, x_{1,1}, \dots)$ から最適な組み合わせを探索することになる。この問題の特徴を活かしながら探索空間を縮小する工夫を加えることにより、 16×16 ノードの 2 次元トラス網において最大約 1.3 倍の高速化を達成した。

この評価実験の過程で、結合網内の輻輳状況の経過を詳細に記録し得られた通信性能と比較した結果、通信時間を最短にするには、結合網内での輻輳の発生を完全に抑えるのではなく、多少の輻輳を許容することが必要であることが明らかになった。

(3). 時間的最適化の深層学習による求解

上記の項目 (2) での評価実験に用いたプログラムを一部改変し、各ノードでのパケット送出時刻を完全にランダムに設定できるようにした。これにより、上記項目と同様の $(x_{0,0}, x_{0,1}, \dots, x_{1,0}, x_{1,1}, \dots)$ の空間中にランダムに散在するサンプルを多数準備した。これらを学習用の教師データおよびテストデータとして使用し、入力層のユニット 64、中間層 100、出力層 2 の神経回路網に学習処理をさせた。その結果、分別しやすいと考えられるテストデータの場合、約 90% 以上の正解率が得られた。分別困難なテストデータについても 70% 以上の正解率を得ることができ、神経回路網の適用も有望であることが示された。

(4). シミュレータの高速化

セルオートマトンの考え方に基づいて相互結合網の内部動作をモデリングすることにより、無駄を極力省いた並列処理を可能とし、シミュレーション速度を飛躍的に高めることができた。民生用 GPU (Graphics Processing Unit) を用いた場合で、従来型のシミュレータと比較して、最大 1264 倍の高速化を達成した。マルチコアプロセッサを用いた場合でも、最大 162 倍の高速化を達成した。このシミュレータは、本研究課題の遂行に際しても (2), (3), (5), (6) の実施項目に適用され、研究効率を高めている。

(5). 空間的最適化の遺伝的アルゴリズム (GA) による求解

問題の性質に合わせた遺伝子表現および突然変異処理を工夫することにより、 32×32 ノード (1,024 ノード) の 2 次元トラス網において、通信パターン 1 種の場合で最大 8.9 倍、2 種の組み合わせの場合で最大 4.9 倍の速度向上を達成した。

(6). 時間的最適化の機械学習による求解 (初期検討)

相互結合網内の各ルータの状態を統合したものを「状態」とし、特定の 1 ノードからのパケット送出の可否を機械学習 (Q 学習) により、通信時間が短くなるように学習させた。本実施項目は、研究の初期段階であるため、Q 学習のアルゴリズムをナイーブに実装し、効果を確認した。その結果、設定した通信パターンの種類や指定するノードの位置により得られる効果の有無があるものの、いくつかの条件によっては、通信時間を改善できることが明らかになった。本課題では、多数ある計算ノードの 1 個のみを対象としたが、これを複数ないし全体に拡張することにより、通信性能向上の効果が得られるものと期待される。

5. 主な発表論文等

〔雑誌論文〕 計40件（うち査読付論文 16件 / うち国際共著 0件 / うちオープンアクセス 2件）

1. 著者名 深澤 勇人, 横田 隆史, 大津 金光	4. 巻 1
2. 論文標題 Q学習による相互結合網制御の初期検討	5. 発行年 2020年
3. 雑誌名 情報処理学会 第82回全国大会 講演論文集	6. 最初と最後の頁 93-94
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 無
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -
1. 著者名 YOKOTA Takashi, OOTSU Kanemitsu, OHKAWA Takeshi	4. 巻 E102.D
2. 論文標題 Accelerating Large-Scale Interconnection Network Simulation by Cellular Automata Concept	5. 発行年 2019年
3. 雑誌名 IEICE Transactions on Information and Systems	6. 最初と最後の頁 52 ~ 74
掲載論文のDOI (デジタルオブジェクト識別子) 10.1587/transinf.2018EDP7131	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -
1. 著者名 YOKOTA Takashi, OOTSU Kanemitsu, OHKAWA Takeshi	4. 巻 E101.D
2. 論文標題 A Genetic Approach for Accelerating Communication Performance by Node Mapping	5. 発行年 2018年
3. 雑誌名 IEICE Transactions on Information and Systems	6. 最初と最後の頁 2971 ~ 2975
掲載論文のDOI (デジタルオブジェクト識別子) 10.1587/transinf.2018PAL0002	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -
1. 著者名 Takashi Yokota, Kanemitsu Ootsu, Takeshi Ohkawa	4. 巻 E100-D
2. 論文標題 A Static Packet Scheduling Approach for Fast Collective Communication by Using PSO	5. 発行年 2017年
3. 雑誌名 IEICE Transactions on Information and Systems	6. 最初と最後の頁 2781-2795
掲載論文のDOI (デジタルオブジェクト識別子) 10.1587/transinf.2017PAP0015	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Takashi Yokota, Kanemitsu Ootsu, Takeshi Ohkawa	4. 巻 1
2. 論文標題 Large-Scale Interconnection Network Simulation Methods Based on Cellular Automata	5. 発行年 2017年
3. 雑誌名 Proc. 5th International Symposium on Computing and Networking (CANDAR'17)	6. 最初と最後の頁 58-67
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/CANDAR.2017.52	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 横田 隆史, 大津 金光, 大川 猛	4. 巻 117
2. 論文標題 大規模NoC向けシミュレーション手法の検討	5. 発行年 2017年
3. 雑誌名 電子情報通信学会 技術報告	6. 最初と最後の頁 105-110
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 無
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Takashi Yokota, Kanemitsu Ootsu, Takeshi Ohkawa	4. 巻 E99-D
2. 論文標題 Enhancing Entropy Throttling: New Classes of Injection Control in Interconnection Networks	5. 発行年 2016年
3. 雑誌名 IEICE Transactions on Information and Systems	6. 最初と最後の頁 2911-2922
掲載論文のDOI (デジタルオブジェクト識別子) 10.1587/transinf.2016PAP0007	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Takashi Yokota, Kanemitsu Ootsu, Takeshi Ohkawa	4. 巻 1
2. 論文標題 Introducing PSO for Optimal Packet Scheduling of Collective Communication	5. 発行年 2016年
3. 雑誌名 Proc. 4th International Symposium on Computing and Networking (CANDAR 2016)	6. 最初と最後の頁 429-435
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/CANDAR.2016.39	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 相澤 和秀, 横田 隆史, 大津 金光, 大川 猛	4. 巻 1
2. 論文標題 神経回路網による相互結合網パケットスケジュール最適化の初期検討	5. 発行年 2017年
3. 雑誌名 情報処理学会 第79回全国大会 講演論文集	6. 最初と最後の頁 55-56
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 無
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 横田 隆史, 大津 金光, 大川 猛	4. 巻 2016-ARC-221
2. 論文標題 一斉同期通信の最短時間パケットスケジューリング PSOによる最適化の試み	5. 発行年 2016年
3. 雑誌名 情報処理学会 研究報告	6. 最初と最後の頁 1-8
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 無
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

〔学会発表〕 計33件 (うち招待講演 0件 / うち国際学会 9件)

1. 発表者名 Takashi Yokota, Kanemitsu Ootsu, Takeshi Ohkawa
2. 発表標題 Large-Scale Interconnection Network Simulation Methods Based on Cellular Automata
3. 学会等名 5th International Symposium on Computing and Networking (CANDAR'17) (国際学会)
4. 発表年 2017年

1. 発表者名 横田 隆史, 大津 金光, 大川 猛
2. 発表標題 大規模NoC向けシミュレーション手法の検討
3. 学会等名 電子情報通信学会コンピュータシステム研究会
4. 発表年 2017年

1. 発表者名 Takashi Yokota, Kanemitsu Ootsu, Takeshi Ohkawa
2. 発表標題 Introducing PSO for Optimal Packet Scheduling of Collective Communication
3. 学会等名 4th International Workshop on Computer Systems and Architectures (国際学会)
4. 発表年 2016年

1. 発表者名 横田 隆史, 大津 金光, 大川 猛
2. 発表標題 一斉同期通信の最短時間パケットスケジューリング PSOによる最適化の試み
3. 学会等名 情報処理学会 システム・アーキテクチャ研究会
4. 発表年 2016年

1. 発表者名 相澤 和秀, 横田 隆史, 大津 金光, 大川 猛
2. 発表標題 神経回路網による相互結合網パケットスケジュール最適化の初期検討
3. 学会等名 情報処理学会第79回全国大会
4. 発表年 2017年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究協力者	大津 金光 (Ootsu Kanemitsu) (00292574)	宇都宮大学・工学部・准教授 (12201)	

6. 研究組織（つづき）

	氏名 (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究 協力者	大川 猛 (Ohkawa Takeshi) (80392596)	宇都宮大学・工学研究科・助教 (12201)	平成30年度まで