様　式　Ｃ－１９、Ｆ－１９－１、Ｚ－１９　（共通）

科学研究費助成事業　　研究成果報告書

令和　 元 年　 ６ 月 ２４ 日現在

機関番号：３２６４１
研究種目：研究活動スタート支援
研究期間：2017～2018
課題番号：１７Ｈ０７１１６

研究課題名（和文）マルコフ連鎖解析に基づく非正則・動的ネットワーク上負荷分散アルゴリズムの理論保証

研究課題名（英文）Analysis of distributed load balancing algorithms on inhomogeneous and dynamic networks

研究代表者
白髪　丈晴（Shiraga, Takeharu）

中央大学・理工学部・助教

研究者番号：５０８０３９９６

交付決定額（研究期間全体）：（直接経費）　　2,100,000円

研究成果の概要（和文）：各頂点に離散値の負荷が与えられた分散ネットワーク上おいて，各頂点における負荷がなるべく均等に配分されるように負荷を再配分する「ネットワーク上の負荷分散」アルゴリズムを考える．本研究ではその単純さ故に多くの研究が為されてきた拡散アルゴリズムの解析を行う．これは，各時刻において各頂点が自身の近傍に負荷を拡散させることを繰り返すことで全体に負荷を分散させるアルゴリズムの総称である．
本研究では新しい確率的な拡散アルゴリズムを提案し，その誤差解析を行った．その結果，任意のグラフに対し次数の劣線形となる上界を与えた．

研究成果の学術的意義や社会的意義
昨今のネットワークの大規模化，P2P ネットワークのような分散型ネットワークの出現により，ネットワーク上の各プロセッサにかかる負荷を均等に配分しなおすことを目指す「ネットワーク上の負荷分散」アルゴリズムは近年その需要が増している．しかし，単純なアルゴリズムに対しても，その解析の困難さより，既存研究では正則グラフをはじめとした限られたグラフ構造上での解析に留まっている．
本研究では主にマルコフ連鎖の過渡解析技法を適用し，既存のモデルを非正則なグラフ上でも動作するように拡張・更にはその誤差に対する制度保証を与えた．非正則なグラフ上での成果はほぼ初であり，その意義は大きい．

研究成果の概要（英文）：Given a distributed network where each vertex has discrete loads, we consider the problem of minimizing the discrepancy between the maximum and minimum loads among all vertices. For this problem, diffusion based algorithms are well studied because of its simplicity. In a diffusion based algorithm, at each synchronous and discrete time step, each vertex is allowed to distribute its loads to each neighbors.
This work is concerned with the ability of natural diffusion based algorithms. We presents a new randomized diffusion algorithm like multiple random walks. Our algorithm achieves a sub linear upper bound of the maximum degree for any graphs.

研究分野：アルゴリズム理論

キーワード：負荷分散　マルコフ連鎖

１．研究開始当初の背景

ネットワーク上の負荷分散問題とは，初期状態として各頂点に離散値の負荷が与えられたグラフ上において，局所的な負荷の遷移の繰り返しによってグラフ全体で負荷を均等に配分しなおすことを目指す問題である．昨今のネットワークの大規模・複雑化により，高速にかつ均等に負荷を分散させるアルゴリズムの需要は急増しており，他の基本的な分散計算との関連も相まって近年積極的に理論研究が行われている．代表的な成果として，``拡散"というシンプルなアイデアに基づいた乱択アルゴリズムに対し正則グラフ上において均等な配分との理論的な誤差保証を与えた[1]や，複数の決定的な拡散アルゴリズムに対しタイトな誤差解析を与えた[2]，そして正則グラフ上において均等な配分との誤差を定数に抑えることが出来る，巧妙な乱択アルゴリズム設計とその解析[3]があげられる．これらの成果は``離散値の拡散"と``連続値の拡散"の誤差に対し理論保証を与えている点で大きなインパクトを与えている．

その一方で，非常に単純なアルゴリズムに対してもその解析は困難であり，既存研究は主に正則グラフなど限られたグラフ構造上でのみ行われていた．正則性は解析の見通しが良くなる仮定の一つではあるが，それが本質的なギャップを誘発しているかどうかは不透明であり，さらに多くの現実世界のネットワークはスケールフリー性を有す非正則な構造をもつことを鑑みても応用・理論双方の視点から非正則なグラフ上における負荷分散アルゴリズムの開発・理論保証が求められていた．

２．研究の目的

以上の背景に基づき，本研究では非正則なグラフ上で動作する負荷分散アルゴリズム設計，及びその理論保証に挑む．

３．研究の方法

まず，最も単純な負荷分散アルゴリズムである，拡散アルゴリズムに対しその誤差解析を行う．拡散アルゴリズムとは自らの負荷を近傍へ拡散させることを繰り返すことでネットワーク上の負荷分散を目指すアルゴリズムの枠組みである．その単純さ，局所性から既存研究で既に誤差解析が行われていた[1]が，解析の困難さから正則グラフ上・また非常に緩い上界しか示されていない．連続値の拡散モデルはマルコフ連鎖そのものであり，正則・非正則問わずその収束性に対し豊富な理論が構築されていることに着目し，本研究ではまずこの拡散モデルに対しマルコフ連鎖の技法を適用し，よりタイトな上界を導出することを目指す．その上で，一般の構造上へ成果の拡張を試み，目標の達成を目指す．

４．研究成果

[成果 1] 非正則なグラフ上における拡散アルゴリズムの理論保証

本研究の主要成果として，まず正則グラフ上において，既存の乱択負荷分散アルゴリズムに対し誤差解析を行い，既存の次数に対する線形オーダーの誤差上界を改善し，次数 d に対し O(¥sqrt{d})の上界を与えた．これは拡散アルゴリズムに対する初めての劣線形上界である．なお，既存研究[2]において，d 正則グラフ上のどのような決定的な拡散アルゴリズムに対しても，誤差が¥Omega(d)となってしまう負荷の初期配置が存在することが示されていた．本研究の劣線形上界は拡散に基づく乱択負荷分散が，決定性負荷分散よりも高性能であることを示した最初の結果ともなっている．解析では，「自身の誤差を近傍へ等分し，余った誤差を確率的に配分する」というアルゴリズムの，負荷の確率的な移動を枝ごとに見積もり，確率不等式を用いて上界の導出を行った．この枝ごとの見積もりにより，既存研究でボトルネックとなっていた次数の上界の洗練，そして定常状態への収束誤差に関わる確率行列の 2 乗和の既存成果を組合せることで所望の上界を得た．

更に，同様の上界を非正則なグラフ上でも達成するアルゴリズムの設計を行った．このアルゴリズムは複数トークンのランダムウォークをアイデアに設計されている．具体的には，各頂点上の負荷１つ１つに独立な「区間」を割り当て，各負荷がその区間内で乱数を発生させ，その乱数に従って近傍への遷移が発生する(全ての負荷の区間が[0,1]の場合，複数トークンのランダムウォークに対応する）．このモデルはシンプルではあるが非正則なグラフ上で一様分布を簡単に達成でき，更には，より一般的に，任意の遷移確率行列に対し適用できるアルゴリズムとなっている．このモデルに対し，先述の拡散アルゴリズムと同様の，次数の劣線形となる誤差上界の導出に成功した．なお，遷移確率行列としてメトロポリスヘイスティング法を用いれば，自身の近傍の次数分布のみの情報で一様分布が達成できる．従って，提案アルゴリズムによって，任意の非正則なグラフ上において最大次数の劣線形オーダーの誤差で均等な誤差配分を達成することが出来る．また，一般の確率行列に対応する拡散モデルとしては，deterministic random walk と呼ばれるランダムウォークを模倣する決定的な拡散過程研究においても研究が行ってきたが，その誤差上界は遷移確率行列の第二固有値に依存していた（雑誌論文(2)も参照)．即ち，エキスパンダーグラフなど完全グラフに近い構造では良い上界を与えるが，ボトルネックがあるグラフでは誤差が大きくなってしまう．これに対し，提案乱択アルゴリズム誤差上界は対応する状態遷移図の最大次数のみに依存しており，ここでも，乱択の威力が発揮される結果となった．

解析においては，各負荷に対し割り当てた区間の独立性を活用し，頂点回りにおいて次数に依

存しない誤差上界を導出することに成功し，更に，遷移確率行列の 2 乗和に対する新たな上界を導出することで上述の上界を示すことに成功した．既存研究では遷移確率行列を正則グラフ上の単純ランダムウォークのものに固定し，その正則性を巧妙に利用することで 2 乗和の上界を導出していた．それに対し，本研究では 2 乗和と遷移確率行列の Dirichlet 形式が密接に関連することを発見した．これにより，遷移確率行列の固有値が非負ならば（lazy）遷移確率行列の 2 乗和が定数に収束することを示すことが出来る．固有値の非負性はマルコフ連鎖の収束解析で頻繁に表れる仮定であるが，既存の拡散ベースの負荷分散アルゴリズムへは適用可能かどうか不明瞭である．一方で，任意の確率行列に適用できる提案アルゴリズムへは容易に仮定でき，これらにより優れた上界の導出に成功した．
上述の研究成果は[4]にまとめられている．

[成果 2] Dispersion process における相転移現象
拡散ベースの負荷分散アルゴリズムを考える際，拡散を繰り返したのち，いかに理想の配分との誤差を小さくするアルゴリズムを設計できるか，という点が問題となる．これに対し，本研究では理想の配分を達成するアルゴリズムに必要な拡散の回数を解析することを目指し，dispersion process と呼ぶ確率過程の収束時間の解析を行った．Dispersion process とは複数トークンによるランダムウォークに似た確率過程である．与えられた有限グラフ上において，各時刻にグラフ上の各トークンは，自身のいる頂点に他のトークンが存在している場合にのみ近傍へ確率的な遷移を行う．即ち，各頂点においてトークン数が 1 以下になったときにプロセスは停止する．Dispersion process はある種の拡散に制限がかかった粒子の挙動を記述するモデルであり，拡散律速凝集（Diffusion Limited Aggregation, DLA)と呼ばれる過程とも関連が深い．
このモデルに対し，本研究ではまず完全グラフ上において停止時間の解析を行った．そして，トークン数が頂点数の半数未満の場合，頂点数の対数時間で停止が起こり，逆にトークン数が頂点数の半数よりも多い場合，頂点数の指数時間かかっても停止が起こらないという相転移現象が発生することを示した．さらに，しきい値は自己遷移確率に依存して決まり，例えば 1/2 の確率で自己遷移が起こるモデルにおいては，頂点数 n に対し，トークン数が(3/4)n より小さければ対数時間で停止し，(3/4)n より大きければ指数時間かかっても停止しないことも示した．この成果は有限グラフ上における誤差 1 以下の拡散モデルは負荷数と頂点数に依存性があることを示唆している．
完全グラフ以外の構造に対しても，（トークン数に対して十分大きな頂点数を持つ）パス，木，超立方体において，停止状態におけるトークン配置の特徴づけを行った．
証明は主に，各時刻における各頂点のトークン数の期待値が，総トークン数が頂点数の半数付近で大きく振る舞いを変えること，そしてトークンの遷移は期待値回りでに強く収束することを確率不等式で示すことから成る．
本成果は雑誌論文(1)にまとめられている．

[成果 3] 投票モデルの合意時間解析
負荷分散研究における遷移確率行列に対する解析技法を用い，投票モデル（意見拡散モデル）と呼ばれる確率過程に対し，その収束時間の解析を行った．各頂点が``意見''を保持しているグラフを考える．投票モデルでは，各頂点は自身の近傍をランダムに選択し，自身の意見を選択された頂点の意見へ更新する．このモデルに対し，既存研究では最終的にグラフ全体が 1 つの意見でまとまるまでにかかる時間・収束時間の解析が積極的に行われている．近年，各頂点が自身の近傍を 3 つランダムに選び，その中で多数派の意見に自身の意見を更新する，という 3-Majority モデルが提案され，研究が行われてきたが，その研究は完全グラフ上に留まり，一般のグラフ構造上で分かっていることはなかった．本研究では一般のグラフ上の 3-Majority モデルの解析を遷移確率行列のコンダクタンスと呼ばれる特徴量に結びつけ解析を行い，既存研究の完全グラフ上での解析を拡張することに成功した．
本成果は学会発表(1)において発表を行った．

参考文献

1. P. Berenbrink, C. Cooper, T. Friedetzky, T. Friedrich, and T. Sauerwald, "Randomized diffusion for indivisible loads," Journal of Computer and System Sciences, 159-185(2015).
2. P. Berenbrink, R. Klasing, A. Kosowski, F. Mallmann-Trenn, and P. Uznanski, "Improved analysis of deterministic load balancing schemes," Proc. PODC 2015, 301-310.
3. T. Sauerwald and H. Sun, "Tight bounds for randomized load balancing on arbitrary network topologies," Proc. FOCS 2012, 341-350.
4. T. Shiraga, "Discrepancy analysis of a new randomized diffusion algorithm," arXiv:1802.06532, 2018.

５．主な発表論文等

〔雑誌論文〕（計　２件）

1. Colin Cooper, Andrew McDowell, Tomasz Radzik, Nicolas Rivera, Takeharu Shiraga, "Dispersion processes," Random Structures and Algorithms, 53(4), 561-585 (2018). DOI: 10.1002/rsa.20822
2. Takeharu Shiraga, Yukiko Yamauchi, Shuji Kijima, Masafumi Yamashita, "Deterministic random walks for rapidly mixing chains," SIAM Journal on Discrete Mathematics, 32(3), 2180–2193. (2018)　DOI: 10.1137/16M1087667

〔学会発表〕（計　２件）

1. "Fast plurality consensus in regular expanders," Colin Cooper, Tomasz Radzik, Nicolas Rivera, Takeharu Shiraga, the 31st International Symposium on Distributed Computing (DISC 2017), Vienna, Austria, Oct.17-19, 2017 (Oct 19).
2. "Analyses of the cover time of deterministic random walks," Takeharu Shiraga, the 21st Conference of the International Federation of Operational Research Societies (IFORS 2017), Quebec, Canada, July 17-21, 2017 (July 18).

〔図書〕（計　０件）

〔産業財産権〕

○出願状況（計　０件）

名称：
発明者：
権利者：
種類：
番号：
出願年：
国内外の別：

○取得状況（計　０件）

名称：
発明者：
権利者：
種類：
番号：
取得年：
国内外の別：

〔その他〕

ホームページ等
https://sites.google.com/view/takeharu-shiraga/

６．研究組織

(1)研究分担者

研究分担者氏名：

ローマ字氏名：

所属研究機関名：

部局名：

職名：

研究者番号（８桁）：

(2)研究協力者

研究協力者氏名：

ローマ字氏名：