

遺伝的アルゴリズムを用いた帯域幅割当における 分散アルゴリズムの設計 — リンク障害への対応 —

小林英博¹, 棟朝雅晴², 赤間清³, 佐藤義治⁴

1,4:北海道大学大学院 工学研究科 数理情報工学講座 情報解析学分野

2,3:北海道大学 情報メディア教育研究総合センター 情報メディアシステム分野

ネットワーク資源は有限であり、特に高速で帯域幅の大きなリンクは高価である。よって帯域幅を効率よく割り当てることによって得られる利益は大きく、有限の資源を無駄なく使用するためには帯域幅割当アルゴリズムが重要となる。これまで遺伝的アルゴリズムを用いたネットワーク帯域幅割当のためのアルゴリズムとして GRA(Genetic Routing Algorithms)が提案されているが、これは集中型のアルゴリズムでありネットワーク障害が発生した場合には割り当てが不可能となる。そこで筆者らは GRA を分散化した D-GRA(Distributed GRA)を提案したが、D-GRA は単なる分散化に留まっており、障害回復処理やリンク障害に対して不十分な点が残されている。そこで本研究では、D-GRA を改善し、リンク障害の回復に対応したアルゴリズムを提案する。

Designing a Distributed Algorithm Using Genetic Algorithms for Bandwidth Allocation — Recovery Mechanism from Link Failure —

Hidehiro Kobayashi¹, Masaharu Munetomo², Kiyoshi Akama³ and Yoshiharu Sato⁴

1,4:Information and Data Analysis, Graduate School of Engineering,

Hokkaido University

2,3:Center for Information and Multimedia Studies, Hokkaido University

Network resources are limited in their capacity, moreover fast high capacity communication links are expensive. Thus bandwidth allocation and efficient usage of limited resources are important. Previously, GRA(Genetic Routing Algorithm) has been proposed as a bandwidth allocation with genetic algorithm. This algorithm is centralized algorithms and execute localizing optimizing, in case of trouble it is impossible to allocate bandwidth. Therefore we suggested distributed-GRA(D-GRA) previously, however this is imperfect about communication links failure and recovering from trouble. Then, in this paper, we will improve D-GRA to recover from communication link failure.

1 はじめに

現在ではネットワークを利用した新たなサービスとして、ビデオ・オン・デマンドや音楽の配信など、一定時間、一定の帯域幅を確保する必要のあるサービスが求められている。帯域幅確保のための枠組が提供されていない既存のネットワーク

環境において帯域幅を確保するためには、リンクの帯域幅を大きくするかあるいは帯域幅を効率良く割り当てる帯域幅割当のアルゴリズムが必要とされる。本研究ではその二つの中から帯域幅割当アルゴリズムに着目した。なぜならネットワーク資源は有限であり、特に高速で帯域幅の大きなリ

リンクは高価であるため、割当の効率化により得られる利益が大きいためである。

帯域幅の割当を最適化するアルゴリズムとして、GRA(Genetic Routing Algorithm)が提案されている [3]。GRA とは、一つのノードが全ノードの通信要求を把握し、グローバルに帯域幅割当を最適化するアルゴリズムである。GRA の問題点は割り当てに用いる情報が特定のノードに集中してしまうため、ネットワークに障害が発生した場合には割り当てを実行できないことである。そこで本研究では GRA を分散化した D-GRA を提案する。分散化することでネットワーク障害が発生した場合でもシステムが破綻することなく帯域幅の割り当てが可能となるアルゴリズムを設計する。

具体的には、割り当てを各ノードごとに独立して行なうことでネットワーク障害への対応を可能としている。しかし、各ノードごとに割り当てを行う場合、ネットワーク全体を考えると割り当てが効率的でないことが考えられる。そこで本研究では全体としての効率も高めるため、各ノードが求めた割当に対して平均遅延による適応度を定める。この適応度を他ノードの適応度と交換し評価して適応度が最大のノードを一つ決定する。続いて適応度最大ノードを除いた中から 2 番目に適応度最大のノードを一つ決定する。この過程を全ノードの適応度が決定するまで繰り返すことでネットワーク全体の適応度を高めていくことになる。

2 GRA

Theeb らにより帯域幅割当を実現するためのアルゴリズム、GRA(Genetic Routing Algorithm)が提案されている [3]。GRA は起点ノード (S) から目的ノード (D) への全通信要求に対し、それぞれの S-D ペアのルーティングテーブルから要求される帯域幅を割り当てる経路を決定するアルゴリズムである。このアルゴリズムではあるノードが各 S-D ペアの通信要求とルーティングテーブル情報を管理し、通信を割り当てる経路をそのノードが一括して計算を実行している。

また経路を求める際にはネットワーク全体の平均パケット遅延が最小となるように、要求帯域幅と実際に割り当てる経路との組合せを GA を使っ

て探索する。ここで平均パケット遅延 T は待ち行列理論により、

$$T = 1/\lambda \sum_{m=1}^M \frac{f_m}{c_m - f_m} \quad (1)$$

という式で求めることができる。(λは平均パケット到着率。Mは総リンク数。c_mは m 番目のリンクの容量。f_mは m 番目のリンクを流れる通信量。)

例えばノード A からノード B へ 50Mbps、ノード C からノード D へ 100Mbps の帯域幅を必要とする通信要求が発生したとする。A-B と C-D のルーティングテーブルは表 1 である。

S-D ペア	A-B	S-D ペア	C-D
Route No.	Link No.	Route No.	Link No.
0	link0	0	link 2
1	link1,link2	1	link3
⋮	⋮	⋮	⋮
j	link3,link4	k	link4,link5

表 1: A-B と C-D のルーティングテーブル

まず A-B より No.j が、C-D より No.k が選択されたとする。これは link3 と link4 をそれぞれ 50Mbps 使用し、link4 と link5 をそれぞれ 100Mbps 使用することになるので、(1) 式により平均遅延は次の式で求めることができる。

$$T = 1/\lambda \left(\frac{50}{c_3 - 50} + \frac{150}{c_4 - 150} + \frac{100}{c_5 - 100} \right)$$

実際にはこの作業を全 S-D ペアに対して行なうことになる。全 S-D ペアに対して選択した経路の一例が表 2 である。

S-D ペア	0-1	0-2	⋯	a-b	⋯
Bandwidth	20	10	⋯	40	⋯
Route No.	2	0	⋯	k	⋯
Encode	010	000	⋯	***	⋯

表 2: 経路の選択とエンコード

全ノード数が P であれば S-D の組合せは $P \times (P-1)$ 通り存在し、各 S-D ペアのルーティングテーブルサイズが Q であるならば、帯域幅を割り当てる経路の組合せは全部で $P \times (P-1) \times Q$ 通りとなる。P の数により組合せが膨大となるので、ここに GA を適用して効率のよい探索を実現する。GA

を適用するためには、選択した Route No. をビット列へとエンコードする。仮にルーティングテーブルのサイズが 8 だとすれば 3 ビットで表現できるので、総ビット数は $P \times (P - 1) \times Q \times 3$ ビットとなる。続いてこのビット列を GA における個体とし、

$$\min T(010\ 000 \dots \dots \dots)$$

を満たす個体を遺伝的操作により求めていく。具体的には候補となる個体を 50 個発生させ適応度 (ここでは平均遅延 T) に基づき選択操作を行なう。選択操作はランキング方式を採用している。この方式は個体を適応度の高い順に並べ、上位の個体を重複を許して選択するというものである。GRA では上位 80% の個体の中から次世代の個体として 50 個の個体を選択している。続いて選択された個体群に対して交叉を行なう。この操作はまず二つの個体を選択し、遺伝子のビット列の中から交叉点を一点決定した後、二つの個体の交叉点から後ろのビット列を交換し新たな二つの個体を発生させる一点交叉である。交叉の回数は交叉パラメータにより制御する。最後に突然変異の操作を行なう。この操作は解が極値解に収束するのを防ぐために行なわれる。ある個体の中から任意に一点を選び、その点のビットを反転させる突然変異である。突然変異の回数も交叉同様にパラメータにより制御する。これら 3 つの遺伝的操作を繰り返すことで最終的に最適な解を求めることが可能となる。

この GRA における最適化は、最適化のあるノードが集中して行なうことでグローバルな帯域幅割当を実現している。ところがグローバルな割当では、ネットワークに障害が発生した場合に割当が不可能となることが考えられる。しかし現実のネットワークでは、その規模が大きくなるほどに障害が発生する可能性が高くなるため、一つのノードに計算を集中させることは望ましいことではない。よって現実へのネットワークへの実装を考えるならば、ネットワーク障害に対応するためのアルゴリズムの分散化が必要となる。

3 D-GRA3

これまでに筆者らは GRA を分散化した D-GRA(Distributed GRA) を提案した [4]。D-GRA

は GRA を単純に分散化したアルゴリズムであり、通信要求は各発信元ノードにおいて帯域幅を割り当てる経路が求められる。各ノードが求めた経路に対しては平均遅延に基づいた適応度を定め、適応度を他ノードと互いに交換し評価することでネットワーク全体の最適化を図っている。しかし D-GRA はノードやリンクが正常に機能している場合での分散化を行なったにすぎず、ネットワーク障害へは未対応であった。現在 D-GRA に改良を加えノード障害に対応したアルゴリズムを D-GRA2 として研究中であるが [6]、本研究ではリンク障害へ対応したアルゴリズムを提案しており、これを D-GRA3 と呼ぶことにする。

3.1 D-GRA3 のアルゴリズム

D-GRA3 の流れは図 2 となる。D-GRA2 では、新たに通信要求が発生した場合、または通信が終了した場合に全ノードでリアロケーションを行っていた。しかし、比較的小さい帯域幅を割り当てるために全ノードがリアロケーションするのではかえって非効率的である。また、使用している経路が頻繁に変更されることは望ましいことではない。D-GRA3 ではそれを改善し、新たな通信が発生した場合に全体をリアロケーションすることはせず、通信の発生したノードにおいてのみ図 3 に示すアルゴリズムにより割り当て計算を実行する。

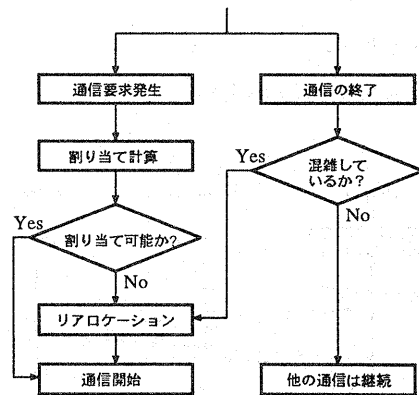


図 2: アルゴリズムの流れ

しかしネットワークが混雑している場合には割り当てる経路が存在しない可能性がある。よって割り当てる経路が見つからなかった場合にはネット

ワークが混雑しているものと考え、全ノードでのリアロケーションを試みる。このリアロケーション・アルゴリズムは図4に示す。

現在の通信が終了した場合はその通信に使用していたリンクの帯域幅が空くだけなので、通信終了の場合のリアロケーションを実行する必要はない。だが場合によってはリアロケーションしたほうが遅延が小さくなることも考えられる。ところがリアロケーションには少なからずコストがかかるためその判断が難しい。そこで本研究では、ネットワークの混雑度が大きい時にはリアロケーションを行ない、小さい時には他の通信はそのまま継続とする。混雑度の判断には平均リンク使用率を利用する。平均遅延はリンクの使用率が50%を超えたあたりから急激に増大するので[6]、平均リンク使用率が50%より大きい場合を混雑している状態と判断する。

3.2 割当アルゴリズムの詳細

D-GRA3で用いるメッセージは以下の5種類である。これらのメッセージはIDで区別され、それぞれ発信元ノードIDとメッセージ内容との組になっている。

- Msg1:リアロケーション要求
- Msg2:割り当て情報の問い合わせ
- Msg3:最新の割り当て情報
- Msg4:適応度
- Msg5:終了メッセージ

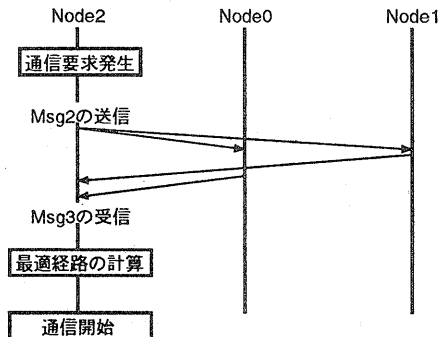


図3: D-GRA3のアルゴリズム

図3が割り当てアルゴリズムの詳細である。ノード2で通信要求が発生した場合の動作を示している。ノード2で通信要求が発生すると、他のノードにMsg2をブロードキャストする。続いてMsg3を受信し、この情報に基づいて最も平均遅延が小さくなるような経路を決定する。ここにGRAと同様の遺伝的操作を適用している。

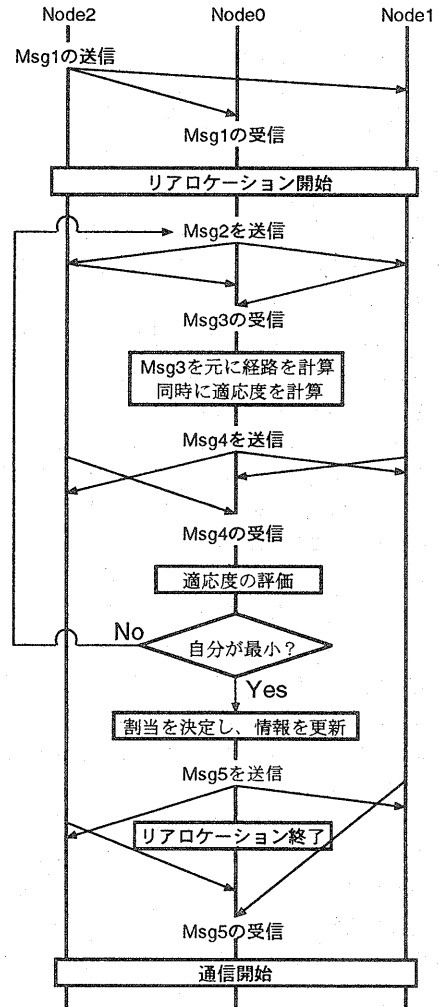


図4: リアロケーション・アルゴリズム

図4がリアロケーションアルゴリズムの詳細である。ノード2において図3に示すアルゴリズムで割り当てる経路が見つからなかった場合、リアロケーション要求がブロードキャストされる。図4はノード0でのリアロケーションの動作を示し

ているが、ノード 2 やノード 1 でも同様の動作が行なわれる。

まずノード 2 よりリアロケーション要求 (Msg1) がブロードキャストされる。Msg1 を受け取ることにより、全ノードが一斉にリアロケーションを実行する。リアロケーションは初めに他ノードに現在の通信状況、すなわちどのリンクを何 Mbps 使用しているかを問い合わせる (Msg2)。続いて割り当て状況の情報が返されてくる (Msg3)。他ノードから受け取った割り当て情報を元に通信を割り当てる経路を GA を適用して求める。求めた経路のネットワーク全体での適性を評価するため、求めた経路の適応度 (Msg4) をブロードキャストする。同時に他ノードから Msg4 を受信する。受信した Msg4 を評価し、適応度が最小のノードが自分であれば、割り当て情報をここで求めた経路に更新する。同時に終了メッセージ (Msg5) をブロードキャストする。適応度最小のノードが自分でなければ、Msg2 を送信するところからもう一度やり直す。これを全ノードの Msg5 を受信するまで繰り返す。実際の通信は全ノードの Msg5 が揃ったところで開始する。

D-GRA3 の目的は、リンクに障害が発生した場合にアルゴリズムが破綻する事態を回避し、確実に割り当てを実行することである。その結果として平均遅延が増加することも考えられるが、まずはシステムの回復を優先と考える。

3.3 リンク障害への対応

割り当て前からリンク障害が発生していた場合、各ノードにおけるルーティングテーブルの更新により障害の発生したリンクを避けることができる。各ノードでは D-GRA3 とは別にルーティングテーブルを更新するためのパケットを一定時間ごとに交換し、常に最新のルーティングテーブルを保持しているからである。しかし割り当ての途中でリンクに障害が発生すると、割り当てのために必要なメッセージの到着が遅れるか、あるいは全く到着しないことになる。つまりリンク障害に対応するためには、各メッセージの待ち時間を設定することが重要となる。待ち時間が短か過ぎるとは割り当ての最適性が低くなるが、長過ぎるとは割り当て計算のためのオーバーヘッドが大きくなるからで

ある。さらにアルゴリズムが破綻しないためには一定時間経過してもメッセージが届かない場合での動作を詳細に考慮することが大切になる。すなわちリンク障害へ対応するために重要なことは、メッセージを待つ時間を定めることと一定時間内にメッセージが届かない場合にどう対応するか、を定めることである。以下でそれぞれのメッセージについて考察する。

Msg1

リアロケーションを要求するメッセージ。リアロケーションが必要と判断されたノードからブロードキャストされる。リアロケーションはある程度同期的に行なわれたほうがより適切な割り当てが行なえるので、Msg1 の到着が遅れることを考慮し Msg1 を送信したノードは Msg1 を送信してから一定時間経過してから計算を実行する。この待ち時間はネットワークの平均リンク容量と平均通信要求量とから (1) 式により求めた値とする。Msg1 を受信したノードは受信と同時に計算を実行する。またリンク障害により Msg1 が届かなかった場合、そのノードではリアロケーションを行なうことはできない。だが他ノードがリアロケーションを行なうと Msg2 が送られてくることがあり得る。その場合には Msg3 を返す。

Msg2

割り当て状況を問い合わせるためのメッセージ。このメッセージを送信すると同時に Msg3 の受信待ちになる。メッセージが往復することになるので Msg1 の 2 倍の時間待つことにする。その間に Msg3 が送られてこない場合には再び Msg2 を送信する。しかしオーバーヘッドの増加を防ぐため Msg2 の再送は数回程度とする。またリンク障害により Msg1 が届かずリアロケーションを行なっていないとしても、Msg2 が届いた時には Msg3 を返す。

Msg3

ノードにおける最新の割り当て情報。すなわち、どのリンクを何 Mbps 使っているかの情報である。リアロケーションはこの情報を元に行なわれるため、割り当てには最も重要なメッセージとなる。

よってこの情報が届かない場合には Msg2 を数回送信し直す。しかし、それでも送られてこない場合には各ノードが持っているそのノードに関する古い情報に基づいて割り当てを計算する。これは最適な割り当てを求めることは不可能になるが、障害からの回復を優先とするためである。

Msg4

各ノードにおいて計算された適応度。(1) 式に基づいて計算した平均遅延を適応度として用いる。ネットワーク全体での最適性を評価するために重要なメッセージである。Msg4 は適応度が求まると同時にブロードキャストされ、Msg4 リクエストは行なわない。リンクの状態によって Msg4 に遅延が発生することが考えられる。さらにノードの性能によっては適応度を求める計算に時間がかかることも考えられるので、Msg4 の待ち時間はリンクによる遅延に幅を持たせ Msg1 の待ち時間の 2 倍の時間、待つことにする。

Msg5

リアロケーションが終了すると他ノードに送信するための割り当て情報 (Msg3) を更新し、同時に終了したことを知らせるため Msg5 をブロードキャストする。このメッセージを送信した段階ではまだ新たな通信は開始されていない。全ノードからの Msg5 を受け取った時点で更新された割り当て情報にしたがって新たな通信を開始するものとする。リアロケーションが終了するまで最も時間のかかるノードはリアロケーション・アルゴリズムをノードの数だけ繰り返すことになるので、リアロケーション・アルゴリズム一回あたりにかかる時間の平均にノード数をかけただけの時間待つことにする。

4 おわりに

本研究ではリンク障害に対応した帯域幅割当アルゴリズム D-GRA3 を提案した。D-GRA3 の特徴は割り当て計算の実行中にリンク障害によりメッセージの到着が遅れた場合、またはメッセージが完全に届かない場合でも割り当てが可能なことである。今後の課題は、D-GRA3 を実装したシミュ

レーション実験を行なうことで、その障害回復機能を確認することである。

さらに、実際にはノード障害・リンク障害が同時に起こることも考えられるので、今後は D-GRA2 と D-GRA3 を組合せより障害に強いアルゴリズムへと発展させ、シミュレーション実験によりその有効性を検証し、現実のネットワークへの実装を目指していく。

参考文献

- [1] Mario Gerla, Jose Augusto Suruagy Monteiro, Rodolfo Pazos: "Topology Design and Bandwidth Allocation in ATM Nets", IEEE Journal on selected areas in communications, Vol.7, No.8, pp.1253-1262 (1989)
- [2] Hong Pan, Irving Y. Wang: "The Bandwidth Allocation of ATM Through Genetic Algorithm", IEEE Global Telecommunication Conference, pp.0125-0129 (1991)
- [3] Theeb A. Al-Qahtani, Mohammed J. Abedin, Syed I. Ahson: "Dynamic Routing in Homogeneous ATM Networks Using Genetic Algorithms", IEEE International Conference on Evolutionary Computation proceedings : IEEE World Congress on Computational Intelligence, pp.114-119 (1998)
- [4] 小林英博, 棟朝雅晴, 佐藤義治: "遺伝的アルゴリズムを用いたネットワーク帯域幅割当", 情報処理学会研究報告, Vol.99, No.94, pp.91-96 (1999)
- [5] 小林英博, 棟朝雅晴, 赤間清, 佐藤義治: "ネットワーク帯域幅割当のための分散アルゴリズム", 情報処理北海道シンポジウム 2000, pp.110-111 (2000)
- [6] 小林英博, 棟朝雅晴, 赤間清, 佐藤義治: "遺伝的アルゴリズムを用いた帯域幅割当における分散アルゴリズムの設計—ノード障害への対応—", 情報処理学会「マルチメディア通信と分散処理ワークショップ」, 発表予定 (2000)