レベル構造に基づいたバンド幅縮小アルゴリズムが苦手とするグラフクラス

梅澤 香織　大舘 陽太　山崎 浩一

群馬大学 工学部 情報工学科
〒376-8515 群馬県 桐生市 天神町 1-5-1

概要　Cuthill-McKee 法に代表されるレベル構造に基づくバンド幅縮小アルゴリズムでは，精度良く縮小できないグラフ (データ) を紹介する．

Hard graphs for level algorithms on the bandwidth minimization problem

Kaori Umezawa　Yota Otachi　Koichi Yamazaki

Department of Computer Science Gunma University
1-5-1 Tenjin-cho, Kiryu zip:376-8515, Gunma, Japan

abstract　In this paper, we introduce a graph class and demonstrate that level algorithms for the bandwidth problem do not perform well on the graph class.

1　はじめに

バンド幅縮小問題は様々な分野に応用を持つため古くからよく研究されている (例えば, [5, 15] 参照). バンド幅縮小問題には二つのとらえ方があり, 一つは行列の問題として, もう一つはグラフの問題としてとらえる方法である. 行列の問題としてとらえた場合, 逆行列や行列式の計算の効率化などの応用が挙げられ, グラフの問題としてとらえた場合, データ伝送や VLSI 設計などへの応用が挙げられる. バンド幅縮小問題に対し, 最適解または精度の良い近似解を短時間で求めることができれば, 上記に挙げた応用分野への貢献に繋がる. 本論文では, バンド幅縮小問題をグラフの問題としてとらえ説明を行う.

理論的な立場から見ると, バンド幅縮小問題は, 計算量的に困難な問題であることが知られている. 例えば, 入力グラフを最大次数が 3 である木に制限したとしても, 対応する判定問題は

NP 完全であることが知られている[1][12]. また, 入力を木に制限しても, 近似率[2]は 4/3 より良くならない [3]. さらに, 入力に何も制限を加えなければ ($P \neq NP$ の仮定の下で) 定数近似は不可能であることが知られている [20]. また, 与えられた正整数 k に対しバンド幅が k 以下であるかの判定問題は実質 brute force で行う方法しかないことが予想されている [4]. 肯定的な結果としては, [7] の結果が知られている.

一方応用的な立場から見ると, 入力グラフによっては比較的短時間で実用上許容できる精度を持った解の出力が期待できる, 様々な発見的手法が考案されている. それらの多くは, 先ず**レベル構造**と呼ばれる頂点集合の列を求め, 次にそのレベル構造を基に解を構成する, という方法を取っている. そのようなアルゴリズムを, 本論文では, **レベルアルゴリズム**と呼ぶ. 代表的なレベルアルゴリズムとして, Cuthill-McKee(CM) 法 [6] や Gibbs Poole Stockmayer(GPS) 法 [11] が挙げられる. 実験結果や確率的解析からも, ランダムグラフに対しては, レベルアルゴリズムがうまく働くことが知られている [19, 10]. またレベルアルゴリズム以外にも, 遺伝的アルゴリズム, タブーサーチ, 焼きなまし法などのメタヒューリスティクの有効性が報告されている [13, 14, 17, 18].

バンド幅縮小問題に対する研究は今も盛んに行われており, 90 年代後半からに関して見れば, Volume Respecting Embedding 法 (VRE) 法 [9] や半正定値計画法を用いた近似アルゴリズム [1] やレベルアルゴリズムに分類される Wonder Bandwidth Reduction Algorithm (WBRA) 法 [8] などが提案されている.

上述の通り, CM 法や GPS 法に代表されるレベルアルゴリズムは, ランダムグラフに対してうまく働くことが知られいる. では逆に, どのようなグラフに対してうまく働かないのであろうか? 現在知られているレベルアルゴリズムが苦手とするグラフクラスを考察することは, レベルアルゴリズムの改善やその限界を知る上において重要である.

本論文では, **Stretched Complete Binary Tree** (略して $SCBT$) と呼ぶグラフクラスを紹介し, そのグラフクラスに対しては, レベルアルゴリズムで従来知られているものでは, 精度の良い解が出せないことを実験的に示す. すなわち, $SCBT$ のある上界を理論的に求め, 一方で, $SCBT$ に GPS 法を適用した実験結果を示し, これら理論値と実験値の比較を行う.

2 準備

本章では, バンド幅とレベル構造に関する諸定義を行う.

G をあるグラフとする. V から $\{1, \ldots, |V|\}$ への全単射の集合を $LO(G)$ で表す. グラフ G のバンド幅 $bw(G)$ は,

$$bw(G) = \min_{f \in LO(G)} \max_{\{u,v\} \in E} |f(u) - f(v)| \tag{1}$$

で定義される.

グラフ $G = (V, E)$ の頂点の部分集合の列 $(L_1, \ldots, L_M)$ が**レベル構造**を成すとは,

1. $\sum_{i=1}^{M} L_i = V$,

2. $\{u, v\} \in E, u \in L_i, v \in L_j \Rightarrow |i - j| \leq 1$

[1]良く知られている問題で, 木に制限しても NP 完全であるものはあまり無い (他には例えば subgraph isomorphism や achromatic number など)

[2]近似率は常に 1 以上で, 1 に近付く程精度が良い

を満たすときをいう. $\max_{1\leq i\leq M}|L_i|$ をレベル構造 $(L_1,\ldots,L_M)$ の幅と呼ぶ. グラフ G のレベル構造での最小の幅を $lw(G)$ で表すと, $lw(G)\leq bw(G)\leq 2lw(G)-1$ という関係が成り立つ. 従って, もし最小の幅を持つレベル構造を見つけることができれば, バンド幅の2倍近似が可能となる. しかしながら, 最小の幅を持つレベル構造を見つけることはNP困難であることが知られている [2]. 従って, 最小の幅は保証しないが, 比較的簡単にレベル構造が実現できる, 距離に基づいてレベルを構成する方法がよく用いられる.

R_0 を頂点のある部分集合とし, R_i を R_0 から距離 i にある頂点の集合とする. このとき $R_0,\ldots,R_h$ (h は R_0 から最も遠い頂点までの距離) はレベル構造を成す (速度向上の理由から R_0 のサイズは小さい場合が多い). 距離に基づいてレベルを構成する方法は, 比較的高速で実装も容易ではあるが, (理論的には) それにより得られるレベル構造の幅が最小の幅 $lw(G)$ よりかなり大きくなることが起こり得る [21]. また距離に基づいて構成されたレベルの中での最小幅を求める問題は, たとえグラフを木に制限したとしても, 4/3 未満に近似率を下げられないことが知られている [22].

3 Stretched complete binary tree

本章では Stretched complete binary tree を紹介し, そのバンド幅の上界を示す.

G をあるグラフとし, u, v を G の頂点とし, S を頂点集合の部分集合とする. G 内の u と v の距離を $d_G(u,v)$ で表す. また, u と S の距離 $\min_{v\in S} d_G(u,v)$ を $d_G(u,S)$ で表す. G が文脈から明らかなときは, 添字の G を省略する.

T をある2分木とし, r を T の根とする. T のある頂点 v が**深さ** i を持つとは, $d(r,v)=i$ のときをいう. 従って, 根の深さは0となる. T のある辺 $e=\{u,v\}$ が**深さ** i を持つとは, $d(r,\{u,v\})=i-1$ のときをいう. 頂点 v および辺 e の深さをそれぞれ $depth(v)$ および $depth(e)$ で表す.

辺にいくつかの頂点を挿入する操作は**細分**と呼ばれるが, 本論文では k 個の頂点を挿入した細分のことを k **細分**と呼ぶことにする. すなわち, ある辺に対し k 細分を行うことは, その辺を長さ $k+1$ のパスに置き換えることを意味する.

深さ d の完全2分木を B_d で表す. r を B_d の根とする. ある2分木 T が**深さ d の Stretched complete binary tree** であるとは, T が, B_d の各辺 e に対し $2^{d-depth(e)}$ 細分を行うことにより得られたグラフと同型であるときをいう. 深さ d の Stretched Complete Binary Tree を $SCBT_d$ で表す (図1参照).

簡単な考察により, 以下の定理を得ることが出来る:

定理 3.1 $bw(SCBT_d)\leq d$.

一方, $tbw(B_d)=\lceil d/2\rceil$ が知られていることより, $\lceil d/2\rceil\leq bw(SCBT_d)$ であることがわかる, ここで $tbw(G)$ は, グラフ G の Topological Bandwidth を表す (詳しくは, [16] を参照) [3].

4 実験結果

本研究では, 深さ5から10までの stretched complete binary tree, 深さ7から13までの complete binary tree, および文献 [8] で扱われている matrix market の Harwell-Boeing Collection

[3] B_d の pathwidth が $\lceil d/2\rceil$ であることからでもわかる

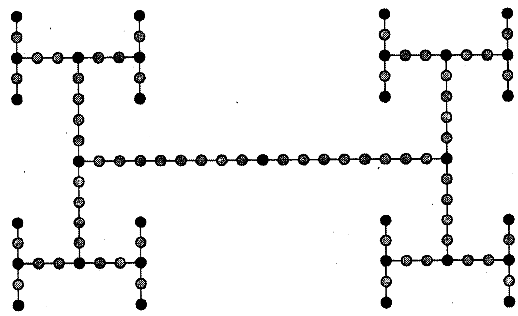

図 1: $SCBT_d$

のデータセット DWT [4]からのデータ 14 個を実験データとし，計算機実験を行った．実験で比較したアルゴリズムは，GPS 法 [11]，GPS 法で得られた解に文献 [13] で私用されている山登り法を適用したもの，VRE 法 [9]，VRE 法で得られた解に同じ山登り法を適用したものの 4 種類である．

各表において，opt は最適解，u.b. はバンド幅の上限，GPS は GPS 法，VRE は VRE 法，HC は山登り法，WBRA は WBRA 法，MatLab は数値計算ソフト MatLab をそれぞれ意味する．

4.1 Stretched complete binary tree に対する結果

表 1 に実験結果を示す．

depth	size	u.b.	GPS	GPS+HC	VRE	VRE+HC
5	223	5	9	9	13.1	10.8
6	511	6	17	17	20.0	17.2
7	1151	7	33	33	27.5	23.3
8	2559	8	65	65	40.3	35.3
9	5631	9	129	129	60.6	52.1
10	12287	10	257	257	82.0	72.5

表 1: Stretched complete binary tree に対する結果

値として小数値が現れているが，これは以下の操作を 10 回行った結果の平均値を記載しているためである．ここでその操作とは，同じ入力グラフ (行列) に対し，入力行列をランダムに置換し，その置換された行列に対しそれぞれの方法を適用するという操作である．GPS 法や GPS+HC 法では，方法の特性と $SCBT$ の対称性より 10 回ともすべて同じ値を得た．

表 1 から分かる通り，頂点数約 1 万 2 千の深さ 10 では GPS+HC 法による出力解の精度は，最適解の上限の 25 倍以上となり極めて悪いと言える．一方，レベルアルゴリズムとは異なる VRE+HC 法では，深さ 10 でも最適解の上限の 8 倍以下に抑えられている．興味深いことに，GPS 法で得られた結果に山登り法をかけても値が変わらなかった．

[4]http://math.nist.gov/MatrixMarket/data/Harwell-Boeing/dwt/dwt.html から入手可能

4.2 Complete binary tree に対する結果

表 2 に実験結果を示す.

depth	size	opt	GPS	GPS+HC	VRE	VRE+HC
7	255	19	34	30	44.2	37.4
8	511	32	66	57	85.2	73.5
9	1023	57	129	98	162.3	141.8
10	2047	103	257	194	305.7	268.7
11	4095	187	513	386	599.9	538.5
12	8191	342	1025	770	1201.2	1109.7
13	16383	631	2049	1538	2280.0	2133.0

表 2: Complete binary tree に対する結果

表 2 から分かる通り, 頂点数約 1 万 6 千の深さ 13 での GPS+HC 法による出力解の精度は, 最適解の 2.5 倍程度で抑えられている. 一方 VRE+HC 法では, 深さ 13 では 3.5 倍程度に抑えられている. *SCBT* と同様にここでも, GPS 法や GPS+HC 法では, 方法の特性と完全 2 分木の対称性より 10 回ともすべて同じ値を得た.

4.3 Harwell-Boeing Collection のデータセット DWT に対する結果

表 3 に実験結果を示す. 表 3 中の WBRA および MatLab の実験値は, 文献 [8] からの引用である.

size	GPS	GPS+HC	VRE	VRE+HC	WBRA	MatLab
59	8.7	8.1	12.7	9.6	7	8
66	3.0	3.0	6.2	4.8	3	3
72	6.0	6.0	11.2	9.2	7	7
87	16.4	14.1	24.8	16.5	13	18
162	13.9	13.5	26.1	20.0	13	16
193	44.0	42.3	80.3	49.2	35	57
209	37.6	35.2	60.3	37.2	30	33
245	39.8	35.2	64.9	43.3	27	55
307	38.8	38.2	64.3	47.3	32	44
361	14.4	14.4	48.3	35.5	14	15
503	61.7	56.5	114.0	75.5	51	59
592	36.1	33.9	68.2	48.1	39	42
869	40.8	39.5	86.0	59.8	39	43
1005	106.4	93.1	138.5	97.5	68	65

表 3: Harwell-Boeing Collection のデータセット DWT に対する結果

表 3 から分かる通り, GPS+HC 法による出力解の精度は, VRE+HC 法のそれよりは若干良

いと言える. 上記3タイプのデータを見る限り, stretched complete binary tree に対しては, レベルアルゴリズムがうまくは働いていないと考えられる.

謝辞

本研究は, 日本学術振興会科学研究費補助金（基盤研究 (C):課題番号 16500008）の助成を受け行った研究成果の一部である.

参考文献

[1] A.Blum, G.Konjevod, R.Ravi and S.Vempala, Semi-Definite Relaxations for Minimum Bandwidth and other Vertex-Ordering problems, Proc. of 30th STOC, (1998), 100-105. (Journal version) Theor. Comput. Sci. 235(1), (2000), 25-42.

[2] H.L.Bodlaender, The complexity of finding uniform emulations on paths and ring networks, Information and Computation 86 (1990) 87-106.

[3] G.Blache and M.Karpinski, On Approximation Intractability of the Bandwidth Problem, ECCC TR98-014 (1998).

[4] H.L.Bodlaender, M.R.Fellows, and M.T.Hallett, Beyond NP-completeness for problems of bounded width (extended abstract): hardness for the W hierarchy, Proc. of 26th STOC, (1994), 449-458.

[5] P.Z.Chinn, J.Chvatalova, A.K.Dewdney, and N.E.Gibbs, The bandwidth problem for graphs and matrices—a survey, J. Graph Theory 6 (1982), 223-254.

[6] Cuthill, E., McKee,J.: Reducing the bandwidth of sparse symmetric matrices. In: Proc. ACM Nat.Conf. New York (1969), 157-172.

[7] K.Engel, S.Guttmann, Testing bandwidth k for k-connected graphs, SIAM J. Discrete Math. 16 (2003), 301-312.

[8] A.Esposito,M.S.F.Catalano, F.Malucelli, and L.Tarricone, A new matrix bandwidth reductionalgorithm, Operations Research Letters 23, (1998), 99-107.

[9] U.Feige, Approximating the bandwidth via volume respecting embeddings, (Extended abstract) STOC (1998) 90-99: (Journal version) Journal of Computer and System Sciences 60(3), (2000), 510–539.

[10] U.Feige and R.Krauthgamer, Improved performance guarantees for bandwidth minimization heuristics, manuscript, (1998).

[11] N. E. Gibbs, W. G. Poole, Jr., and P. K. Stockmeyer, An algorithm for reducing the bandwidth and profile of a sparse matrix, SIAM J. Numer. Anal. (1976), 13:236-250.

[12] M.Garey, R.Graham, D.Johnson, and D.Knuth, Complexity Results For Bandwidth Minimization, SIAM J. Appl. Math. 34 (1978), 477-495.

[13] A.Lim, B.Rodrigues, and F.Xiao, A Genetic Algorithm with Hill Climbing for the bandwidth minimization, preprint, 2003.

[14] A.Lim, B.Rodrigues, and F.Xiao, Integrated Genetic Algorithm with Hill Climbing for Bandwidth Minimization Problem", Genetic and Evolutionary Computation Conference 2003, (Chicago, USA).

[15] Y. Lai and K. Williams, A survey of solved problems and applications on bandwidth, edgesum, and profile of graphs, J. Graph Theory 31 (1999), 75-94.

[16] F.S.Makedon, C.H.Papadimitriou, and I.Sudborough, Topological bandwidth, SIAM Journal on Algebraic and Discrete Methods, 6 (1985), 418-444.

[17] R.Martí,, M.Laguna, F.Glover, and V.Campos, Reducing the Bandwidth of a Sparse Matrix with Tabu Search, European Journal of Operational Research 135(2), (2001), 211-220.

[18] E.Piñana, I.Plana, V.Camposy, and R.Martí, GRASP and Path Relinking for the Matrix Bandwidth Minimization, European Journal of Operational Research 153, (2004), 200-210.

[19] J.S.Turner, On the Probable Performance of Heuristics for Bandwidth Minimization, SIAM J. Comput. 15(2) (1986), 561-580.

[20] W.Unger, The Complexity of the Approximation of the Bandwidth Problem, Proc. 39th FOCS (1998), 82–91.

[21] M.Wiegers and B.Monien, Bandwidth and Profile Minimization, Proceedings of the 14th International Workshop on Graph-Theoretic Concepts in Computer Science, (1988) 378-393.

[22] K.Yamazaki, On Approximation Intractability of the Path-Distance-Width Problem, Discrete Applied Mathematics Vol.110 (2001), 317-325.