Technische Universität Berlin

FACHBEREICH 3
MATHEMATIK

Between Min Cut and Graph Bisection

by

Dorothea Wagner[1], Frank Wagner[2]

No. 307/1991

Between Min Cut and Graph Bisection

Dorothea Wagner*[1], Frank Wagner[2]

[1] Fachbereich Mathematik, Technische Universität Berlin, Straße des 17. Juni 136, W-1000 Berlin 12, Germany, wagner@math.tu-berlin.de
[2] Institut für Informatik, Fachbereich Mathematik, Freie Universität Berlin, Arnimallee 2-6, W-1000 Berlin 33, Germany, wagner@math.fu-berlin.de

Abstract. We investigate a class of graph partitioning problems whose two extreme representatives are the well-known Min Cut and Graph Bisection problems. The former is known to be efficiently solvable by flow techniques, the latter to be $\mathcal{NP}$-complete. The results presented in this paper are

- a monotony result of the type "The more balanced the partition we look for has to be, the harder the problem".
- a complexity result clarifying the status of a large part of intermediate problems in the class.

Thus we show the existence and partly localize an "efficiency border" between the two extremes.

1 Introduction

Partitioning the vertex set of a graph such that the two parts are connected by few cut edges is of great interest since it is (or would be) the foundation for divide-and-conquer algorithms for a lot of problems. For the special case of planar graphs [Lipton and Tarjan]'s Planar Separator Theorem (which deals with the closely related problem of partitioning a graph by taking away few vertices) is a key tool for a large class of algorithms on planar graphs. For general graphs two main results are known:

Using flow methods [Ford and Fulkerson] showed how to partition a graph efficiently such that two given vertices, a source and a sink, are in different parts and the number of cut edges is minimum . For divide-and-conquer approaches this Min Cut algorithm is not very helpful since it yields only a method to cut a graph into pieces of possibly very different sizes resulting in a linear recursion depth.

A bisection of the graph into two equal-sized parts with provably few cut edges would be the most desirable thing, but the corresponding decision problem MINIMUM BISECTION was shown to be $\mathcal{NP}$-hard by [Garey, Johnson and Stockmeyer]. In order to have a logarithmical recursion depth it would suffice to guarantee that both pieces contain at least a constant fraction of the vertices. The complexity status of this weaker version of the bisection problem was open so far.

In this paper we discuss the family of all such problems defined formally as:

* This work was done while the authors were with Lehrstuhl für angewandte Mathematik insbesondere Informatik, Rheinisch-Westfälische Technische Hochschule Aachen.

Definition 1. Let f be a function from the positive integers to the positive reals. An *f-balanced-bipartition (f-BB)* of a graph is a pair V_1, V_2 of subsets of its vertex set such that

- $V_1 \cup V_2 = V$, $V_1 \cap V_2 = \emptyset$ and
- $|V_i| \geq f(|V|)$.

The *(cut-)size* $|V_1, V_2|$ of such an f-BB is the number of edges between V_1 and V_2. An f-BB is *minimum* if its cutsize is minimum. f is called the *balancing function*.

Remarks: Throughout the paper we assume $f(n) \leq \frac{n}{2}$, since else an f-BB cannot exist.
We assume that the balancing function is efficiently computable. This is no real restriction for who wants to know a balanced bipartition, if she or he can't even compute the balancing function?

The decision problem MIN f-BB is:

Instance: A graph G and a bound K.
Question: Does G have an f-BB of size at most K?

In this context MINIMUM BISECTION equals MIN $\frac{n}{2}$-BB and MIN CUT (without specified source and sink) is exactly MIN 1-BB.

In Section 2 we will present a monotony result of the form "The more balanced the partition we look for has to be, the harder the problem" .

This motivates us to present in Sections 3 and 4 the results of our search for the "efficiency border":

Even MIN αn^{ϵ}-BB is $\mathcal{NP}$-hard for arbitrarily small positive constants α and ϵ; this includes, for $\epsilon = 1$, the problem of cutting off at least a constant fraction of the vertices. MIN C-BB is efficiently solvable for every constant C and the complexity status of MIN $\log n$-BB is left as an open problem.

2 A Monotony Result

A precise formulation of the dependency between the size of the balancing function and the complexity of the corresponding problem is the following theorem.

Theorem 2 Monotony. *Let g be a balancing function with the property*

$$0 \leq g(n)\text{-}g(n\text{-}1) \leq 1.$$

Then $f(n) \leq g(n)$ for all n implies
MIN f-BB is polynomially reducible to MIN g-BB.

Remark: The technical condition $0 \leq g(n) - g(n-1) \leq 1$ guarantees that g is monotonically increasing, i.e. the larger the graph the larger the pieces we look for, and that the speed of increase is bounded. There are a lot of functions which fulfill this property e. g. constants, $\log n$, n^{ϵ} (for $\epsilon \in [0, 1]$). Of course, we have to consider the admissible balancing functions, which correspond to these functions, i.e. if F is of one of these types we consider $f := \min\left\{\frac{n}{2}, F\right\}$.

Proof. Let G, K be an instance of f-BB. By adding an appropriate number δ of isolated vertices we reduce it to an instance H, K of g-BB.

$$V_H := V_G \cup V_\delta \text{ with } |V_\delta| = \delta, \quad E_H := E_G$$

δ is chosen such that (if $|V_G| =: n$ and $|V_H| =: m$)

$$m - \lceil g(m) \rceil = n - \lceil f(n) \rceil.$$

Such a δ always exists, since for $r(\ell) := \ell - \lceil g(\ell) \rceil$ we have

(i) $r(n) = n - \lceil g(n) \rceil \leq n - \lceil f(n) \rceil$
(ii) $r(2n) = 2n - \lceil g(2n) \rceil \geq n \geq n - \lceil f(n) \rceil$
(iii) $r(\ell) - r(\ell - 1) = 1 - (\lceil g(\ell) \rceil - \lceil g(\ell - 1) \rceil) \in \{0, 1\}$

So, there is an $m \in [n, \ldots, 2n]$ with $m - \lceil g(m) \rceil = r(m) = n - \lceil f(n) \rceil$. We always choose the smallest such m.

Here we need that the balancing functions are efficiently computable. Because of this the reduction is polynomial.

Now, let G have an f-BB V_G^1, V_G^2 of size at most K. Let w.l.o.g. $|V_G^1| \geq |V_G^2|$:

Case 1: $|V_G^2| \geq g(m)$.
Then $V_H^1 := V_G^1 \cup V_\delta, \quad V_H^2 := V_G^2$ defines a g-BB of the same size.
Case 2: $|V_G^2| < g(m)$.
Then $V_H^1 := V_G^1 \cup V_\delta^1, \quad V_H^2 := V_G^2 \cup V_\delta^2$ is a g-BB of the kind we look for if $|V_\delta^2| = \delta_2 := \lceil g(m) \rceil - |V_G^2|$, since:

a) $\delta_2 \leq \lceil g(m) \rceil - \lceil f(n) \rceil = m - n = \delta$.
b) By construction $|V_H^2| \geq g(m)$.
c) It remains to show that $|V_H^1| \geq g(m)$:
Assume otherwise, then $m - \lceil g(m) \rceil = |V_H^1| < g(m) \leq \frac{m}{2}$, so $\lceil g(m) \rceil > \frac{m}{2}$.
But this is only possible if m is odd and $\lceil g(m) \rceil = \frac{m+1}{2}$.
Now $\frac{m-1}{2} \geq \lceil g(m-1) \rceil \geq \lceil g(m) \rceil - 1 = \frac{m-1}{2}$, so $\lceil g(m-1) \rceil = \frac{m-1}{2}$.
Then $m - 1 - \lceil g(m-1) \rceil = m - \lceil g(m) \rceil$, which because of the minimality of m and δ means that $m = n$, so $\lceil f(n) \rceil = \frac{n+1}{2} > \frac{n}{2}$. But this contradicts the existence of an f-BB.

Let on the other hand (H, K) be a YES-instance of MIN g-BB. Then $V_G^i := V_H^i \cap V_G$ is an f-BB of G of the same size, since
$|V_G^i| \leq |V_H^i| \leq m - \lceil g(m) \rceil = n - \lceil f(n) \rceil \quad (i = 1, 2)$
$\Longrightarrow |V_G^i| \geq \lceil f(n) \rceil \quad (i = 1, 2)$.

Notice that this theorem is useful in two directions:

- If we show the $\mathcal{NP}$-completeness of MIN f-BB for a concrete function f the same holds for every larger function.
- If we can decide MIN g-BB efficiently, we can do so for every smaller balancing function. This part is even constructive in the sense that we can efficiently construct a g-BB of minimum size, if we can do so for the balancing function f.

3 Approaching the efficiency border from below

Finding a minimum 1-BB can be done by solving the classical Max Flow problem [Ford and Fulkerson] for all pairs of source/sink. This gives an $\mathcal{O}(n^5)$ polynomial algorithm.

A much faster solution was given by [Hao and Orlin] who essentially reduce it to one application of the preflow push algorithm by [Goldberg and Tarjan] yielding a $\mathcal{O}(nm \log n^2/m)$ algorithm.

A minimum C-BB (C constant) can be constructed by taking C of the vertices as one sink and C as one source and solving the Max Flow problem on the (slightly) smaller graphs for all possible choices of the $2C$ vertices. Since there are $\mathcal{O}(n^{2C})$ ways of choosing the source/sink-vertices this gives a $\mathcal{O}(n^{2C+3})$ polynomial algorithm. This method stops to work for MIN $\log n$-BB as it would lead to an $\mathcal{O}(n^{2\log n+3})$ algorithm which is superpolynomial.

4 ... from above

The smallest balancing function f for which we are able to show the $\mathcal{NP}$-completeness of MIN f-BB is an arbitrary small but positive power of n. An interesting point is that the reduction is from MINIMUM BISECTION, i.e. it runs in a sense in the opposite direction of Theorem 2.

Theorem 3 Upper bound. *MIN αn^ϵ-BB is $\mathcal{NP}$-complete for $\alpha, \epsilon > 0$.*

Proof. We assume that $\epsilon \leq 1$ and $\alpha \leq \frac{1}{4}$. Theorem 2 shows that the theorem also holds for larger α as long as αn^ϵ fulfills the technical condition.

Given an instance G, K of MINIMUM BISECTION we reduce it to an instance H, L of MIN f-BB such that

G, K is a YES-instance of MINIMUM BISECTION

$\Leftrightarrow H, L$ is a YES-instance of MIN f-BB.

First we handle two extreme cases: If n is odd (i.e. no bisection can exist) let $H = K_1$ and $L = 0$; if n is even and $K \geq \frac{n^2}{4}$ (i.e. such a bisection always exist since the maximum cut size is at most $\frac{n^2}{4}$) let $H = K_2$ and $L = 1$.

So let n be even and $K < \frac{n^2}{4}$: H is now constructed from G by substituting each vertex of G by a K_{n^2}, a "small" clique on n^2 vertices. Each edge of G is simulated by an edge between the corresponding cliques. To ensure the close relation between bisections of G and f-BB's of H we add as an enforcer an additional "large" clique on $\left\lfloor \left(\frac{n^3}{2\alpha}\right)^{1/\epsilon} \right\rfloor - n^3$ vertices. The small cliques are called $H_1, \ldots, H_n$ — corresponding to $v_1, \ldots, v_n$, the vertices of G — the large one H_0. H_0 is connected to each of the small cliques by n additional edges which are chosen avoiding parallel edges, such that H is also a simple graph. The bound L is chosen as $K + \frac{n^2}{2}$. Is H_0 large enough to avoid parallel edges?

$$\begin{aligned} |V_{H_0}| &= \left\lfloor \left(\frac{n^3}{2\alpha}\right)^{1/\epsilon} \right\rfloor - n^3 \geq \left\lfloor \left(2n^3\right)^{1/\epsilon} \right\rfloor - n^3 \\ &\geq 2n^3 - n^3 = n^3 \end{aligned}$$

This is of course large enough. The construction of H guarantees that a bisection of G induces an $\alpha|V_H|^\epsilon$-BB of H where the smaller part contains exactly $\lceil\alpha|V_H|^\epsilon\rceil$ vertices:

A bisection of G where one half consists of the vertices $v_{j_1},\ldots,v_{j_{n/2}}$ induces a partition of V_H into, say V_H^1 and V_H^2 with

$$V_H^1 = \bigcup_{j=1}^{\frac{n}{2}} V_{H_{j_i}} \text{ and } V_H^2 = V_H \backslash V_H^1.$$

Since

$$\alpha|V_H|^\epsilon = \alpha\left\lfloor\left(\frac{n^3}{2\alpha}\right)^{1/\epsilon}\right\rfloor^\epsilon$$

we have

$$\frac{n^3}{2} \geq \alpha|V_H|^\epsilon \geq \alpha\left(\left(\frac{n^3}{2\alpha}\right)^{1/\epsilon} - 1\right)^\epsilon$$

$$\geq \alpha\left(\frac{n^3}{2\alpha} - 1\right),$$

since $(x-1)^\epsilon \geq x^\epsilon - 1$, if $\epsilon \leq 1$ and $x \geq 1$,

$$= \frac{n^3}{2} - \alpha.$$

Thus $\lceil\alpha|V_H|^\epsilon\rceil = \frac{n^3}{2}$, since $\alpha < 1$.

So a bisection of G of size at most K induces an $\alpha|V_H|^\epsilon$-BB of H of size $K + \frac{n}{2}\cdot n$ (the simulated edges of G plus the edges between the $V_{H_{j_i}}$ and V_{H_0}) which is at most L.

On the other hand assume that H, L is a YES-instance for MIN $\alpha|V_H|^\epsilon$-BB. Let V_H^1, V_H^2 be a $\alpha|V_H|^\epsilon$-BB of size at most L. Observe that none of the cliques (small or large) of H can have vertices in both V_H^1 and V_H^2, because otherwise the cut-size would be at least n^2-1 (every clique has at least n^2 vertices) which is larger than

$$L = K + \frac{n^2}{2} < \frac{n^2}{4} + \frac{n^2}{2} = \frac{3}{4}n^2.$$

Thus V_H^1, V_H^2 induces canonically a bipartition V_G^1, V_G^2 of G. We will modify V_H^1, V_H^2 without increasing its size such that the induced bipartition is a bisection:

If $|V_G^1| = |V_G^2|$ we are done, so let w.l.o.g. $|V_G^2| > |V_G^1|$. It follows that $V_{H_0} \subseteq V_H^1$, since else $|V_H^1| < \frac{n}{2}\cdot n^2 = \lceil\alpha|V_H|^\epsilon\rceil$ in contradiction to our assumption.

Define M as $\{i \in \{1,\ldots,n\} \mid V_{H_i} \subseteq V_H^2\}$, let M_2 be an arbitrary subset of M of order $n/2$ and let M_1 be $M \setminus M_2$. Now we move the small cliques with indices from M_1 from V_H^2 to V_H^1 thus producing a new bipartition W_H^1, W_H^2 which obviously induces a bisection of G. What about the cut size?

Let $C := \bigcup_{i \in M_1} V_{H_i}$ be the set of moved vertices, then

$$|W_H^1, W_H^2| - |V_H^1, V_H^2| = |C, W_H^2| - |V_H^1, C| =: \alpha.$$

Now

$$|C, W_H^2| \leq \frac{|C|}{n^2} \cdot \frac{|W_H^2|}{n^2} = |M_1| \cdot \frac{n}{2},$$

since the only edges between C and W_H^2 (both consisting only of small cliques) are the simulated edges of G.

In addition we have $|V_H^1, C| \geq |V_{H_0}, C| = |M_1| \cdot \frac{n}{2}$ and thus $\alpha \leq 0$, which means that the cutsize was not increased by our modification.

For the size of the induced bisection W_G^1, W_G^2 of G we have

$$\begin{aligned} |W_G^1, W_G^2| &= |W_H^1 \setminus V_{H_0}, W_H^2| \\ &= |W_H^1, W_H^2| - |V_{H_0}, W_H^2| \\ &\leq L - \frac{n^2}{2} = K. \end{aligned}$$

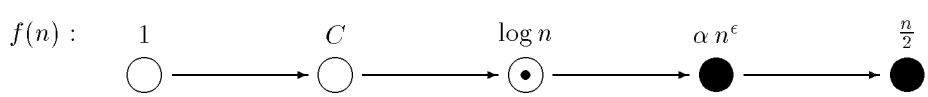
Thus we have shown the claimed existence of a bisection of G of the right size. What about the efficiency of the reduction. The only critical point is the size of the large clique, which is $\mathcal{O}\left(n^{3/\epsilon}\right)$ which is a polynomial in the input size as long as ϵ is constant.

If we try to show the $\mathcal{NP}$-completeness for smaller balancing functions along the same lines as in the proof of Theorem 3, we fail since an analogous reduction would no longer be polynomial; adding a large clique for the case of MIN $\log n$-BB would yield a graph H of size $e^{n^3/2}$.

5 Conclusion

Using similar methods as [Bui, Chaudhuri, Leighton and Sipser] we can extend the hardness results to three-regular graphs.

The concluding figure is a graphical summary of the results in this paper.



where ○ stands for $\in \mathcal{P}$,

● for $\mathcal{NP} - complete$

and ⊙ for open.

References

[Bui, Chaudhuri, Leighton and Sipser] T. N. Bui, S. Chaudhuri, F. T. Leighton, M. Sipser: *Graph Bisection Algorithms With Good Average Case Behaviour*, Combinatorica **7** (1987) 171–191

[Ford and Fulkerson] L. R. Ford, D. R. Fulkerson: *Maximal Flow Through a Network*, Canadian J. Math. **8** (1956) 399–404

[Garey, Johnson and Stockmeyer] M. J. Garey, D. S. Johnson, L. Stockmeyer: *Some Simplified $\mathcal{NP}$-complete Graph Problems*, Theoretical Computer Science **1** (1976) 237–267

[Goldberg and Tarjan] A. V. Goldberg, R. E. Tarjan: *A New Approach to the Maximum Flow Problem*, Journal of the ACM **35** (1988) 921–940

[Hao and Orlin] J. Hao, J. B. Orlin: *A Faster Algorithm for Finding the Minimum Cut in a Graph*, Proceedings of the third ACM-SIAM Symposium on Discrete Algorithms (SODA'91) (1991) 165–174

[Lipton and Tarjan] R. J. Lipton, R. E. Tarjan: *A Planar Separator Theorem*, SIAM J. on Applied Math. **36** (1979) 177–189

Reports from the group

"Algorithmic Discrete Mathematics"

of the Department of Mathematics, TU Berlin

373/1994 *Andreas S. Schulz:* A Note on the Permutahedron of Series–Parallel Posets

368/1993 *Stefan Felsner, Rudolf Müller, Lorenz Wernisch:* Optimal Algorithms for Trapezoid Graphs

367/1993 *Dorothea Wagner:* Simple Algorithms for Steiner Trees and Paths Packing Problems in Planar Graphs, erscheint in *"Special Issue on Disjoint Paths" (eds. B. Gerards, A. Schrijver) in CWI Quarterly*

365/1993 *Rolf H. Möhring:* Triangulating Graphs without Asteroidal Triples

359/1993 *Karsten Weihe:* Multicommodity Flows in Even, Planar Networks

358/1993 *Karsten Weihe:* Non-Crossing Path Packings in Planar Graphs with Applications

357/1993 *Heike Ripphausen-Lipa, Dorothea Wagner, Karsten Weihe:* Linear-Time Algorithms for Disjoint Two-Face Paths Problems in Planar Graphs

354/1993 *Dorothea Wagner, Karsten Weihe:* CRoP: A Library of Algorithms for the Channel Routing Problem

351/1993 *Jens Gustedt:* Finiteness Theorems for Graphs and Posets Obtained by Compositions

350/1993 *Jens Gustedt, Angelika Steger:* Testing Hereditary Properties Efficiently

349/1993 *Stefan Felsner:* 3-Interval Irreducible Partially Ordered Sets

348/1993 *Stefan Felsner, Lorenz Wernisch:* Maximum k-Chains in Planar Point Sets: Combinatorial Structure and Algorithms

345/1993 *Paul Molitor, Uwe Sparmann, Dorothea Wagner:* Two-Layer Wiring with Pin Preassignments

344/1992 *Dorothea Wagner, Karsten Weihe:* A Linear Time Algorithm for Multicommodity Flow in Planar Graphs

343/1992 *Dorothea Wagner, Frank Wagner:* Modeling Hypergraphs by Graphs with the same Mincut Properties

341/1992 *Klaus Jansen, Petra Scheffler, Gerhard Woeginger:* The Disjoint Cliques Problem

337/1992 *Rudolf Müller:* On the Transitive Acyclic Subdigraph Polytope

336/1992 *Michel Habib, Rolf H. Möhring:* Treewidth of cocomparability graphs and a new order-theoretic parameter

333/1992 *Dirk Möller, Rudolf Müller:* A Concept for the Representation of Data and Algorithms

324/1992 *Heike Ripphausen-Lipa, Dorothea Wagner, Karsten Weihe:* The Vertex-Disjoint Menger Problem in Planar Graphs

323/1992 *Rolf H. Möhring, Dorothea Wagner, Frank Wagner:* VLSI Network Design

322/1992 *Jens Gustedt, Michel Morvan:* N-free Orders and Minimal Interval Extensions

320/1992 *Michel Habib, David Kelly, Rolf H. Möhring:* Comparability Invariance of Geometric Notions of Order Dimension

319/1992 *Klaus Jansen, Petra Scheffler:* Generalized coloring for tree-like graphs

318/1992 *Dorothea Wagner:* Optimal Routing Through Dense Channels

317/1992 *Majid Sarrafzadeh, Dorothea Wagner, Frank Wagner, Karsten Weihe:* Wiring Knock-Knee Layouts: A Global Approach

309/1991 *Stefan Felsner:* Tolerance Graphs and Orders

308/1991 *Stefan Felsner, William T. Trotter:* Colorings of Diagrams of Interval Orders and α-Sequences of Sets

307/1991 *Dorothea Wagner, Frank Wagner:* Between Min Cut and Graph Bisection

295/1991 *Stefan Felsner, Dorothea Wagner:* On the Complexity of Partial Order Properties

291/1991 *Stefan Felsner, Jens Gustedt, Michel Morvan, Jean-Xavier Rampon:* Constructing Colorings for Diagrams

290/1991 *Jens Gustedt:* Well Quasi Ordering Finite Posets and Formal Languages

285/1991 *Stefan Felsner, Michel Habib, Rolf H. Möhring:* On the Interplay between Interval Dimension and Dimension

279/1991 *Dorothea Wagner:* A New Approach to Knock-Knee Channel Routing

276/1991 *Christoph Meinel, Stephan Waack:* Separating complexitiy classes related to bounded alternating ω-branching programs

275/1991 *Stefan Felsner:* Bounds for the Jump Number of Partially Ordered Sets

Reports may be requested from:

D. Wagner
Fachbereich Mathematik, MA 6–1
TU Berlin
Straße des 17. Juni 136
D-10623 Berlin
e-mail: wagner@math.TU-Berlin.DE

This article was processed using the LaTeX macro package with LLNCS style