The Complexity of the Bottleneck Graph Bipartition Problem

F. Shahrokhi
Department of Computer Science
University of North Texas
Denton, Texas, 76203

L. A. Székely†
Institut für Ökonometrie und Operations Research
Rheinische Friedrich-Wilhelms Universitat
5300 Bonn, Germany
and
Department of Computer Science
Eötvös Loránd University
Budapest, Hungary

Abstract. In this paper we prove the NP-hardness of the bottleneck graph bipartition problem and study the complexity status of possible approximation algorithms for some related problems.

Our standard reference to graph theory is [CL86]. For standard concepts and definitions on approximating combinatorial problems, the reader should consult [PS82]; throughout this paper, however, we will give a brief description of the concepts which are relevant to our work.

For a graph $\langle V, E\rangle$, and $B \subseteq V$, we denote by $e(B)$ the edge set induced by B. For disjoint $A, B \subseteq V$, we denote by $e(A, B)$ the set of all edges with exactly one end point in A and one end point in B. A partition $(B, \bar{B})$ of V is called a cut.

The bottleneck bipartition problem (BBNP) is to find a cut $(B, \bar{B})$ in $G = \langle V, E\rangle$ with the smallest possible $\max\{|e(B)|, |e(\bar{B})|\}$. We denote this number, which was introduced by Entringer [En88], by $\gamma(G)$ and observe that $\gamma(G) = 0$ for bipartite graphs. Erdős [Er88] conjectured that $\gamma(G) \leq |E|/4 + O(\sqrt{|E|})$. This was recently proved by Porter [Po89] in a nonconstructive and nonprobabilistic fashion. Clark [Cl88] proved a weaker version of this conjecture using a probabilistic argument. Clark, Shahrokhi and Szekély [CSS92] gave an algorithm to approximate the BBNP and verified constructively a weaker version of the Erdős' conjecture.

Define $\overline{\gamma}(G) = \min_{(B,\bar{B})}\{|e(B)| + |e(\bar{B})|\}$. Observe that $\overline{\gamma}(G)$ is the minimum number of edges in G whose removal leaves a bipartite subgraph of G, and if $\overline{\gamma}(G) = |e(C)| + |e(\bar{C})|$ for some cut $(C, \bar{C})$, then $(C, \bar{C})$ is a maximum cut.

†Research supported by the A. v. Humboldt-Stiftung

The problem of computing $\overline{\gamma}(G)$ is called the minimum edge deletion bipartite subgraph problem (MEDBP) [Se89]. It is easy to verify that,

$$\frac{1}{2}\overline{\gamma}(G) \leq \gamma(G) \leq \overline{\gamma}(G). \tag{1}$$

The maximum cut problem (MXCP) is to find a cut $(B,\bar{B})$ in $G = \langle V,E\rangle$ with the largest possible $|e(B,\bar{B})|$. Although MXCP is NP-hard, [GJ78], the special case in which the underlying graph is planar, can be solved in polynomial time [Ha75], [KCS88].

Lemma 1. *Let $(C,\bar{C})$ be a maimum cut in G with $|e(C)| = |e(\bar{C})|$, then*

$$\frac{1}{2}\overline{\gamma}(G) = \gamma(G).$$

Proof. Observe that for the cut $(C,\bar{C})$, $|e(C)| = |e(\bar{C})| > \gamma(G)$ and $|e(C)| = |e(\bar{C})| = \frac{1}{2}\overline{\gamma}(G)$ so that $\gamma(G) < \frac{1}{2}\overline{\gamma}(G)$ while $\gamma(G) \geq \overline{\gamma}(G)$ by (1). Consequently, $\gamma(G) = \frac{1}{2}\overline{\gamma}(G)$. ∎

Theorem 2. *Computing $\gamma(G)$ is NP-hard.*

Proof. We reduce the maximum cut problem to the problem of computing γ. For $G = \langle V,E\rangle$ a graph, construct the cartesian product G' of K_2 and G (see [CL86]). Note that G' consists precisely of two isomorphic copies of G such that identical vertices in these copies are joined by an edge. An edge e in G' is called a cross edge if the end points of e are located in different copies of G (in G'). Observe that there are exactly $|V|$ cross edges in G'.

Let $(C,\bar{C})$ be a maximum cut in G, and $(D,\bar{D})$ be any cut in G', we claim that

$$|e(D,\bar{D})| \leq 2|e(C,\bar{C})| + |V|.$$

To see this, note that $D = A \cup B$ and $\bar{D} = \bar{A} \cup \bar{B}$, where $(A,\bar{A})$ and $(B,\bar{B})$ are cuts in two different copies of G in G'. It is easy to verify that,

$$e(D,\bar{D}) = e(A,\bar{A}) \cup e(B,\bar{B}) \cup e(A,\bar{B}) \cup e(B,\bar{A}).$$

However, $|e(A,\bar{A})| \leq |e(C,\bar{C})|$, and $|e(B,\bar{B})| \leq |e(C,\bar{C})|$, since $(C,\bar{C})$ is a a maximum cut in G. Furthermore, $|e(A,\bar{B})| + |e(B,\bar{A})| \leq |V|$, since any edge in $e(A,\bar{B}) \cup e(B,\bar{A})$ is a cross edge in G'. It follows that $|e(D,\bar{D})| \leq 2|e(C,\bar{C})| + |V|$, as we claimed. Next, consider a cut $(K,\bar{K})$ in G' by letting K be the union of C from one copy of G and C from the other copy of G. Then

$$|e(K,\bar{K})| = 2|e(C,\bar{C})| + |V|. \tag{2}$$

Therefore, $(K,\bar{K})$ is a maximum cut in G'. However, $(K,\bar{K})$ has $|e(K)| = |e(\bar{K})| = |e(C)| + |e(\bar{C})|$ so that by Lemma 1

$$\bar{\gamma}(G') = 2\gamma(G'). \tag{3}$$

Now assume that (G') is computed in polynomial time. Then (3) implies that $\bar{\gamma}(G')$ is also computed in polynomial time. Therefore, $|e(K,\bar{K})|$ for a maximum cut $(K,\bar{K})$ in G' is computed in polynomial time. However, then (2) implies that $|e(C,\bar{C})|$ for a maximum cut in G is computed in polynomial time which verifies the result. ∎

Theorem 2 emphasizes the significance of the estimating $\gamma(G)$ by deriving upper bounds such as the upper bound of Erdős proved by Porter [Po92]. Theorem 2 also suggests that since it is less likely to compute $\gamma(G)$ in polynomial time, one should concentrate on approximating $\gamma(G)$ in polynomial time. In the rest of the paper we explore the relative degree of difficulty with respect to ϵ-approximation of BBNP, MEDBP, and MXCP.

Let $\mathcal{P}$ be a combinatorial optimization problem with a positive integral cost (or objective) function; denote the optimal cost value for any instance I of $\mathcal{P}$ by $\hat{c}(I)$. Assume $\mathcal{A}$ is an algorithm which, given an instance I of $\mathcal{P}$, returns a feasible solution to I with cost value $c_{\mathcal{A}}(I)$. Let $\epsilon > 0$ be fixed; we say $\mathcal{A}$ is an ϵ-approximate algorithm for $\mathcal{P}$, if

$$\frac{|\hat{c}(I) - c_{\mathcal{A}}(I)|}{\hat{c}(I)} \leq \epsilon \tag{4}$$

for any instance I of P. If (4) holds for every fixed $\epsilon > 0$ for all instances I of $\mathcal{P}$, then $\mathcal{A}$ is called an ϵ-approximate scheme. An ϵ-approximate algorithm whose running time is polynomially bounded in the problem size I is termed a polynomial time approximate algorithm. Similarly, an ϵ-approximate scheme whose computing time is bounded in the problem size is called a polynomial time approximation scheme (PTAS).

A very strong type of approximation for a combinatorial optimization problem $\mathcal{P}$ occurs when we wish to satisfy (4) for any $\epsilon > 0$ within a time bounded by a polynomial in both ϵ^{-1} and the input size. More formally, we say that algorithm 4 is a fully polynomial approximation scheme (FPAS) for $\mathcal{P}$, if and only if given any instance I of $\mathcal{P}$ and any $\epsilon > 0$, $\mathcal{A}$ computes within a time bounded by a polynomial in both ϵ^{-1} and the size of I, a feasible solution to $\mathcal{P}$ such that (4) holds.

Theorem 3.

(i) *For MXCP, BBNP or MEDBP there is no FPAS unless $P = NP$.*

(ii) *Any ϵ-approximate algorithm $\mathcal{A}$ for MEDBP is also an ϵ-approximate algorithm for MXCP.*

(iii) *Let $\mathcal{A}$ be any ϵ-approximate algorithm for BBNP, then there is an approximate algorithm for MEDBP with the same time complexity as $\mathcal{A}$.*

Proof. Part (i) follows from Theorem 17.12 of [PS82; p.430].

To verify (ii), note that for the cut $(B, \bar{B})$ produced by $\mathcal{A}$ and a maximum cut $(C, \bar{C})$ in G. We have

$$\left|e(C,\bar{C})\right| - \left|e(B,\bar{B})\right| = \left|e(B)\right| + \left|e(\bar{B})\right| - \left|e(C)\right| - \left|e(\bar{C})\right|.$$

Next, observe that any maximum cut in G contains at least half of the edges of G. Thus,

$$\left|e(C,\bar{C})\right| \geq \frac{|E|}{2} \geq \left|e(c)\right| + \left|e(\bar{C})\right|.$$

It follows that

$$\frac{\left|e(C,\bar{C})\right| - \left|e(B,\bar{B})\right|}{\left|e(C,\bar{C})\right|} \leq \frac{\left|e(B)\right| + \left|e(\bar{B})\right| - \left|e(C)\right| - \left|e(\bar{C})\right|}{\left|e(C)\right| + \left|e(\bar{C})\right|}.$$

However,

$$\frac{\left|e(B)\right| + \left|e(\bar{B})\right| - \left|e(C)\right| - \left|e(\bar{C})\right|}{\left|e(C)\right| + \left|e(\bar{C})\right|} \leq \epsilon,$$

since $(B, \bar{B})$ is the cut constructed by algorithm $\mathcal{A}$. Consequently,

$$\frac{\left|e(C,\bar{C})\right| - \left|e(B,\bar{B})\right|}{\left|e(C,\bar{C})\right|} \leq \epsilon.$$

Finally to verify (iii), consider the cartesian product G' of K_2 and G as in the proof of Theorem 2. Let $(K, \bar{K})$ and $(C, \bar{C})$ be maximum cuts in G' and G, respectively. We have $|e(K,\bar{K})| = 2|e(C,\bar{C})| + |V|$, by (2), so that

$$\overline{\gamma}(G') = 2|E| + |V| - \left|e(K,\bar{K})\right| = 2\left(|E| - \left|e(C,\bar{C})\right|\right) = 2\overline{\gamma}(G).$$

By (3) (in Theorem 2), we have $\overline{\gamma}(G') = 2\overline{\gamma}(G)$. It follows that

$$\gamma(G') = \overline{\gamma}(G). \tag{5}$$

Now let $\mathcal{A}'$ be an algorithm which first constructs G' from G (in $O(|E| + |V|)$ time) and then applies algorithm $\mathcal{A}$ to G'. It is clear that A' has the same time complexity as $\mathcal{A}$. For the cut $(D, \bar{D})$ in G' produced by $\mathcal{A}'$ we have

$$\max\left\{\left|e(D)\right|, \left|e(D)\right|\right\} \leq (1+\epsilon)\gamma(G').$$

Employing (5), we get

$$\max\left\{\left|e(D)\right|, \left|e(\bar{D})\right|\right\} \leq (1+\epsilon)\overline{\gamma}(G). \tag{6}$$

However, $D = A_1 \cup B_1$ and $\bar{D} = \bar{A}_1 \cup \bar{B}_1$ where $(A_1, \bar{A}_1)$ and $(B_1, \bar{B}_1)$ are cuts in the two different copies of G in G'. Assuming (with no loss of generality that) $|e(A)| + |e(\bar{A})| \leq |e(B)| + |e(\bar{B})|$, and using (6), we get

$$|e(A_1)| + |e(\bar{A}_1)| \leq \frac{(|e(D)| + |e(\bar{D})|)}{2}$$
$$\leq \max\{|e(D)|, |e(\bar{D})|\} \leq (1+\epsilon)\overline{\gamma}(G).$$

Thus $(A_1, \bar{A}_1)$ is a desirable cut. ∎

Part (i) of Theorem 3 justifies the significance of weaker types of approximation algorithms than FPAS for our problems. Parts (ii) and (iii) of Theorem 3, however, establish the relative degree of difficulty between these problems with respect to ϵ-approximation.

For the MXCP efficient $\frac{1}{2}$-approximate algorithms are known [PY88], [CSS92]. We have not been able to find an efficient ϵ-approximate algorithm for the MXCP with $\epsilon < 1/2$, neither have we been able to show that one does not exist (unless $P = NP$). We note here, that effective algebraic upper bounds for MXCP by Delorme and Poljack [DP91], are conjectured to be within a multiplicative factor of 1.14 from the max cut, although this method did not yield yet provably good polynomial time approximation algorithms.

For MEDBP an efficient 1-approximate algorithm has been discovered [Se89]. No ϵ-approximate algorithms for BBPN are known, for any ϵ. We indicate that the algorithm in [CSS92] computes a good upper bound on $\gamma(G)$. However, this algorithm is not an ϵ-approximate algorithm. We believe that finding an ϵ-approximate algorithm for BBPN (if possible) is important and at the same time difficult, since by Theorem 3 such algorithm would imply new ϵ-approximate algorithms for MXCP and MEDBP.

References

[CL86] G. Chartrand and L. Lesniak, "Graph and Digraph", 1986. Wadsworth and Brooks/Cole Mathematics Series.

[C188] L. H. Clark. Personal communication.

[CSS92] L. H. Clark, F. Shahrokhi, L. A. Szekély, *A linear-time algorithm for graph partition problems*, Inform. Process. Letters. 42 (1992), 19–24.

[DP91] C. Delorme, S. Poljak, *The performance of an eigenvalue bound on the maxcut problem in some classes of graphs.* Forchungintitut fr dikrete Mathematik, Bonn, Report No. 91705-0R.

[En88] R. C. Entringer. Personal communication.

[Er88] P. Erdős. Personal communication.

[GJ78] M. R. Garey, D. S. Johnson, "Computer and Intractability, a Guide to the Theory of NP-Completene", W. H. Freeman and Co., San Francisco, 1978.

[GJS76] M. R. Garey, D. J. Johnson, L. Stockmeyer, *Some simplified NP-complete graph problems*, Theor. Computer Science 1 (1976), 237–267.

[Ha75] F. Hadlock, *Finding a maximum cut of a planar graph in polynomial time*, SIAM J. Computing 4(3) (1975), 221–225.

[KCS88] Y. S. Kuo, T. C. Chern and Wei-Kuan Shih, *Fast algorithms for optimal layer assignment*, Proceedings of Design Automation Conference (1988), 554–558.

[Po89] D. Porter, *On a bottleneck bipartition conjecture of Erdős.* To appear in Combinatorica 12 (1992).

[PS82] C. H. Papadimitriou and K. Steiglitz, "Combinatorial Optimization: Algorithm and Complexity", Prentice-Hall, Englewood Cliffs, N.J., 1982.

[PY88] C. H. Papadimitriou and M. Yannalcakis, *Optimization, Approximation and Complexity*, Proceeding of STOC (1988), 229–234.

[Se89] R. K. Sen. An approximate algorithm for the minimum edge deletion bipartite subgraph problem, mancript.