Computing 16, 263—270 (1976)

The NP-Completeness of the Bandwidth Minimization Problem*

Ch. H. Papadimitriou, Princeton, New-Jersey

Received February 14, 1975; revised June 31, 1975

Abstract — Zusammenfassung

The NP-Completeness of the Bandwidth Minimization Problem. The Problem of minimizing the bandwidth of the nonzero entries of a sparse symmetric matrix by permuting its rows and columns and some related combinatorial problems are shown to be NP-Complete.

Die NP-Geschlossenheit des Problems der minimalen Bandbreite. Es wird gezeigt, daß das Problem, die minimale Bandbreite der von Null verschiedenen Elemente einer schwach besetzten symmetrischen Matrix durch Umstellung der Reihen und Spalten zu finden, und einige verwandte Probleme der Kombinatorik NP-geschlossene sind.

I. Introduction

In computations involving large sparse matrices it can be desirable to have a matrix stored in such a way, so that all nonzero terms are as close to the diagonal as possible. Such a matrix can be manipulated with sometimes considerable savings in both storage and time. The absolute value of the difference of the coordinates of the most off-diagonal non-zero term is called the bandwidth of the matrix. In the past the problem of effectively decreasing (see, for example, [1]—[3]) or minimizing ([4], [5]) the bandwidth of a sparse, symmetric matrix by permuting its rows and columns has been considered, and a graph-theoretic approach to the problem has been developed by Harary [6].

In the present paper it is proved that the problem of minimizing the bandwidth of a matrix is NP-complete in the sense of [7], a result that suggests that, unless $P=NP$, the task of obtaining an exact solution to it is inherently hard. An immediate consequence is that the problem of minimizing the bandwidth with the first row fixed is P-complete. ("P-Complete Problems" are those problems which are NP-Complete under a wider definition of polynomial time reducibility (see [8], Ch. 10.2). NP-Completeness implies P-Completeness, but the converse is an un-

* This work was supported in part by NSF Grant GK-42048 and the U.S. Army Research Office – Durham under Contract DAHC04-69-C0012.

settled conjecture.) In fact, the algorithm S given by [4] for this particular problem fails to solve it in a polynomial number of steps.

Like other combinatorial optimization problems such as the Maximal Clique Problem and the Traveling Salesman Problem, the problem of minimizing the bandwidth of a matrix is transformed to a language recognition problem, that is, a question that can be answered by "yes" or "no". It can be argued that, if the language recognition problem is inherently hard, then the corresponding optimization problem is also inherently hard. Nevertheless some recent results [9] indicate that the task of finding good approximate solutions to P-complete combinatorial optimization problems is not necessarily hard. In fact, the Bandwidth Minimization Problem seems to be an example of an NP-Complete Problem for which good approximate solutions can be obtained by a reasonable computational effort.

Finally it should be mentioned, that the NP-completeness of another combinatorial optimization problem concerning sparse matrices, the Minimum Profile Problem, [2], is implicit in [10] (Theorem 1.5).

II. Definitions — Notation

Let $G=[V,E]$ be a graph with $V=\{1,2,\ldots,n\}$. The chromatic number of G, $\chi[G]$, is the minimal number of colors that can be assigned to vertices of V so that no adjacent vertices have the same color. The *bandwidth* of G is defined by

$$\beta[G]=\min_{\pi\in S_n}\ \max_{(i,j)\in E}\ (|\pi(i)-\pi(j)|)$$

where S_n is the symmetric group of n objects.

The following indicates that the bandwidth provides us with an upper bound for the chromatic number.

Proposition: In any graph G, $\chi[G]\le\beta[G]+1$.

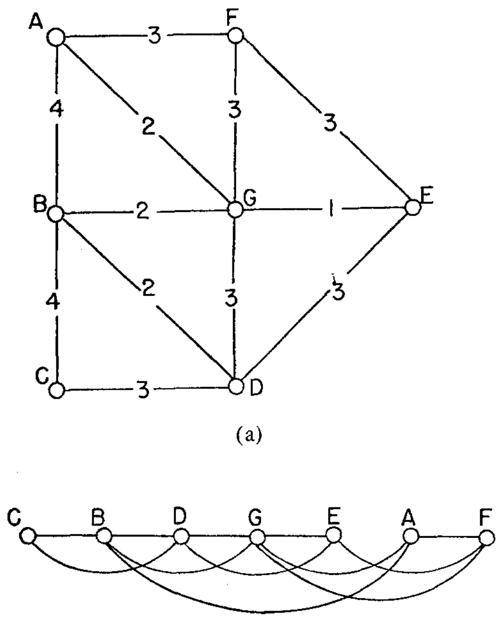
Proof: Let π be the permutation by which the minimum bandwidth is obtained. The following is an admissible coloring of V: Assign to vertex j the $c[j]$-th color where $1\le c[j]\le\beta[G]+1$ and $c[j]\equiv\pi[j]\ (\mathrm{mod}\ \beta[G]+1)$.||

Note that the minimal bandwidth of a symmetric matrix as defined in the introduction, is equal to the bandwidth of $G=(V,E)$, where V is the set of the columns of the matrix and $(i,j)\in E$ iff the (i,j)-th entry is nonzero.

The Bandwidth Minimization Problem (*BMP*) is the following: Given a graph G and an integer $b>0$, is $\beta[G]\le b$?

The Linear Array Problem (*LAP*) is the following: Given a graph G and an integer valued labeling function l over its edges, can we arrange the vertices of G in a linear array with unit distances between consecutive vertices, such that no two adjacent vertices have a distance greater than the label of the edge joining them? Equivalently, is there $\pi\in S_n$ such that $(i,j)\in E\Rightarrow|\pi(u)-\pi(j)|\le l[(i,j)]$?

If G is a labeled graph, by $LAP[G]$ we shall denote the LAP induced by G. As an illustration the LAP of the graph in Fig. 1 a is solvable by the arrangement of its vertices shown in Fig. 1 b.



(a)

(b)

Fig. 1

Since answers to both the LAP and the BMP can be given by examining the connected components of the graph separately, henceforth all graphs will be assumed connected.

The *satisfiability problem* of Boolean expressions in conjunctive normal form with *at most* (resp. *exactly*) three distinct literals per clause will be abbreviated by *At Most* 3-*SAT* (resp. *Exactly* 3-*SAT*).

Boolean expressions will be written with $\wedge$ replaced by . and $\vee$ replaced by $+$.

A partition of the set $S=\{x_i, \bar{x}_i \mid i=1, \ldots, n\}$ of n Boolean variables and their negations into 2 sets P, Q with n elements each is called consistent if for all i not both x_i and $\bar{x}_i$ belong to P. A consistent partition $S=P \cup Q$ (with P considered to be distinguished) *induces* the following truth assignment:

$$t[x_i]=1 \text{ iff } \bar{x}_i \in P.$$

Finally *Restricted LAP* is a LAP where the only permissible labels of the edges are b and $2b-1$ for some positive integer b. Note that the BMP is also a special case of the LAP, with $l[e]=b$ for all $e \in E$.

III. The NP-Completeness of the BMP

Lemma 1: *The following problems are in NP:*

a) Exactly 3-SAT

b) LAP

c) Restricted LAP, BMP.

Proof:

a) The Exactly 3-*SAT* problem is a special case of the *SAT* problem which is in *NP* (see, for example, [8], Theorem 10.2).

b) A Nondeterministic Turing Machine could test simultaneously all possible permutations of V in polynomial time.

c) These problems are special cases of the *LAP*. ||

Lemma 2: *The Exactly 3-SAT problem is NP-Complete.*

Proof: We shall reduce to it the At Most 3-SAT problem which is known to be *NP*-Complete [7]. Given a Boolean expression $B=\prod_{i=1}^{r} F_i$ with at most three literals per clause for each F_i, $i=1, \dots, r$, let

$$F_i' = \begin{cases} F_i & \text{if } F_i \text{ has 3 literals} \\ F_i+y & \text{if } F_i \text{ has 2 literals} \\ F_i+y+z & \text{if } F_i \text{ has 1 literal} \end{cases}$$

and set

$$B' = \left[\prod_{i=1}^{r} F_i'\right] \cdot (\bar{y}+\bar{\alpha}+\beta)\cdot(\bar{y}+\alpha+\bar{\beta})\cdot(\bar{y}+\bar{\alpha}+\bar{\beta})\cdot(\bar{y}+\alpha+\beta)\cdot(\bar{z}+\bar{\alpha}+\beta) \cdot(\bar{z}+\alpha+\bar{\beta})\cdot(\bar{z}+\bar{\alpha}+\bar{\beta})\cdot(\bar{z}+\alpha+\beta),$$

where y, z, α and β are new variables. B' has exactly 3 literals per clause and B' is satisfiable iff B is. Moreover the construction of B' is achievalbe in polynomial time. ||

Theorem 1: *The LAP is NP-Complete.*

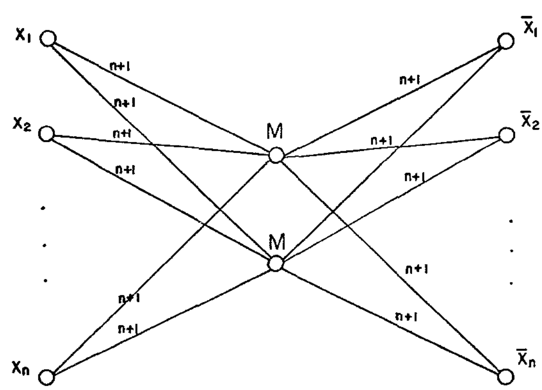


Fig. 2

Proof: Consider an instance B of the Exactly 3-SAT with variables $x_1, \ldots, x_n$ and clauses $F_1, \ldots, F_r$. We shall construct a labeled graph G such that B is satisfiable iff $LAP\,[G]$ is solvable.

Consider the graph H of Fig. 2. The $LAP\,[H]$ is solvable by any arrangement of the nodes of H so that M and M' occupy the $n+1$-st and $n+2$-nd position. Thus any solution of the LAP corresponds to a partition of $S=\{x_i, \bar{x}_1 \mid i=1, \ldots, n\}$, into sets P and Q of size n, with M, M' serving as endmarkers.

The Graph G will contain $n+r+1$ copies of H (denoted by $H_1, \ldots, H_{n+r+1}$) and $n+r$ other vertices A_i, $i=1, \ldots, n+r$. Each A_i is joined with both M and M' of H_i and H_{i+1} by edges of length $n+2$. Note that, so far, the $LAP\,[G]$ can be solved by arranging H_1, then A_1, then $H_2, \ldots, A_{n+r} H_{n+r+1}$, where the copies H_i of H are each arranged according to some partition of S into two subsets having n elements each.

Now we connect the corresponding literals of consecutive copies of H with edges of length $2n+5$. This forces the partitions in all copies of H to be the same, although some "mobility" within the sets is permitted. We denote this unique partition by $S=P \cup Q$ assuming that Q is the set of literals in H_i nearest to A_i (Fig. 3).

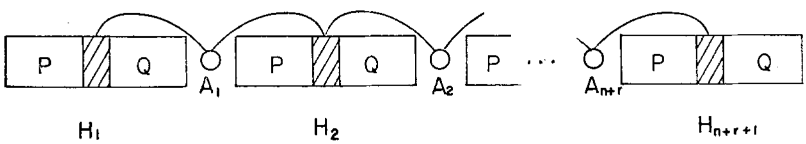


Fig. 3

The first n copies of H force the partition to be consistent. This is done by joining A_i to the copies of x_i, $\bar{x}_i$ in H_i by edges of length $n+3$. Suppose that both x_i and $\bar{x}_i$ are in P for some i, or equivalently, that the partition is inconsistent. Then A_i is at least $n+3$ away from at least one of x_i, $\bar{x}_i$ in H_i, hence this is not an acceptable solution to the LAP.

In the next r copies of H the partition $P \cup Q$ (rather the induced truth assignment t) is forced to satisfy B. Assume that the literals in P are given the value 0. The node A_{n+i} is connected with the copies in H_{n+i} of all 3 literals occurring in F_i with edges of length $n+4$. If t does not satisfy B, then for some $j \leq r$ $F_j(t)=0$ i.e. all the 3 literals of F_j are in P, hence in the $n+j$-th copy of H at least one of these literals will be $n+5$ or more away from A_{n+j}. On the other hand, if t satisfies B, a partition can be found so that t is induced by it; in the $n+r$ copies of H the literals can be arranged (due to the mobility mentioned above) so that the $LAP\,[G]$ is solved.

Hence B is satisfiable iff $LAP\,[G]$ is solvable. Moreover the constructions employed in the proof are all of polynomial time complexity. ||

In the construction of G for the proof of Theorem 1 there are some permissible variations.

Instead of 2 nodes M and M' in the i-th copy of H, a block of $m_i \geq 2$ nodes can be used. All of them are connected with A_i and $m_i' \leq m_i$ of them are connected with A_{i-1} for $i > 1$. Similarly, instead of A_i we can use a block of a_i nodes. All of them are connected with the m_{i+1}' nodes of H_{i+1}, $a_i' \leq a_i$ of them are connected with all m_i nodes of H_i and $a_i'' \leq a_i$ with the literals in H_i, according to the construction of G. Also m_i' nodes of the M-block of H_i are connected with all nodes in A_i and a_i' nodes of A_i are connected with all m_{i+1} nodes in H_{i+1}.

The length of the edges is modified the obvious way so that for a solution of the LAP [G] all positions are uniquely determined up to permutations within blocks.

Also define the complete graph K_p with p vertices and length of edges $p-1$. If K_p is a subgraph of any graph G and p is greater than all lengths of edges in G, it is clear that, in any solution of the LAP [G], K_p will occupy the first (or last) p places in the arrangement of the nodes. We can replace the $(n+r+1)$-st copy of H (whose only function was to keep A_{n+r} in the right place) with K_p; all but $a_{n+r}+1$ nodes of K_p are connected with the a_{n+r} nodes of A_{n+r} with edges of length $p-1$. Another copy of K_p can be placed before H_1 by connecting all m_1 nodes of H_1 with $p-n-m_1-1$ nodes of this copy of K_p with edges of length $p-1$.

What is worth noticing is that, if the parameters have the values

$$m_i = 4,\ a_i' = 2 \qquad 1 \leq i \leq n$$

$$m_i = 3,\ a_i' = 3 \qquad n+1 \leq i \leq n+r$$

$$m_i' = 2,\ a_i = 3,\ a_i'' = 1 \ \text{ for all } i$$

$$p = 2n+8$$

the resulting graph has labels with values either $n+4$ or $2n+7$. Hence we have proved:

Corollary: *The Restricted LAP is NP-Complete.* ||

Note that the graph G constructed in the proof of the Corollary has number of nodes n and parameter b satisfying the inequality

$$N > 2b^2 + 2b \tag{1}$$

given that $r \geq 5$, which is a trivial constraint. In fact inequality (1) can be included in the definition of the restricted LAP, and this problem is still NP-Complete.

In addition to the graph K_p defined above, we define K_p^q for $p \leq q$ to be the graph with vertices $\{v_1, \ldots, v_q\}$ and edges $\{(v_i, v_j) : |i-j| < p\}$. The length of all edges is $p-1$.

Like K_p, if the graph K_p^q is a subgraph of any graph G and q is larger than any length of edge in G, in any solution of the LAP [G] K_p^q will be placed at one end of the arrangement of the nodes.

Lemma 3: *Any Restricted LAP with $m > 0$ edges of length $2b-1$ can be transformed in polynomial time into a Restricted LAP with $m-1$ edges of length $2b'-1$, where b' is the parameter of the new Restricted LAP.*

Proof: Let $G=(V, E, l)$ with $\| V \| = N$ and let $(i,j) \in E$ with $l((i,j)) = 2b-1$. Let $k = [N/b]$..

We obtain G' from G by the following transformation

1. Add a node c and the edges (i, c), (j, c) with length $b+1$. Delete the edge (i, j).
2. Increase the length of all edges in G from b to $b+1$, and from $2b-1$ to $2b+1$.
3. Add two copies of the Graph K_{b+1}^{2b+1} with nodes $\{v_1. v_2, \dots\}$ and $\{v'_1, v'_2, \dots\}$ resp.
4. Add the "chains"

$$C = \{c = c_0, c_1, \dots, c_k = v_1\}, \quad C' = \{c = c'_0, c'_1, \dots, c'_k = v'_1\}$$

with edges (c_{i-1}, c_i), (c'_{i-1}, c'_i) of length $b+1$, $i = 1, \dots, k$.

Suppose that $LAP\,[G']$ is solvable. Then the two copies of K_{b+1}^{2b+1} are at the two ends of the arrangement and hence the two chains C, C' traverse G. Consequently in any interval of length $b+1$ there is at least one node of these chains. Hence by simply removing the two chains we obtain a solution to the $LAP\,[G]$.

Now assume that the $LAP\,[G]$ has a solution. We can assume, that in worst case, i and j are exactly $2b-1$ places apart. Then a solution of the $LAP\,[G']$ is obtained by placing c in the middle between i and j and ordering the rest of the chains exactly $b+1$ apart until the whole graph G is traversed. The copies of K_{b+1}^{2b+1} are placed at the two ends of the arrangement.

Consequently the $LAP\,[G]$ is solvable iff the $LAP\,[G']$ is solvable. Moreover the graph G' can be constructed in polynomial time and has exactly $m-1$ edges of length $2b+1$. ||

Theorem 2: *The BMP is NP-complete.*

Proof: By reducing the Restricted *LAP* to it.

We can apply the construction of Lemma 3 m times so that we finally obtain a graph with no $2b-1$ edges. Hence it remains to prove that this construction is polynomial.

Suppose that the graph G has N nodes. Then G' has $f[N] = N + 2\left[\frac{N}{b}\right] + 4b + 1$ nodes. Since we are particularly interested in the graph resulting in the proof of the Corollary, we can assume that, by (1)

$$f[N] < N\left[1 + \frac{4}{b}\right].$$

Note that (1) will hold in G', too. Also note that $m < N^2$ and that the construction of G', given G, as in Lemma 3, can be done in less than N^2 steps. Denote by $f^L\,[N]$ the number of nodes of the graph resulting after L applications of the construction of Lemma 3. Then the computational complexity T of the whole construction is bounded by

$$T < \sum_{i=0}^{m-1} (f^i\,[N])^2 \tag{2}$$

For $k \geq 4$ we have

$$f^k(N) < N \prod_{i=0}^{k-1} \left(1 + \frac{4}{b+i}\right) = \frac{N(b+k-1)(b+k-2)(b+k-3)(b+k-4)}{b(b+1)(b+2)(b+3)} \leq N\,k^4$$

Consequently for all k $0 \leq k \leq m-1$

$$f^k[N] < N^9$$

and hence by (2)

$$T < m(N^9)^2 < N^{20}.$$

Hence the construction requires a polynomial number of steps and the Theorem is proved. ||

Acknowledgement

I am grateful to Professor K. Steiglitz for his encouragement, participation and many helpful remarks.

References

[1] Cuthill, E., McKee, J.: Reducing the Bandwidth of Sparse Symmetric Matrices. Proc. 24-th Nat. Conf. ACM, pp. 157—172 (1969).

[2] Cuthill, E.: Several Strategies for Reducing the Bandwidth of Matrices, in: Sparse Matrices and their Applications (Rose, D., Willoughby, R., eds.). Plenum Press 1972.

[3] Tewarson, R. P.: Sparse Matrices. Chapter 3.8. Academic Press 1973.

[4] Chen, K. Y.: Minimizing the Bandwidth of Sparse Symmetric Matrices. Computing *11*, 103—110 (1973).

[5] Chen, K. Y.: Note on Minimizing the Bandwidth of Sparse Symmetric Matrices. Computing *11*, 27—30 (1973).

[6] Harary, F.: Sparse Matrices in Graph Theory, in: Large Sparse Sets of Linear Equations (Reidl, J. K., ed.), pp. 139—150. Academic Press 1970.

[7] Karp, R. M.: Reducibility among Combinatorial Problems, in: Complexity of Computer Computations (Miller, R., Thatcher, J., eds.). Plenum Press 1972.

[8] Aho, A. V., Hopcroft, J. E., Ullman, J. D.: The Design and Analysis of Computer Algorithms. Addison-Wesley 1974.

[9] Sahni, S., Gonzalez, T.: P-Complete Problems and Approximate Solutions. Proc. 15th SWAT Symposium, 1974.

[10] Garey, M. R., Johnson, D. S., Stockmeyer, L.: Some Simplified NP-Complete Problems. Proc. 6th Annual ACM Symposium on Theory of Computing, pp. 47—63 (1974).

Ch. H. Papadimitriou
Department of Electrical Engineering
Princeton University
Princeton, NJ 08540
U.S.A.