PROFILE MINIMIZATION PROBLEM FOR MATRICES AND GRAPHS*†

LIN YIXUN (林诒勋) YUAN JINJIANG (原晋江)

(Department of Mathematics, Zhengzhou University, Zhengzhou 450052, China)

1. Introduction and Basic Results

For a simple graph G with n vertices, a bijection (1-1 mapping) $f : V \to \{1, 2, \cdots, n\}$ will be called a numbering (or labelling) of G. For a numbering f, the profile width of vertex v is defined as

$$w_f(v) = f(v) - \min_{x \in N^*(v)} f(x),$$

where $N^*(v) = \{x \in V \mid x = v \text{ or } (x, v) \in E(G)\}$ is the closed neighbor set of v. The profile of numbering f for G is defined as

$$P_f(G) = \sum_{v \in V} w_f(v) = \sum_{v \in V} \Big(f(v) - \min_{x \in N^*(v)} f(x)\Big).$$

Finally, the profile of G is the minimum value

$$P(G) = \min_f P_f(G),$$

where f runs through all numberings of G. A numbering f that attains the minimum will be called an optimal numbering.

In the area of numerical analysis, a number of profile, as well as bandwidth, reduction algorithms have been developed[1-4]. In graph theory, there has been a strong interest in the bandwidth problem[5,6], but less concern in the profile problem so far.

A class of graphs, called the interval graphs, will play an important role in our study. Let $J_1, J_2, \cdots, J_n$ be intervals on a line. We define a graph G with vertex set $\{J_1, J_2, \cdots, J_n\}$, called an interval graph, by connecting two intervals if and only if they have a point in common. The characterization of interval graphs can be found in the literature (e.g. [7]). Here is one.

Lemma 1.1[7]. A graph G is an interval graph if and only if it does not contain any of the following graphs as an induced subgraph.

*Received January 17, 1991. Revised June 25, 1992.

†This project is supported by the National Natural Science Foundation of China.

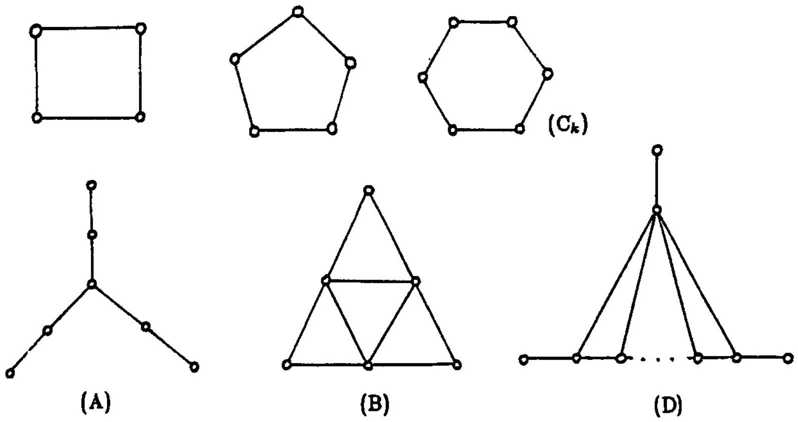

Fig. 1

From this, it is easily seen that a tree T is an interval graph if and only if it is a caterpillar (i.e. a tree which yields a path when all its pendant vertices are removed).

We have another characterization of interval graphs as follows.

Lemma 1.2. A graph G is an interval graph if and only if there exists a numbering f such that if $f(x) < f(y) < f(z)$ and $xz \in E(G)$ then $y, z \in E(G)$.

Theorem 1.3. For any graph G, $P(G) \geq |E(G)|$; and $P(G) = |E(G)|$ if and only if G is an interval graph.

Corollary 1.3.1. For any tree T, $P(T) \geq n-1$; and $P(T) = n-1$ if and only if T is a caterpillar.

Theorem 1.4. For any graph G, $P(G)$ is the minimum number of edges of an interval supergraph of G.

This theorem leads to a problem of extremal graph theory determining the minimum number of edges of an interval supergraph. This is called "the interval graph completion problem", which is NP-complete in general (see [9, p.198]). However, in some special graphs, we can obtain exact solutions.

2. Profile of Special Graphs

In this section, we shall derive some lower bounds of the profile, and determine $P(G)$ for special graphs.

For a given numbering f, the profile width $w_f(v) = f(v) - \min_{x \in N^*(v)} f(x)$ can be regarded as the weight of vertex v. Further, the weight of subgraph H of G can be defined as

$$w_f(H) = \sum_{v \in V(H)} w_f(v).$$

Especially, $P_f(G) = w_f(G)$.

Lemma 2.1. Let P be a (u,v)-path in G with $f(u) < f(v)$. Then

$$w_f(P) \geq f(v) - f(u) + w_f(u).$$

Proof. Suppose that $P = v_0 v_1 v_2 \cdots v_k$ where $v_0 = u$, $v_k = v$. From

$$w_f(v_i) = f(v_i) - \min_{x \in N^*(v_i)} f(x) \geq f(v_i) - f(v_{i-1}),$$

it follows that

$$w_f(P) = \sum_{i=0}^{k} w_f(v_i) \geq \sum_{i=1}^{k} [f(v_i) - f(v_{i-1})] + w_f(v_0)$$
$$= f(v_k) - f(v_0) + w_f(v_0).$$

Lemma 2.2. For any vertex $v \in V(G)$, $P(G-v) \leq P(G) - d_G(v)$.

Theorem 2.3. If G is k-connected, then $P(G) \geq \frac{k}{2}(2n-k-1)$.

Proof. For a given numbering f, suppose that $f(u_0) = 1$, $f(v_1) = n$, $f(v_2) = n-1, \cdots, f(v_k) = n-k+1$. We add a new vertex v_0 to G, and join v_0 to $v_1, v_2, \cdots, v_k$. Since G is k-connected, so is $G + v_0$. By Menger's theorem ([8], Theorem 11.7), there are k internally-disjoint (u_0, v_0)-paths in $G + v_0$. Namely, we have (u_0, v_i)-paths $P^{(i)}$ in G, $i = 1, 2, \cdots, k$, such that they have only one vertex u_0 in common. By Lemma 2.1,

$$w_f(P^{(i)}) - w_f(u_0) \geq f(v_i) - f(u_0) = n - i, \qquad 1 \leq i \leq k.$$

Noticing that $w_f(u_0) = 0$, we have

$$P_f(G) \geq \sum_{i=1}^{k} w_f(P^{(i)}) \geq \sum_{i=1}^{k} (n-i) = \frac{k}{2}(2n-k-1).$$

Corollary 2.3.1. Let C_n be a cycle with n vertices. Then $P(C_n) = 2n-3$.

Corollary 2.3.2. Let W_n be a wheel with n vertices. Then $P(W_n) = 3n-6$.

Corollary 2.3.3. For the complete bipartite graph $K_{m,n}$, $m \leq n$, $P(K_{m,n}) = mn + \frac{1}{2}m(m-1)$.

This result can be generalized to a decomposition theorem as follows. The join $G_1 \vee G_2$ of disjoint graphs G_1 and G_2 is a graph obtained from $G_1 \cup G_2$ by joining each vertex of G_1 to each vertex of G_2.

Theorem 2.4. If $G = G_1 \vee G_2$, $|V(G_1)| = m$, $|V(G_2)| = n$, then

$$P(G) = \min\left\{P(G_1) + mn + \frac{1}{2}n(n-1),\ P(G_2) + mn + \frac{1}{2}m(m-1)\right\}.$$

3. A Result on Trees with $D(T)=4$

We consider a tree T with diameter 4, as shown in Figure 2. Here, denote the center of T by v_0; and denote the other non-pendant vertices by $v_1, v_2, \cdots, v_k$ $(k \geq 2)$. Each v_i and its neighbors constitute a star, called the v_i-branch (relative to the center v_0). Note that the pendant edges incident to v_0 will not be called branches.

Let $d(v_i)$ be the degree of v_i. Then there are $d(v_i) - 1$ leaves (pendant vertices) in the v_i-branch $(1 \leq i \leq k)$. We may assume that

$$d(v_1) \geq d(v_2) \geq \cdots \geq d(v_k) \geq 2. \tag{1}$$

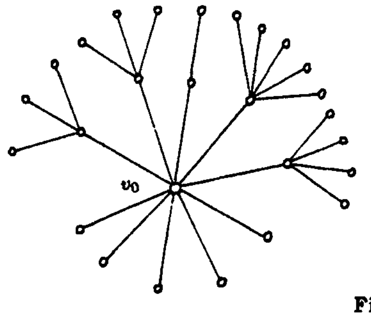

Fig. 2

Theorem 3.1. For a tree T with diameter 4,

$$P(T) = |E(T)| + \sum_{i=3}^{k} (d(v_i) - 1). \tag{2}$$

Proof. When $k = 2$, T is a caterpillar; by Corollary 1.3.1, the conclusion holds. Assume now that $k \geq 3$. Let G be an interval supergraph of T with a minimum number of edges. Denote $E^* = E(G)\backslash E(T)$, the set of additional edges. By Lemma 1.1, G does not contain the forbidding graphs C_k ($k \geq 4$), A, B and D as an induced subgraph. Among them, the only one tree is A. For any subtree A of T, there must be some additional edges in the induced subgraph $G[A]$.

Case 1. No additional edges cross two branches. In other words, every additional edge is from the center v_0 to a leaf of some v_i-branch. We can see that there are at least $k-2$ branches in which all leaves are connected by additional edges. Otherwise, there would be a forbidding graph A in G. On the other hand, by the minimality of E^*, there are exact $k-2$ branches having such a property. Also, by the minimality of E^* and the assumption (1), the two exceptional branches must be those of v_1 and v_2. Hence $|E^*| = \sum_{i=3}^{k}(d(v_i)-1)$.

Case 2. There are additional edges crossing two branches.

Case 2.1. An edge $e \in E^*$ connects two leaves x, y of distinct branches, as shown in Figure 3(a) (the additional edges are depicted by dotted lines).

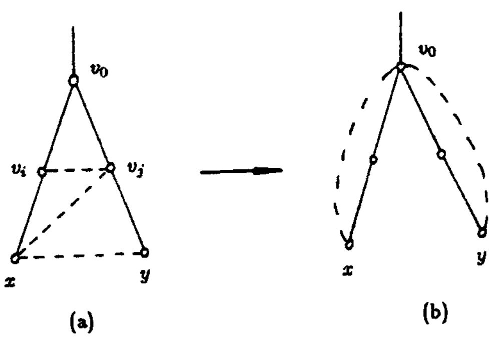

Fig. 3

Since C_k ($k \geq 4$) are forbidding graphs, there must be two chords in the cycle $v_0v_ixyv_jv_0$. That is, there are three additional edges within these two branches. We may change these edges in (a) into the ones in (b). After this local transformation, the resulting graph G' is still an interval supergraph of T. This contradicts the minimality of G. Therefore, this case is impossible.

Case 2.2. An edge $e \in E^*$ joins some v_i and a leaf x of another branch, as shown in Figures 4(a). Similarly to the previous case, we may change (a) to (b) as follows.

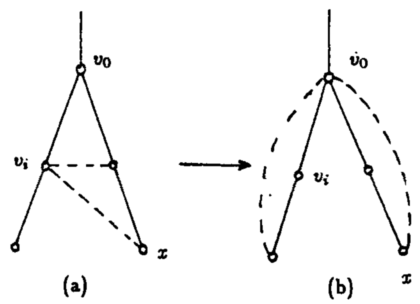

Fig. 4

After that, the resulting graph G' is still interval. And this reduces to Case 1.

Case 2.3. An edge $e \in E^*$ joins some v_i and v_j $(1 \leq i < j \leq k)$. Suppose that Case 2.1 and 2.2 do not occur. We have the following situations:

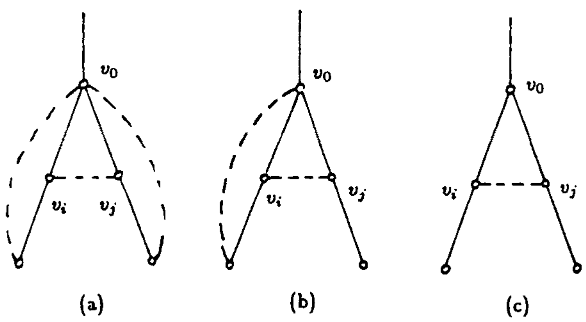

Fig. 5

In the cases of (a) and (b), the additional edge $e = v_iv_j$ can be deleted from G. It is easy to see that $G - e$ is still an interval graph (by Lemma 1.1). This contradicts the minimality of G. In the case of (c), there will be a forbidding graph D (of Figure 1) in G. Any way, Case 2.3 is impossible.

References

[1] R.P. Tewarson, Sparse Matrices, Academic Press, New York, 1973.

[2] O.C. Zienkiewicz, The Finite Element Method in Engineering Science, McGraw–Hill, London, 1971.

[3] 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.*, 13 (1976), 235–251.

[4] G.C. Everstine, A Comparison of Three Resequencing Algorithm for the Reduction of Matrix Profile and Wave-Front. *Internat. J. Number. Methods in Enge.*, 14 (1979), 837–863.

[5] P.Z. Chinn, J. Chvatalova, A.K. Dewdney, N.E. Gibbs, The Bandwidth Problem for Graphs and Matrices—a Survey, *J. Graph Theory*, 6 (1982), 223–254.

[6] F.R.K. Chung, Labelling of Graphs, in: Selected Topics in Graph Theory, Vol. 3 (L.W. Beineke and R.J. Wilson, Eds), 1988, 151–168.

[7] L. Lovasz, Perfect Graphs, in: Selected Topics in Graph Theory, Vol. 2 (L.W. Beineke and R.J. Wilson, Eds), 1983, 55–88.

[8] J.A. Bondy and U.S.R. Murty, Graph Theory with Applications, Macmillan Press, New York, 1976.

[9] M.R. Garey and D.S. Johnson, Computers and Intractability: A Guide to the Theory of NP-completeness, Freeman, San Francisco, CA, 1979.