Linear algorithms for w-medians of graphs^{*}

Hai-Yen Lee
Department of International Trade
Chung Kuo Institute of Technology and Commerce
56, Section 3, Hsing-Lung Road
Wen-Shan District, Taipei, Taiwan

Gerard J. Chang
Department of Applied Mathematics
National Chiao Tung University
Hsinchu 30050, Taiwan
Email: gjchang@math.nctu.edu.tw

Abstract

Suppose G = (V, E) is a graph in which every vertex v has a nonnegative real number w(v) as its weight. The w-distance sum of v is $D_{G,w}(v) = \sum_{u \in V} d(v,u)w(u)$. The w-median $M_w(G)$ of G is the set of all vertices v with minimum w-distance sum $D_{G,w}(v)$. This paper gives linear-time algorithms for computing the w-medians of interval graphs and block graphs.

Dedicated to Prof. Stephen T. Hedetniemi on the occasion of his 60th birthday.

1 Introduction

The concept of center and median arise from facility location problems, which deal with the job of choosing a site subject to certain criterion. These distance related concepts have been extensively studied, see the book by Buckley and Harary [4]. In particular, algorithms have been developed for

^{*}Supported in part by the National Science Council under grant NSC81-0208-M-009-26.

them, see [2, 5, 8, 10, 11, 12, 16, 17, 23, 24]. The purpose of this paper is to study w-medians of graphs from an algorithmic point of view.

All graphs in this paper are simple, i.e., undirected, loopless and without multiple edges. In a graph G = (V, E), the distance $d_G(u, v)$ between two vertices u and v is the minimum number of edges in a u-v path; $d_G(u, v) = \infty$ if there is no u-v path. Suppose every vertex v has a non-negative real number w(v) as its weight. The w-distance sum of a vertex v in G is

$$D_{G,w}(v) = \sum_{u \in V} d_G(v,u)w(u).$$

The w-median $M_{\mathbf{w}}(G)$ of G is the set

$$M_{w}(G) = \{v \in V : D_{G,w}(v) \leq D_{G,w}(u) \text{ for all } u \in V\}.$$

For the case in which w(v) = 1 for all vertices v, the w-distance sum $D_{G,w}(v)$ is called the distance sum $D_G(v)$ and the w-median $M_w(G)$ is the median M(G).

Slater [22] showed that for every (not necessarily connected) graph H there exists a graph G such that H is the subgraph of G induced by the median M(G). Lee and Chang [14] generalized this result to w-medians. Zelinka [27] showed that the median of a tree is a clique of size one or two. This also follows from a more general result obtained by Truszczyński [25] that says the median of a connected graph lies in a block of the graph. Slater [22] showed that the median of a 2-tree is a clique of size at most three. Nieminen [18] and Yushmanov [26] proved that the median of a Ptolemaic graph is a clique. Lee and Chang [13] showed that the w-median of a strongly chordal graph is a clique if the weight function w is positive. Note that trees are block graphs, block graphs are Ptolemaic, and Ptolemaic graphs are strongly chordal.

In general, the w-median of a graph G=(V,E) can be computed by finding the distances between all pairs of vertices. A standard breadth-first-search, which costs O(|V||E|) time, does the job. This paper employs the idea of the proof in [13] to obtain an algorithm for finding the w-median of a strongly chordal graph. This algorithm is then adapted to linear-time algorithms for the w-median problem in interval graphs and block graphs, which are both strongly chordal.

2 Strongly chordal graphs

A graph is *chordal* (or *triangulated*) if every cycle with more than three vertices has a *chord*, i.e., an edge joining two non-contiguous vertices of the cycle. A *p-sun* is a chordal graph with a vertex set $\{x_1, x_2, ..., x_p, y_1, y_2, ..., y_p\}$

such that $\{y_1,y_2,...,y_p\}$ is an independent set, $(x_1,x_2,...,x_p,x_1)$ is a cycle and each vertex y_i has exactly two neighbors x_{i-1} and x_i , where $x_0=x_p$. A graph G is strongly chordal if it is chordal and contains no p-sun for $p\geq 3$. A vertex v is simple if for any two vertices $x,y\in N_G[v]$ either $N_G[x]\subseteq N_G[y]$ or $N_G[y]\subseteq N_G[x]$, where $N_G(x)=\{y:xy\in E(G)\}$ and $N_G[x]=\{x\}\cup N_G(x)$. Note that if v is a simple vertex, then G-v is a distance-invariant induced subgraph of G, i.e., $d_{G-v}(x,y)=d_G(x,y)$ for all vertices x and y in G-v. A maximal neighbor of a simple vertex v is a vertex $x\in N_G[v]$ such that $N_G[y]\subseteq N_G[x]$ for all $y\in N_G[v]$. Farber [6] proved that every strongly chordal graph that is not a complete graph has two non-adjacent simple vertices. Furthermore, a graph G=(V,E) is strongly chordal if and only if it has a simple ordering, i.e., an ordering (v_1,v_2,\cdots,v_n) of V such that v_i is a simple vertex of the graph $G-\{v_1,v_2,\cdots,v_{i-1}\}$. The main result of [13] is:

Theorem 1 [13] The w-median of a connected strongly chordal graph is a clique if w is a positive weight function.

We first sketch the proof of Theorem 1. In order to prove Theorem 1, [13] introduced the following more general concept. Every vertex v in G = (V, E) has a non-negative weight w(v) and a non-negative cost c(v). The cost w-distance sum of v (with respect to w and c) is

$$D_{G,w,c}(v) = \sum_{u \in V} d_G(v,u)w(u) - c(v).$$

The cost w-median $M_{w,c}(G)$ of G is

$$M_{w,c}(G) = \{ v \in V : D_{G,w,c}(v) \le D_{G,w,c}(u) \text{ for all } u \in V \}.$$

It is easy to see that $M_{w,c}(G) = M_w(G)$ when c(v) = 0 for all vertices $v \in V$. However, unless we choose c properly, it is not the case that one can modify Theorem 1 to get a cost w-median result for an arbitrary cost function c. The proof of Theorem 1 requires an inductive approach starting with a connected strongly chordal graph G = (V, E) with a positive weight function w and a cost function $c \equiv 0$. For induction, it uses the following two terms.

First, for any vertex $x \in V$, there exists a set $S_x = \{x \equiv x_0, x_1, \dots, x_{n(x)}\}$ $\subseteq N_G[x]$ such that (C1) and (C2) hold.

(C1)
$$N_G[x_0] \subseteq N_G[x_1] \subseteq \cdots \subseteq N_G[x_{n(x)}].$$

(C2) If
$$c(x) > 0$$
, then $n(x) \ge 1$ and $c(x) < \sum_{i=1}^{n(x)} w(x_i)$ and $c(x) < \sum_{i=1}^{j} w(x_i) + c(x_j)$ for $1 \le j \le n(x) - 1$.

Initially, each $S_x = \{x \equiv x_0\}$ and so conditions (C1) and (C2) hold. Secondly, it uses a poset (partially ordered set) P whose elements are precisely the vertices of G and y < z in P if $y = x_i$ and $z = x_j$ for some $x \in V$ with $0 \le i < j \le n(x)$. Note that P is not necessarily a poset if $\{S_x : x \in V\}$ is not chosen properly. However, initially each $|S_x| = 1$ and so P is simply a poset in which each pair of distinct elements is incomparable.

Theorem 1 is clear when G is a complete graph. Suppose G is not a complete graph. Choose a pair of non-adjacent simple vertices u and v (see [6]). Without loss of generality, we may assume that u and v are chosen so that they are as small as possible in the poset P. Suppose u is not a minimal element in P. Then $u=x_i$ for some $x\in V$ and $x_i\in S_x$ with $i\geq 1$. Since $N_G[x_0]\subseteq N_G[x_i]$ and x_i is a simple vertex not adjacent to v, x_0 is also a simple vertex not adjacent to v. But then x_0 is smaller than $u=x_i$ in P, which contradicts our choice of u. So, u is a minimal element in P. Similarly, v is minimal in P. Without loss of generality, we may assume that $w(v)+c(v)\leq w(u)+c(u)$.

Now choose a maximal neighbor m of v in G. Without loss of generality, we may assume that m is chosen so that it is as large as possible in the poset P. Suppose m is not a maximal element in P. Then $m = x_i$ for some $x \in V$ and $x_i \in S_x$ with i < n(x). Since $N_G[x_i] \subseteq N_G[x_{n(x)}]$ and x_i is a maximal neighbor of v, $x_{n(x)}$ is also a maximal neighbor of v. But then $x_{n(x)}$ is larger than $m \equiv x_i$ in P, which contradicts our choice of m. So, m is maximal in P.

Keeping all these results in mind, we now consider the distance-invariant subgraph G - v of G, denoted by G' = (V', E'), which is also connected strongly chordal. We define the new weight function w' and the new cost function c' on V' as:

$$w'(x) = w(x) + w(v)$$
 if $x = m$ and $w'(x) = w(x)$ otherwise,

$$c'(y) = c(y) + w(v)$$
 if $y \in N_G(v) - \{m\}$ and $c'(y) = c(y)$ otherwise.

It remains true that w' is positive and c' is non-negative. We also update $\{S_x : x \in V\}$ and P as follows. Since m is a maximal element of the poset P, for any vertex $x \in V$, either $m \notin S_x$ or $m = x_{n(x)}$. Let

$$S'_x = S_x \cup \{m\}$$
 if $x \in N_G(v)$ with $m \notin S_x$ and $S'_x = S_x$ otherwise.

Now the poset P' contains elements of V'. Some new relationships are also added to P' when $S'_x = S_x \cup \{m\}$ for some $x \in V'$. However, since m is a maximal element in P, P' remains a poset even when new relationships are added to it.

Theorem 1 then follows from induction and the following lemmas.

Lemma 2 [13]
$$M_{w,c}(G) = M_{w',c'}(G')$$
.

Lemma 3 [13] $\{S'_x : x \in V'\}$ satisfies conditions (C1) and (C2) of G'.

Lemma 4 If G is a complete graph, then $M_{w,c}(G) = \{u \in V : w(u) + c(u) \geq w(v) + c(v) \text{ for all } v \in V\}.$

Proof. Lemma 4 follows from that fact that for any vertices v and u of G,

$$D_{w,c}(v) - D_{w,c}(u) = \left(\sum_{x \neq v} w(x) - c(v)\right) - \left(\sum_{x \neq u} w(x) - c(u)\right)$$
$$= (w(u) + c(u)) - (w(v) + c(v)).$$

Q.E.D.

To implement the idea of the proof for Theorem 1, we need not to keep the sets S_x and the poset P. Instead, a "mark" is given to each vertex v that ensures the maximal neighbor of a simple vertex v is the last deleted vertex in $N_G(v)$. More precisely, we have the following algorithm.

Algorithm MS. Compute the w-median of a connected strongly chordal graph.

Input: A connected strongly chordal graph G = (V, E) in which every vertex v has a positive weight w(v).

Output: The w-median $M_w(G)$.

Method:

```
begin
```

```
c(v) \leftarrow 0 for all v \in V;

mark(v) \leftarrow 0 for all v \in V;

while \ V is not a clique do

find two non-adjacent simple vertices u and v in G with the smallest

mark(u) + mark(v), W.L.O.G., assume w(v) + c(v) \leq w(u) + c(u);

choose a maximal neighbor m of v with the largest mark(m);

w(m) \leftarrow w(m) + w(v);

c(y) \leftarrow c(y) + w(v) for all y \in N(v) - \{m\};

mark(m) \leftarrow max\{mark(y) : y \in N(v)\} + 1;

G \leftarrow G - v;

enddo;

M_w(G) \leftarrow \{u \in V : w(u) + c(u) \geq w(v) + c(v) \text{ for all } v \in V\};

output(M_w(G))

end.
```

Since finding two simple vertices during any iteration is costly for a general strongly chordal graph, the time complexity of this algorithm is greater than O(|V||E|). However, in the next two sections, we shall modify the algorithm to obtain linear-time algorithms for the problem in interval graphs and block graphs.

3 Interval graphs

A graph G=(V,E) is an *interval graph* if there exists a family $\{I_v:v\in V\}$ of intervals such that two distinct vertices u and v are adjacent in G if and only if $I_u\cap I_v\neq\emptyset$; such a family $\{I_v:v\in V\}$ is referred to as an *interval representation* of G.

Gilmore and Hoffman [7] showed that a graph is an interval graph if and only if its maximal cliques can be linearly ordered into C_1, C_2, \cdots, C_m , such that for every vertex v, the maximal cliques containing v occur contiguously. Suppose for every vertex v, i_v (respectively, j_v) is the minimum (respectively, maximum) index i such that $v \in C_i$. $\{[i_v, j_v] : v \in V\}$ is then an interval representation of G, which we call a canonical representation. Booth and Lueker [3] gave an O(|V| + |E|)-time algorithm for an arbitrary graph G = (V, E) that tests whether G is an interval graph. In the case in which G is an interval graph, the algorithm also gives an ordering $C_1, C_2, ..., C_m$ of its maximal cliques, and so, a canonical interval representation.

Roberts [21], Ramalingam and Pandu Rangan [20], and Olariu [19] gave another characterization in which a graph G is an interval graph if and only if an interval ordering (v_1, v_2, \dots, v_n) of V exists such that i < j < k, and $v_i v_k \in E$ imply $v_j v_k \in E$, or equivalently, $i \leq j \leq k$ and $v_i \in N_G[v_k]$ imply $v_j \in N_G[v_k]$.

Lemma 5 Any interval ordering (v_1, v_2, \dots, v_n) of a graph G is a simple ordering. Consequently, an interval graph is a strongly chordal graph.

Proof. We only need to show that v_1 is a simple vertex of G. Suppose $v_i, v_j \in N_G[v_1]$. Assume $i \leq j$. We shall prove that $N_G[v_i] \subseteq N_G[v_j]$. Suppose $v_k \in N_G[v_i]$. For the case in which $j \leq k$, since $i \leq j \leq k$ and $v_i \in N_G[v_k]$, $v_j \in N_G[v_k]$ and so $v_k \in N_G[v_j]$. For the case in which k < j, since $1 \leq k \leq j$ and $v_1 \in N_G[v_j]$, $v_k \in N_G[v_j]$. Thus, v_1 is a simple vertex of G.

Q.E.D.

Suppose $\{I_v=[a_v,b_v]:v\in V\}$ is an interval representation of an interval graph G. Sort the right end-point b_v 's of the intervals I_v 's into $b_{v_{r(1)}}\leq b_{v_{r(2)}}\leq \cdots \leq b_{v_{r(n)}}$. It is straightforward to check that $R=(v_{r(1)},v_{r(2)},\cdots,v_{r(n)})$ is an interval ordering of G. Similarly, if we sort the left end-point a_v 's into $a_{v_{l(1)}}\geq a_{v_{l(2)}}\geq \cdots \geq a_{v_{l(n)}}$, then $L=(v_{l(1)},v_{l(2)},\cdots,v_{l(n)})$ is an interval ordering of G. Note that if the interval representation is canonical, then we can use bucket sorts to sort the interval end-points allowing E and E to be computed in linear time. Suppose E is the graph E and E to be computed in linear time. Suppose E is the graph E and E to be computed in linear time. Suppose E is the graph E and E to be computed in linear time. Suppose E is the graph E and E is the suppose of E in the sum of E and the following lemma, can be used for efficient implementation of Algorithm MS for interval graphs.

Lemma 6 If $v_{r(i)}$ and $v_{l(j)}$ are in G_{ij} and $v_{r(i)} \in N_{G_{ij}}[v_{l(j)}]$, then G_{ij} is a complete graph.

Proof. Since $I_{v_{r(i)}}$ $(I_{v_{l(j)}})$ has the smallest (largest) right (left) endpoints among all vertices in G_{ij} and $v_{r(i)} \cap v_{l(j)} \neq \emptyset$, $a_{v_{r(i)}} \leq a_{v_{l(j)}} \leq b_{v_{r(i)}} \leq b_{v_{l(j)}}$. If $x = v_{r(i')} = v_{l(j')}$ is a vertex in G_{ij} , then $a_x \leq a_{v_{l(j)}} \leq b_{v_{l(j)}} \leq b_x$ and so I_x contains $a_{v_{l(j)}}$. Thus G_{ij} is a complete graph. Q. E. D.

We are now able to modify Algorithm MS to get a linear-time algorithm for the w-median of an interval graph.

Algorithm MI. Compute the w-median of a connected interval graph. Input: A connected interval graph G = (V, E) in which every vertex v has a positive weight w(v). Two interval orderings R and L as above. Output: The w-median $M_w(G)$ of G.

```
begin
```

Method:

```
i \leftarrow 1;
     j \leftarrow 1;
     c(v) \leftarrow 0 \text{ for all } v \in V;
     while v_{r(i)} \notin N_{G_{ij}}[v_{l(j)}] do
          if w(v_{r(i)}) + c(v_{r(i)}) \le w(v_{l(j)}) + c(v_{l(j)}) then
                choose a maximal neighbor m of v_{r(i)};
                w(m) \leftarrow w(m) + w(v_{r(j)});
                c(y) \leftarrow c(y) + w(v_{r(j)}) for all y \in N(v_{r(j)}) - \{m\};
                G \leftarrow G - v_{r(i)};
         else
                choose a maximal neighbor m of v_{l(i)};
                w(m) \leftarrow w(m) + w(v_{l(i)});
                c(y) \leftarrow c(y) + w(v_{l(j)}) for all y \in N(v_{l(j)}) - \{m\};
                G \leftarrow G - v_{l(i)};
         endif;
         while v_{r(i)} not in G do i \leftarrow i+1;
         while v_{l(j)} not in G do j \leftarrow j+1;
     enddo;
    M_w(G) \leftarrow \{u \in V : w(u) + c(u) \ge w(v) + c(v) \text{ for all } v \in V\};
    \operatorname{output}(M_w(G))
end.
```

4 Block graphs

The concept of a block graph was introduced by Harary [9], who defined the block graph B(G) of a graph G as the intersection graph of blocks of G.

He then proved that a graph is the block graph of some graph if and only if all of its blocks are complete graphs. So, we may define a *block graph* as a graph whose blocks are complete graphs.

A graph with one or more cut-vertices contains at least two blocks, each of which contains exactly one cut-vertex; we call them end blocks (see [1]). If a graph has vertices u and v that are not in the same block, then any path from u to v must pass through a unique sequence of blocks B_1, B_2, \ldots, B_n , where B_i and B_{i+1} , $i = 1, 2, \ldots, n-1$, have a common cut-vertex that is a vertex of the path. Moreover, for any graph G containing m blocks B_1, B_2, \ldots, B_m and n cut-vertices c_1, c_2, \ldots, c_n , consider the graph $G^* = (V^*, E^*)$, which we call the block-cut-vertex structure of G, where

$$V^* = \{B_1, B_2, \dots, B_m, c_1, c_2, \dots, c_n\} \text{ and }$$

$$E^* = \{(B_i, c_i) : 1 \le i \le m, 1 \le j \le n, c_i \in B_i\}.$$

Then G^* is a forest whose leaves are exactly the end blocks of G and whose isolated vertices are exactly those blocks without cut-vertices in G. The block-cut-vertex structure G^* of a graph G can be constructed in linear time by using a depth-first search.

An end vertex of a block graph is a vertex in some end block but is not a cut-vertex. It is easy to show that an end vertex of a block graph is a simple vertex with the cut-vertex in the end block containing it being its maximal neighbor. Consequently, block graphs are strongly chordal. Note that if the block-cut-vertex structure is found then the two non-adjacent end vertices can be found in a constant time. Therefore, we now modify Algorithm MS to get a linear-time algorithm for the w-median problem in block graphs.

Algorithm MB. Compute the w-median of a connected block graph.

Input: A connected block graph G = (V, E) in which every vertex v has a positive weight w(v) and its block-cut-vertex structure T.

Output: The w-median $M_w(G)$ of G. Method:

begin

```
c(v) \leftarrow 0 for all v \in V(G);

while |V(T)| > 1 do

find two end vertices u and v from different end blocks B_i and

B_j of T, W.L.O.G., assume w(v) + c(v) \leq w(u) + c(u);

w(m) \leftarrow w(m) + w(v) where m is the cut vertex of B_i;

c(y) \leftarrow c(y) + w(v) for all y \in B_i - \{m\};

G \leftarrow G - v;

B_i \leftarrow B_i - \{v\};

if B_i = \{m\} then
```

```
T \leftarrow T - B_i; if m is a leaf of T then T \leftarrow T - m endif; endif; enddo; M_w(G) \leftarrow \{u \in V(G) : w(u) + c(u) \geq w(v) + c(v) \text{ for all } v \in V(G)\}; output M_w(G) end.
```

References

- M. Behzad, G. Chartrand, and L. Lesniak-Foster, Graph and Digraphs, Wadworth International Group, Belmont, CA (1979).
- [2] T. Beyer, S. M. Hedetniemi, and S. T. Hedetniemi, A linear algorithm for finding the center of a unicyclic graph, Dept. of Computer Information Science, Univ. of Oregon, Tech. Rept. CIS-TR-80-12 (1980).
- [3] K. S. Booth and G. S. Lueker, Testing for the consecutive ones property, interval graphs, and planarity testing using PQ-tree algorithm, J. Comput. Systems Sci. 13 (1976) 335-379.
- [4] F. Buckley and F. Harary, Distance in Graphs, Addison-Wesley, New York (1989).
- [5] E. J. Cockayne, S. T. Hedetniemi, and S. L. Mitchell, Linear algorithms for mean distance in trees and unicyclic graphs, Dept. of Computer Information Science, Univ. of Oregon, Tech. Rept. CS-TR-79-22 (1979).
- [6] M. Farber, Characterizations of strongly chordal graphs, Disc. Math. 43 (1983) 173-189.
- [7] P. C. Gilmore and A. J. Hoffman, A characterization of comparability graphs and of interval graphs, Canadian. J. Math. 16 (1964) 539-548.
- [8] G. Y. Handler and P. B. Mirchandani, Location on Networks: Theory and Algorithms, The MIT Press, Cambridge, Mass. (1979).
- [9] F. Harary, A characterization of block graphs, Canad. Math. Bull. 6 (1963) 1-6.
- [10] S. M. Hedetniemi, E. J. Cockayne, and S. T. Hedetniemi, Linear algorithms for the Jordan centre and path centre of a tree, J. Transportation Sci. 15 (1981) 98-114.
- [11] S. M. Hedetniemi, S. T. Hedetniemi, and P. J. Slater, Centers and median of C_n -trees, Utilitas Math. 21 (1982) 225-234.

- [12] S. Hedetniemi and S. Mitchell, Centers of recursive graphs, Dept. of Computer Information Science, Univ. of Oregon, Tech. Rept. CS-TR-79-11 (1979).
- [13] H.-Y. Lee and G. J. Chang, The w-median of a connected strongly chordal graph, J. Graph Theory 18 (1994) 673-680.
- [14] H.-Y. Lee and G. J. Chang, Medians of graphs and kings of tournaments, Taiwanese J. Math. 1 (1997) 103-110.
- [15] S. C. Liaw and G. J. Chang, Wide diameters of butterfly networks, Taiwanese J. Math. 3 (1999) 83-88.
- [16] E. Minika, Optimization Algorithms for Networks and Graphs, Wiley, New York (1978).
- [17] C. A. Morgan and P. J. Slater, A linear algorithm for the core of a tree, J. Algorithms 1 (1980) 247-258.
- [18] J. Nieminen, The center and the distance center of a Ptolemaic graph, Oper. Research Letters 7 (1988) 91-94.
- [19] S. Olariu, A simple linear-time algorithm for computing the center of an interval graph, Intern. J. Computer Math. 34 (1990) 121-128.
- [20] G. Ramalingam and C. Pandu Rangan, A unified approach to domination problems on interval graphs, Inform. Process. Letters 27 (1988) 271-274.
- [21] F. S. Roberts, On the comparability between a graph and a simple order, J. Comb. Theory 11 (1971) 28-38.
- [22] P. J. Slater, Medians of arbitrary graphs, J. Graph Theory 4 (1980) 389-392.
- [23] P. J. Slater, On locating a facility to service areas within a network, Oper. Research 29 (1981) 523-531.
- [24] P. J. Slater, Locating central paths in a graph, J. Transportation Sci. 16 (1982) 1-18.
- [25] M. Truszczyński, Centers and centroids of unicyclic graphs, Math. Slovaca 35 (1985) 223-228.
- [26] S. V. Yushmanov, On metric properties of chordal graphs and Ptolemaic graph, Soviet Math. Dokl. 37 (1988) 665-668.
- [27] B. Zelinka, Medians and peripherians of trees, Arch. Math. (Brno) 4 (1968) 87-95.