A BOUND INVOLVING THE CENTROID AND WIENER INDEX OF A TREE

Thomas Dale Porter[1]

Department of Mathematics
Southern Ilinois University
Carbondale, IL 62901 USA

Abstract. In this short note it is shown that for any tree T of order n, $\frac{\sigma(T)}{\sigma_T(c)} \geq n - bw(c)$, where $\sigma(T)$ is the Wiener index of T , c is in its centroid, and $bw(c)$ is the branch weight at c. It provides a companion result to the upper bound $\frac{\sigma(T)}{\sigma_T(c)} \leq n-1$, shown by Barefoot, Entringer, and Székely.

1. Introduction.

The distance of a vertex u in a connected graph G is defined by $\sigma_G(u) = \sum\limits_{v \epsilon V(G)} d_G(u,v)$. The distance of the graph G is defined by $\sigma(G) = \frac{1}{2} \sum\limits_{u \epsilon V(G)} \sigma_G(u)$, $\sigma(G)$ is also called the Wiener index of G [5]. The problem of finding a spanning tree of minimum Wiener index is known to be NP-hard [3].

The following properties and definitions are found in many of the referenced papers. A nice summary is found in [EKS] and we include that here.

A maximal subtree containing a vertex v of a tree T as an endvertex is called a *branch* of T at v. The *weight* of a branch B, denoted $bw(B)$, is the number of edges in it. The *branch weight of a vertex* v, denoted $bw(v)$, is the maximum weight of a branch at v. The *centroid* of a tree T is the set of vertices of T with minimum branch weight $bw(v)$.

We have the following useful characteristics of the centroid.

Theorem A (Jordan [4]). *If $C = C(T)$ is the centroid of a tree T of order n then one of the following holds:*

(i) $C = \{c\}$ and $bw(c) \leq (n-1)/2$,

(ii) $C = \{c_1, c_2\}$ and $bw(c_1) = bw(c_2) = n/2$.

In both cases, if $v \epsilon V(T) \setminus C$ then $bw(v) > n/2$.

[1]Research partially supported by NSA Grant MDA 904-92-H-3050

Theorem B (Zelinka [6]). *The set of vertices with minimum distance in a tree is the centroid of T.*

Combining these results gives the following. A vertex v of a tree has minimum distance iff $bw(v) \leq n/2$.

2. The branch weight at the centroid.

We have the following Theorem.

Theorem A (Barefoot, Entringer, and Székely [ ][ ]).
If T is a tree of order n, and c is in the centroid of t, then

$$\frac{\sigma(T)}{\sigma_T(c)} \leq n-1.$$

We conclude with a lower bound on the optimal tree involving the branch weight at the centroid. An edge e is said to be *contracted* if it is deleted and its ends are identified; the resulting graph is denoted by $G \cdot e$.

Theorem. *If T is a tree of order n, and c is in the centroid of T, then $\sigma(T) \geq \sigma_T(c)(n - bw(c))$.*

Proof. The proof is by induction on $bw(c)$. For convenience let $d = bw(c)$. For the ground case, if $d = 1$, then $T = K_{1,n-1}$, and it is easy to check $\sigma(T) = (n-1)^2$ and $\sigma_T(c) = n-1$.

Now assume $d \geq 2$. Let c be in the centroid of T, i.e., $c \epsilon C(T)$. Now suppose there are $k \geq 1$ branches $B_1, B_2, \cdots, B_k$ at c with $bw(B_i) = d$. Let $e_i = cu_i$ denote the edge in $B_i, 1 \leq i \leq k$, containing c as an endvertex.

Consider the new tree T', resulting from contracting each edge e_i in T, and then reattaching this edge $e_i = cu_i$ to c, where now u_i is a leaf in T'. (See Figure 2.) We have by Jordan's Theorem c is still in the centroid of T', i.e., $c \epsilon C(T')$, also we have $bw_{T'}(c) < d$, i.e., the maximum branch weight at c in T' is now reduced by at least one from that in T. Applying the inductive hypothesis to T' yields; $\sigma(T') \geq (n-d+1)\sigma_{T'}(c)$. We have the following:

(i) $\sigma_T(c) = \sigma_{T'}(c) + k(d-1)$

(ii) $\sigma(T) = \sigma(T') + k(d-1)(n-d-1)$.

(i) occurs since the only changes to distances from c are to vertices in $B_i - u_i$, $1 \leq i \leq k$, and if y is one of the $k(d-1)$ vertices of the branches $B_i \cdot e_i$ and $y \neq c$, then $d_T(c, y) = d_{T'}(c, y) + 1$.

(ii) occurs since the distance changes $\sigma(T) - \sigma(T')$ come from,

(a) between vertices in $B_i \cdot e_i$ and $B_j \cdot e_j$ where $i \neq j$, or

(b) between vertices in $B_i \cdot e_i$ and vertices in $V - \{c\} \setminus \bigcup_{j=1}^{k} B_j$.

In (a), we have for x in $B_i \cdot e_i$ and y in $B_j \cdot e_j$, $d_T(x, y) = d_{T'}(x, y) + 2$, hence (a) contributes $2\binom{k}{2}(d-1)^2$ to $\sigma(T) - \sigma(T')$.

In (b), for any x in $B_i \cdot e_i$ and $y \epsilon V - \{c\} \setminus \bigcup_{j=1}^{k} B_j$, $d_T(x,y) = d_{T'}(x,y) + 1$, hence (b) contributes $k(d-1)[n-(d+1)-(k-1)(d-1)]$ to $\sigma(T) - \sigma(T')$.

Summarily, the contribution of (a) and (b) in $\sigma(T) - \sigma(T')$ is $k(d-1)(n-d-1)$ as in (ii).

We now show,

$$\frac{\sigma_T(c)}{\sigma(T)} = \frac{\sigma_{T'}(c) + k(d-1)}{\sigma(T') + k(d-1)(n-d-1)} \leq \frac{1}{n-d}$$

to establish the inductive step. We have the following sequence of inequalities.

$$\begin{aligned}
&\sigma_{T'}(c) \geq k(d-1) \qquad \text{(since there are at least } kd \text{ vertices in } T' \\
&-\sigma_{T'}(c) \leq -k(d-1) \\
&-\sigma_{T'}(c) + k(d-1)(n-d) \leq k(d-1)(n-d-1) \\
&(\sigma(T') - \sigma_{T'}(c)) + k(d-1)(n-d) \leq \sigma(T') + k(d-1)(n-d-1).
\end{aligned}$$

By the inductive hypothesis, i.e., $\frac{\sigma_{T'}(c)}{\sigma(T')} \leq \frac{1}{n-d+1}$, we have $\sigma_{T'}(c)(n-d) \leq \sigma(T') - \sigma_{T'}(c)$, hence

$$\sigma_{T'}(c)(n-d) + k(d-1)(n-d) \leq \sigma(T') + k(d-1)(n-d-1) \tag{1}$$

Consequently (1) gives

$$\frac{\sigma_T(c)}{\sigma(T)} = \frac{\sigma_{T'}(c) + k(d-1)}{\sigma(T') + k(d-1), (n-d-1)} \leq \frac{1}{n-d}.$$

□

References

1. D. Johnson, J. Lenstra, and A. Rinnooy-Kan, *The complexity of the network design problem, Networks 8* (1978), 279–285.
2. C. Jordan, *Sur les assemblages de lignes* (1869), J. Reine Angew Math 70, 185–190.
3. H. Wiener, *Structural determination of paraffin boiling points*, J. Am. Chem. Soc. **69** (1947), 17–20.
4. B. Zelinka, *Medians and periperhians of trees*, Arch. Math. (Brno) **4** (1968), 87–95.