

Note on “hook-length” as a graph invariant of trees*

HUA WANG

The hook-length of a vertex v in a rooted tree T , analogous to that defined in the Ferrers diagrams of integer partitions, is the number of descendants of v (including v itself) in T . In this note we consider two different types of “average” hook-length of vertices in a rooted tree, yielding a graph invariant that has interesting correlation to the distance functions in trees. This correlation is observed and used throughout the study of related properties and extremal questions.

1. Introduction

The *hook-length* of a vertex v in a rooted tree T is defined as the number of descendants of v including itself and denoted by $h_T(v)$. This concept is comparable with the hook-length of a cell v in the Ferrers diagram of integer partitions, defined as the number of cells, in the same row and column as v , above or to the right of v including v itself. Many recent studies examine the hook-length formulas from various aspects. See, for instance, [5, 7, 8, 9, 12, 13, 23] and the references therein. It is natural to consider the hook-length as a function on vertices of a tree and examine the corresponding graph invariant. In this note, we start with two different concepts of the “average” hook-length of a vertex.

For a tree T , the *mean hook-length at v* is defined as

$$mh_T(v) := \frac{1}{|V(T)|} \sum_{u \in V(T)} h_{T_v}(u)$$

where T_v is the tree T rooted at v , i.e., mean hook-length of all vertices when v is the root. Intuitively, the function $mh_T(\cdot)$ provides some indication of how “centered” a vertex is in T . Indeed we will see some evidence of this intuition from the studies in this note. A closely related concept, named the

*This work was partially supported by grants from the Simons Foundation (#245307).

average hook-length of v , is defined as

$$ah_T(v) := \frac{1}{|V(T)|} \sum_{u \in V(T)} h_{T_u}(v).$$

From the definitions it is natural to conjecture that these two functions are somewhat correlated. This correlation will be shown through various observations in this note. Considering $mh_T(\cdot)$ and $ah_T(\cdot)$ as local functions on each single vertex, the corresponding global functions are denoted by

$$mh(T) := \sum_{v \in V(T)} mh_T(v)$$

and

$$ah(T) := \sum_{v \in V(T)} ah_T(v).$$

One of the most well known distance based graph invariants is the sum of distances, also known as the Wiener index for its applications in biochemistry [22]. Restricting our attention to trees, the Wiener index of a tree T is defined as

$$W(T) = \frac{1}{2} \sum_{v \in V(T)} d(v)$$

where

$$d(v) := d_T(v) = \sum_{u \in U(T)} d(u, v)$$

is the sum of distances from all other vertices to v .

Characterizing extremal trees that maximize or minimize a certain graph invariant has been an active area of research. For $W(T)$, the extremal trees maximizing the Wiener index has been vigorously studied for trees with given maximum degree [6], trees with given degree sequence [4, 14, 20, 24], etc. Such studies lead to some interesting observations on the correlation between the Wiener index and the number of subtrees, whose extremal structures coincide [11, 15, 25]. The correlations between different topological indices are studied in [18].

While the above mentioned extremal results deal with the global function, the local functions corresponding to different invariants very well define the “middle part” of a tree that maximize or minimize the value of such a function. The examination of different middle parts and related extremal

problems are studied for both distance functions and the number of subtrees [1, 2, 16, 17].

In this note, we first study the properties of $mh_T(\cdot)$ and $ah_T(\cdot)$ in Section 2. An interesting connection with the distance functions of trees are found and discussed. With these observations, the related extremal problems are considered in Section 3. In Section 4, we point out the extremal structures regarding the global functions $mh(T)$ and $ah(T)$ as immediate consequences of their correlation with the distance function. In Section 5, the comparison of extremal structures are presented. This results in the characterizations of many extremal structures as corollaries.

2. Middle parts defined by $mh_T(\cdot)$ and $ah_T(\cdot)$

First recall that the *centroid* of a tree T , denoted by $C(T)$, is the set of vertices in a tree T minimizing $d(\cdot)$. It is known that $C(T)$, along with other middle parts of trees defined on the number of subtrees and distances, has the interesting property that it contains one or two adjacent vertices. For $C(T)$, this follows from the strict concavity of the function $d(\cdot)$ along any path of the tree [10]. That is,

Proposition 2.1. *For any three vertices $x, y, z \in V(T)$ such that $xy, yz \in E(T)$, we have*

$$2d(y) < d(x) + d(z).$$

Proposition 2.1 implies that, on any path of T , there are at most two adjacent vertices with the smallest $d(\cdot)$ (since the sequence of values of $d(\cdot)$ is strictly unimodal). Consequently one can conclude that $C(T)$ contains at most two adjacent vertices with minimum $d(\cdot)$.

We consider the analogous question for $mh_T(\cdot)$ and $ah_T(\cdot)$.

2.1. Behavior of $mh_T(\cdot)$ and $ah_T(\cdot)$ along any path of a tree

First we present the following simple but important observation showing the close relation between $mh_T(\cdot)$ and $d(\cdot)$.

Proposition 2.2. *For a tree T of order n and any vertex $v \in V(T)$, we have*

$$n \cdot mh_T(v) = d(v) + n.$$

Proof. By definition, we have

$$n \cdot mh_T(v) = \sum_{u \in V(T)} h_{T_v}(u) = \sum_{w \in V(T)} D_v(w)$$

where $D_v(w)$ is the number of ancestor of w in T_v . That is, the sum of hook-lengths of all vertices is the sum of the numbers of times each vertex contribute to the hook-length of its ancestor. Noting that

$$D_v(w) = d(w, v) + 1,$$

we have

$$n \cdot mh_T(v) = \sum_{w \in V(T)} D_v(w) = \sum_{w \in V(T)} d(w, v) + n = d(v) + n. \quad \square$$

Consequently, with given T of order n , $mh_T(\cdot)$ behaves in exactly the same way as $d(\cdot)$ along any path of T .

Corollary 2.3. *For any three vertices $x, y, z \in V(T)$ such that $xy, yz \in E(T)$ (Figure 1), we have*

$$2mh_T(y) < mh_T(x) + mh_T(z).$$

Hence the values of $mh_T(\cdot)$ along any path of T is strictly concave.

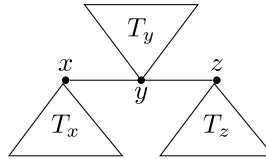


Figure 1: The tree T with vertices x, y, z .

Interestingly, $ah_T(\cdot)$ behaves in exactly the opposite way.

Proposition 2.4. *For any three vertices $x, y, z \in V(T)$ such that $xy, yz \in E(T)$, we have*

$$2ah_T(y) > ah_T(x) + ah_T(z).$$

Hence the values of $ah_T(\cdot)$ along any path of T is strictly convex.

Proof. Let T_x, T_y, T_z be the connected components containing (respectively) x, y, z in $T - \{xy\} - \{yz\}$ (Figure 1). For convenience we consider $n \cdot ah_T(\cdot)$ in our arguments.

Now compare $n \cdot ah_T(y)$ and $n \cdot ah_T(x)$. Note that, for any $v \in V(T_x)$, we have

$$h_{T_v}(y) = h_{T_v}(x) - 1.$$

Similarly,

$$h_{T_v}(y) = h_{T_v}(x) + 1$$

for any $v \in V(T_y) \cup V(T_z)$. Thus

$$n \cdot ah_T(y) - n \cdot ah_T(x) = |V(T_y)| + |V(T_z)| - |V(T_x)|.$$

Similarly,

$$n \cdot ah_T(y) - n \cdot ah_T(z) = |V(T_y)| + |V(T_x)| - |V(T_z)|.$$

Consequently

$$2n \cdot ah_T(y) - (n \cdot ah_T(x) + n \cdot ah_T(z)) = 2|V(T_y)| > 0. \quad \square$$

2.2. Middle parts of a tree with respect to $mh_T(\cdot)$ and $ah_T(\cdot)$

Let $C_{mh}(T)$ and $C_{ah}(T)$ denote the set of vertices of T that (respectively) minimize $mh_T(\cdot)$ and maximize $ah_T(\cdot)$. Corollary 2.3 and Proposition 2.4 imply that, along any path of a tree T , the value of $mh_T(\cdot)$ ($ah_T(\cdot)$) is minimized (maximized) in the middle at exactly one or two adjacent vertices and maximized (minimized) at an end vertex of the path. Applying this fact to all paths of T we easily obtain the following conclusion analogous to those for other “middle parts” of a tree.

Corollary 2.5. *For a tree T , $C_{mh}(T)$ and $C_{ah}(T)$ each contain one or two adjacent vertices. On the other hand, the maximum value of $mh_T(\cdot)$ and minimum value of $ah_T(\cdot)$ are always obtained at some leaf vertex.*

Given the similar (but opposite) behavior of $mh_T(\cdot)$ and $ah_T(\cdot)$, it is natural to ask if $C_{mh}(T)$ and $C_{ah}(T)$ are always the same for a given tree T . We see, from the following lemmas, that this is indeed the case.

Lemma 2.6. *Given a tree T and $v \in C_{mh}(T)$, let u be a neighbor of v . We have*

$$n_{uv}v \geq n_{uv}u$$

where $n_{uv}v$ denotes the number of vertices closer to v than u in T , with equality if and only if u is also in $C_{mh}(T)$.

Lemma 2.7. *Given a tree T and $v \in C_{ah}(T)$, let u be a neighbor of v . We have*

$$n_{uv}v \geq n_{uv}u$$

with equality if and only if u is also in $C_{ah}(T)$.

Remark 1. Noting the connection between $mh_T(\cdot)$ and $d(\cdot)$, Lemma 2.6 follows from exactly the argument for the same statement on $d(\cdot)$. See, for instance, [19]. Lemma 2.7 follows the same idea and is an immediate consequence of the proof of Proposition 2.4.

As stated by Lemmas 2.6 and Lemma 2.7, $C_{mh}(T)$ and $C_{ah}(T)$ are characterized by the same conditions. Further noting that $n_{uv}v \geq n_{uv}u$ implies $n_{uv}u > n_{uv}w$ for a neighbor w (different from v) of u , there are at most a pair of adjacent vertices u and v with $n_{uv}v = n_{uv}u$ and strict inequality holds in this condition for any other pair of vertices. Hence we have the following.

Theorem 2.8. For any given tree T , $C_{mh}(T)$ and $C_{ah}(T)$ are exactly the same.

3. Extremal problems on $mh_T(\cdot)$ and $ah_T(\cdot)$

As a first step of studying the related extremal problems, we consider the characterizations of trees and the corresponding vertices that obtain the maximum and minimum values of $mh_T(\cdot)$ and $ah_T(\cdot)$ among trees of given order.

3.1. Among general trees

Among general trees, such extremal structures coincide with those with respect to many other graph invariants, namely the star and the path.

Proposition 3.1. Among all trees of order n , the star T (with v being its center) is the unique tree that obtains the minimum $mh_T(\cdot)$ and maximum $ah_T(\cdot)$.

Proof. First consider $n \cdot mh_T(v) = \sum_{u \in V(T)} h_{T_v}(u)$, note that each $h_{T_v}(u)$ is at least 1 in the summation for $u \neq v$ and $h_{T_v}(v) = n$, we have

$$n \cdot mh_T(v) = \sum_{u \in V(T)} h_{T_v}(u) \geq 2n - 1$$

with equality if and only if each $u \neq v$ is a leaf. This is exactly when T is a star centered at v .

Similarly, in the summation $n \cdot ah_T(v) = \sum_{u \in V(T)} h_{T_u}(v)$, $h_{T_u}(v)$ is at most $n - 1$ unless $u = v$ (in which case $h_{T_u}(v) = n$). Hence

$$n \cdot ah_T(v) = \sum_{u \in V(T)} h_{T_u}(v) \leq n^2 - n + 1$$

with equality if and only if T is a star centered at v . □

Considering the analogous question on maximizing $d(\cdot)$ among trees of given order n , it is easy to see the following. We include the proof for completeness.

Proposition 3.2. *Among all trees of order n , the path T (with v being one of its end vertices) is the unique tree that obtains the maximum $d(v)$ and hence $mh_T(\cdot)$.*

Proof. Consider $d(u, v)$ for any $u \neq v$, $d(u, v) = k \geq 1$ if and only if there is a vertex w (possibly v) such that $d(w, v) = k - 1$. Hence

$$d(v) \leq 1 + 2 + \dots + (n - 1) = \frac{n(n - 1)}{2}$$

with equality if and only if T is a path with v as a leaf.

Correspondingly,

$$mh_T(v) \leq \frac{1}{n} \left(\frac{n(n - 1)}{2} + n \right)$$

with equality if and only if T is a path with v as a leaf. □

Note that $h_{T_u}(v) = 1$ for any leaf v and any $u \neq v$, and $h_{T_u}(v) = n$ if $u = v$, we have the following simple observation. Unlike other cases, the extremal tree and vertex minimizing $ah_T(v)$ is by no means unique.

Proposition 3.3. *Among trees of order $n \geq 2$, the minimum*

$$ah_T(v) = \frac{2n - 1}{n}$$

is obtained by any tree T with v being one of the leaves.

3.2. Among trees of given degree sequence

Extremal trees of given degree sequence (the non-increasing sequence of degrees of all vertices) have been of great interests because of related applications and the generalization it provides for other categories of trees.

Among trees of given degree sequence, the caterpillar (a tree whose removal of leaves yields a path) has been known as the analogue of the path among general trees. In particular, it is known to maximize the Wiener index and minimize the number of subtrees [14, 20, 24]. However, the specific characterization of the extremal caterpillar usually depends on the degree sequence [4].

Proposition 3.3 also answers the same question for trees with a given degree sequence.

Corollary 3.4. *Among all trees with a given degree sequence, the minimum $ah_T(\cdot)$ is obtained by any such tree with v being one of the leaves.*

To maximize $mh_T(\cdot)$ among trees with given degree sequence, considering the distance version of the problem yields that the caterpillar is indeed extremal. Furthermore, the extremal caterpillar can be characterized.

Proposition 3.5. *Among trees with given degree sequence*

$$(d_1, d_2, \dots, d_k, 1, \dots, 1)$$

(with k internal vertices and $d_1 \geq d_2 \geq \dots \geq d_k \geq 2$), the maximum $d(\cdot)$ and hence $mh_T(\cdot)$ is uniquely obtained by a “biased” caterpillar T (defined below) with v being one of the leaves.

Here T is formed by a longest path $v_0v_1v_2 \dots v_kv$ and pendant edges such that $\deg(v_i) = d_i$ for $1 \leq i \leq k$. Figure 2 shows an example of such a “biased” caterpillar with the corresponding vertex v .

Proof. By Proposition 2.2, we only need to consider $d(v)$ in T . Since the maximum value must be obtained by a leaf, let v be such a leaf and $P(u, v)$ be a longest path from v with u being a leaf. Supposing for contradiction that T is not a caterpillar, there exists a vertex v' on $P(u, v)$ such that a neighbor w of v' is neither a leaf nor on $P(u, v)$. Let W denote the component containing w in $T - \{v'w\}$ (Figure 3).

Now consider a new tree T' resulted from T by detaching W from w and reattaching to u (Figure 3), i.e., $T' = T - \{v'w\} + \{uw\}$. Through this operation the distances from v to any other vertex stay the same but the distances from v to any vertex in W strictly increases (note that v' is at least distance 2 away from u), thus a contradiction.

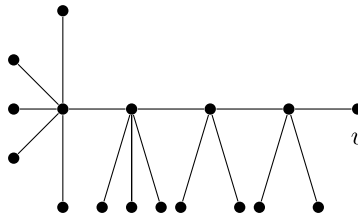


Figure 2: A “biased” caterpillar with degree sequence $(6, 5, 4, 4, 1, \dots, 1)$.

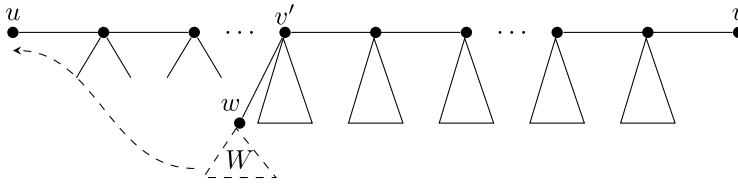


Figure 3: The path $P(u, v)$ and the vertices v, v' and w .

Hence the maximum $d(v)$ must be obtained by a caterpillar with a leaf v that is one end of the longest path. It is obvious that, to maximize $d(v)$, larger degrees are assigned to internal vertices further away from v . This results in exactly the structure of a “biased” caterpillar. \square

We now consider trees with given degree sequence that minimize $mh_T(\cdot)$. We first define the following *greedy tree*, known to be extremal with respect to many other graph invariants such as (minimizing) the Wiener index or (maximizing) the number of subtrees [14, 20, 24].

Definition 1 (Greedy trees). *Given a tree degree sequence deg , the greedy tree is achieved through the following “greedy” algorithm:*

- i) Start with a single vertex v as the root and give v the appropriate number of children so that it has the largest degree;*
- ii) Label the neighbors of v as v_1, v_2, \dots , assign to them the largest available degrees such that $\text{deg}(v_1) \geq \text{deg}(v_2) \geq \dots$;*
- iii) Label the neighbors of v_1 (except v) as v_{11}, v_{12}, \dots such that they take all the largest degrees available and that $\text{deg}(v_{11}) \geq \text{deg}(v_{12}) \geq \dots$, then do the same for v_2, v_3, \dots ;*
- iv) Repeat (iii) for all the newly labeled vertices, always start with the neighbors of the labeled vertex with largest degree whose neighbors are not labeled yet.*

For example, Fig. 4 shows a greedy tree with degree sequence

$$(4, 4, 4, 3, 3, 3, 3, 3, 3, 3, 2, 2, 1, \dots, 1).$$

We will see that such a greedy tree and its root v indeed minimize $d(v)$ (and hence $mh_T(v)$) but far less constrained structures can achieve the same.

Definition 2 (Semi-Greedy trees). *With a given tree degree sequence deg , a rooted tree is semi-greedy if $\text{deg}(u) \geq \text{deg}(w)$ for any two vertices u and w where the height of w is larger than that of u .*

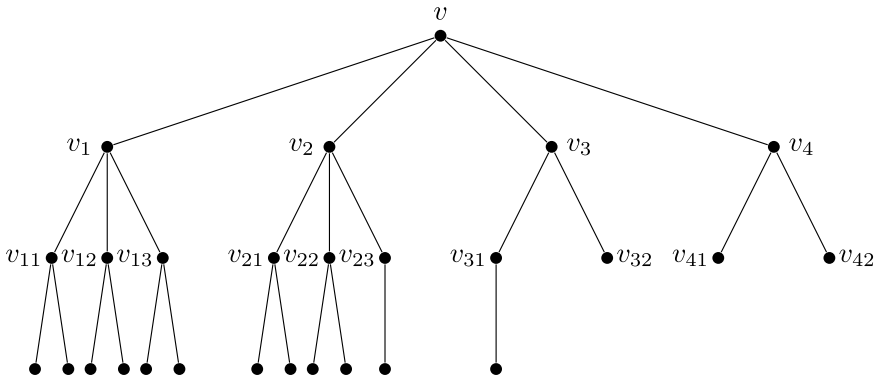


Figure 4: A greedy tree.

Remark 2. Note that the only difference between the semi-greedy trees and the greedy tree is that we do not require the uniform ordering of degrees of vertices of each “level” (height). It is easy to see that a greedy tree is semi-greedy but not vice versa. Fig. 5 shows such an example.

Proposition 3.6. Among trees with given degree sequence, the minimum $d(v)$, and hence minimum $mh_T(v)$ is obtained by a semi-greedy tree and its root. Such extremal trees are generally not unique.

Proof. Again consider the distance version of the question. To minimize $d(v)$ in the tree rooted at v , one simply wants to accommodate as many vertices as possible at height 1, then 2, etc. This is exactly what the conditions of a semi-greedy tree achieves (i.e., assigning larger degrees to vertices closer to the root so that we have the largest possible number of vertices of height 1, height 2, etc.). \square

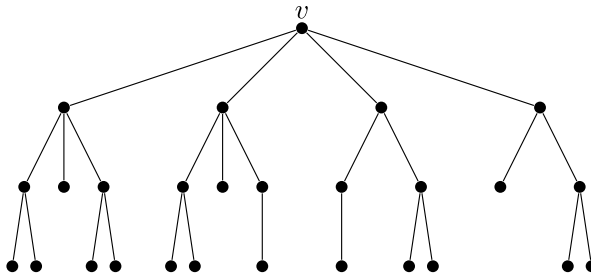


Figure 5: A Semi-greedy tree that is not greedy.

Next we consider trees of given degree sequence that maximize $ah_T(\cdot)$. For this purpose, consider the extremal tree T as rooted at v where the maximum $ah_T(v)$ is obtained. Let T_1, T_2, \dots, T_s denote the connected components in $T - \{v\}$ and let $n_i = |V(T_i)|$ for $1 \leq i \leq s$. We have the following.

Theorem 3.7. *Under the above notations, among all trees of given degree sequence, the maximum $ah_T(v)$ is obtained when the value $\sum_{i=1}^s n_i^2$ is minimized. As a consequence, v has the largest degree in such an extremal tree.*

Proof. First it is easy to see that

$$\sum_{i=1}^s n_i = n - 1 = \sum_{j=1}^k d_j - k + 1$$

where d_1, d_2, \dots, d_k are the degrees of internal vertices. See Figure 6.

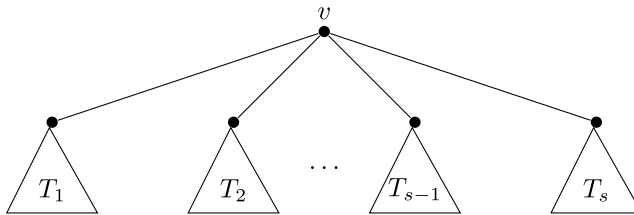


Figure 6: The extremal tree T and the components T_i 's.

Note that, in the summation $n \cdot ah_T(v) = \sum_{u \in V(T)} h_{T_u}(v)$, a vertex $w \in V(T_i)$ contribute to the term $h_{T_u}(v)$ if and only if $u \in V(T - T_i)$. Thus

$$n \cdot ah_T(v) = \sum_{u \in V(T)} h_{T_u}(v) = \sum_{i=1}^s n_i(n - n_i) + n = n^2 - \sum_{i=1}^s n_i^2$$

is maximized when $\sum_{i=1}^s n_i^2$ is minimized.

We also claim that v must be of the largest degree in such an extremal tree, otherwise, let $deg(v) = s < r = deg(w)$ for some vertex $w \in V(T_1)$ (without loss of generality). Let the children of w be w_i for $1 \leq i \leq r - 1$, denote by n'_i be the order of the subtree induced by w_i and its descendants for $1 \leq i \leq r - 1$.

We now create a new tree T' by “switching” the degrees of v and w through “moving” w_i and its descendants from w to v (Figure 7) for $1 \leq i \leq d := r - s$. That is,

$$T' = T - \{ww_1, ww_2, \dots, ww_d\} + \{vw_1, vw_2, \dots, vw_d\}$$

is a tree with the same degree sequence as T .

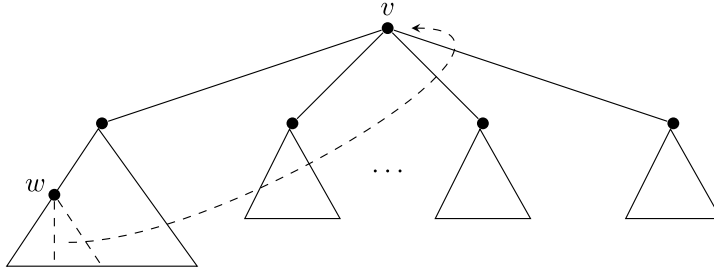


Figure 7: Generating T' from T .

Then

$$\begin{aligned} n \cdot ah_{T'}(v) &= n^2 - \sum_{i=2}^s n_i^2 - \left(n_1 - \sum_{i=1}^d n'_i \right)^2 - \sum_{i=1}^d (n'_i)^2 \\ &> n^2 - \sum_{i=2}^s n_i^2 - n_1^2 \\ &= n \cdot ah_T(v), \end{aligned}$$

contradicting to the optimality of T . □

Remark 3. *As one can see from Theorem 3.7, the complete characterization of the extremal trees that achieve the maximum $ah_T(v)$ depends on the specific degree sequence. However, since the internal vertex degrees in each T_i completely determines n_i , Theorem 3.7 reduces the problem to a discrete optimization question with specifically given degree sequences.*

4. Extremal trees with respect to $mh(T)$ and $ah(T)$

If $mh_T(\cdot)$ and $ah_T(\cdot)$ are considered as “local” functions on hook-length, then $mh(T)$ and $ah(T)$ are the natural “global” versions. As aforementioned, extremal questions on other graph invariants such as the Wiener index have been vigorously studied.

First note the simple observation that

$$\begin{aligned}
 & n \cdot ah(T) \\
 = & \sum_{v \in V(T)} (n \cdot ah_T(v)) = \sum_{v \in V(T)} \sum_{u \in V(T)} h_{T_u}(v) = \sum_{v \in V(T)} (n \cdot mh_T(v)) \\
 = & n \cdot mh(T).
 \end{aligned}$$

From Proposition 2.2 and the definition of $W(T)$, we easily obtain the following observation.

Proposition 4.1. *Given any tree T of order n , we have*

$$n \cdot mh(T) = 2W(T) + n^2.$$

Thus the extremal structures with respect to $mh(T)$ and $ah(T)$ follow from the well known facts on $W(T)$ (see, for instance, [6, 20]).

Corollary 4.2. *Among all trees of given order, $mh(T) = ah(T)$ is minimized by the star and maximized by the path; Among all trees of given degree sequence, $mh(T) = ah(T)$ is minimized by the greedy tree and maximized by a caterpillar.*

5. Comparison between trees of different degree sequence

As a final remark on the extremal structures among trees with given degree sequence, we consider their ordering with different degree sequences. Such results in previous works on other indices yield many characterizations of extremal structures as immediate corollaries.

Definition 3. *For two nonincreasing sequences $\pi = (d_0, \dots, d_{n-1})$ and $\pi' = (d'_0, \dots, d'_{n-1})$, π' is said to majorize π if for $k = 0, \dots, n - 2$*

$$\sum_{i=0}^k d_i \leq \sum_{i=0}^k d'_i \quad \text{and} \quad \sum_{i=0}^{n-1} d_i = \sum_{i=0}^{n-1} d'_i.$$

This is denoted by

$$\pi \triangleleft \pi'.$$

Lemma 5.1. [21] *Let $\pi = (d_0, \dots, d_{n-1})$ and $\pi' = (d'_0, \dots, d'_{n-1})$ be two nonincreasing graphic degree sequences. If $\pi \triangleleft \pi'$, then there exists a series of graphic degree sequences π_1, \dots, π_m such that $\pi \triangleleft \pi_1 \triangleleft \dots \triangleleft \pi_m \triangleleft \pi'$, where π_i and π_{i+1} differ at exactly two entries, say d_j (d'_j) and d_k (d'_k) of π_i (π_{i+1}), with $d'_j = d_j + 1$, $d'_k = d_k - 1$ and $j < k$.*

With Lemma 5.1, it is easy to compare $mh_T(v)$ of two semi-greedy trees or biased caterpillars with degree sequences differing by only two entries, i.e., consider

$$\pi = (d_0, \dots, d_{n-1}) \triangleleft (d'_0, \dots, d'_{n-1}) = \pi'$$

with $d'_j = d_j + 1$, $d'_l = d_l - 1$ for some $j < l$ and all other entries the same.

For two biased caterpillars with degree sequences π and π' , it is obvious that the one with degree sequence π' has one vertex further away from v than that with π while all other vertices stay the same. Hence we have the following.

Proposition 5.2. *Let T_1 be a biased caterpillar of order n with degree sequence π_1 and T_2 of order n with degree sequence π_2 . If $\pi_1 \triangleleft \pi_2$, then the corresponding maximum value of $d(\cdot)$ (and hence $mh_T(\cdot)$) of T_2 is at least as large as that of T_1 .*

Similarly, the semi-greedy tree with degree sequence π' has some vertex whose distance from v is at most that in the semi-greedy tree with degree sequence π while all other vertices maintain the same distance from v . Thus we have the following.

Proposition 5.3. *Let T_1 be a semi-greedy tree of order n with degree sequence π_1 and T_2 of order n with degree sequence π_2 . If $\pi_1 \triangleleft \pi_2$, then the corresponding minimum value of $d(\cdot)$ (and hence $mh_T(\cdot)$) of T_2 is at most as large as that of T_1 .*

For instance, it is obvious that the star (path) has degree sequences majorizing (being majorized by) all other degree sequences among general trees. The statements regarding $mh_T(\cdot)$ in Propositions 3.1 and 3.2 are then simple consequences of the above observations. Similarly, if the number of leaves of a tree is given, we have the following.

Corollary 5.4. *Among trees with given order n and number of leaves l , the “star-like tree” (a tree formed by joining ends of paths, whose lengths differ by at most 1; Figure 8) achieves the minimum $mh_T(\cdot)$ and the “comet” (a tree formed by attaching pendant edges at one end of a path, also called a “broom”; Figure 9) achieves the maximum $mh_T(\cdot)$.*

Proof. Indeed, the degree sequence

$$(l, 2, \dots, \underbrace{2, 1, \dots, 1}_{l \text{ 1's}})$$

majorizes all other degree sequences under the condition that there are exactly l 1's. The corresponding semi-greedy tree is a star-like tree and the corresponding biased caterpillar is a comet. \square

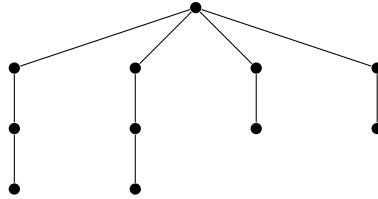


Figure 8: A star-like tree with $n = 11$ and $l = 4$.

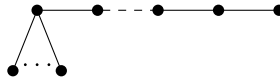


Figure 9: A comet or broom.

The analogous statement regarding the greedy trees of different degree sequences that minimize $n \cdot mh(T) = n \cdot ah(T) = 2W(T) + n^2$, although not stated explicitly before, easily follows from the properties of $d(\cdot)$ in greedy trees. See for instance [3, 20].

Theorem 5.5. *Let T_1 be a greedy tree of order n with degree sequence π_1 and T_2 of order n with degree sequence π_2 . If $\pi_1 \triangleleft \pi_2$, then*

$$n \cdot mh(T_2) = n \cdot ah(T_2) = 2W(T_2) + n^2 \leq n \cdot mh(T_1) = n \cdot ah(T_1) = 2W(T_1) + n^2.$$

As an example of the application, the “extended good k -ary trees” is defined as the greedy tree but with all vertex degrees equal to k in the recursive construction until one runs out of vertices (hence there is at most one vertex of degree strictly between 1 and k , see Figure 10). For trees with given order n and maximum vertex degree k , it is obvious that the degree sequence

$$(k, \dots, k, 1 \leq r < k, 1, \dots, 1)$$

majorizes all other degree sequences. The corresponding greedy tree yields an extended good k -ary tree, showing the following. Interested readers can check [3, 25] for many more such applications.

Corollary 5.6. *Among all trees of order n and maximum vertex degree k , the extended good k -ary tree minimizes $W(T)$ and hence $mh(T)$ and $ah(T)$.*

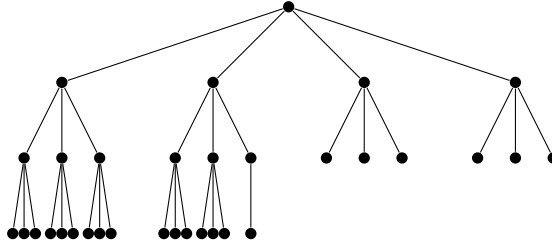


Figure 10: An extended good 4-ary tree.

Acknowledgements

The author is in debt to the anonymous referee whose valuable suggestions on both mathematical reasoning and presentation of the paper greatly improved the quality of this manuscript. The author also thanks L. A. Székely for introducing to him the concept of hook-length.

References

- [1] A. Ádám. (1974). The centrality of vertices in trees. *Studia Sci. Math. Hung.* **9** 285–303. [MR0406840](#)
- [2] C. A. Barefoot, R. C. Entringer, L. A. Székely. (1997). Extremal values for ratios of distances in trees. *Discrete Appl. Math.* **80** 37–56. [MR1489059](#)
- [3] M. Bartlett, E. Krop, C. Magnant, F. Mutiso, H. Wang. (2014). Variations of distance-based invariants of trees. *J. Combin. Math. Combin. Comput.* **91** 19–29. [MR3287704](#)
- [4] E. Cela, N. S. Schmuck, S. Wimer, G. J. Woeginger. (2011). The Wiener maximum quadratic assignment problem. *Discrete Optim.* **8** 411–416. [MR2811629](#)
- [5] W. Y. C. Chen, O. X. Q. Gao, P. L. Guo. (2011). Hook length formulas for trees by Han’s expansion. *Electronic Journal of Combinatorics* **16** #R62. [MR2505104](#)

- [6] M. Fischermann, A. Hoffmann, D. Rautenbach, L. A. Székely, L. Volkman. (2002). Wiener index versus maximum degree in trees. *Discrete Appl. Math.* **122** 127–137. [MR1907827](#)
- [7] G. N. Han. (2010). New hook length formulas for binary trees. *Combinatorica* **30** 253–256. [MR2676840](#)
- [8] G. N. Han. (2009). Yet another generalization of Postnikovs hook length formula for binary trees. *SIAM J. Discrete Mathematics* **23** 661–664. [MR2496909](#)
- [9] G. N. Han. (2008). Discovering hook length formulas by an expansion technique. *Electronic Journal of Combinatorics*. **15** #R133. [MR2448883](#)
- [10] C. Jordan. (1869). Sur les assemblages de lignes. *J. Reine Angew. Math.* **70** 185–190. [MR1579443](#)
- [11] R. Kirk, H. Wang. (2008). Largest number of subtrees of trees with a given maximum degree. *SIAM J. Discrete Mathematics* **22**(3) 985–995. [MR2424834](#)
- [12] M. Kuba, A. Panholzer. (2013). A unifying approach for proving hook-length formulas for weighted tree families. *Graphs and Combinatorics* **29** 1839–1865. [MR3119942](#)
- [13] B. E. Sagan. (2009). Probabilistic proofs of hook length formulas involving trees. *Séminaire Lotharingien de Combinatoire* **61A** Article B61Ab. [MR2529393](#)
- [14] N. Schmuck, S. Wagner, H. Wang. (2012). Greedy trees, caterpillars, and Wiener-type graph invariants. *MATCH Commun. Math. Comput. Chem.* **68**(1) 273–292. [MR2986487](#)
- [15] L. A. Székely, H. Wang. (2005). On subtrees of trees. *Advances in Applied Mathematics* **34** 138–155. [MR2102279](#)
- [16] L. A. Székely, H. Wang. (2013). Extremal values of ratios: distance problems vs. subtree problems in trees. *Electronic J. Combin.* **20**(1) paper 67, 20 pp. [MR3040629](#)
- [17] L. A. Székely, H. Wang. (2014). Extremal values of ratios: distance problems vs. subtree problems in trees II. *Discrete Math.* **322** 36–47. [MR3164035](#)
- [18] S. Wagner. (2007). Correlation of graph-theoretical indices. *SIAM J. Discrete Mathematics* **21**(1) 33–46. [MR2299692](#)

- [19] H. Wang. (2015). Centroid, leaf-centroid, and internal-centroid. *Graphs Combin.* **31** 783–793. [MR3338033](#)
- [20] H. Wang. (2008). The extremal values of the Wiener index of a tree with given degree sequence. *Discrete Appl. Math.* **156** 2647–2654; Corrigendum (2009). *Discrete Appl. Math.* **157** 3754. [MR2451087](#)
- [21] W. D. Wei. (1982). The class $\mathcal{U}(R, S)$ of $(0,1)$ matrices. *Discrete Mathematics* **39** 201–205. [MR0676194](#)
- [22] H. Wiener. (1947). Structural determination of paraffin boiling point. *J. Amer. Chem. Soc.* **69** 17–20.
- [23] L. L. M. Yang. Generalizations of Han’s Hook Length Identities. [arXiv:0805.0109](#).
- [24] X. D. Zhang, Q. Y. Xiang, L. Q. Xu, R. Y. Pan. (2008). The Wiener index of trees with given degree sequences. *MATCH Commun. Math. Comput. Chem.* **60** 623–644. [MR2457876](#)
- [25] X. M. Zhang, X. D. Zhang, D. Gray, H. Wang. (2013). The number of subtrees of trees with given degree sequence. *J. Graph Theory.* **73**(3) 280–295. [MR3062797](#)

HUA WANG

DEPARTMENT OF MATHEMATICAL SCIENCES

GEORGIA SOUTHERN UNIVERSITY

STATESBORO, GA 30460

USA

E-mail address: hwang@georgiasouthern.edu

RECEIVED 27 MARCH 2015