The Number of Subtrees of Trees with Given Degree Sequence

Xiu-Mei Zhang,1,2 Xiao-Dong Zhang,1 Daniel Gray,3 and Hua Wang4

1 DEPARTMENT OF MATHEMATICS, SHANGHAI JIAO TONG UNIVERSITY
800 DONGCHUAN ROAD, SHANGHAI 200240, P. R. CHINA
E-mail: xiaodong@sjtu.edu.cn

2 DEPARTMENT OF MATHEMATICS, SHANGHAI SANDAU UNIVERSITY
2727 JINHAI ROAD, SHANGHAI 201209, P. R. CHINA
E-mail: wfluckyzxm@163.com

3 DEPARTMENT OF MATHEMATICS, UNIVERSITY OF FLORIDA GAINESVILLE,
FLORIDA 32611
E-mail: dgray1@ufl.edu

4 DEPARTMENT OF MATHEMATICAL SCIENCES, GEORGIA SOUTHERN UNIVERSITY STATEBORO
GEORGIA 30460
E-mail: hwang@georgiasouthern.edu

Received March 2, 2011; Revised April 20, 2012

Published online 6 August 2012 in Wiley Online Library (wileyonlinelibrary.com).
DOI 10.1002/jgt.21674

Contract grant sponsor: National Natural Science Foundation of China; Contract grant number: 10971137; Contract grant sponsor: the National Basic Research Program (973) of China; Contract grant number: 2006CB805900; Contract grant sponsor: Science and Technology Commission of Shanghai Municipality (STCSM); Contract grant number: 09XD1402500.

Journal of Graph Theory
© 2012 Wiley Periodicals, Inc.
Abstract: This article investigates some properties of the number of subtrees of a tree with given degree sequence. These results are used to characterize trees with the given degree sequence that have the largest number of subtrees, which generalize the recent results of Kirk and Wang (SIAM J Discrete Math 22 (2008), 985–995). These trees coincide with those which were proven by Wang and independently Zhang et al. (2008) to minimize the Wiener index. We also provide a partial ordering of the extremal trees with different degree sequences, some extremal results follow as corollaries.

Keywords: tree; subtree; degree sequence; majorization

MSC 2010: 05C05, 05C30

1. INTRODUCTION

All graphs in this article will be finite, simple, and undirected. A tree $T = (V, E)$ is a connected, acyclic graph where $V(T)$ and $E(T)$ denote the vertex set and edge set, respectively. We refer to vertices of degree 1 of $T$ as leaves. The unique path connecting two vertices $u, v$ in $T$ will be denoted by $P_T(u, v)$. The number of edges on $P_T(u, v)$ is called distance $dist_T(u, v)$, or for short $dist(u, v)$ between them. We call a tree $(T, r)$ rooted at the vertex $r$ (or just by $T$ if it is clear what the root is) by specifying a vertex $r \in V(T)$. The height of a vertex $v$ of a rooted tree $T$ with root $r$ is $h_T(v) = dist_T(r, v)$. For any two different vertices $u, v$ in a rooted tree $(T, r)$, we say that $v$ is a successor of $u$ and $u$ is an ancestor of $v$ if $P_T(r, u) \subset P_T(r, v)$. Furthermore, if $u$ and $v$ are adjacent to each other and $dist_T(r, u) = dist_T(r, v) - 1$, we say that $u$ is the parent of $v$ and $v$ is a child of $u$. Two vertices $u, v$ are siblings of each other if they share the same parent. A subtree of a tree will often be described by its vertex set.

The number of subtrees of a tree has received much attention. It is well known that the path $P_n$ and the star $K_{1, n-1}$ have the most and least subtrees among all trees of order $n$, respectively. The binary trees that maximize or minimize the number of subtrees are characterized in [6, 8]. Formulas are given to calculate the number of subtrees of these extremal binary trees. These formulas use a new representation of integers as a sum of powers of 2. Number theorists have already started investigating this new binary representation [2]. Also, the sequence of the number of subtrees of these extremal binary trees (with $2l$ leaves, $l = 1, 2, \ldots$) appears to be new [5]. Later, a linear-time algorithm to count the subtrees of a tree is provided in [12].

In a related article [7], the number of leaf-containing subtrees is studied for binary trees. The results turn out to be useful in bounding the number of acceptable residue configurations. For a given phylogenetic tree from the DNA or RNA sequences, label each vertex with either a residue or a gap, and denote that a residue configuration for the tree. Further, a residue configuration is called to be acceptable if (1) the vertices labeled by residues form a subtree and (2) at least one leaf has a residue. See [4] for details.

An interesting fact is that among binary trees of the same size, the extremal one that minimizes the number of subtrees is exactly the one that maximizes some chemical indices such as the well-known Wiener index, and vice versa. In [3], subtrees of trees with
given order and maximum vertex degree are studied. The extremal trees coincide with the ones for the Wiener index as well. Such correlations between different topological indices of trees are studied in [9].

Recently, in [14] and [10], respectively, extremal trees are characterized regarding the Wiener index with a given degree sequence. Moreover, extremal trees are also characterized for the adjacency [1] and Laplacian [13] spectral radii with a given degree sequence. Then it is natural to consider the following question.

**Problem 1.1.** Given the degree sequence and the number of vertices of a tree, find the upper bound for the number of subtrees, and characterize all extremal trees that attain this bound.

It will not be a surprise to see that such extremal trees coincide with the ones that attain the minimum Wiener index. Along this line, we also provide an ordering of the degree sequences according to the largest number of subtrees. With our main results, Theorems 2.3 and 2.4, one can deduce extremal graphs with the largest number of subtrees in some classes of graphs. This generalizes the results of [3, 6], etc.

The rest of this article is organized as follows. In Section 2, some notations and the main theorems are stated. In Section 3, we present some observations regarding the structure of the extremal trees. In Section 4, we present the proofs of the main theorems. In Section 5, we show, as corollaries, characterizations of the extremal trees in different categories of trees including previously known results.

## 2. PRELIMINARIES

For a nonincreasing sequence of positive integers \( \pi = (d_0, \ldots, d_{n-1}) \) with \( n \geq 3 \), let \( T_\pi \) denote the set of all trees with \( \pi \) as its degree sequence. We can construct a special tree \( T_\pi^* \in T_\pi \) by using breadth-first search method as follows. First, label the vertex with the largest degree \( d_0 \) as \( v_{01} \) (the root). Second, label the neighbors of \( v_0 \) as \( v_{11}, v_{12}, \ldots, v_{1d_0} \) from left to right and let \( d(v_{1i}) = d_i \) for \( i = 1, \ldots, d_0 \). Then repeat the second step for all newly labeled vertices until all degrees are assigned. For example, if \( \pi = (4, 4, 3, 3, 3, 3, 2, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1) \), \( T_\pi^* \) is shown in Figure 1. There is a vertex \( v_{01} \) (the root) in layer 0 with the largest degree 4; its four neighbors are labeled as \( v_{11}, v_{12}, v_{13}, v_{14} \) in layer 1, with degrees 4, 3, 3, 3 from left

![FIGURE 1. An example of \( T_\pi^* \).](image-url)
to right; nine vertices $v_{21}, v_{22}, \ldots, v_{29}$ in layer 2; five vertices $v_{31}, v_{32}, v_{33}, v_{34}, v_{35}$ in
layer 3. The number of vertices in each layer $i$, denoted by $s_i$ can be easily calculated as $s_0 = 1, s_1 = d_0 = 4, s_2 = d_1 + d_2 + d_3 + d_4 - s_1 = 4 + 3 + 3 + 3 - 4 = 9$, and $s_3 = d_5 + \cdots + d_{13} - s_2 = 5$.

To explain the structure and properties of $T_\pi^*$, we need the following notation from [13].

**Definition 2.1** ([13]). Let $T = (V, E)$ be a tree with root $v_0$. A well ordering $\prec$ of the vertices is called a breadth-first search ordering with nonincreasing degrees (a BFS-ordering for short) if $\prec$ satisfies the following properties:

1. If $u, v \in V$, and $u \prec v$, then $h(u) \leq h(v)$ and $d(u) \geq d(v)$;
2. If there are two edges $uv, v_1 \in E(T)$ and $v_1 \in E(T)$ such that $u \prec v$, $h(u) = h(u_1) - 1$ and $h(v) = h(v_1) - 1$, then $u_1 \prec v_1$.

We call trees that have a BFS-ordering of its vertices a BFS-tree.

It is easy to see that $T_\pi^*$ has a BFS-ordering and any two BFS-trees with degree sequence $\pi$ are isomorphic (e.g., see [13]).

Let $\pi = (d_0, \ldots, d_{n-1})$ and $\pi' = (d'_0, \ldots, d'_{n-1})$ be two nonincreasing sequences. If $\sum_{i=0}^k d_i \leq \sum_{j=0}^k d'_j$ for $k = 0, \ldots, n-2$ and $\sum_{i=0}^{n-1} d_i = \sum_{i=0}^{n-1} d'_i$, then the sequence $\pi'$ is said to majorize the sequence $\pi$ and denoted by $\pi \diamond \pi'$. It is known that the following holds (e.g., see [11] or [13]).

**Proposition 2.2.** (Wei [11]) Let $\pi = (d_0, \ldots, d_{n-1})$ and $\pi' = (d'_0, \ldots, d'_{n-1})$ be two nonincreasing graphic degree sequences. If $\pi \prec \pi'$, then there exists a series of graphic degree sequences $\pi_1, \ldots, \pi_k$ such that $\pi \prec \pi_1 \prec \cdots \prec \pi_k \prec \pi'$, where $\pi_i$ and $\pi_{i+1}$ differ at exactly two entries, say $d_j$ ($d'_j$) and $d_k$ ($d'_k$) of $\pi_i$ ($\pi_{i+1}$), with $d'_j = d_j + 1$, $d'_k = d_k - 1$ and $j < k$.

The main results of this article can be stated as follows.

**Theorem 2.3.** With a given degree sequence $\pi$, $T_\pi^*$ is the unique tree with the largest number of subtrees in $T_\pi$.

**Theorem 2.4.** Given two different degree sequences $\pi$ and $\pi_1$. If $\pi \diamond \pi_1$, then the number of subtrees of $T_\pi^*$ is less than the number of subtrees of $T_{\pi_1}$.

### 3. SOME OBSERVATIONS

In order to prove Theorems 2.3 and 2.4, we need to introduce some more terminologies. For a vertex $v$ of a rooted tree $(T, r)$, let $T(v)$, the subtree induced by $v$, denote the subtree of $T$ (rooted at $v$) that is induced by $v$ and all its successors. For a tree $T$ and vertices $v_1, v_2, \ldots, v_{m-1}, v_m$ of $T$, let $f_T(v_1, v_2, \ldots, v_{m-1}, v_m)$ denote the number of subtrees of $T$ that contain the vertices $v_1, v_2, \ldots, v_{m-1}, v_m$. In particular, $f_T(v)$ denotes the number of subtrees of $T$ that contain $v$. Let $\varphi(T)$ denote the number of nonempty subtrees of $T$.

Let $W$ be a tree and $x, y$ be two vertices of $W$. The path $P_W(x, y)$ from $x$ to $y$ can be denoted by $x_m x_{m-1} \cdots x_2 x_1 y_1 y_2 \cdots y_{m-1} y_m$ for odd dist($x, y$) or $x_m x_{m-1} \cdots x_2 x_1 y_1 y_2 \cdots y_{m-1} y_m$ for even dist($x, y$), where $x_m = x, y_m = y$. Let $G_1$ be the graph resulted from $W$ by deleting all edges in $P_W(x, y)$. The connected components (in $G_1$) containing $x_i, y_i$, and $z$ are denoted by $X_i, Y_i, Z_i$, respectively, for $i = 1, 2, \ldots, m$. 

*Journal of Graph Theory* DOI 10.1002/jgt
We also let $X_{\geq k}$ be the connected component of $W$ containing $x_k$ after deleting the edge $x_{k-1}x_k$ and $Y_{\geq k}$ be the connected component of $W$ containing $y_k$ after deleting the edge $y_{k-1}y_k$, for $k = 1, \ldots, m$. Clearly, $X_{\geq m} = X_m$ and $Y_{\geq m} = Y_m$. Figure 2 shows such a labeling according to a path of odd length (without $x$).

We need the next two lemmas from [3] to proceed.

**Lemma 3.1.** [3] Let $W$ be a tree with a path $P_W(x_m, y_m) = x_mx_m-1 \ldots x_2x_1(z)y_1y_2 \ldots y_{m-1}y_m$ from $x_m$ to $y_m$. If $f_X(x_i) \geq f_Y(y_i)$ for $i = 1, 2, \ldots, m$, then $f_W(x_m) \geq f_W(y_m)$. Furthermore, if this inequality holds, then $f_W(x_m) = f_W(y_m)$ if and only if $f_X(x_i) = f_Y(y_i)$ for $i = 1, 2, \ldots, m$.

Now let $X$ and $Y$ be two rooted trees with roots $x$ and $y$. Let $T$ be a tree containing vertices $x'$ and $y'$. Then we can build $T'$ by identifying the root $x$ of $X$ with $x'$ of $T$ and the root $y$ of $Y$ with $y'$ of $T$, and $T''$ by identifying the root $x$ of $X$ with $x'$ of $T$ and the root $y$ of $Y$ with $x'$ of $T$ (see Figure 3).

**Lemma 3.2.** [3] Let $T$ be a tree with two vertices $x, y$. If $f_T(x) \geq f_T(y)$ and $f_X(x) \leq f_Y(y)$, then $\phi(T'') \geq \phi(T')$ with equality if and only if $f_T(x) = f_T(y)$ or $f_X(x) = f_Y(y)$.

From Lemmas 3.1 and 3.2, we immediately achieve the following observation. We leave the proof to the reader.

**Lemma 3.3.** Let $T$ be a tree in $T_\pi$ and $P(x_m, y_m) = x_mx_{m-1} \ldots x_2x_1(z)y_1y_2 \ldots y_{m-1}y_m$ be a path of $T$. Let $T'$ be the tree from $T$ by deleting the two edges $x_{k}x_{k+1}$ and $y_{k}y_{k+1}$ and adding two edges $x_{k+1}y_{k}$ and $y_{k+1}x_{k}$. If $f_X(x_i) \geq f_Y(y_i)$ for $i = 1, \ldots, k$ and $1 \leq k \leq m-1$, and $f_{X_{\geq k+1}}(x_{k+1}) \leq f_{Y_{\geq k+1}}(y_{k+1})$, then

$$\phi(T) \leq \phi(T')$$

with equality if and only if $f_{X_{\geq k+1}}(x_{k+1}) = f_{Y_{\geq k+1}}(y_{k+1})$ or $f_X(x_i) = f_Y(y_i)$ for $i = 1, \ldots, k$.

For convenience, we refer to trees that maximize the number of subtrees as optimal. In terms of the structure of the optimal tree, we have the following version of Lemma 3.3.

**Corollary 3.4.** Let $T$ be an optimal tree in $T_\pi$ and $P(x_m, y_m) = x_mx_{m-1} \ldots x_2x_1(z)y_1y_2 \ldots y_{m-1}y_m$ be a path of $T$. If $f_X(x_i) \geq f_Y(y_i)$ for $i = 1, \ldots, k$ with at least one strict inequality and $1 \leq k \leq m-1$, then $f_{X_{\geq k+1}}(x_{k+1}) \geq f_{Y_{\geq k+1}}(y_{k+1})$. .
Lemma 3.5. Let $T$ be an optimal tree in $T_n$ and $P(x_m, y_m) = x_m x_{m−1} \ldots x_2 x_1 (z) y_1 y_2 \ldots y_m−1 y_m$ be a path of $T$. If $f_{X_i}(x_i) \geq f_{Y_i}(y_i)$ for $i = 1, \ldots, k$ with at least one strict inequality and $1 \leq k \leq m − 1$, then $f_{X_{k+1}}(x_{k+1}) \geq f_{Y_{k+1}}(y_{k+1})$.

Proof. If $k = m − 1$, then by Corollary 3.4, the assertion holds since $f_{X_{≥m}}(x_m) = f_{X_n}(x_m)$ and $f_{Y_{≥m}}(y_m) = f_{Y_n}(y_m)$. Hence, we assume that $1 \leq k \leq m − 2$. Suppose that $f_{X_{k+1}}(x_{k+1}) < f_{Y_{k+1}}(y_{k+1})$. Denote by $M$ the number of subtrees of $T$ not containing vertices $x_k$ and $y_k$. Let $W$ be the connected component of $T$ by deleting the two edges $x_k x_{k+1}$ and $y_k y_{k+1}$ containing vertices $x_k$ and $y_k$. Then

$$\varphi(T) = \left\{1 + f_{X_{k+1}}(x_{k+1})[1 + f_{X_{≥k+2}}(x_{k+2})]\right\} [f_w(x_k) − f_w(x_k, y_k)]$$
$$+ \left\{1 + f_{Y_{k+1}}(y_{k+1})[1 + f_{Y_{≥k+2}}(y_{k+2})]\right\} [f_w(y_k) − f_w(x_k, y_k)]$$
$$+ \left\{1 + f_{X_{k+1}}(x_{k+1})[1 + f_{X_{≥k+2}}(x_{k+2})]\right\} [1 + f_{Y_{k+1}}(y_{k+1})[1$$
$$+ f_{Y_{≥k+2}}(y_{k+2})]\right\} f_w(x_k, y_k) + M.$$ 

On the other hand, let $T'$ be the tree from $T$ by deleting four edges $x_k x_{k+1}$, $x_{k+1} x_{k+2}$, $y_k y_{k+1}$, and $y_{k+1} y_{k+2}$ and adding four edges $x_k y_{k+1}$, $y_{k+1} x_{k+2}$, $y_{k+2} y_{k+1}$, and $x_{k+1} y_{k+2}$. Clearly, $T' \in T_n$ and

$$\varphi(T') = \left\{1 + f_{Y_{k+1}}(y_{k+1})[1 + f_{X_{≥k+2}}(x_{k+2})]\right\} [f_w(x_k) − f_w(x_k, y_k)]$$
$$+ \left\{1 + f_{X_{k+1}}(x_{k+1})[1 + f_{Y_{≥k+2}}(y_{k+2})]\right\} [f_w(y_k) − f_w(x_k, y_k)]$$
$$+ \left\{1 + f_{Y_{k+1}}(y_{k+1})[1 + f_{X_{≥k+2}}(x_{k+2})]\right\} [1 + f_{X_{k+1}}(x_{k+1})[1$$
$$+ f_{Y_{≥k+2}}(y_{k+2})]\right\} f_w(x_k, y_k) + M.$$ 

Hence,

$$\varphi(T') − \varphi(T) = [f_{Y_{k+1}}(y_{k+1}) − f_{X_{k+1}}(x_{k+1})][1 + f_{X_{≥k+2}}(x_{k+2})][f_w(x_k) − f_w(x_k, y_k)]$$
$$− [1 + f_{Y_{≥k+2}}(y_{k+2})][f_w(y_k) − f_w(x_k, y_k)] + f_w(x_k, y_k)[f_{X_{≥k+2}}(x_{k+2})$$
$$− f_{Y_{≥k+2}}(y_{k+2})].$$ 

Obviously, we have $f_w(y_k) > f_w(x_k, y_k)$ and $f_w(x_k) > f_w(x_k, y_k)$. By Lemma 3.1, we have $f_w(x_k) > f_w(y_k)$. Further by Corollary 3.4, we have $f_{X_{≥k+1}}(x_{k+1}) \geq f_{Y_{≥k+1}}(y_{k+1})$. Since $f_{X_{≥k+1}}(x_{k+1}) = f_{X_{k+1}}(x_{k+1})(1 + f_{X_{≥k+2}}(x_{k+2}))$ and $f_{Y_{≥k+1}}(y_{k+1}) = f_{Y_{k+1}}(y_{k+1})(1 + f_{Y_{≥k+2}}(y_{k+2}))$, we have $f_{X_{≥k+2}}(x_{k+2}) \geq f_{Y_{≥k+2}}(y_{k+2})$ since we assumed $f_{X_{k+1}}(x_{k+1}) < f_{Y_{k+1}}(y_{k+1})$. Therefore, $\varphi(T') > \varphi(T) > 0$, contradicting to the optimality of $T$. So the assertion holds.

Lemma 3.6. Let $P$ be a path of an optimal $T$ in $T_n$ whose end vertices are leaves.

(i) If the length of $P$ is odd $(2m − 1)$, then the vertices of $P$ can be labeled as $x_m x_{m−1} \ldots x_1 y_1 y_2 \ldots y_m$ such that

$$f_{X_1}(x_1) \geq f_{Y_1}(y_1) \geq f_{X_2}(x_2) \geq f_{Y_2}(y_2) \geq \cdots \geq f_{X_m}(x_m) = f_{Y_m}(y_m) = 1.$$ 

(ii) If the length of $P$ is even $(2m)$, then the vertices of $P$ can be labeled as $x_m x_{m−1} x_{m−2} \ldots x_1 y_1 y_2 \ldots y_m$ such that

$$f_{X_1}(x_1) \geq f_{Y_1}(y_1) \geq f_{X_2}(x_2) \geq f_{Y_2}(y_2) \geq \cdots \geq f_{X_m}(x_m) \geq f_{Y_m}(y_m) = f_{X_{m+1}}(x_{m+1}) = 1.$$ 

Journal of Graph Theory DOI 10.1002/jgt
Proof. We provide the proof of part (i), part (ii) can be shown in a similar manner.

Let \( P = u_1 u_2 \ldots u_{2m} \) and \( f_{U_i}(u_i) = \max\{f_{U_j}(u_i), 1 \leq i \leq 2m\} \). If \( f_{U_i}(u_i) > \max\{f_{U_{i+1}}(u_{i+1}), f_{U_{i-1}}(u_{i-1})\} \), then assume that \( f_{U_{i+1}}(u_{i+1}) > f_{U_{i-1}}(u_{i-1}) \). Otherwise there exist \( l_1, l_2 \geq 0 \) and \( l_1 + l_2 \geq 1 \) such that \( f_{U_{i-l_1}}(u_{i-l_1}) = \cdots = f_{U_i}(u_i) = \cdots = f_{U_{i+l_2}}(u_{i+l_2}) \).

Hence, without loss of generality, there exists a \( 1 \leq k \leq m \) such that the vertices of \( P \) may be labeled as \( x_1x_{r-1} \ldots x_1y_1y_2 \ldots y_s \) and

\[
\begin{align*}
&f_{X_1}(x_1) > f_{Y_1}(y_1), \quad f_{Y_1}(y_1) = f_{X_2}(x_2), \ldots, \ f_{X_k}(x_k) = f_{Y_k}(y_k) > f_{X_{k+1}}(x_{k+1}), \\
&\quad \text{or} \\
&f_{X_1}(x_1) = f_{Y_1}(y_1) = \cdots = f_{X_k}(x_k) = f_{Y_k}(y_k) > f_{X_{k+1}}(x_{k+1}) \geq f_{Y_{k+1}}(y_{k+1}),
\end{align*}
\]

(1)

where \( r + s = 2m \). If (1) holds, then we must have

\[
r = s = m.
\]

Otherwise, if \( r < s \), then by Lemma 3.5, \( f_{X_i}(x_i) \geq f_{Y_i}(y_i) \) for \( i = 1, \ldots, r \). Hence, by Corollary 3.4, we have \( f_{X_r}(x_r) \geq f_{Y_r}(y_r) \). On the other hand, it is clear that \( f_{X_r}(x_r) = 1 \) and \( f_{Y_r}(y_r) \geq 2 \), contradiction.

If \( r > s \), then \( r \geq s + 2 \) since \( r + s = 2m \). Now we consider the path from vertex \( x_{s+1} \) to \( y_s \). By Lemma 3.5, we have \( f_{Y_i}(y_i) \geq f_{X_{i+1}}(x_{i+1}) \) for \( i = 1, \ldots, s \). Further, by Corollary 3.4, we have \( f_{Y_s}(y_s) \geq f_{X_{s+1}}(x_{s+1}) \). Similarly, since \( f_{Y_s}(y_s) = 1 \) and \( f_{X_{s+1}}(x_{s+1}) \geq 2 \), contradiction. Therefore, \( r = s = m \).

Now by Lemma 3.5 applied to the path from \( x_m \) to \( y_m \), we have \( f_{X_i}(x_i) \geq f_{Y_i}(y_i) \) for \( i = 1, \ldots, m \). On the other hand, by Lemma 3.5 applied to the path from \( y_{m-1} \) to \( x_m \), we have \( f_{Y_i}(y_i) \geq f_{X_{i+1}}(x_{i+1}) \) for \( i = 1, 2, \ldots, m - 1 \). Hence, the assertion holds.

Case 2 or 3 can be handled in the same manner, we omit the details here. \( \square \)

Following the conditions in Lemma 3.6, we have the following.

Lemma 3.7.

(i) If case (i) of Lemma 3.6 holds, then

\[
f_T(x_1) \geq f_T(y_1) > f_T(x_2) \geq f_T(y_2) > \cdots > f_T(x_m) \geq f_T(y_m).
\]

Moreover, if \( f_T(x_k) = f_T(y_k) \) for some \( 1 \leq k \leq m \), then \( f_T(x_i) = f_T(y_i) \) for \( i = k, \ldots, m \).

(ii) If case (ii) of Lemma 3.6 holds, then

\[
f_T(x_1) > f_T(y_1) \geq f_T(x_2) > f_T(y_2) \geq \cdots \geq f_T(x_m) > f_T(y_m) \geq f_T(x_{m+1}).
\]

Moreover, if \( f_T(y_k) = f_T(x_{k+1}) \) for some \( 1 \leq k \leq m \), then \( f_T(y_i) = f_T(x_{i+1}) \) for \( i = k, \ldots, m \).

Proof. We only prove part (i), part (ii) is similar.

For any \( 2 \leq k \leq m \), let \( V_{k-1} \) be the connected component of \( T \) containing vertices \( x_{k-1} \) and \( y_{k-1} \) after removing the edges \( x_{k-1}x_k \) and \( y_{k-1}y_k \). For \( k = 1 \) and \( k = m \), it is easy to
see
\[ f_T(x_1) - f_T(y_1) = f_{X_1}(x_1)(1 + f_{X_{\geq 2}}(x_2)) - f_{Y_1}(y_1)(1 + f_{Y_{\geq 2}}(y_2)) \]
and
\[ f_T(x_m) - f_T(y_m) = f_{X_m}(x_m)(1 + f_{W_{m-1}}(x_{m-1})) - f_{Y_m}(y_m)(1 + f_{W_{m-1}}(y_{m-1})). \]
Moreover,
\[ f_T(x_k) = f_{X_k}(x_k)(1 + f_{X_{k+1}}(x_{k+1}))(1 + f_{W_{k-1}}(x_{k-1}) + f_{W_{k-1}}(x_{k-1}, \ldots, y_{k-1}))f_{Y_k}(y_k) \]
\[ (1 + f_{Y_{\geq 1}}(y_{k+1}))) \]  
(4)
and
\[ f_T(y_k) = f_{Y_k}(y_k)(1 + f_{Y_{\geq 1}}(y_{k+1}))(1 + f_{W_{k-1}}(y_{k-1}) + f_{W_{k-1}}(y_{k-1}, \ldots, x_{k-1})f_{X_k}(x_k) \]
\[ (1 + f_{X_{k+1}}(x_{k+1}))). \]  
(5)
By Equations (4) and (5), we have
\[ f_T(x_k) - f_T(y_k) = f_{X_k}(x_k)(1 + f_{W_{k-1}}(x_{k-1})) (1 + f_{X_{k+1}}(x_{k+1})) \]
\[ -f_{Y_k}(y_k)(1 + f_{W_{k-1}}(y_{k-1})) (1 + f_{Y_{\geq 1}}(y_{k+1}))). \]  
(6)
Now we claim that for \(1 \leq k \leq m - 1\),
\[ f_{X_{k+1}}(x_{k+1}) \geq f_{Y_{k+1}}(y_{k+1}). \]  
(7)
If there is at least one strict inequality in \(f_{X_i}(x_i) \geq f_{Y_i}(y_i)\) for \(i = 1, \ldots, k\), then by Lemma 3.5, (7) holds.
If \(f_{X_i}(x_i) = f_{Y_i}(y_i)\) for \(i = 1, \ldots, k\) and there exists a \(k < l < m\) such that \(f_{X_i}(x_i) = f_{Y_i}(y_i)\) for \(i = 1, \ldots, l - 1\) and \(f_{X_i}(x_i) > f_{Y_l}(y_l)\). Then by Lemma 3.5, we have \(f_{X_{l+1}}(x_{l+1}) \geq f_{Y_{l+1}}(y_{l+1})\). Moreover,
\[ f_{X_{k+1}}(x_{k+1}) = \sum_{j=k+1}^{l} \prod_{i=k+1}^{j} f_{X_i}(x_i) + f_{X_{l+1}}(x_{l+1}) \prod_{i=k+1}^{l} f_{X_i}(x_i) \]  
(8)
and
\[ f_{Y_{k+1}}(y_{k+1}) = \sum_{j=k+1}^{l} \prod_{i=k+1}^{j} f_{Y_i}(y_i) + f_{Y_{l+1}}(y_{l+1}) \prod_{i=k+1}^{l} f_{Y_i}(y_i). \]  
(9)
By Equations (8) and (9), the claim holds.
If \(f_{X_i}(x_i) = f_{Y_i}(y_i)\) for \(i = 1, \ldots, m\), then by Equations (8) and (9), we have \(f_{X_{k+1}}(x_{k+1}) = f_{Y_{k+1}}(y_{k+1})\) and the claim holds.
Hence, (7) is proved.
On the other hand, by Lemma 3.1, we have \(f_{W_{k-1}}(x_{k-1}) \geq f_{W_{k-1}}(y_{k-1})\). Together with (7), we see that (6) \(\geq 0\). Then \(f_T(x_k) \geq f_T(y_k)\).
Now we prove \(f_T(y_k) \geq f_T(x_{k+1})\) for any \(1 \leq k \leq m - 1\). Let \(U_k\) be the connected component of \(T\) containing vertex \(x_k\) after removing the edges \(y_{k-1}y_k\) (if \(k = 1\), let \(y_0 = x_1\)) and \(x_kx_{k+1}\). Then
\[ f_T(y_k) = f_{Y_k}(y_k)(1 + f_{Y_{\geq 1}}(y_{k+1}))(1 + f_{U_k}(y_{k-1}) + f_{U_k}(y_{k-1}, \ldots, x_k)f_{X_{k+1}}(x_{k+1}) \]
\[ (1 + f_{X_{\geq 1}}(x_{k+2}))) \]  
(10)
and
\[ f_T(x_{k+1}) = f_{X_{k+1}}(x_{k+1}) + f_{X_{k+2}}(x_{k+2}) + f_{U_1}(x_k) + f_{U_2}(x_k, \ldots, x_{k-1}) \]
\[ \times (1 + f_{Y_{k+1}}(y_{k+1})). \]  
(11)

Similar to 7, we can show that \( f_{Y_{k+1}}(y_{k+1}) \geq f_{X_{k+2}}(x_{k+2}) \). By Lemma 3.1, we have \( f_{U_1}(y_{k-1}) \geq f_{U_1}(x_k) \). Hence, (10) and (11) imply that
\[ f_T(y_k) - f_T(x_{k+1}) = f_{Y_{k+1}}(y_k)(1 + f_{U_2}(x_k, y_k)) + f_{Y_{k+1}}(y_k) \]
\[ - f_{X_{k+2}}(x_{k+2})(1 + f_{U_2}(x_k)) \]
\[ = f_{Y_{k+1}}(y_k)(1 + f_{U_2}(y_k)) - f_{X_{k+2}}(x_{k+2})(1 + f_{U_2}(x_k)) \geq 0. \]  
(12)

Moreover, if \( f_T(x_k) = f_T(y_k) \) for some \( 1 \leq k \leq m \), then by (6), we have
\[ f_{X_k}(x_k) = f_{Y_k}(y_k), \quad f_{X_{k+1}}(x_k) = f_{Y_{k+1}}(y_k), \quad f_{W_{k-1}}(x_k) = f_{W_{k-1}}(y_{k-1}). \]  
(13)

Since \( f_{X_k} = f_{X_{k+2}}(x_{k+2}) \) and \( f_{Y_k} = f_{Y_{k+2}}(y_{k+2}) \), we have
\[ f_{X_{k+1}}(x_{k+1}) = f_{Y_{k+1}}(y_{k+1}) \]  
(14)

Therefore, we have \( f_T(x_k) = f_T(y_i) \) for \( i = k, \ldots, m \).

Finally, we prove that \( f_T(y_i) > f_T(x_{i+1}) \) for \( i = 1, \ldots, m-1 \). Suppose that \( f_T(y_k) = f_T(x_{k+1}) \) for some \( 1 \leq k \leq m-1 \). Then by Equation (12), we have \( f_{Y_k}(y_k) = f_{X_{k+1}}(x_{k+1}) \) and
\[ f_{Y_{k+1}}(y_k) = f_{X_{k+2}}(x_{k+2}). \]
Moreover,
\[ f_{Y_{k+1}}(y_{k+1}) = f_{X_{k+2}}(x_{k+2}) + f_{X_{k+2}}(x_{k+2}) \]
\[ f_{Y_{k+1}}(y_{k+1}) = f_{X_{k+2}}(x_{k+2}) \]
\[ = f_{X_{k+2}}(x_{k+2}) + f_{X_{k+2}}(x_{k+2}). \]
Hence, \( f_{Y_{k+1}}(y_{k+1}) = f_{X_{k+2}}(x_{k+2}) \). Continuing this way in an inductive manner, we have
\[ f_{Y_{k+1}}(y_{m-1}) = f_{X_{k+2}}(x_{m}). \]
But \( f_{Y_{k+1}}(y_{m-1}) \geq 2 \) and \( f_{X_{k+2}}(x_{m}) = 1 \), contradiction.

Combining the above results, we have proved part (i).

The next Lemma relates the number of subtrees to the structure of the tree.

**Lemma 3.8.** For a path \( P(x_m, y_m) = x_m x_{m-1} \ldots x_2 x_1 (z) y_1 y_2 \ldots y_{m-1} y_m \) in an optimal tree \( T \), if \( f_T(x_i) \geq f_T(y_i) \) for \( i = 1, \ldots, k, 1 \leq k \leq m-1 \), then \( d(x_i) \geq d(y_i). \)

Moreover, if \( f_T(x_i) = f_T(y_i) \) for \( i = 1, \ldots, k, 1 \leq k \leq m-1 \), then \( d(x_i) = d(y_i). \)

**Proof.** Suppose that \( d(x_k) < d(y_k) \), let \( r = d(y_k) - d(x_k) \geq 1 \) and \( y_k u_i \in Y_{\geq k} \) for \( i = 1, \ldots, r \).

Further, let \( W \) be the connected component of \( T \) containing vertices \( x_k \) and \( y_k \) after removing the \( r \) edges \( y_k u_1, \ldots, y_k u_r \). Let \( X \) be the single vertex \( x_k \) and let \( Y \) be the connected component of \( T \) containing vertex \( y_k \) after removing all edges incident to \( y_k \) except for the \( r \) edges \( y_k u_1, \ldots, y_k u_r \). Since \( f_T(x_i) \geq f_T(y_i) \) for \( i = 1, \ldots, k \), it is easy
to see that \( f_w(x_k) > f_w(y_k) \) and \( f_v(x_k) = 1 < 2 \leq f_v(y_k) \). By Lemma 3.2, there exists another tree \( T' \in T_\pi \) such that \( \varphi(T) < \varphi(T') \), contradicting to the optimality of \( T \).

Therefore, the assertion holds. The case of equality is similar.

From Lemmas 3.6, 3.7, and 3.8, we have the following Lemma that decides the “center” of the optimal tree.

**Lemma 3.9.** Let \( T \) be an optimal tree in \( T_\pi \). If \( f_T(v_0) = \max\{f_T(v), v \in V(T)\} \), then \( d(v_0) = \max\{d(v), v \in V(T)\} \).

**Proof.** The assertion clearly holds for small trees, so we assume that \( |V(T)| \geq 4 \). Suppose that \( d(v_0) < \max\{d(v), v \in V(T)\} \). Then there exists a vertex \( w \) such that \( d(v_0) < d(w) \). By Theorem 9.1 in [6], \( f_T(v) \) is maximized at one or two adjacent vertices of \( T \). Thus, we have \( f_T(v_0) > f_T(v) \) for \( v \in V(T) \setminus \{v_0\} \), or \( f_T(v_0) = f_T(v_1) > f_T(v) \) for \( v \in V(T) \setminus \{v_0, v_1\} \) and \( v_0v_1 \in E(T) \).

**Case 1:** \( f_T(v_0) > f_T(v) \) for any \( v \in V(T) \setminus \{v_0\} \). Hence, \( f_T(v_0) > f_T(w) \). It is easy to see that \( v_0 \) is not a leaf (otherwise, let \( u \) be a neighbor of \( v_0 \) and we have \( f_T(u) > f_T(v_0) \)). Let \( P \) be a path containing vertex \( v_0 \) and \( w \) whose end vertices are leaves. Let the length of \( P \) be \( 2m - 1 \) (the even length case is similar). Then by Lemma 3.6, the vertices of \( P \) can be labeled as \( P = x_m \ldots x_1y_1 \ldots y_m \) such that

\[
f_{x_1}(x_1) \geq f_{x_1}(y_1) \geq f_{x_2}(x_2) \geq f_{x_2}(y_2) \geq \cdots \geq f_{x_n}(x_m) = f_{x_n}(y_m) = 1.
\]

Hence, by Lemma 3.7, we have

\[
f_T(x_1) \geq f_T(y_1) \geq f_T(x_2) \geq f_T(y_2) \geq \cdots \geq f_T(x_m) \geq f_T(y_m).
\]

Therefore, \( x_1 \) must be \( v_0 \) and \( w \) must be \( x_k \) for \( 2 \leq k \leq m \) or \( y_j \) for \( 1 \leq j \leq m \). By Lemma 3.8, we have \( d(v_0) = d(x_1) \geq d(x_k) = d(w) \) or \( d(v_0) = d(x_1) \geq d(y_j) = d(w) \), contradiction. Hence, the assertion holds.

**Case 2:** \( f_T(v_0) = f_T(v_1) > f_T(v) \) for \( v \in V(T) \setminus \{v_0, v_1\} \) and \( v_0v_1 \in E(T) \). If \( w = v_1 \), then by Lemma 3.8, we have \( d(w) = d(v_1) = d(v_0) < d(w) \), contradiction.

Hence, we assume that \( w \neq v_1 \). First note that \( v_0 \) and \( v_1 \) are not leaves. Let \( P \) be a path containing vertices \( v_0, v_1 \), and \( w \) whose end vertices are leaves. Let the length of \( P \) be \( 2m - 1 \) (the even case is similar), then by Lemma 3.6, the vertices of \( P \) can be labeled as \( P = x_m \ldots x_1y_1 \ldots y_m \) such that

\[
f_{x_1}(x_1) \geq f_{x_1}(y_1) \geq f_{x_2}(x_2) \geq f_{x_2}(y_2) \geq \cdots \geq f_{x_n}(x_m) \geq f_{x_n}(y_m) = 1.
\]

Hence, by Lemma 3.7, we have

\[
f_T(x_1) \geq f_T(y_1) \geq f_T(x_2) \geq f_T(y_2) \geq \cdots \geq f_T(x_m) \geq f_T(y_m).
\]

Therefore, \( \{x_1, y_1\} = \{v_0, v_1\} \) and \( w \) must be \( x_k \) or \( y_k \) for \( 1 \leq k \leq m \). By Lemma 3.8, \( d(v_0) \geq d(w) \) and \( d(v_1) \geq d(w) \), contradiction.

Combining cases (1) and (2), the assertion is proved.

**Lemma 3.10.** Let \( T \) be an optimal tree in \( T_\pi \). If there is a path \( P = u_1u_{l-1} \ldots u_1v_0v_1 \ldots v_k \) with \( f_T(v_0) = \max\{f_T(v) : v \in V(P)\} \), \( f_T(u_l) \geq f_T(v_1) \), and \( l = k \) (or \( l = k + 1 \)), then

\[
f_T(u_1) \geq f_T(v_1) \geq f_T(u_2) \geq \cdots \geq f_T(u_k) \geq f_T(v_k) \) (or \( \geq f_T(u_{k+1}) \))
\]
and
\[ d(u_1) \geq d(v_1) \geq d(u_2) \geq \cdots \geq d(u_k) \geq d(v_k) \] (or \( \geq d(u_{k+1}) \)).

**Proof.** Clearly, there exists a path \( Q \) that contains the path \( P \) and its end vertices are leaves. We assume \( l = k \) (the \( l = k + 1 \) case is similar).

Let the length of \( Q \) be \( 2m - 1 \) (the even length case is similar). By Lemmas 3.7 and 3.8, the vertices of \( Q \) can be labeled as \( Q = x_mx_{m-1} \cdots x_1y_1 \cdots y_m \) such that
\[ f_T(x_1) \geq f_T(y_1) > f_T(x_2) \geq f_T(y_2) > \cdots > f_T(x_m) \geq f_T(y_m) \]
and
\[ d(x_1) \geq d(y_1) \geq d(x_2) \geq d(y_2) \geq \cdots \geq d(x_m) = d(y_m) = 1. \]

**Case 1:** \( v_0 = x_1 \). We must have \( u_1 = y_1 \) and \( v_1 = x_2 \). Then \( u_i = y_i \) and \( v_i = x_{i+1} \) for \( i = 1, \ldots, k \). Hence, the assertion holds.

**Case 2:** \( v_0 = x_i \) for \( i > 1 \). Then \( f_T(v_0) \geq f_T(x_1) \geq f_T(y_1) \geq f_T(x_2) = f_T(v_0) \), which implies \( f_T(x_1) = f_T(y_1) = f_T(v_0) \) and contradicts to Theorem 9.1 in [6].

**Case 3:** \( v_0 = y_i \). Then \( i = 1 \) and \( f_T(x_1) = f_T(y_1) = f_T(v_0) \). We must have \( u_1 = x_1 \) and \( v_1 = y_2 \). Then \( u_i = x_i \) and \( v_i = y_{i+1} \) for \( i = 1, \ldots, k \). So the assertion holds.

Now for an optimal tree \( T \) in \( T\pi \), let \( v_0 \in V(T) \) be the root of \( T \) with \( f_T(v_0) = \max \{ f_T(v) : v \in V(T) \} \) and \( d(v_0) = \max \{ d(v) : v \in V(T) \} \).

**Corollary 3.11.** If there is a path \( P = u_k \ldots u_1wv_1v_2 \ldots v_k \) with \( \text{dist}(u_k, v_0) = \text{dist}(v_k, v_0) = \text{dist}(w, v_0) + k \) and \( f_T(u_1) \geq f_T(v_1) \), then
\[ f_T(u_1) \geq f_T(v_1) \geq f_T(u_2) \geq \cdots \geq f_T(u_k) \geq f_T(v_k) \]
and
\[ d(u_1) \geq d(v_1) \geq d(u_2) \geq \cdots \geq d(u_k) \geq d(v_k). \]

If there is a path \( P = u_{k+1} \ldots u_1wv_1v_2 \ldots v_k \) with \( \text{dist}(u_{k+1}, v_0) = \text{dist}(v_k, v_0) + 1 = \text{dist}(w, v_0) + k + 1 \) and \( f_T(u_1) \geq f_T(v_1) \), then
\[ f_T(u_1) \geq f_T(v_1) \geq f_T(u_2) \geq \cdots \geq f_T(u_k) \geq f_T(v_k) \geq f_T(u_{k+1}) \]
and
\[ d(u_1) \geq d(v_1) \geq d(u_2) \geq \cdots \geq d(u_k) \geq d(v_k) \geq d(u_{k+1}). \]

**Proof.** If \( w = v_0 \), then the assertion follows from Lemma 3.10. If \( w \neq v_0 \), then there exists a path \( Q \) containing vertices \( u_k, \ldots, u_1, w, v_0 \) whose end vertices are leaves. By Lemma 3.10, we have \( f_T(w) \geq f_T(u_1) \geq \cdots \geq f_T(u_k) \). Similarly, there exists a path \( R \) containing vertices \( v_k, \ldots, v_1, w, v_0 \) whose end vertices are leaves and we have \( f_T(w) \geq f_T(v_1) \geq \cdots \geq f_T(v_k) \). Therefore, \( f_T(w) = \max \{ f_T(v) : v \in V(P) \} \), the assertion follows from Lemma 3.10.

*Journal of Graph Theory* DOI 10.1002/jgt
4. PROOFS OF THEOREMS 2.3 AND 2.4

Now we are ready to prove Theorems 2.3 and 2.4.

**Proof.** of Theorem 2.3. Let $T$ be an optimal tree in $T_v$. By Lemma 3.9, there exists a vertex $v_0$ such that $f_T(v_0) = \max \{ f_T(v) : \ v \in V(T) \}$ and $d(v_0) = \max \{ d(v) : \ v \in V(T) \}$. Let $v_0$ be the root of $T$ and put $V_i = \{ v : \ \text{dist}(v, v_0) = i \}$ for $i = 0, \ldots, p + 1$ with $V(T) = \bigcup_{i=0}^{p+1} V_i$. Denote by $|V_i| = s_i$ for $i = 1, \ldots, p + 1$. We now can relabel the vertices of $V(T)$ by the recursion method. For $v_0$, relabel $v_0$ by $v_{01}$ as the root of tree $T$. The vertices of $V_1$ (consisting of all neighbors $v_{01}$) are relabeled as $v_{11}, \ldots, v_{1,s_1}$, satisfying:

$$f_T(v_{11}) \geq f_T(v_{12}) \geq \cdots \geq f_T(v_{1,s_1})$$

and

$$f_T(v_{1i}) = f_T(v_{1j}) \implies d(v_{1i}) \geq d(v_{1j}) \text{ for } 1 \leq i < j \leq s_1.$$  

Generally, we assume that all vertices of $V_i$ are relabeled as $\{ v_{i1}, \ldots, v_{is_i} \}$ for $i = 1, \ldots, t$. Now consider all vertices in $V_{i+1}$. Since $T$ is tree, it is easy to see that $s_1 = d(v_{01})$ and

$$s_{i+1} = |V_{i+1}| = d(v_{11}) + \cdots + d(v_{is_i}) - s_i.$$  

Hence, for $1 \leq r \leq s_i$, all neighbors in $V_{i+1}$ of $v_{ir}$ are relabeled as

$$v_{i+1,d(v_{i+1})+\cdots+d(v_{i+1})+(r-1)+1}, \ldots, v_{i+1,d(v_{i+1})+\cdots+d(v_{is_i})-r}$$

and satisfy the conditions:

$$f_T(v_{i+1,i}) \geq f_T(v_{i+1,j})$$  

and

$$f_T(v_{i+1,i}) = f_T(v_{i+1,j}) \implies d(v_{i+1,i}) \geq d(v_{i+1,j})$$

for $d(v_{1i}) + \cdots + d(v_{i+1,j}) - (r - 1) + 1 \leq i < j \leq d(v_{1i}) + \cdots + d(v_{i+1,j}) - r$. In this way, we have relabeled all vertices of $V(T) = \bigcup_{i=0}^{p+1} V_i$. Therefore, we are able to define a well ordering of vertices in $V(T)$ as follows:

$$v_{ik} < v_{jl}, \text{ if } 0 \leq i < j \leq p + 1 \text{ or } i = j \text{ and } 1 \leq k < l \leq s_i.$$  

We need to prove that this well ordering is a BFS-ordering of $T$. In other words, $T$ is isomorphic to $T_v$.

We first prove, for $t = 0, \ldots, p + 1$, the following inequalities.

$$f_T(v_{i1}) \geq f_T(v_{i2}) \geq \cdots \geq f_T(v_{is_i}) \geq f_T(v_{i+1,1})$$

and

$$d(v_{i1}) \geq d(v_{i2}) \geq \cdots \geq d(v_{is_i}) \geq d(v_{i+1,1}).$$

For any two vertices $v_{il}$ and $v_{lj}$ with $1 \leq i < j \leq s_l$, there exists a path $P = v_{il} \cdots v_{k+1,l} v_{k+1,r} \cdots v_{lj}$ with $l < r$, where $\text{dist}(v_{il}, v_{01}) = \text{dist}(v_{lj}, v_{01}) = \text{dist}(v_{kl}, v_{01}) + t - k$. Then we have $f_T(v_{k+1,l}) \geq f_T(v_{k+1,r})$, $f_T(v_{i1}) \geq f_T(v_{lj})$ and $d(v_{i1}) \geq d(v_{lj})$ by Corollary 3.11. On the other hand, we consider the path $Q = v_{i+1,1} v_{i1} \cdots v_{i1} v_{01} v_{i1} \cdots v_{i1}$.
Then \( f_T(v_{i,s}) \geq f_T(v_{i+1,1}) \) and \( d(v_{i,s}) \geq d(v_{i+1,1}) \) by Corollary 3.11. Therefore, (18) and (19) hold for \( t = 0, \ldots, p + 1 \). That is
\[
f_T(v_{01}) \geq f_T(v_{11}) \geq \cdots \geq f_T(v_{1,s_1}) \geq f_T(v_{21}) \geq \cdots \geq f_T(v_{p+1,s_{p+1}})
\]
and
\[
d(v_{01}) \geq d(v_{11}) \geq d(v_{1,s_1}) \geq d(v_{21}) \geq \cdots \geq d(v_{p+1,s_{p+1}}).
\]
(20)

By (17), (20), and (21), it is easy to see that this well ordering satisfies all conditions in Definition 2.1. Hence, \( T \) has a BFS-ordering. Further, by Proposition 2.2 in [13], \( T \) is isomorphic to \( T^{\pi}_n \). So \( T^{\pi}_n \) is the unique optimal tree in \( T_n \) having the largest number of subtrees.

Proof. of Theorem 2.4. By proposition 2.2, without loss of generality, we assume that \( \pi = (d_0, d_1, \ldots, d_i, \ldots, d_{n-1}) \) and \( \pi_1 = (d_0, d_1, \ldots, d_i + 1, \ldots, d_j - 1, \ldots, d_{n-1}) \) with \( i < j \), then we have \( \pi \prec \pi_1 \). Let \( T^{\pi}_n \) be the optimal tree in \( T_n \).

By the proof of Theorem 2.3, the vertices of \( T^{\pi}_n \) can be labeled as the \( V = \{v_0, \ldots, v_{n-1}\} \) such that
\[
f_T^{\pi_n}(v_0) \geq f_T^{\pi_n}(v_1) \geq \cdots \geq f_T^{\pi_n}(v_{n-1})
\]
and
\[
d(v_0) \geq d(v_1) \geq \cdots \geq d(v_{n-1}),
\]
where \( d(v_l) = d_l \) for \( l = 0, \ldots, n - 1 \). Moreover, \( v_0 \) is the root of \( T^{\pi}_n \). There exists a vertex \( v_k \) such that \( v_j v_k \in E(T^{\pi}_n) \) with \( k > j \). Let \( W \) be the tree achieved from \( T^{\pi}_n \) by removing the subtree induced by \( v_k \). Moreover, let \( X \) be the single vertex \( v_j \) and \( Y \) be the subtree induced by \( v_k \) with the edge \( v_j v_k \) added, respectively. Clearly, \( f_T(v_j) = f_W(v_j) + f_W(v_j)(f_Y(v_j) - 1) \) and \( f_T(v_j) = f_W(v_j) + f_Y(v_j)(f_Y(v_j) - 1) \). Hence, by \( f_W(v_j) < f_W(v_j) \) and \( f_W(v_j) \geq f_T(v_j) \), we have \( f_W(v_j) = f_Y(v_j) \). On the other hand, let \( T_1 \) be the tree from \( T \) by deleting the edge \( v_j v_k \) and adding the edge \( v_j v_k \). Then the degree sequence of \( T_1 \) is \( \pi_1 \). By Lemma 3.2, we have \( \varphi(T^{\pi}_n) < \varphi(T_1) \). Hence, \( \varphi(T^{\pi}_n) < \varphi(T^{\pi_1}_n) \).

The assertion is then proved.

5. APPLICATIONS OF THE MAIN THEOREMS

In the end, we use Theorems 2.3 and 2.4 to achieve extremal graphs with the largest number of subtrees in some classes of graphs. As corollaries, we provide proofs to some results in [3, 6], etc.

Let \( T_{n,\Delta}^{(1)} \) be the set of all trees of order \( n \) with the largest degree \( \Delta \), \( T_{n,s}^{(2)} \) be the set of all trees of order \( n \) with \( s \) leaves, \( T_{n,\alpha}^{(3)} \) be the set of all trees of order \( n \) with the independence number \( \alpha \) and \( T_{n,\beta}^{(4)} \) be the set of all trees of order \( n \) with the matching number \( \beta \).

Corollary 5.1. [6] Let \( T \) be any tree of order \( n \). Then
\[
\left( \frac{n+1}{2} \right) \leq \varphi(T) \leq 2^{n-1} + n - 1
\]

Journal of Graph Theory DOI 10.1002/jgt
with left equality if and only if \( T \) is a path of order \( n \) and the right equality if and only if \( T \) is the star \( K_{1,n-1} \).

**Proof.** Let \( T \) be a tree of order \( n \) with degree sequence \( \tau \). Let \( \pi_1 = (2, \ldots, 2, 1, 1) \) and \( \pi_2 = (n-1, 1, \ldots, 1) \) with \( n \) terms. Clearly, the path \( P \) of order \( n \) is the only tree with the degree sequence \( \pi_1 \) and the star \( K_{1,n-1} \) of order \( n \) is the only tree with degree sequence \( \pi_2 \). Furthermore, \( \pi_1 \preceq \tau \preceq \pi_2 \). Hence, by Theorems 2.3 and 2.4, the assertion holds.

**Corollary 5.2.** \([3]\) There is only one optimal tree \( T_\Delta^* \) in \( T_{n,\Delta}^{(1)} \) with \( \Delta \geq 3 \), where \( T_\Delta^* \) is \( T_\Delta^* \) with degree sequence \( \pi \) as follows: denote \( p = \lceil \log_{\Delta-1} \frac{n(\Delta-2)+2}{\Delta} \rceil - 1 \) and \( n - \frac{\Delta(\Delta-1)^{p-1}}{\Delta-2} = (\Delta - 1)r + q \) for \( 0 \leq q < \Delta - 1 \). If \( q = 0 \), put \( \pi = (\Delta, \ldots, \Delta, 1, \ldots, 1) \) with the number \( \frac{\Delta(\Delta-1)^{p-1}}{\Delta-2} + r \) of degree \( \Delta \). If \( q \geq 1 \), put \( \pi = (\Delta, \ldots, \Delta, q, 1, \ldots, 1) \) with the number \( \frac{\Delta(\Delta-1)^{p-1}}{\Delta-2} + r \) of degree \( \Delta \).

**Proof.** For any tree \( T \) of order \( n \) with the largest degree \( \Delta \), let \( \pi_1 = (d_0, \ldots, d_{n-1}) \) be the nonincreasing degree sequence of \( T \). Assume that \( T_\Delta^* \) has \( p + 2 \) layers. Then there is a vertex in layer 0 (i.e., the root), there are \( \Delta \) vertices in layer 1, there are \( \Delta(\Delta-1) \) vertices in layer 2, \ldots, there are \( \Delta(\Delta-1)^{p-1} \) vertices in layer \( p \), there are at most \( \Delta(\Delta-1)^{p} \) vertices in layer \( p + 1 \). Hence,\[
1 + \Delta + \Delta(\Delta-1) + \cdots + \Delta(\Delta-1)^{p-1} < n \leq 1 + \Delta + \Delta(\Delta-1) + \cdots + \Delta(\Delta-1)^{p}.
\]
Thus\[
\frac{\Delta(\Delta-1)^{p} - 2}{\Delta - 2} < n \leq \frac{\Delta(\Delta-1)^{p+1} - 2}{\Delta - 2}.
\]
Hence,\[
p = \lceil \log_{\Delta-1} \frac{n(\Delta-2)+2}{\Delta} \rceil - 1
\]
and there exist integers \( r \) and \( 0 \leq q < \Delta - 1 \) such that\[
n - \frac{\Delta(\Delta-1)^{p} - 2}{\Delta - 2} = (\Delta - 1)r + q.
\]
Therefore, the degrees of all vertices from layer 0 to layer \( p - 1 \) are \( \Delta \) and there are \( r \) vertices in layer \( p \) with degree \( \Delta \). Denote by \( m = \frac{\Delta(\Delta-1)^{p-1}}{\Delta-2} + r - 1 \). Then there are \( m + 1 \) vertices with degree \( \Delta \) in \( T_\Delta^* \). Hence, the degree sequence of \( T_\Delta^* \in T_{n,\Delta} \) is \( \pi = (d_0', \ldots, d_{n-1}') \) with \( d_0' = \cdots = d_m' = \Delta, d_{m+1}' = \cdots = d_{n-1}' = 1 \) for \( q = 0 \); and is \( \pi = (d_0', \ldots, d_{n-1}') \) with \( d_0' = \cdots = d_q' = \Delta, d_{q+1}' = q, d_{q+2}' = \cdots = d_{n-1}' = 1 \) for \( q > 0 \). It follows from \( d_i \leq \Delta \) that \( \sum_{i=0}^{k} d_i \leq \sum_{i=0}^{k} d_i' \) for \( k = 0, \ldots, m \). Further, by \( d_i' = 1 \leq d_i \) for \( k = m + 2, \ldots, n - 1 \), we have\[
\sum_{i=0}^{k} d_i = 2(n - 1) - \sum_{i=k+1}^{n-1} d_i \leq 2(n - 1) - \sum_{i=k+1}^{n-1} d_i' = \sum_{i=0}^{k} d_i'
\]
for \( k = m + 1, \ldots, n - 1 \). Thus \( \pi_1 \preceq \pi \). Hence, by Theorems 2.3 and 2.4, \( \varphi(T) \leq \varphi(T_\Delta^*) \) with equality if and only if \( T = T_\Delta^* \). ■
Remark. If \( \Delta = 3 \) in Corollary 5.2, then the result is precisely Theorem 2.1 in [8].

**Corollary 5.3.** There is only one optimal tree \( T^*_s \) in \( T^{(2)}_{n,t} \) where \( T^*_s \) is obtained from \( t \) paths of order \( q + 2 \) and \( s - t \) paths of order \( q + 1 \) by identifying one end of the \( s \) paths. Here \( n - 1 = sq + t, 0 \leq t < s \). In other words, for any tree of order \( n \) with \( s \) leaves,

\[
\varphi(T) \leq (q + 2)^t (q + 1)^{s-t} + \frac{(q + 1)(qs + 2t)}{2}
\]

with equality if and only if \( T \) is \( T^*_s \).

**Proof.** Let \( T \) be any tree in \( T^{(2)}_{n,t} \) with the nonincreasing degree sequence \( \pi_1 = (d_0, \ldots, d_{n-1}) \). Thus \( d_{n-s-1} > 1 \) and \( d_{n-s} = \cdots = d_{n-1} = 1 \). Let \( T^*_s \) be a BFS-tree with degree sequence \( \pi = (s, 2, \ldots, 1) \), where there are the number \( s \) of \( 1 \)'s in \( \pi \). It is easy to see that \( \pi_1 \ll \pi \). By Theorem 2.4, the assertion holds. \( \square \)

**Corollary 5.4.** There is only one optimal tree \( T^*_s \) in \( T^{(3)}_{n,t} \), where \( T^*_s \) is \( T^*_s \) with degree sequence \( \pi = (\alpha, 2, \ldots, 2, 1, \ldots, 1) \) with numbers \( n - \alpha - 1 \) of \( 2 \)'s and \( \alpha \) of \( 1 \)'s, i.e., \( T^*_s \) is obtained from the star \( K_{1,\alpha} \) by adding \( n - \alpha - 1 \) pendant edges to \( n - \alpha - 1 \) leaves of \( K_{1,\alpha} \). In other words, for any tree of order \( n \) with the independence number \( \alpha \),

\[
\varphi(T) \leq 2^{2n-\alpha+1}3^{n-\alpha-1} + 2n + \alpha - 2
\]

with equality if and only if \( T \) is \( T^*_s \).

**Proof.** For any tree \( T \) of order \( n \) with the independence number \( \alpha \), let \( I \) be an independent set of \( T \) with size \( \alpha \) and \( \tau = (d_0, \ldots, d_{n-1}) \) be the degree sequence of \( T \). If there exists a leaf \( u \) with \( u \notin I \), then there exists a vertex \( v \in I \) with \( (u, v) \in E(T) \). Hence, \( I \cup \{u\} \setminus \{v\} \) is an independent set of \( T \) with size \( \alpha \). Therefore, one can always construct an independent set of \( T \) with size \( \alpha \) that contains all leaves of \( T \). Hence, there are at most \( \alpha \) leaves. Then \( d_{n-\alpha-1} \geq 2 \) and \( \tau \ll \pi \). By Theorems 2.3 and 2.4, the assertion holds. \( \square \)

**Corollary 5.5.** There is only one optimal tree \( T^*_s \) in \( T^{(4)}_{n,\beta} \), where \( T^*_s \) is \( T^*_s \) with degree sequence \( \pi = (n - \beta, 2, \ldots, 2, 1, \ldots, 1) \). Here, the number of \( 1 \)'s is \( n - \beta \). That is, \( T^*_s \) is obtained from the star \( K_{1,n-\beta} \) by adding \( \beta - 1 \) pendant edges to \( \beta - 1 \) leaves of \( K_{1,n-\beta} \). In other words, for any \( T \in T^{(4)}_{n,\beta} \),

\[
\varphi(T) \leq 2^{n-2\beta+1}3^{\beta-1} + n - \beta - 2
\]

with equality if and only if \( T \) is \( T^*_s \).

**Proof.** For any tree \( T \) of order \( n \) with matching number \( \beta \), let \( \tau = (d_0, \ldots, d_{n-1}) \) be the degree sequence of \( T \). Let \( M \) be a matching of \( T \) with size \( \beta \). Since \( T \) is connected, there are at least \( \beta \) vertices in \( T \) such that their degrees are at least two. Hence, \( d_{\beta-1} \geq 2 \) and \( \tau \ll \pi \). By Theorems 2.3 and 2.4, the assertion holds. \( \square \)

**ACKNOWLEDGMENTS**

The authors would like to thank the anonymous referees very much for valuable suggestions, corrections, and comments.

*Journal of Graph Theory* DOI 10.1002/jgt
REFERENCES