The Number of Subtrees of Trees with Given Diameter

Zichong CHEN
Jinling High School, Hexi Campus, Nanjing 210019, P.R. China

Abstract: A tree is a connected acyclic graph. A subtree of a tree T is a tree whose vertex set is the subset of the vertex set of T. Let $\mu(T)$ denote the number of subtrees of a tree T. Székely and Wang [*On subtrees of trees*, *Advances in Applied Mathematics*, 34(2005), 138-155] showed that $\mu(P_n) \leq \mu(T) \leq \mu(S_n)$ for any tree of order n, where P_n and S_n are a path and a star of order n, respectively.

In this paper, we consider the same problem with the condition that the diameter of a tree is given. Let $\mathscr{T}_{n,d}$ denote the set of all trees of order n with diameter d. We obtain the following three new results:

(1) For any $T \in \mathscr{T}_{n,d}$,

$$\mu(T) \leq \left\lceil \frac{d-2}{2} \right\rceil \left\lfloor \frac{d}{2} \right\rfloor + \left(\left\lfloor \frac{d}{2} \right\rfloor + 1 \right) \left(\left\lceil \frac{d}{2} \right\rceil + 1 \right) 2^{n-d-1} + n - 1;$$

(2) If $n \geq 8$ and $d = 4$, then

$$\mu(T) \geq 3^{\ell} 5^{\frac{n-2\ell-1}{3}} + 2n - 2 - \ell,$$

where $n \equiv r \pmod 3$ with $0 \leq r < 3$, and $4 - r \equiv \ell \pmod 3$ with $0 \leq \ell < 3$;
(3) If $n \geq 6$, $d = 5$ and $n \equiv r_i \pmod i$ for $i = 2, 3, 6$ with $0 \leq r_i < i$, then

$$\mu(T) \geq 3^{2-r_3} 5^{\frac{n+2r_3-6}{3}} + (6 - 2r_3) 5^{\frac{n+2r_3-6}{6}} + 2n - 6 + r_3$$

if $r_2 = 0$,

$$\mu(T) \geq 3^{2-r_3} 5^{\frac{n+2r_3-6}{3}} + 5^{\frac{n+2r_3-3}{6}} + 3^{2-r_3} 5^{\frac{n+2r_3-9}{6}} + 2n - 6 + r_3$$

if $r_2 = 1$.

All trees such that the equalities hold in (1), (2) or (3) are completely characterized.
Key words: Tree, The number of subtrees, Diameter

1. Introduction

A tree is a connected acyclic graph. Let T be a tree with vertex set $V(T)$ and edge set $E(T)$. For $v \in V(T)$, the *neighborhood* of v is defined as $N(v) = \{u \mid u \text{ is adjacent to } v\}$ and $N[v] = N(v) \cup \{v\}$. The *degree* of v is $d(v) = |N(v)|$. For $S \subseteq V(T)$, $G[S]$ denotes the subgraph induced by S and $T - S$ the subgraph obtained from T by deleting all the vertices in S and the edges with at least one end-vertex in S. If $G[S]$ is a tree, then $G[S]$ is called a *subtree* of T. An edge connecting two vertices u and v is denoted by uv. If $uv \in E(T)$, then $T - uv$ is a graph obtained from T by deleting the edge uv and if $uv \notin E(T)$, then $T + uv$ is a graph obtained from T by adding the edge uv to T. For two trees T and T', $T \cup T'$ is the vertex disjoint union of T and T', and ℓT is the vertex disjoint union of ℓ copies of T. A *path* is a tree in which the degree of each vertex is at most two and a *star* is a tree which has one vertex adjacent to all other vertices. A path and a star on n vertices are denoted by P_n and S_n, respectively. A single vertex is also called a path or a star. A *caterpillar* is a tree whose vertices of degree at least two induces a path. For $u, v \in V(T)$, the *distance* of them, denoted by $d(u, v)$, is the length of the only path connecting u and v in T, and the *diameter* of T, denoted by $diam(T)$, is the maximum of distances taken over all pairs of vertices in T.

Let T be a tree and $v \in V(T)$. We use $\mu(T)$, $\mu(T, v)$ and $\mu(T, \overline{v})$ to denote the number of the subtrees of T, the subtrees containing v in T and the subtrees not containing v in T, respectively. Clearly, $\mu(T) = \mu(T, v) + \mu(T, \overline{v})$ for any $v \in V(T)$. Figure 1 is an example illustrating all the subtrees of a given tree T on 5 vertices, where T-k denotes the subtrees on k vertices of T. This tree T has $5 + 4 + 4 + 3 + 1 = 17$ subtrees.

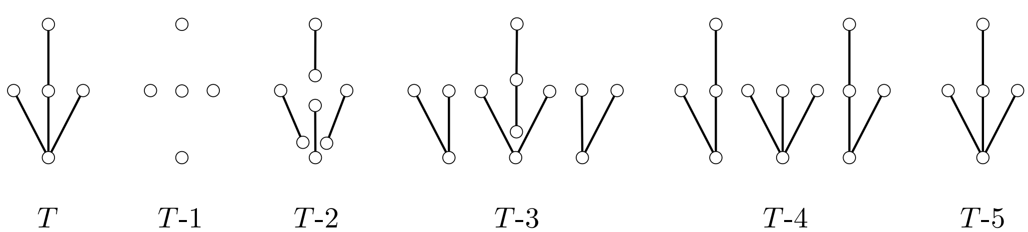

Figure 1. A tree T and all its subtrees

For a given tree on n vertices, Székely and Wang first established the best possible bounds for the number of the subtrees of it.

Theorem 1 (Székely and Wang [2]). Let T be any tree of order n, then $\mu(P_n) \leq \mu(T) \leq \mu(S_n)$.

Except the bounds for all trees on n vertices, there are many other results concerning the bounds for the number of subtrees of trees in some given subclasses on n vertices, see for instance [1, 3]. In this paper, we investigate the maximum and minimum values of the number of subtrees of a tree on n vertices with diameter d. Let $\mathscr{T}_{n,d}$ denote the set of all trees of order n with diameter d. To state our main results, we first define three classes of trees as follows.

- $T^*_{n,d}$: the caterpillar of order n with diameter path $P = v_1v_2\cdots v_{d+1}$ such that all the vertices not in P are adjacent to $v_{\lfloor d/2\rfloor+1}$.
- T_n^r: a tree of order n obtained from ℓ copies of P_3 and $\frac{n-2\ell-1}{3}$ copies of S_4, by taking one vertex of degree one in each copy, and then identifying the $\ell+\frac{n-2\ell-1}{3}$ chosen vertices into one vertex, where $n \equiv r \pmod 3$, $4-r \equiv \ell \pmod 3$ and $0 \le r, \ell < 3$, see Figure 1. The identifying vertex is called the root of T_n^r.

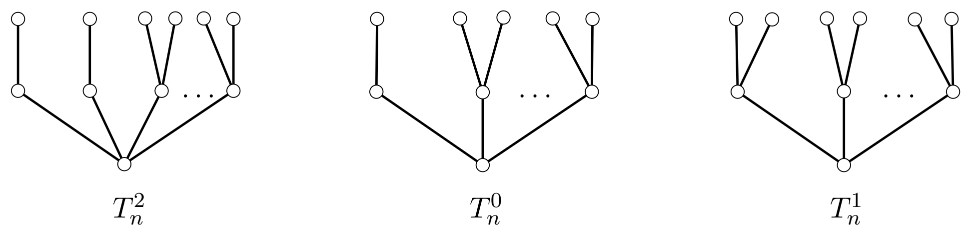

Figure 1. Extremal Trees in Theorem 3

- $H_n(n_1)$: a tree of order n obtained from a tree $T_{n_1}^{r_1}$ and a tree $T_{n-n_1}^{r_2}$, by adding a new edge connecting their roots.

The main results of this paper are as follows.

Theorem 2. Let $T \in \mathscr{T}_{n,d}$. Then

$$\mu(T) \le \left\lceil \frac{d-2}{2} \right\rceil \left\lfloor \frac{d}{2} \right\rfloor + \left(\left\lfloor \frac{d}{2} \right\rfloor + 1 \right) \left(\left\lceil \frac{d}{2} \right\rceil + 1 \right) 2^{n-d-1} + n - 1$$

and the equality holds if and only if $T \cong T^*_{n,d}$.

Since $S_n \cong T^*_{n,2}$ is the only tree of order n with diameter 2, and $\mu(T^*_{n,d})$ is a decreasing function of d, by Theorem 2, we have the following corollaries which were obtained in [2].

Corollary 1. $\mu(S_n) = 2^{n-1} + n - 1$.

Corollary 2. For any tree T of order n, $\mu(T) \le \mu(S_n)$.

We are not able to establish the sharp lower bound for general d. If $d = 2$, then $\mathscr{T}_{n,2}$ has only one element and so there is nothing to do. If $d = 3$, say $v_1v_2v_3v_4$ is a diameter path of T, then $T - v_2v_3$ are two stars T_1 with $v_2 \in V(T_1)$ and T_2 with $v_3 \in V(T_2)$. Assume that $|V(T_i)| = n_i$ for $i = 1, 2$. Then $\mu(T) = \mu(T_1, v_2)\cdot\mu(T_2, v_3) + \mu(T_1) + \mu(T_2)$.

By Corollary 1, $\mu(T) = 2^{n_1-1} + 2^{n_2-1} + 2^{n-2} + n - 2$, which takes its minimum if $|n_1 - n_2| = 0$ or 1. In this paper, we will give the sharp lower bound for $\mu(T)$ when $d = 4$ or 5, and characterize all extremal trees.

Theorem 3. Let $T \in \mathscr{T}_{n,4}$, $n \geq 8$, $n \equiv r \pmod 3$ with $0 \leq r < 3$ and $4 - r \equiv \ell \pmod 3$ with $0 \leq \ell < 3$. Then

$$\mu(T) \geq 3^{\ell} 5^{\frac{n-2\ell-1}{3}} + 2n - 2 - \ell$$

and the equality holds if and only if $T \cong T_n^r$.

Theorem 4. Let $T \in \mathscr{T}_{n,5}$ and $n \equiv r_i \pmod i$ for $i = 2, 3, 6$ with $0 \leq r_i < i$. Then

$$\mu(T) \geq 3^{2-r_3} 5^{\frac{n+2r_3-6}{3}} + (6 - 2r_3) 5^{\frac{n+2r_3-6}{6}} + 2n - 6 + r_3$$

if $r_2 = 0$,

$$\mu(T) \geq 3^{2-r_3} 5^{\frac{n+2r_3-6}{3}} + 5^{\frac{n+2r_3-3}{6}} + 3^{2-r_3} 5^{\frac{n+2r_3-9}{6}} + 2n - 6 + r_3$$

if $r_2 = 1$, and the equalities hold if and only if $T \cong H_n(\lceil \frac{n}{2} \rceil + \lfloor \frac{r_6}{4} \rfloor)$.

2. The number of subtrees of caterpillars in $\mathscr{T}_{n,d}$

In this section, we will establish the sharp upper bound for the number of subtrees of a caterpillar and characterize all extremal trees. Let $T \in \mathscr{T}_{n,d}$ be a caterpillar with diameter path $P = v_1 v_2 \cdots v_{d+1}$ and $R(T) = V(T) - V(P)$.

If $R(T) = \emptyset$, then T is a path and the number of subtrees of a path is given in [2].

Lemma 1(Székely and Wang [2]). $\mu(P) = (d + 1)(d + 2)/2$.

If $R(T) \neq \emptyset$, we let $\mathscr{S}(T)$ be the set of subtrees of T that contain all the vertices in $R(T)$ and at least one vertex in P, and set $s(T) = |\mathscr{S}(T)|$. In the following lemma, we will establish the upper bound for $s(T)$.

Lemma 2. For any caterpillar T of order m with diameter d, if $R(T) \neq \emptyset$, then

$$s(T) \leq \left(\left\lfloor \frac{d}{2} \right\rfloor + 1 \right) \left(\left\lceil \frac{d}{2} \right\rceil + 1 \right)$$

and the equality holds if and only if $T \cong T^*_{m,d}$.

Proof. Let $I = \{i \mid v_i$ is adjacent to at least one vertex in $R(T)\}$. Since $R(T) \neq \emptyset$, we have $I \neq \emptyset$. Assume that x and y are the minimum and maximum integer in I, respectively. Then any subtree of T in $\mathscr{S}(T)$ certainly contains the path $v_x v_{x+1} \cdots v_y$.

So the only difference between the trees in $\mathscr{S}(T)$ is the number of vertices they contain in $\{v_1, \ldots, v_{x-1}\}$ and in $\{v_{y+1}, \ldots, v_{d+1}\}$. Notice that for any $T' \in \mathscr{S}(T)$, the vertices that T' contain in $\{v_1, \ldots, v_{x-1}\}$ form a subpath of $P' = v_{x-1} \cdots v_1$ starting at v_{x-1}, and those in $\{v_{y+1}, \ldots, v_{d+1}\}$ form a subpath of $P'' = v_{y+1} \cdots v_{d+1}$ starting at v_{y+1}, so there are $|P'| + 1 = x$ different choices in P' and $|P''| + 1 = d - y + 2$ different ones in P''. According to the multiplication rule, we have

$$s(T) = (|P'| + 1)(|P''| + 1) = x(d - y + 2),$$

where $2 \le x \le \ y \le d$. Obviously, $|P'| + |P''| \le d$. Thus, by the well known inequality, we have

$$(|P'| + 1)(|P''| + 1) \le \left(\frac{|P'| + |P''|}{2} + 1\right)^2 \le \left(\frac{d}{2} + 1\right)^2.$$

The equality on the right holds when $|P'| + |P''| = d$, which is equivalent to $x = y$ and certainly holds when $s(T)$ reaches its maximum. (If not, then $x \le y - 1$, we can replace x with $x + 1$ to make $s(T)$ larger.) The equality on the left holds when $|P'| = |P''|$, so when d is even, both conditions can be satisfied at the same time. In this case, we have

$$s(T) \le \left(\frac{d}{2} + 1\right)^2$$

and the equality holds when $x = y = \frac{d}{2} + 1$, which means, $T \cong T^*_{m,d}$. When d is odd, the equality on the left cannot hold, since $|P'|$ and $|P''|$ are integers. In such a case, $s(T)$ reaches its maximum when $||P'| - |P''|| = 1$. That is to say, $x = y = \frac{d \pm 1}{2} + 1$. In this case, we have

$$s(T) \le \left(\frac{d - 1}{2} + 1\right)\left(\frac{d + 1}{2} + 1\right)$$

and the equality holds when $x = y = \frac{d \pm 1}{2} + 1$. In both cases, we have

$$s(T) \le \left(\left\lfloor \frac{d}{2} \right\rfloor + 1\right)\left(\left\lceil \frac{d}{2} \right\rceil + 1\right)$$

and the equality holds when $x = y = \left\lfloor \frac{d}{2} \right\rfloor + 1$ or $\left\lceil \frac{d}{2} \right\rceil + 1$, which is equivalent to $T \cong T^*_{m,d}$. This completes the proof. ∎

The following theorem is the main result of this section, which tells us the sharp upper bound of the number of subtrees of a caterpillar and when the upper bound can be reached.

Theorem 5. Let T be a caterpillar of order n with diameter d. Then

$$\mu(T) \leq \left\lceil \frac{d-2}{2} \right\rceil \left\lfloor \frac{d}{2} \right\rfloor + \left(\left\lfloor \frac{d}{2} \right\rfloor + 1 \right) \left(\left\lceil \frac{d}{2} \right\rceil + 1 \right) 2^{n-d-1} + n - 1$$

and the equality holds if and only if $T \cong T^*_{n,d}$.

Proof. Let $P = v_1 v_2 \cdots v_{d+1}$ be a diameter path of T and $R(T) = V(T) - V(P)$. Set $\mathscr{T}_i$ be the set of subtrees of T that contain exactly i vertices in $R(T)$ and $t_i = |\mathscr{T}_i|$, where $0 \leq i \leq n-d-1$. Then the set of the subtrees of T is

$$\mathscr{T} = \bigcup_{i=0}^{n-d-1} \mathscr{T}_i,$$

and the number of the subtrees of T is

$$\mu(T) = \sum_{i=0}^{n-d-1} t_i.$$

By Lemma 1, we have

$$t_0 = \mu(P) = (d+1)(d+2)/2.$$

Noting that $\mathscr{T}_1$ consists of subtrees of order one which is a vertex in $R(T)$, and subtrees of order at least two which contain exactly one vertex in $R(T)$ and at least one vertex in P, by Lemma 2 we have

$$t_1 \leq \binom{n-d-1}{1} \left[\left(\left\lfloor \frac{d}{2} \right\rfloor + 1 \right) \left(\left\lceil \frac{d}{2} \right\rceil + 1 \right) + 1 \right]$$

and

$$t_i \leq \binom{n-d-1}{i} \left(\left\lfloor \frac{d}{2} \right\rfloor + 1 \right) \left(\left\lceil \frac{d}{2} \right\rceil + 1 \right)$$

for $2 \leq i \leq n-d-1$, and the equality holds if and only if $T' \cong T^*_{d+1+i,d}$ for each $T' \in \mathscr{T}_i$, which implies that $T \cong T^*_{n,d}$ if $i \geq 2$. Therefore, we have

$$\begin{aligned}
\mu(T) &= t_0 + t_1 + \sum_{i=2}^{n-d-1} t_i \\
&\leq \frac{(d+1)(d+2)}{2} + \binom{n-d-1}{1} + \sum_{i=1}^{n-d-1} \binom{n-d-1}{i} \left(\left\lfloor \frac{d}{2} \right\rfloor + 1 \right) \left(\left\lceil \frac{d}{2} \right\rceil + 1 \right) \\
&= \left\lceil \frac{d-2}{2} \right\rceil \left\lfloor \frac{d}{2} \right\rfloor + \left(\left\lfloor \frac{d}{2} \right\rfloor + 1 \right) \left(\left\lceil \frac{d}{2} \right\rceil + 1 \right) \sum_{i=0}^{n-d-1} \binom{n-d-1}{i} + n - 1
\end{aligned}$$

$$= \left\lceil \frac{d-2}{2} \right\rceil \left\lfloor \frac{d}{2} \right\rfloor + \left(\left\lfloor \frac{d}{2} \right\rfloor + 1 \right) \left(\left\lceil \frac{d}{2} \right\rceil + 1 \right) 2^{n-d-1} + n - 1.$$

Noticing that the equality above holds if and only if $T \cong T^*_{n,d}$, we see that the result follows. ∎

3. Upper bound for the number of subtrees of trees in $\mathscr{T}_{n,d}$

In this section, we will give the proof of Theorem 2.

By Theorem 5, we need only to consider the maximum number of the subtrees of a tree $T \in \mathscr{T}_{n,d}$ which is not a caterpillar. The main idea of dealing with this case is to show that, for every tree of this kind, there exists a caterpillar that has the same order and diameter as but more subtrees than it, so that in trees of given order n with given diameter d, the one that has the most subtrees must be a caterpillar. To do so, we first define a transform on a tree in $\mathscr{T}_{n,d}$ which is not a caterpillar as follows.

Let $T \in \mathscr{T}_{n,d}$ not be a caterpillar and $P = v_1 v_2 \cdots v_{d+1}$ a diameter path of T. Then there exists some vertex u such that $dist(u, P) = \min\{d(u, v) \mid v \in V(P)\} \geq 2$. Assume that $dist(u_1, P) \geq 2$ and $u_1 u_2 \cdots u_\ell v_i$ is the path from u_1 to P in which $u_i \notin V(P)$ for $1 \leq i \leq \ell$. Clearly, $\ell \geq 2$. Now, a transform on T is to contract the edge $u_\ell v_i$ and then add a new vertex u and connect u and v_i. Denoted by T' the resulting graph, we have $T' \in \mathscr{T}_{n,d}$ by the definition of a transform. A transform on T is shown in Figure 2.

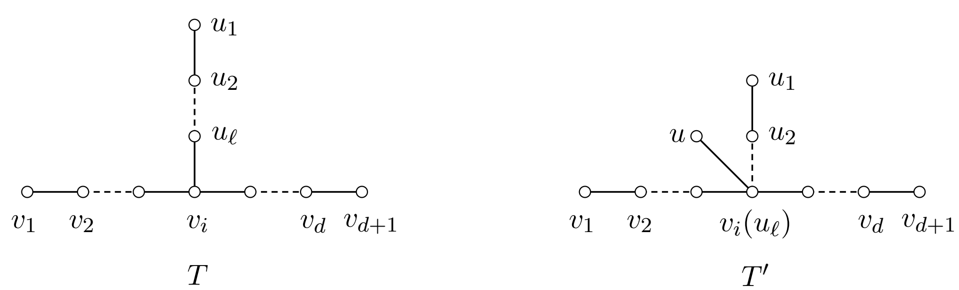

Figure 2. A transform on T

Lemma 3. Let $T \in \mathscr{T}_{n,d}$ not be a caterpillar and T' a tree obtained from T by a transform. Then $\mu(T) < \mu(T')$.

Proof. Let T_{v_i} and T_{u_ℓ} be the component of $T - u_\ell v_i$ that contain v_i and u_ℓ, respectively. Clearly, $T'_{u_\ell} = T'[V(T_{u_\ell}) \cup \{v_i\} - u_\ell] \cong T_{u_\ell}$. Define

$$\begin{aligned}
\mathscr{A}_0 &= \{\text{the subtrees of } T \text{ that contain neither } u_\ell \text{ nor } v_i\},\\
\mathscr{A}_1 &= \{\text{the subtrees of } T \text{ that contain both } u_\ell \text{ and } v_i\},\\
\mathscr{A}_2 &= \{\text{the subtrees of } T \text{ that contain } u_\ell \text{ but no } v_i\},
\end{aligned}$$

$\mathscr{A}_3 = \{$the subtrees of T that contain v_i but no $u_\ell\}$.

Then $\mathscr{A}_i \cap \mathscr{A}_j = \emptyset$ for $0 \le i < j \le 3$ and the set of all subtrees of T is

$$\bigcup_{i=0}^{3} \mathscr{A}_i.$$

Set

$\mathscr{B}_0 = \{$the subtrees of T' that contain neither u nor $v_i\}$,
$\mathscr{B}_1 = \{$the subtrees of T' that contain both u and $v_i\}$,
$\mathscr{B}_2 = \{$the subtrees of T'_{u_ℓ} that contain $v_i\}$,
$\mathscr{B}_3 = \{$the subtrees of T' that contain v_i but no u, and are not in $\mathscr{B}_2\} \cup \{u\}$.

Then $\mathscr{B}_i \cap \mathscr{B}_j = \emptyset$ for $0 \le i < j \le 3$, and the set of all subtrees of T' is

$$\bigcup_{i=0}^{3} \mathscr{B}_i.$$

Let f be a mapping from $V(T)$ to $V(T')$ defined as follows:

$$f(z) = \begin{cases} z, & if\ z \in V(T) - \{v_i, u_\ell\}, \\ u, & if\ z = v_i, \\ v_i, & if\ z = u_\ell. \end{cases}$$

Then it is easy to check that f is an isomorphic mapping from $T - \{u_\ell, v_i\}$ to $T' - \{u, v_i\}$, and from T_{u_ℓ} to T'_{u_ℓ}, so we have $|\mathscr{A}_0| = |\mathscr{B}_0|$ and $|\mathscr{A}_2| = |\mathscr{B}_2|$.

Let g be a mapping from $V(T)$ to $V(T')$ defined as follows:

$$g(z) = \begin{cases} z, & if\ z \in V(T) - \{u_\ell\}, \\ u, & if\ z = u_\ell. \end{cases}$$

Let T_1 be any subtree, $V(T_1) = \{z_1, z_2, \ldots, z_m\}$ and $T_1' = T'[\{g(z_1), g(z_2), \ldots, g(z_m)\}]$. If $T_1 \in \mathscr{A}_1$, then T_1' can be obtained by a transform on T_1: contracting $u_\ell v_i$ and add a new vertex u then connect u and v_i, so T_1' is a tree in $\mathscr{B}_1$. If $T_1 \in \mathscr{A}_3$ and $T_1 \ne v_i$, then T_1 is a subtree of T_{v_i} that contains v_i, so T_1' is a tree in $\mathscr{B}_3$. If $T_2 \in \mathscr{A}_1$ or $\mathscr{A}_3 - \{v_i\}$ and $T_2 \ne T_1$, then $V(T_1) \ne V(T_2)$, so $T'[\{g(z) \mid z \in V(T_1)\}] \ne T'[\{g(z) \mid z \in V(T_2)\}]$. Thus, the 1-1 mapping g induces a single mapping from $\mathscr{A}_1$ or $\mathscr{A}_3 - \{v_i\}$ to $\mathscr{B}_1$ or $\mathscr{B}_3 - \{u\}$, respectively. Therefore, we have $|\mathscr{B}_1| \ge |\mathscr{A}_1|$ and $|\mathscr{B}_3| \ge |\mathscr{A}_3|$. On the other hand, since $T' - u \in \mathscr{B}_3$ is a tree of order $n-1$ and each tree in $\mathscr{A}_3$ contains no $u_1, u_2, \ldots, u_\ell$, no tree in $\mathscr{A}_3$ can be mapped onto $T' - u$ by the single mapping induced by g, hence $|\mathscr{B}_3| > |\mathscr{A}_3|$. Thus, we have

$$\mu(T) = \sum_{i=0}^{3} |\mathscr{A}_i| < \sum_{i=0}^{3} |\mathscr{B}_i| = \mu(T'),$$

which completes the proof. ∎

Now, we are in position to prove Theorem 2.

Proof of Theorem 2. Let $T \in \mathscr{T}_{n,d}$. If T is not a caterpillar, then T can be transformed into some caterpillar T' by a series of transforms. By Lemma 3, $\mu(T) < \mu(T')$. Thus, the result follows by Theorem 5. ∎

4. Lower bound for the number of subtrees of trees in $\mathscr{T}_{n,4}$

In this section, our main goal is to establish the sharp lower bound for the number of subtrees of a tree $T \in \mathscr{T}_{n,4}$. That is, to prove Theorem 3.

In order to prove Theorem 3, we need the following lemmas.

Lemma 4. Let $T \in \mathscr{T}_{n,4}$, $v_1v_2v_3v_4v_5$ be a diameter path of T with $d(v_2) \geq d(v_4)$, and $u \in N(v_3)$ with $d(u) = 1$. Set $T' = T - uv_3 + uv_4$. Then $T' \in \mathscr{T}_{n,4}$, $\mu(T) > \mu(T')$ and $\mu(T, v_3) > \mu(T', v_3)$.

Proof. Obviously, $T' \in \mathscr{T}_{n,4}$. Let n_0, n_1, n_2 denote the number of subtrees of $T - u = T' - u$ which contain v_3 but not v_4, v_4 but not v_3 and both v_3 and v_4, respectively. Then $T - u = T' - u$ has $n_0 + n_2$ subtrees containing v_3 and $n_1 + n_2$ subtrees containing v_4. It is not difficult to see that T has $\mu(T - u) + 1$ subtrees not containing the edge uv_3 and $n_0 + n_2$ subtrees containing the edge uv_3. Similarly, the number of subtrees in T' not containing the edge uv_4 is $\mu(T' - u) + 1$ and that containing the edge uv_4 is $n_1 + n_2$. That is to say, $\mu(T) = \mu(T - u) + 1 + (n_0 + n_2)$ and $\mu(T') = \mu(T' - u) + 1 + (n_1 + n_2)$. Thus, in order to show $\mu(T) > \mu(T')$, it is sufficient to prove $n_0 > n_1$. Let ℓ be the number of subtrees containing v_2 but not v_3, then since $d(v_2) \geq d(v_4)$, we have $\ell \geq n_1$. Because each subtree containing v_2 but not v_3 together with v_3 can form a subtree that contains v_3 but not v_4, we have $n_0 = \ell + 1 > n_1$, and hence $\mu(T) > \mu(T')$.

Since $T - v_3$ and $T' - v_3$ are disjoint unions of stars, and except the stars $T[N[u] - \{v_3\}]$, $T[N[v_4] - \{v_3\}]$ in T and $T'[N[v_4] - \{v_3\}]$ in T', all other stars are the same in T and T', we can see that

$$\mu(T', \overline{v_3}) - \mu(T, \overline{v_3}) = \mu(S_{d(v_4)+1}) - (\mu(S_{d(v_4)}) + 1) = 2^{d(v_4)-1} > 0,$$

which implies that

$$\mu(T, v_3) = \mu(T) - \mu(T, \overline{v_3}) > \mu(T') - \mu(T', \overline{v_3}) = \mu(T', v_3),$$

as required. ∎

Lemma 5. Let $T \in \mathscr{T}_{n,4}$ be a tree as shown in Figure 3, where $N(v_0) = \{v_1, v_2, ..., v_k\}$, $d(v_i) \geq 2$ for $1 \leq i \leq k-1$ and $d(v_k) \geq 4$. Set $T' = T - u_1v_k - u_2v_k + v_0u_1 + u_1u_2$. Then $T' \in \mathscr{T}_{n,4}$, $\mu(T) > \mu(T')$ and $\mu(T, v_0) \geq \mu(T', v_0)$ with the equality holds if and only if $d(v_k) = 4$.

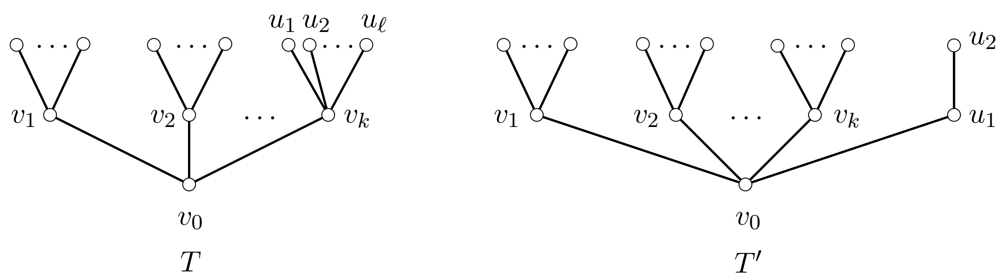

Figure 3

Proof. Obviously, $T' \in \mathscr{T}_{n,4}$. Let $n_i = d(v_i) - 1$ for $1 \leq i \leq k$ where $v_i \in V(T)$, then $T[N[v_i]] \cong S_{n_i+2}$ with v_0 being a vertex of degree one. For each i with $1 \leq i \leq k$, $T[N[v_i]]$ has $\mu(S_{n_i+2}) - \mu(S_{n_i+1}) = 2^{n_i} + 1$ subtrees containing v_0 by Corollary 1. By the multiplication rule, we have

$$\mu(T, v_0) = \prod_{i=1}^{k} (2^{n_i} + 1).$$

Similarly, noting that $T[N[v_i]] = T'[N[v_i]]$ for $1 \leq i \leq k-1$, $T'[N[v_k]] = S_{n_k-2}$ and $T'[N[u_1]] = S_3$, we have

$$\mu(T', v_0) = 3(2^{n_k-2} + 1) \prod_{i=1}^{k-1} (2^{n_i} + 1).$$

Thus,

$$\begin{aligned}\mu(T, v_0) - \mu(T', v_0) &= \prod_{i=1}^{k} (2^{n_i} + 1) - 3(2^{n_k-2} + 1) \prod_{i=1}^{k-1} (2^{n_i} + 1) \\ &= [(2^{n_k} + 1) - 3(2^{n_k-2} + 1)] \prod_{i=1}^{k-1} (2^{n_i} + 1)\end{aligned}$$

$$= (2^{n_k-2} - 2) \prod_{i=1}^{k-1} (2^{d(v_i)} + 1).$$

Since $d(v_k) \geq 4$, we have $n_k \geq 3$, which means

$$\mu(T, v_0) \geq \mu(T', v_0) \tag{1}$$

and the equality holds if and only if $n_k = 3$, which is equivalent to $d(v_k) = 4$.

Because $T - v_0$ and $T' - v_0$ are disjoint unions of k and $k+1$ stars, respectively, by Corollary 1 we have

$$\mu(T, \overline{v_0}) = \sum_{i=1}^{k} \mu(S_{n_i+1}) = \sum_{i=1}^{k} (2^{n_i} + n_i)$$

and

$$\mu(T', \overline{v_0}) = \sum_{i=1}^{k-1} \mu(S_{n_i+1}) + \mu(S_{n_k-1}) + \mu(S_2) = \sum_{i=1}^{k-1} (2^{n_i} + n_i) + (2^{n_k-2} + n_k - 2) + 3.$$

Hence,

$$\begin{aligned} &\mu(T) - \mu(T') \\ &= \left[\sum_{i=1}^{k} (2^{n_i} + n_i) + \mu(T, v_0)\right] - \left[\sum_{i=1}^{k-1} (2^{n_i} + n_i) + (2^{n_k-2} + n_k - 2) + 3 + \mu(T', v_0)\right] \\ &= (3 \cdot 2^{n_k-2} - 1) + (\mu(T, v_0) - \mu(T', v_0)). \end{aligned}$$

Since $n_k \geq 3$, we have $3 \cdot 2^{n_k-2} - 1 > 0$. By (1), $\mu(T, v_0) - \mu(T', v_0) \geq 0$, hence

$$\mu(T) > \mu(T'),$$

which completes the proof. ∎

Lemma 6. Let T_7' and T_7'' be the trees on 7 vertices as shown Figure 4. Then $\mu(T_7', v_0) = 25$, $\mu(T_7'', v_0) = 27$, $\mu(T_7') = 37$ and $\mu(T_7'') = 36$.

Figure 4

Proof. Let $P = v_0v_1v_2$ be a path, $S = S_4$ with $v_0 \in V(S)$ and $d(v_0) = 1$. Obviously, $\mu(P, v_0) = 3$. By Corollary 1, $\mu(S, v_0) = \mu(S_4) - \mu(S_3) = 5$. Since T_7' can be obtained from identifying v_0 of two copies of S, by the multiplication rule, we have $\mu(T_7', v_0) = 5 \times 5 = 25$. Similarly, T_7'' can be obtained from identifying v_0 of three copies of P, we have $\mu(T_7'', v_0) = 3 \times 3 \times 3 = 27$.

On the other hand, since $T_7' - v_0 = 2S_3$ has $2(2^2+3-1) = 12$ subtrees and $T_7'' - v_0 = 3P_2$ has $3 \times 3 = 9$ subtrees, we have $\mu(T_7') = 25 + 12 = 37$ and $\mu(T_7'') = 27 + 9 = 36$. ∎

Lemma 7. Let T be a tree of order at least 2 with $v_0 \in V(T)$, T_*', T_*'' denote the trees obtained by identifying v_0 of T and T_7', T and T_7'', respectively. Then $\mu(T_*') < \mu(T_*'')$.

Proof. It is easy to see that

$$\mu(T_*') = \mu(T_*', v_0) + \mu(T_*', \overline{v_0}) = \mu(T, v_0) \cdot \mu(T_7', v_0) + \mu(T, \overline{v_0}) + \mu(T_7', \overline{v_0})$$

and

$$\mu(T_*'') = \mu(T_*'', v_0) + \mu(T_*'', \overline{v_0}) = \mu(T, v_0) \cdot \mu(T_7'', v_0) + \mu(T, \overline{v_0}) + \mu(T_7'', \overline{v_0}).$$

By Lemma 6, we have

$$\mu(T_*'') - \mu(T_*') = 2\mu(T, v_0) - 3.$$

Since T is of order at least 2, $\mu(T, v_0) \geq 2$. Hence we have $\mu(T_*') < \mu(T_*'')$. ∎

Now, we begin to prove Theorem 3.

Proof of Theorem 3. Let $T \in \mathscr{T}_{n,4}$ such that $\mu(T)$ is minimum and $u_1u_2v_0u_3u_4$ be a path of length 4 in T with $N(v_0) = \{v_1, v_2, ..., v_k\}$ and $d(v_1) \leq \cdots \leq d(v_k)$. By Lemma 4, we have $d(v_1) \geq 2$. By Lemma 5, we have $d(v_k) \leq 3$. This implies that $k \geq 3$ since $n \geq 8$. By Lemma 7, we have $d(v_3) = \cdots = d(v_k) = 3$. Thus, we have $d(v_1) = d(v_2) = 2$ if $r = 2$, which means $T \cong T_n^2$; $d(v_1) = 2$ and $d(v_2) = 3$ if $r = 0$, which means $T \cong T_n^0$; and $d(v_1) = d(v_2) = 3$ if $r = 1$, which means $T \cong T_n^1$.

Let $P = P_3$ with $u \in V(P)$ and $d(u) = 1$, $S = S_4$ with $v \in V(S)$ and $d(v) = 1$. Clearly, $\mu(P, u) = 3$. By Corollary 1, $\mu(S, v) = \mu(S_4) - \mu(S_3) = 5$. Since T_n^r is the tree obtained from ℓ copies of P and $\frac{n-2\ell-1}{3}$ copies of S, by identifying ℓ copies of u and $\frac{n-2\ell-1}{3}$ copies of v into one vertex v_0, by the multiplication rule, we can obtain that

$$\mu(T_n^r, v_0) = 3^\ell 5^{\frac{n-2\ell-1}{3}}.$$

On the other hand, because $T_n^r - v_0 = \ell P_2 \cup \frac{n-2\ell-1}{3} S_3$, P_2 has 3 subtrees and S_3 has 6 subtrees by Corollary 1, we have

$$\mu(T_n^r, \overline{v_0}) = 3\ell + 6 \cdot \frac{n-2\ell-1}{3} = 2n - 2 - \ell.$$

Therefore,

$$\mu(T_n^r) = 3^\ell 5^{\frac{n-2\ell-1}{3}} + 2n - 2 - \ell.$$

This completes the proof of Theorem 3. ∎

5. Lower bound for the number of subtrees of trees in $\mathscr{T}_{n,5}$

In this section, we will give the proof of Theorem 4, which establishes the sharp lower bound for the number of subtrees of trees that belong to $\mathscr{T}_{n,5}$.

Lemma 8. Let x_1, x_2, n be positive integers with $x_1 + x_2 = n \geq 6$ and $f(x_1, x_2)$ a function defined as

$$f(x_1, x_2) = \prod_{i=1}^{2} 3^{\ell_i} 5^{\frac{x_i - 2\ell_i - 1}{3}} + \sum_{i=1}^{2} \left(3^{\ell_i} 5^{\frac{x_i - 2\ell_i - 1}{3}} + 2x_i - 2 - \ell_i \right),$$

where, $x_i \equiv t_i \pmod 3$ with $0 \leq t_i < 3$ and $4 - t_i \equiv \ell_i \pmod 3$ with $0 \leq \ell_i < 3$ for $i = 1, 2$. Set $n \equiv k \pmod 6$ with $0 \leq k < 6$. Then

$$f(x_1, x_2) \geq f\left(\left\lceil \frac{n}{2} \right\rceil + \left\lfloor \frac{k}{4} \right\rfloor, \left\lfloor \frac{n}{2} \right\rfloor - \left\lfloor \frac{k}{4} \right\rfloor \right)$$

and the equality holds if and only if $|x_1 - x_2| = \lceil \frac{n}{2} \rceil - \lfloor \frac{n}{2} \rfloor + 2\lfloor \frac{k}{4} \rfloor$.

Proof. Assume that $f(n_1, n_2)$ is the minimum value of $f(x_1, x_2)$. If $|n_1 - n_2| \geq 4$, we may assume that $n_1 \geq n_2 + 4$. Let $n_1' = n_1 - 3$ and $n_2' = n_2 + 3$. Set $n_i' \equiv t_i' \pmod 3$ with $0 \leq t_i' < 3$ and $4 - t_i' = \ell_i'$ with $0 \leq \ell_i' < 3$ for $i = 1, 2$. Then $t_i' = t_i$ and $\ell_i' = \ell_i$ for $i = 1, 2$. Thus,

$$f(n_1, n_2) - f(n_1', n_2') = 4 \cdot 3^{\ell_1} 5^{\frac{n_1 - 2\ell_1 - 4}{3}} - 4 \cdot 3^{\ell_2} 5^{\frac{n_2 - 2\ell_2 - 1}{3}}.$$

Noticing that $n_1 \geq n_2 + 4$ and $0 \leq \ell_1, \ell_2 \leq 2$, we always have $f(n_1, n_2) - f(n_1', n_2') > 0$, a contradiction. Hence we have

$$|n_1 - n_2| \leq 3.$$

By the symmetry of n_1 and n_2, we assume that $n_1 \geq n_2$.

(1) If $k = 0, 2, 4$, then n is even, so we have $n_1 = n_2 = \frac{n}{2}$ or $n_1 = n_2 + 2 = \frac{n}{2} + 1$. Let $\frac{n}{2} \equiv r \pmod 3$ with $0 \leq r < 3$ and $4 - r \equiv \ell \pmod 3$ with $0 \leq \ell < 3$. Then

$$f\left(\frac{n}{2}, \frac{n}{2}\right) = 3^{2\ell} 5^{\frac{n-4\ell-2}{3}} + 2 \cdot 3^\ell 5^{\frac{n-4\ell-2}{6}} + 2n - 2\ell - 4,$$

and

$$f\left(\frac{n}{2} + 1, \frac{n}{2} - 1\right) = 3^{\ell_1 + \ell_2} 5^{\frac{n - 2(\ell_1 + \ell_2) - 2}{3}} + 3^{\ell_1} 5^{\frac{n - 4\ell_1}{6}} + 3^{\ell_2} 5^{\frac{n - 4\ell_2 - 4}{6}} + 2n - (\ell_1 + \ell_2) - 4.$$

For a given n, it is easy to see that $r, t_1, t_2, \ell, \ell_1, \ell_2$ can be determined completely in the two equalities above. So, if $k = 0$, then $(r, t_1, t_2, \ell, \ell_1, \ell_2) = (0, 1, 2, 1, 0, 2)$ and if $k = 2$, then $(r, t_1, t_2, \ell, \ell_1, \ell_2) = (1, 2, 0, 0, 2, 1)$. Replacing ℓ, ℓ_1, ℓ_2 with (1,0,2) and (0,2,1) in the two equalities above, we find that $f(\frac{n}{2}+1, \frac{n}{2}-1) > f(\frac{n}{2}, \frac{n}{2})$. If $k = 4$, then $(r, t_1, t_2, \ell, \ell_1, \ell_2) = (2, 0, 1, 2, 1, 0)$. Replacing ℓ, ℓ_1, ℓ_2 with (2,1,0) in the two equalities above, we get that $f(\frac{n}{2}, \frac{n}{2}) > f(\frac{n}{2}+1, \frac{n}{2}-1)$.

(2) If $k = 1, 3, 5$, then n is odd, so $n_1 = n_2 + 1 = \frac{n+1}{2}$ or $n_1 = n_2 + 3 = \frac{n+3}{2}$. Let $\frac{n-3}{2} \equiv r \pmod 3$ with $0 \le r < 3$ and $4 - r \equiv \ell \pmod 3$ with $0 \le \ell < 3$. Then

$$f\left(\frac{n+1}{2}, \frac{n-1}{2}\right) = 3^{\ell_1+\ell_2}5^{\frac{n-2(\ell_1+\ell_2)-2}{3}} + 3^{\ell_1}5^{\frac{n-4\ell_1-1}{6}} + 3^{\ell_2}5^{\frac{n-4\ell_2-3}{6}} + 2n - (\ell_1+\ell_2) - 4,$$

and

$$f\left(\frac{n+3}{2}, \frac{n-3}{2}\right) = 3^{2\ell}5^{\frac{n-4\ell-2}{3}} + 3^{\ell}5^{\frac{n-4\ell+1}{6}} + 3^{\ell}5^{\frac{n-4\ell-5}{6}} + 2n - 2\ell - 4.$$

If $k = 1, 3$, then $(r, t_1, t_2, \ell, \ell_1, \ell_2) = (2, 1, 0, 2, 0, 1)$ and $(0, 2, 1, 1, 2, 0)$, respectively. Replacing ℓ, ℓ_1, ℓ_2 with (2,0,1) and (1,2,0) in the two equalities above, we have $f(\frac{n+1}{2}, \frac{n-1}{2}) < f(\frac{n+3}{2}, \frac{n-3}{2})$. If $k = 5$, then $(r, t_1, t_2, \ell, \ell_1, \ell_2) = (1, 0, 2, 0, 1, 2)$. Replacing ℓ, ℓ_1, ℓ_2 with (0,1,2) in the two equalities above, we get that $f(\frac{n+1}{2}, \frac{n-1}{2}) > f(\frac{n+3}{2}, \frac{n-3}{2})$.

The proof of Lemma 8 is complete. ∎

Proof of Theorem 4. Let $T \in \mathscr{T}_{n,5}$ with $\mu(T)$ minimum and $v_1v_2v_3v_4v_5v_6$ be a diameter path of T. Assume that $T - v_3v_4 = T_1 \cup T_2$ with $v_3 \in V(T_1)$ and $v_4 \in V(T_2)$. Set $|V(T_i)| = n_i$, $n_i \equiv t_i \pmod 3$ with $0 \le t_i < 3$ and $4 - t_i \equiv \ell_i \pmod 3$ with $0 \le \ell_i < 3$, where $i = 1, 2$.

Since T has $\mu(T_1, v_3) \cdot \mu(T_2, v_4)$ subtrees containing the edge v_3v_4 and $\mu(T_1) + \mu(T_2)$ subtrees not containing the edge v_3v_4, we have

$$\mu(T) = \mu(T_1, v_3) \cdot \mu(T_2, v_4) + \mu(T_1) + \mu(T_2). \tag{2}$$

Clearly, $2 \le diam(T_i) \le 4$ for $i = 1, 2$.

Claim 1. If $diam(T_i) = 2$, then $T_i \cong S_3$ or S_4.

Proof. For convenience, let $T_i = T_1$. Since $diam(T_1) = 2$, $d_{T_1}(v_1) = d_{T_1}(v_3) = 1$. We will prove that $d(v_2) \le 3$. If not, let $N(v_2) - \{v_1, v_3\} = \{u_1, u_2, ..., u_\ell\}$ with $\ell \ge 2$ and $T_1' = T_1 + v_3u_1 + u_1u_2 - v_2u_1 - v_2u_2$. Noting that $d(u_j) = 1$ for $1 \le j \le \ell$, we have

$$\mu(T_1) - \mu(T_1') = \mu(S_{\ell+3}) - (\mu(S_\ell, v_2) \cdot \mu(P_3, v_3) + \mu(S_\ell) + \mu(P_3))$$

$$= 2^{\ell+1} - 3 > 0,$$

and

$$\mu(T_1, v_3) - \mu(T_1', v_3) = (\mu(S_{\ell+3}) - \mu(S_{\ell+2})) - (\mu(P_3, v_3)(\mu(S_\ell, v_2) + 1))$$
$$= 2^{\ell-1} - 2 \geq 0.$$

Let T' be a tree obtained from T by replacing T_1 with T_1', then we have $\mu(T') < \mu(T)$, a contradiction. Hence, $d(v_2) \leq 3$. That is, $T_i \cong S_3$ or S_4. ▮

Claim 2. If $diam(T_i) = 3$, then $T_i \cong P_4$.

Proof. Assume without loss of generality that $i = 1$. Since $\mu(T)$ takes its minimum, by Lemma 3, v_3 has at most one neighbor of degree 1 in T, and so it is in T_1. Given $diam(T_1) = 3$, we know that $d(v_3) = 3$ and v_3 has exactly one neighbor u with $d(u) = 1$. If $d(v_2) \geq 3$, we let $N(v_2) - \{v_1, v_3\} = \{u_1, u_2, ..., u_\ell\}$ with $\ell \geq 1$ and $T_1' = T_1 + uu_\ell - v_2u_\ell$. Noting that $d(u_j) = 1$ for $1 \leq j \leq \ell$, we have

$$\mu(T_1) - \mu(T_1') = (\mu(S_{\ell+2}, v_2) \cdot \mu(P_2, v_3) + \mu(S_{\ell+2}) + \mu(P_2))$$
$$- (\mu(S_{\ell+1}, v_2) \cdot \mu(P_3, v_3) + \mu(S_{\ell+1}) + \mu(P_3))$$
$$= 2^{\ell+1} - 2 > 0,$$

and

$$\mu(T_1, v_3) - \mu(T_1', v_3) = \mu(P_2, v_3)(\mu(S_{\ell+2}, v_2) + 1) - \mu(P_3, v_3)(\mu(S_{\ell+1}, v_2) + 1)$$
$$= 2^\ell - 1 > 0.$$

Replacing T_1 with T_1' in T, we get a new tree T' such that $\mu(T') < \mu(T)$, which is impossible, so $d(v_2) = 2$. That is, $T_i \cong P_4$. ▮

Claim 3. If $diam(T_i) = 4$, then $T_i \cong T_{n_i}^{t_i}$ with root v_{i+2}.

Proof. If $n_i \geq 8$, then by Lemma 4, Lemma 5 and (2), we have $T_i \cong T_{n_i}^{t_i}$. Noting that $diam(T_i) = 4$ implies that $n \geq 5$, we now need only to consider the case where $5 \leq n_i \leq 7$. If $n_i = 5$, then $T_i \cong P_5 \cong T_5^2$ with v_{i+2} being its root. If $n_i = 6$, then by Lemma 4, $T_i \cong T_6^0$ with root v_{i+2}. If $n_i = 7$, then by Lemma 4, $T_i \cong T_7'$ or T_7''. By Lemma 7 and the minimality of $\mu(T)$, we have $T_i \cong T_7' \cong T_7^1$ with root v_{i+2}. Thus, $T_i \cong T_{n_i}^{t_i}$ with root v_{i+2} in any case. ▮

If $diam(T_i) = 3$ for some $i = 1, 2$, say $diam(T_1) = 3$, then by Claim 2, we may assume that $N(v_3) = \{v_2, v_4, u\}$ with $d(u) = 1$. Let $T' = T + uv_2 - uv_3$ and $T_1' =$

$T_1 + uv_2 - uv_3$. Then $T' \in \mathscr{T}_{n,5}$ and

$$\mu(T') = \mu(T_1', v_3) \cdot \mu(T_2, v_4) + \mu(T_1') + \mu(T_2) = 5\mu(T_2, v_4) + 11 + \mu(T_2).$$

Since

$$\mu(T) = \mu(T_1, v_3) \cdot \mu(T_2, v_4) + \mu(T_1) + \mu(T_2) = 6\mu(T_2, v_4) + 10 + \mu(T_2)$$

and $\mu(T_2, v_4) \geq 3$, we have

$$\mu(T') < \mu(T)$$

which is impossible. Therefore, $diam(T_i) = 2$ or 4.

Noting that $S_3 \cong T_3^0$ and $S_4 \cong T_4^1$, by Claim 1 and Claim 3, we can see that $T_i \cong T_{n_i}^{t_i}$ with root v_{i+2} for $i = 1, 2$. Thus, by Theorem 3 and (2) we have

$$\mu(T) = f(n_1, n_2).$$

By Lemma 8,

$$\mu(T) \geq \mu\left(H_n\left(\left\lceil \frac{n}{2} \right\rceil + \left\lfloor \frac{r_6}{4} \right\rfloor\right)\right)$$

and the equality holds if and only if $T \cong H_n(\lceil \frac{n}{2} \rceil + \lfloor \frac{r_6}{4} \rfloor)$.
Since

$$\mu\left(H_n\left(\left\lceil \frac{n}{2} \right\rceil + \left\lfloor \frac{r_6}{4} \right\rfloor\right)\right) = 3^{2-r_3}5^{\frac{n+2r_3-6}{3}} + (6 - 2r_3)5^{\frac{n+2r_3-6}{6}} + 2n - 6 + r_3$$

if $r_2 = 0$ and

$$\mu\left(H_n\left(\left\lceil \frac{n}{2} \right\rceil + \left\lfloor \frac{r_6}{4} \right\rfloor\right)\right) = 3^{2-r_3}5^{\frac{n+2r_3-6}{3}} + 5^{\frac{n+2r_3-3}{6}} + 3^{2-r_3}5^{\frac{n+2r_3-9}{6}} + 2n - 6 + r_3$$

if $r_2 = 1$, we can see that the result follows. ∎

References

[1] R. Kirk and H. Wang, Largest number of subtrees of trees with a given maximum degree, *SIAM Journal on Discrete Mathematics*, 22(2008), 985-995.

[2] L.A. Székely and H. Wang, On subtrees of trees, *Advances in Applied Mathematics*, 34(2005), 138-155.

[3] X.M. Zhang, X.D. Zhang, D. Gray and H. Wang, The number of subtrees of trees with given degree sequence, *Journal of Graph Theory*, 3(2013), 280-295.