Proc. Indian Acad. Sci. (Math. Sci.) (2019) 129:7
https://doi.org/10.1007/s12044-018-0452-x

Maximizing distance between center, centroid and subtree core of trees

DHEER NOAL SUNIL DESAI[1,2] and KAMAL LOCHAN PATRA[1,2,*]

[1]School of Mathematical Sciences, National Institute of Science Education and Research (NISER), Bhubaneswar P.O. – Jatni, District – Khurda, Odisha 752 050, India
[2]Homi Bhabha National Institute (HBNI), Training School Complex, Anushakti Nagar, Mumbai 400 094, India
*Corresponding author.
E-mail: dheer.nsd@niser.ac.in; klpatra@niser.ac.in

MS received 17 May 2017; revised 20 September 2017; accepted 22 September 2017; published online 17 December 2018

Abstract. For $n \geq 5$ and $2 \leq g \leq n-3$, consider the tree $P_{n-g,g}$ on n vertices which is obtained by adding g pendant vertices to one end vertex of the path P_{n-g}. We call the trees $P_{n-g,g}$ as path-star trees. The *subtree core* of a tree T is the set of all vertices v of T for which the number of subtrees of T containing v is maximum. We prove that over all trees on $n \geq 5$ vertices, the distance between the center (respectively, centroid) and the subtree core is maximized by some path-star trees. We also prove that the tree P_{n-g_0,g_0} maximizes both the distances among all path-star trees on n vertices, where g_0 is the smallest positive integer satisfying $2^{g_0} + g_0 > n - 1$.

Keywords. Tree; center; centroid; subtree core; distance.

Mathematics Subject Classification. 05C05, 05C12.

1. Introduction

Let T be a tree with vertex set $V = V(T)$ and edge set $E = E(T)$. We denote by $d(v)$ the degree of a vertex $v \in V$. A vertex of degree one is called a *pendant* vertex of T. For $u, v \in V$, the *length* of the u-v path in T is the number of edges in that path, and the *distance* between u and v in T, denoted by $d_T(u, v)$, is the length of the u-v path. For subsets U and W of V, the distance $d_T(U, W)$ between U and W is defined by

$$d_T(U, W) = \min_{u \in U, w \in W} d_T(u, w).$$

For $v \in V$, the *eccentricity* $e(v)$ of v is defined by $e(v) = \max\{d_T(u, v) : u \in V\}$. The *radius* $\mathrm{rad}(T)$ of T is defined by $\mathrm{rad}(T) = \min\{e(v) : v \in V\}$ and the *diameter* $\mathrm{diam}(T)$ of T is defined by $\mathrm{diam}(T) = \max\{e(v) : v \in V\}$. It is clear that $\mathrm{diam}(T) = \max\{d_T(u, v) : u, v \in V\}$. We say that v is a *central vertex* of T if $e(v) = \mathrm{rad}(T)$. The *center* of T, denoted by $C = C(T)$, is the set of all central vertices of T.

In a tree T, for any vertex v, $d_T(u, v)$ is maximum only when u is a pendant vertex. Using this observation, the following result is proved (see [1, Theorem 4.2]).

Theorem 1.1. *The center of a tree consists of either one vertex or two adjacent vertices.*

From the proof of the above result as given in [1, Theorem 4.2], it is clear that, for any tree T, $C(T)$ is same as the center of any u-v path in T of length $\mathrm{diam}(T)$.

For $v \in V$, a *branch* (rooted) at v is a maximal subtree containing v as a pendant vertex. Note that the number of branches at v is $d(v)$. The *weight* of v, denoted by $\omega(v) = \omega_T(v)$ is the maximum number of edges contained in a branch at v. We say that v is a *centroid vertex* of T if $\omega(v) = \min_{u \in V} \omega(u)$. The *centroid* of T, denoted by $C_d = C_d(T)$, is the set of all centroid vertices of T.

The following result for the centroid of a tree is analogous to Theorem 1.1 (see [1, Theorem 4.3]).

Theorem 1.2. *The centroid of a tree consists of either one vertex or two adjacent vertices.*

Let T be a tree on n vertices. If $|C_d(T)| = 2$ and $C_d(T) = \{u, v\}$, then n must be even and $\omega(u) = \omega(v) = \frac{n}{2}$. Also, among the branches at u (respectively, at v), the branch containing v (respectively, u) has the maximum number of edges. If $n \geq 3$, then observe that neither the center nor the centroid of T contain pendant vertices. In general, there is no relation between the center and the centroid of a tree with regard to the number of vertices or to their location.

Like center and centroid, many researchers have defined middle part of a tree in several other ways. In [5], Zelinka defined the notion 'median' and proved that it coincides with the centroid for a tree. In [2], Mitchell defined the 'telephone center' of a tree and proved that it also coincides with the centroid. In 2005, Szekely and Wang defined in [4] a new middle part, the so-called 'subtree core' of a tree which does not coincide with either the center or the centroid, in general.

For a given tree T and a vertex v of T, let $f_T(v)$ be the number of subtrees of T containing v. The *subtree core* of T, denoted by $S_c = S_c(T)$, is defined as the set of all vertices v for which $f_T(v)$ is maximum.

In the spirit of Theorems 1.1 and 1.2, Szekely and Wang proved the following result in [4, Theorem 9.1].

Theorem 1.3 [4]. *The subtree core of a tree consists of either one vertex or two adjacent vertices.*

While proving the above theorem, the authors used the fact that the function f_T is strictly concave in the following sense.

Lemma 1.4. *If u, v, w are three vertices of a tree T with $\{u, v\}, \{v, w\} \in E(T)$, then* $2f_T(v) - f_T(u) - f_T(w) > 0$.

We shall use the above lemma frequently.

Remark 1.5. Like the center and centroid, for any tree T on $n \geq 3$ vertices, the subtree core $S_c(T)$ does not contain any pendant vertex.

This remark can be seen as follows. Let v be a pendant vertex of T and let $\{v, w\} \in E(T)$. Consider the tree $T' = T - v$. There is only one subtree of T, namely $\{v\}$, containing v

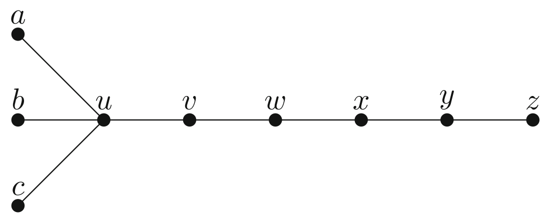

Figure 1. Tree with different center, centroid and subtree core.

but not w. The number of subtrees of T containing both v and w is equal to $f_{T'}(w)$. So $f_T(v) = 1 + f_{T'}(w)$. A similar argument gives $f_T(w) = 2f_{T'}(w)$. Since $n \geq 3$, we have $f_{T'}(w) \geq 2$ and hence $f_T(w) > f_T(v)$.

We denote by P_n the path on the n vertices $1, 2, \ldots, n$, where 1 and n are pendant vertices, and for $i = 2, 3, \ldots, n-1$, vertex i is adjacent to vertices $i-1$ and $i+1$. The center, centroid and subtree core coincide for a path. More precisely, for $n = 2m$, we have

$$C(P_{2m}) = C_d(P_{2m}) = S_c(P_{2m}) = \{m, m+1\}$$

and for $n = 2m+1$, we have

$$C(P_{2m+1}) = C_d(P_{2m+1}) = S_c(P_{2m+1}) = \{m+1\}.$$

We denote by $K_{1,n-1}$ the star on the n vertices $1, 2, \ldots, n$, where n is the only non-pendant vertex. Then

$$C(K_{1,n-1}) = C_d(K_{1,n-1}) = S_c(K_{1,n-1}) = \{n\}.$$

We now give an example of a tree in which the center, centroid and subtree core are different.

Example 1.6. Consider the tree T as in figure 1. We have $e(w) = 3$ and the eccentricity of any vertex other than w is at least 4. So $C(T) = \{w\}$. We have $\omega(v) = 4$ and the weight of any vertex other than v is at least 5. So $C_d(T) = \{v\}$. Finally, $f_T(u) = 48$ and $f_T(k) < 48$ for any vertex k other than u ($f_T(a) = f_T(b) = f_T(c) = 25$, $f_T(v) = 45$, $f_T(w) = 40$, $f_T(x) = 33$, $f_T(y) = 24$, $f_T(z) = 13$). So $S_c(T) = \{u\}$.

For a given tree T, we denote by $d_T(C, C_d)$ (respectively, $d_T(C_d, S_c)$, $d_T(C, S_c)$) the distance between the center and the centroid (respectively, the centroid and the subtree core, the center and the subtree core) of T. It is clear that the minimum of $d_T(C, C_d)$ (respectively, $d_T(C_d, S_c)$, $d_T(C, S_c)$) among all trees T on n vertices is zero. The maximum of $d_T(C, C_d)$ among all trees T on n vertices has been studied in [3], which we describe below.

1.1 *Path-star trees*

Let $P_{n-g,g}$, $n \geq 2$, $1 \leq g \leq n-1$ denote the tree on n vertices which is obtained from the path P_{n-g} by adding g pendant vertices to the vertex $n-g$ (see figure 2). Such a tree $P_{n-g,g}$ is called a *path-star tree*.

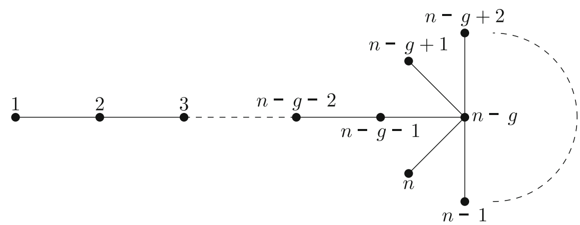

Figure 2. Path-star tree.

Note that $P_{n-1,1}$ is a path, and $P_{1,n-1}$ and $P_{2,n-2}$ are stars. Any tree on less than or equal to 4 vertices is a star or a path. The exact location of the center, the centroid and the subtree core of paths and stars have already been mentioned. Therefore, for a path-star tree, we assume throughout that

$$n \geq 5 \text{ and } 2 \leq g \leq n-3.$$

We denote by Γ_n the class of all path-star trees $P_{n-g,g}$ with the above restrictions on n and g. Then $|\Gamma_n| = n-4$. For the tree $P_{3,n-3}$, we have $C(P_{3,n-3}) = \{2, 3\}$ and $C_d(P_{3,n-3}) = \{3\} = S_c(P_{3,n-3})$. Hence

$$d_{P_{3,n-3}}(C, C_d) = d_{P_{3,n-3}}(C_d, S_c) = d_{P_{3,n-3}}(C, S_c) = 0.$$

Any tree T_5 on 5 vertices is either a path, or a star, or isomorphic to $P_{3,2}$. Therefore, $d_{T_5}(C, C_d) = d_{T_5}(C_d, S_c) = d_{T_5}(C, S_c) = 0$.

In [3, Theorems 2.3, 3.5], the following results are obtained regarding the maximum distance between the center and the centroid among all trees on n vertices.

Theorem 1.7 **[3]**. *Among all trees in Γ_n, the distance between the center and the centroid is maximized when $g = \lfloor \frac{n}{2} \rfloor$. If T is a tree on $n \geq 5$ vertices, then*

(1) $d_{P_{n-g,g}}(C, C_d) \geq d_T(C, C_d)$ *for some* $2 \leq g \leq n-3$.
(2) $d_T(C, C_d) \leq \left\lfloor \frac{n-3}{4} \right\rfloor$.

In this paper, we study the problem of maximizing the distances $d_T(C, S_c)$ and $d_T(C_d, S_c)$ among all trees T on n vertices, in which path-star trees would also play an important role. More precisely, we prove the following.

Theorem 1.8. *Let T be a tree on $n \geq 5$ vertices and let g_0 be the smallest positive integer such that $2^{g_0} + g_0 > n-1$. Then*

(i) $d_T(C, S_c) \leq \lfloor \frac{n-g_0}{2} \rfloor - 1$.
(ii) $d_T(C_d, S_c) \leq \lfloor \frac{n-1}{2} \rfloor - g_o$.

Further, these bounds are attained by the path-star tree P_{n-g_0,g_0}.

2. Center and subtree core

For given vertices $v_1, v_2, \ldots, v_k$ in a tree T, we denote by $f_T(v_1, v_2, \ldots, v_k)$ the number of subtrees of T containing $v_1, v_2, \ldots, v_k$.

Lemma 2.1. *Let T be a tree and $w, y \in V(T)$, where y is a pendant vertex not adjacent to w. Let $\tilde{T}$ be the tree obtained by detaching y from T and adding it as a pendant vertex adjacent to w. Then $f_{\tilde{T}}(a) = f_T(a) - f_T(a, y) + f_T(a, w) - f_T(a, w, y)$ for any $a \in V(T - y)$.*

Proof. Write $T' = T - y = \tilde{T} - y$. Observe that, for any $a \in V(T')$, the number of subtrees of $\tilde{T}$ containing a but not y is equal to $f_{T'}(a)$, and the number of subtrees of $\tilde{T}$ containing both a and y is equal to $f_{T'}(a, w)$. So

$$f_{\tilde{T}}(a) = f_{T'}(a) + f_{T'}(a, w). \tag{1}$$

The set of subtrees of T containing a is a disjoint union of the subtrees of T containing a but not y, and the subtrees of T containing both a and y. This gives

$$f_{T'}(a) = f_T(a) - f_T(a, y), \tag{2}$$

since $f_{T'}(a)$ is equal to the number of subtrees of T containing a but not y. Similarly, we get

$$f_{T'}(a, w) = f_T(a, w) - f_T(a, w, y). \tag{3}$$

Then $f_{\tilde{T}}(a) = f_T(a) - f_T(a, y) + f_T(a, w) - f_T(a, w, y)$ follows from (1), (2) and (3). □

The following lemma compares the subtree core of two trees when one is obtained from the other by some graph perturbation.

Lemma 2.2. *Let T be a tree, $v \in S_c(T)$ and y be a pendant vertex of T not adjacent to v. If $\tilde{T}$ is the tree obtained by detaching y from T and adding it as a pendant vertex adjacent to v, then $S_c(\tilde{T}) = \{v\}$.*

Proof. We show that $f_{\tilde{T}}(v) > f_{\tilde{T}}(b)$ for every $b \in V(\tilde{T}) \setminus \{v\}$. Let $u_1, u_2, \ldots, u_{k+1} = y$ be the vertices adjacent to v in $\tilde{T}$. As y is a pendant vertex, Remark 1.5 implies that $y \notin S_c(\tilde{T})$. Again, by Lemma 1.4, it is enough to show that $f_{\tilde{T}}(v) - f_{\tilde{T}}(u_i) > 0$ for $1 \leq i \leq k$. Taking $a = u_i$ and $w = v$ in Lemma 2.1, we have

$$f_{\tilde{T}}(u_i) = f_T(u_i) - f_T(u_i, y) + f_T(u_i, v) - f_T(u_i, v, y).$$

Again, taking $a = w = v$ in Lemma 2.1, we have

$$f_{\tilde{T}}(v) = f_T(v) - f_T(v, y) + f_T(v) - f_T(v, y).$$

Then, for $1 \leq i \leq k$, we have

$$\begin{aligned} f_{\tilde{T}}(v) - f_{\tilde{T}}(u_i) = {} & [f_T(v) - f_T(u_i)] + f_T(v) - 2f_T(v, y) + f_T(u_i, y) \\ & - f_T(u_i, v) + f_T(u_i, v, y). \end{aligned}$$

Since $v \in S_c(T)$, $f_T(v) - f_T(u_i) \geq 0$. Note that here equality may happen for one i if $S_c(T) = \{u_i, v\}$. Thus, it is enough to prove that

$$f_T(v) - 2f_T(v, y) + f_T(u_i, y) - f_T(u_i, v) + f_T(u_i, v, y) > 0,$$

for $i = 1, 2, \ldots, k$.

Let A_i and B_i be the components of T obtained by deleting the edge $\{v, u_i\} \in E(T)$. We may assume that A_i contains v and B_i contains u_i. Then

$$f_T(v) = f_{A_i}(v) + f_{A_i}(v) f_{B_i}(u_i), \tag{4}$$
$$f_T(u_i, v) = f_{A_i}(v) f_{B_i}(u_i). \tag{5}$$

We shall consider two cases depending on whether $y \in A_i$ or $y \in B_i$.

Case 1. $y \in A_i$. Here we have the following:

$$\begin{aligned} f_T(v, y) &= f_{A_i}(v, y) + f_{A_i}(v, y) f_{B_i}(u_i), \\ f_T(u_i, y) &= f_{A_i}(v, y) f_{B_i}(u_i), \\ f_T(u_i, v, y) &= f_{A_i}(v, y) f_{B_i}(u_i). \end{aligned}$$

Using the above three equations together with (4) and (5), we get

$$\begin{aligned} &f_T(v) - 2f_T(v, y) + f_T(u_i, y) - f_T(u_i, v) + f_T(u_i, v, y) \\ &\quad = f_{A_i}(v) - 2f_{A_i}(v, y). \end{aligned}$$

Let y' be the (unique) vertex adjacent to y in T. Then $y' \in A_i$ and $2f_{A_i}(v, y) = f_{A_i}(v, y')$. Therefore, $f_{A_i}(v) - 2f_{A_i}(v, y) = f_{A_i}(v) - f_{A_i}(v, y') > 0$ as $v \neq y'$.

Case 2. $y \in B_i$. In this case, we have

$$\begin{aligned} f_T(v, y) &= f_{A_i}(v) f_{B_i}(u_i, y), \\ f_T(u_i, y) &= f_{B_i}(u_i, y) + f_{A_i}(v) f_{B_i}(u_i, y), \\ f_T(u_i, v, y) &= f_{A_i}(v) f_{B_i}(u_i, y). \end{aligned}$$

Using the above three equations together with (4) and (5), we get

$$\begin{aligned} &f_T(v) - 2f_T(v, y) + f_T(u_i, y) - f_T(u_i, v) + f_T(u_i, v, y) \\ &\quad = f_{A_i}(v) + f_{B_i}(u_i, y) > 0. \end{aligned}$$

This completes the proof. □

We now prove the following result which says that, among all trees on n vertices, the distance between the center and the subtree core is maximized by a path-star tree.

Theorem 2.3. *Let T be any tree on $n \geq 5$ vertices. Then there exists a path-star tree $P_{n-g,g}$, for some g, with $d_{P_{n-g,g}}(C, S_c) \geq d_T(C, S_c)$.*

Proof. We may assume that $d_T(C, S_c) \geq 1$. Let $C(T) = \{w_1, w_2\}$ and $S_c(T) = \{v_1, v_2\}$, where $w_1 = w_2$ if $|C(T)| = 1$ and $v_1 = v_2$ if $|S_c(T)| = 1$. We may also assume that $d_T(C, S_c) = d_T(w_2, v_1)$.

Let $B_1, B_2, \ldots, B_m$ be the branches of T at v_2. Without loss, we may assume that $C(T), S_c(T) \subseteq V(B_1)$. For $i \in \{2, 3, \ldots, m\}$, if there are pendant vertices y of T contained in B_i but not adjacent to v_2, then detach y from T and add it as a pendant vertex adjacent to v_2. It may happen that a given non-pendant vertex x in B_i becomes a pendant after deletion of certain pendant vertices. Apply the same procedure to x as well (we shall use this graph operation more frequently). Continue this process till all the vertices of T, not in B_1, are attached to v_2 as pendants. We denote by $\tilde{T}$ the new tree obtained from T in this way. By Lemma 2.2, $S_c(\tilde{T}) = \{v_2\}$. If the vertices of T, not in B_1, are all pendants, then we take $\tilde{T} = T$ and proceed with $S_c(\tilde{T}) = \{v_1, v_2\}$.

We now study the position of $C(\tilde{T})$. We know that the center of a tree is the center of any longest path in it. If there is a longest path in T which does not contain v_2, then that path is contained in B_1. In that case, $\text{diam}(T) = \text{diam}(\tilde{T})$ and hence $C(T) = C(\tilde{T})$. Otherwise, every longest path in T contains v_2. Then $\text{diam}(T) \geq \text{diam}(\tilde{T})$ and so $C(\tilde{T})$ may move away from $S_c(T)$ with respect to a path in $\tilde{T}$ containing $C(T)$ and v_2. Therefore, $d_{\tilde{T}}(C, S_c) \geq d_T(C, S_c)$.

If $\tilde{T}$ is a path-star tree, then we are done. Otherwise, let P be a longest path in B_1 containing both $C(\tilde{T})$ and $S_c(\tilde{T})$. We now proceed by detaching pendant vertices of B_1 but not in P, and add them as pendant vertices adjacent to v_2, one after another, till we are left with only the path P in B_1. Call the new tree obtained in this way from $\tilde{T}$ as $\bar{T}$. By Lemma 2.2, $S_c(\bar{T}) = \{v_2\}$. Clearly, $\bar{T}$ is a path-star tree. Since $\text{diam}(\bar{T}) \leq \text{diam}(\tilde{T})$, it follows that $d_{\bar{T}}(C, S_c) \geq d_{\tilde{T}}(C, S_c)$. Thus, we have a path-star tree $\bar{T}$ on n vertices with $d_{\bar{T}}(C, S_c) \geq d_{\tilde{T}}(C, S_c) \geq d_T(C, S_c)$. $\square$

We next try to find a relation between n and g for which the distance $d_{P_{n-g,g}}(C, S_c)$ is maximum. We first look for the position of the subtree core in any $P_{n-g,g}$. For $1 \leq i \leq n-g$, we have

$$f_{P_{n-g,g}}(i) = i(n-g-i) + i(2^g). \tag{6}$$

Here the first term denotes the number of subtrees of $P_{n-g,g}$ containing the vertex i but not $n-g$, while the second term counts the number of subtrees of $P_{n-g,g}$ containing both i and $n-g$. For $n-g+1 \leq i \leq n$, we have

$$f_{P_{n-g,g}}(i) = 1 + (n-g)(2^{g-1}). \tag{7}$$

Here 1 accounts for the number of subtrees of $P_{n-g,g}$ containing i but not $n-g$, while the second term is the number of subtrees of $P_{n-g,g}$ containing both i and $n-g$.

The subtree core of $P_{n-g,g}$ lies in the path from 2 to $n-g$, as it does not contain any pendant vertex. We have $S_c(P_{n-g,g}) = \{n-g\}$ if and only if $f_{P_{n-g,g}}(n-g) - f_{P_{n-g,g}}(n-g-1) > 0$. Using equation (6), $f_{P_{n-g,g}}(n-g) - f_{P_{n-g,g}}(n-g-1) = (n-g)(2^g) - (n-g-1)(1) - (n-g-1)(2^g) = g - n + 1 + 2^g$. Therefore,

$$S_c(P_{n-g,g}) = \{n-g\} \quad \text{if and only if } 2^g + 1 > n-g.$$

Now suppose that $2^g + 1 \leq n-g$. Then the subtree core of $P_{n-g,g}$ intersects the path connecting the vertex 2 to $n-g-1$. So, for $j \in \{2, 3, \ldots, n-g-1\}$ with $j \in S_c(P_{n-g,g})$ and $j-1 \notin S_c(P_{n-g,g})$, we have $f_{P_{n-g,g}}(j) - f_{P_{n-g,g}}(j-1) > 0$ and $f_{P_{n-g,g}}(j+1) - f_{P_{n-g,g}}(j) \leq 0$ (with equality if and only if $j+1 \in S_c(P_{n-g,g})$). By using equation (6), we have $f_{P_{n-g,g}}(j) - f_{P_{n-g,g}}(j-1) = n - g - 2j + 1 + 2^g > 0$. So, $j < \frac{n-g+1+2^g}{2}$. Also,

$$
\begin{aligned}
&f_{P_{n-g,g}}(j+1) - f_{P_{n-g,g}}(j)\\
=&(j+1)(n-g-j-1) + (j+1)(2^g) - j(n-g-j) - j(2^g)\\
=&n-g-2j-1+2^g.
\end{aligned}
$$

If $j+1 \notin S_c(P_{n-g,g})$, then $n-g-2j-1+2^g < 0$ and so $j > \frac{n-g-1+2^g}{2}$. Then

$$\frac{n-g-1+2^g}{2} < j < \frac{n-g+1+2^g}{2}$$

gives $S_c(P_{n-g,g}) = \{j\} = \{\frac{n-g+2^g}{2}\}$. It follows that $n-g$ must be even. Now, if $j+1 \in S_c(P_{n-g,g})$, then $n-g-2j-1+2^g = 0$ and so $j = \frac{n-g-1+2^g}{2}$. Therefore, $S_c(P_{n-g,g}) = \{j, j+1\}$, where $j = \frac{n-g-1+2^g}{2}$. In this case, $n-g$ must be odd. Thus we have the following.

Theorem 2.4. *The subtree core of the path-star tree $P_{n-g,g}$ is given by*

$$
S_c(P_{n-g,g}) = \begin{cases} \begin{cases} \left\{\frac{n-g+2^g}{2}\right\}, & \text{if } n-g \text{ is even} \\ \left\{\frac{n-g-1+2^g}{2}, \frac{n-g+1+2^g}{2}\right\}, & \text{if } n-g \text{ is odd} \end{cases}, & \text{if } 2^g+1 \le n-g, \\ \{n-g\}, & \text{if } 2^g+1 > n-g. \end{cases}
$$

In Example 1.6, the tree T is the path-star tree $P_{6,3}$ and the vertex u of T is identified with the vertex 6 of $P_{6,3}$. The number of subtrees of $P_{6,3}$ containing the vertex 6 but not the vertex 5 is 8. For $1 \le i \le 5$, the number of subtrees of $P_{6,3}$ containing the vertex 6 and $6-i$ but not $6-i-1$ is 8. Hence the total number of subtrees of $P_{6,3}$ containing 6 is 48. So $f_T(u) = f_{P_{6,3}}(6) = 48$ and for any vertex x other than 6, $f_{P_{6,3}}(x) < 48$. In this case, $2^g+1 > n-g$. So, $S_c = \{n-g\} = \{u\}$. But if $2^g+1 \le n-g$, then $S_c(P_{6,3})$ moves away from the vertex $n-g$. For example, in $P_{10,3}$, $f_{P_{10,3}}(9) = 81$ and $f_{P_{10,3}}(x) \le 80$ for any vertex x other than 9. Hence, $S_c(P_{10,3}) = \{9\}$ which is different from the vertex $n-g = 10$.

The position of the center of $P_{n-g,g}$ can also be expressed in terms of $n-g$. The following result is straight-forward.

Theorem 2.5. *The center of the path-star tree $P_{n-g,g}$ is given by*

$$
C(P_{n-g,g}) = \begin{cases} \left\{\frac{n-g+2}{2}\right\}, & \text{if } n-g \text{ is even,} \\ \left\{\frac{n-g+1}{2}, \frac{n-g+3}{2}\right\}, & \text{if } n-g \text{ is odd.} \end{cases}
$$

Theorem 2.6. *The distance between the center and the subtree core of the path-star tree $P_{n-g,g}$ is given by*

$$
d_{P_{n-g,g}}(C, S_c) = \begin{cases} \begin{cases} 2^{g-1}-1, & \text{if } n-g \text{ is even} \\ 2^{g-1}-2, & \text{if } n-g \text{ is odd} \end{cases}, & \text{if } 2^g+1 \le n-g \\ \begin{cases} \frac{n-g-2}{2}, & \text{if } n-g \text{ is even} \\ \frac{n-g-3}{2}, & \text{if } n-g \text{ is odd} \end{cases}, & \text{if } 2^g+1 > n-g. \end{cases}
$$

Proof. From Theorems 2.4 and 2.5, we have

$$d_{P_{n-g,g}}(C,S_c)=\begin{cases}\begin{cases}\frac{n-g+2^g-n+g-2}{2}, & \text{if } n-g \text{ is even}\\ \frac{n-g-1+2^g-n+g-3}{2}, & \text{if } n-g \text{ is odd}\end{cases}, & \text{if } 2^g+1\le n-g\\ \begin{cases}\frac{2n-2g-n+g-2}{2}, & \text{if } n-g \text{ is even}\\ \frac{2n-2g-n+g-3}{2}, & \text{if } n-g \text{ is odd}\end{cases}, & \text{if } 2^g+1> n-g.\end{cases}$$

From the above, the result follows. □

For a given $n \geq 5$, we now try to find a $g \in \{2,\ldots,n-3\}$ for which $d_{P_{n-g,g}}(C,S_c)$ is maximum among all trees on n vertices. Let g_0 denote the smallest value of g with respect to n, for which $2^{g_0}+g_0>n-1$ (that is, $2^{g_0}+1>n-g_0$). By Theorem 2.4, $S_c(P_{n-g_0,g_0})=\{n-g_0\}$. Then by Theorem 2.6, we have

$$d_{P_{n-g_0,g_0}}(C,S_c)=\begin{cases}\frac{n-g_0-2}{2}, & \text{if } n-g_0 \text{ is even}\\ \frac{n-g_0-3}{2}, & \text{if } n-g_0 \text{ is odd.}\end{cases} \tag{8}$$

PROPOSITION 2.7

Among all trees on $n \geq 5$ vertices, the path-star tree P_{n-g_0,g_0} maximizes the distance between the center and the subtree core.

Proof. By Theorem 2.3, we need to prove the following:

$$d_{P_{n-g_0,g_0}}(C,S_c)\geq d_{P_{n-g_0+k,g_0-k}}(C,S_c) \text{ for } k\in\{1,2,\ldots,g_0-2\}$$

and

$$d_{P_{n-g_0,g_0}}(C,S_c)\geq d_{P_{n-g_0-l,g_0+l}}(C,S_c) \text{ for } l\in\{1,2,\ldots,n-g_0-3\}.$$

First, assume that $n-g_0$ is odd. Then, for $k \in \{1,2,\ldots,g_0-2\}$, we have

$$\begin{aligned}d_{P_{n-g_0,g_0}}(C,S_c)-d_{P_{n-g_0+k,g_0-k}}(C,S_c) &\geq \frac{n-g_0-3}{2}-2^{g_0-k-1}+1\\ &\geq \frac{n-g_0-3}{2}-2^{g_0-1-1}+1\\ &=\frac{n-g_0-3-2^{g_0-1}+2}{2}\\ &=\frac{n-g_0-1-2^{g_0-1}}{2}.\end{aligned} \tag{9}$$

By the definition of g_0, we have $2^{g_0-1}\leq n-(g_0-1)-1=n-g_0$. Therefore,

$$d_{P_{n-g_0,g_0}}(C,S_c)-d_{P_{n-g_0+k,g_0-k}}(C,S_c)\geq \frac{n-g_0-1-2^{g_0-1}}{2}\geq\frac{-1}{2}.$$

Since both $d_{P_{n-g_0,g_0}}(C, S_c)$ and $d_{P_{n-g_0+k,g_0-k}}(C, S_c)$ are integers, their difference must be an integer. Hence

$$d_{P_{n-g_0,g_0}}(C, S_c) - d_{P_{n-g_0+k,g_0-k}}(C, S_c) \geq 0.$$

Now, for $l \in \{1, 2, \ldots, n - g_0 - 3\}$, we have

$$\begin{aligned} d_{P_{n-g_0,g_0}}(C, S_c) - d_{P_{n-g_0-l,g_0+l}}(C, S_c) &\geq \frac{n - g_0 - 3}{2} - \left(\frac{n - g_0 - l - 2}{2}\right) \\ &\geq \frac{n - g_0 - 3}{2} - \left(\frac{n - g_0 - 1 - 2}{2}\right) \\ &= 0. \end{aligned} \tag{10}$$

For the case $n - g_0$ even, the above arguments can be used again as $\frac{n-g_0-2}{2} > \frac{n-g_0-3}{2}$. Thus,

$$d_{P_{n-g_0,g_0}}(C, S_c) \geq d_{P_{n-g,g}}(C, S_c)$$

for all $g \in \{1, 2, \ldots, n - 3\}$. This completes the proof. □

As a consequence of Proposition 2.7, we have the following.

COROLLARY 2.8

Let T be tree on $n \geq 5$ vertices and let g_0 be the smallest positive integer such that $2^{g_0} + g_0 > n - 1$. Then $d_T(C, S_c) \leq \lfloor \frac{n-g_0}{2} \rfloor - 1$.

3. Centroid and subtree core

In this section, we prove results similar to Theorem 2.3 and Proposition 2.7 in the case of centroid and subtree core. We first prove the following lemma.

Lemma 3.1. *Let T be a tree, $v \in S_c(T)$ and B be a branch at v. Let u be the vertex in B adjacent to v and x be a pendant vertex of T in B. Suppose that B is not a path. Let y be the closest vertex to x with $d(y) \geq 3$ and let $[y, y_1, y_2, \ldots, y_m = x]$ be the path connecting y and x. Let $z \neq y$ be a vertex of B such that the path from v to z contains y but not y_1. Let $\tilde{T}$ be the tree obtained from T by detaching the path $[y_1, y_2, \ldots, y_m]$ from y and attaching it to z. Then $f_{\tilde{T}}(v) > f_{\tilde{T}}(u)$.*

Proof. Let T' be the tree obtained from T by removing the path $[y_1, y_2, \ldots, y_m]$. Let $a \in V(T')$. Then

$$f_{\tilde{T}}(a) = f_{T'}(a) + m f_{T'}(a, z). \tag{11}$$

Here the first term represents the number of subtrees of $\tilde{T}$ containing a but not y_1, and the second term represents the number of subtrees of $\tilde{T}$ containing both a and y_1. Similarly, we have $f_T(a) = f_{T'}(a) + mf_{T'}(a, y)$ and $f_T(a, y) = (m+1)f_{T'}(a, y)$. It follows that

$$f_{T'}(a) = f_T(a) - \frac{m}{m+1} f_T(a, y). \tag{12}$$

Again, $f_T(a, z) = f_{T'}(a, z) + mf_{T'}(a, z, y)$ and $f_T(a, z, y) = (m+1)f_{T'}(a, z, y)$ imply

$$f_{T'}(a, z) = f_T(a, z) - \frac{m}{m+1} f_T(a, z, y). \tag{13}$$

Using equations (12) and (13) in equation (11), we get

$$f_{\tilde{T}}(a) = f_T(a) - \frac{m}{m+1} f_T(a, y) + m\left[f_T(a, z) - \frac{m}{m+1} f_T(a, z, y)\right], \tag{14}$$

for any $a \in V(T')$. Considering $a = v$ and $a = u$ in equation (14), we get

$$\begin{aligned}
f_{\tilde{T}}(v) - f_{\tilde{T}}(u) &= f_T(v) - \frac{m}{m+1} f_T(v, y) + mf_T(v, z) - \frac{m^2}{m+1} f_T(v, z, y) \\
&\quad - f_T(u) + \frac{m}{m+1} f_T(u, y) - mf_T(u, z) \\
&\quad + \frac{m^2}{m+1} f_T(u, z, y) \\
&\geq \frac{m}{m+1}(f_T(u, y) - f_T(v, y)) + m(f_T(v, z) - f_T(u, z)) \\
&\quad + \frac{m^2}{m+1}(f_T(u, z, y) - f_T(v, z, y)) \\
&= \frac{m}{m+1}(f_T(u, y) - f_T(v, y)) + m(f_T(v, z) - f_T(u, z)) \\
&\quad + \frac{m^2}{m+1}(f_T(u, z) - f_T(v, z)) \\
&= \frac{m}{m+1}(f_T(u, y) - f_T(v, y)) \\
&\quad + \frac{m}{m+1}(f_T(v, z) - f_T(u, z)).
\end{aligned} \tag{15}$$

In the above, the first inequality holds as $f_T(v) - f_T(u) \geq 0$ and the second last equality holds since any subtree containing u and z must contain y. Now, let X be the component of T containing u after deleting the edge $\{u, v\}$. Then it can be seen that

$$f_T(u, y) = f_X(u, y) + f_T(v, y) \text{ and } f_T(u, z) = f_X(u, z) + f_T(v, z).$$

Putting these values of $f_T(u, y)$ and $f_T(u, z)$ in equation (15), we get

$$f_{\tilde{T}}(v) - f_{\tilde{T}}(u) \geq \frac{m}{m+1}(f_X(u, y) - f_X(u, z)).$$

Since $y \neq z$, we have $f_X(u, y) > f_X(u, z)$ and so $f_{\tilde{T}}(v) - f_{\tilde{T}}(u) > 0$. $\square$

Next, we prove a result analogous to Theorem 2.3. It says that, among all trees on n vertices, the distance between the centroid and the subtree core is maximized by a path-star tree.

Theorem 3.2. *Let T be any tree on $n \geq 5$ vertices. Then there exists a path-star tree $P_{n-g,g}$ for some g, with $d_{P_{n-g,g}}(C_d, S_c) \geq d_T(C_d, S_c)$.*

Proof. We may assume that $d_T(C_d, S_c) \geq 1$. Let $C_d(T) = \{w_1, w_2\}$ and $S_c(T) = \{v_1, v_2\}$, where $w_1 = w_2$ if $|C_d(T)| = 1$ and $v_1 = v_2$ if $|S_c(T)| = 1$. We may also assume that $d_T(C_d, S_c) = d_T(w_2, v_1)$.

Let $B_1, B_2, \ldots, B_m$ be the branches at v_2. We assume that the branch B_1 contains $C_d(T)$ and $S_c(T)$. Using the same graph operations recursively as in the proof of Theorem 2.3, construct a new tree $\tilde{T}$ from T by attaching each vertex of B_i, $i \in \{2, 3, \ldots, m\}$, non-adjacent with v_2 in T as a pendant vertex adjacent to v_2. By Lemma 2.2, $S_c(\tilde{T}) = \{v_2\}$. If the vertices of T, not in B_1, are all pendants, then we take $\tilde{T} = T$ and proceed with $S_c(\tilde{T}) = \{v_1, v_2\}$.

We now study the position of $C_d(\tilde{T})$. Note that B_1 remains a branch in $\tilde{T}$ at v_2. For any $v \in V(\tilde{T}) \setminus B_1$, v is a pendant vertex in $\tilde{T}$ and so $\omega_{\tilde{T}}(v) = n - 1 \geq \omega_T(v)$. Also $\omega_{\tilde{T}}(x) = \omega_T(x)$ for every $x \in B_1 \setminus \{v_2\}$, in particular, we have

$$\omega_{\tilde{T}}(w_1) = \omega_T(w_1) = \omega_T(w_2) = \omega_{\tilde{T}}(w_2).$$

The weight of v_2 in $\tilde{T}$ corresponds to the branch B_1 at v_2. If the weight of v_2 in T corresponds to the branch B_1, then $\omega_{\tilde{T}}(v_2) = \omega_T(v_2) > \omega_T(w_2) = \omega_{\tilde{T}}(w_2)$. If the weight of v_2 in T corresponds to a branch B_j, $j \neq 1$, then $\omega_T(w_2)$ must correspond to a branch at w_2 which does not contain v_2 and it follows that $\omega_{\tilde{T}}(v_2) > \omega_{\tilde{T}}(w_2)$. So $C_d(T) = C_d(\tilde{T})$ and

$$d_{\tilde{T}}(C_d, S_c) \geq d_T(C_d, S_c).$$

Now consider the path $P = [w_1, a_1, \ldots, a_k, v_2]$ from w_1 to v_2 in $\tilde{T}$, where $a_1 = w_2$ if $w_1 \neq w_2$ and $a_k = v_1$ if $v_1 \neq v_2$. Suppose that $d_{\tilde{T}}(a_i) \geq 3$ for some i and B is a branch at a_i which contains neither w_1 nor v_2. Applying the graph operation as before (starting with the pendant vertices), attach the vertices of $B \setminus \{a_i\}$ as pendants to v_2. Do this for all branches at a_i, $1 \leq i \leq k$, with $d(a_i) \geq 3$. Name the new tree thus obtained as $\hat{T}$. By Lemma 2.2, $S_c(\hat{T}) = \{v_2\}$. If $d(a_i) = 2$ for all $1 \leq i \leq k$, then take $\hat{T} = \tilde{T}$ and continue with $S_c(\hat{T}) = S_c(\tilde{T})$.

We now study the position of $C_d(\hat{T})$. Any pendant vertex in $\hat{T}$ has weight $n - 1$. Also, $\omega_{\tilde{T}}(x) = \omega_{\hat{T}}(x)$ for any vertex x in a branch at w_1 in $\hat{T}$ which does not contain v_2. In particular, $\omega_{\tilde{T}}(w_1) = \omega_{\hat{T}}(w_1)$. It remains to consider the vertices $a_1, \ldots, a_k$ and $a_{k+1} = v_2$.

Case I. The weight of w_1 in $\tilde{T}$ corresponds to a branch $\bar{B}$ at w_1 not containing v_2 (in this case, observe that we must have $w_1 = w_2$). The branch in $\hat{T}$ at a_i containing w_1 contains $\bar{B}$. So

$$\omega_{\hat{T}}(a_i) > |E(\bar{B})| = \omega_{\tilde{T}}(w_1) = \omega_{\hat{T}}(w_1),$$

for all $1 \leq i \leq k + 1$.

Case II. The weight of w_1 in $\tilde{T}$ corresponds to the branch $\tilde{B}$ at w_1 containing v_2. Since $\omega_{\tilde{T}}(w_1) \leq \omega_{\tilde{T}}(a_i)$, the weight of a_i in $\tilde{T}$ must correspond to the branch at a_i containing w_1. Let $\tilde{B}_1$ denote the branch in $\tilde{T}$ at a_1 containing w_1. Note that $\tilde{B}_1$ is also a branch at a_1 in $\hat{T}$ containing w_1. Now, for $i = 1$,

$$\omega_{\hat{T}}(a_1) = |E(\tilde{B}_1)| = \omega_{\tilde{T}}(a_1) \geq \omega_{\tilde{T}}(w_1) = \omega_{\hat{T}}(w_1).$$

Then, for $2 \leq i \leq k+1$, we have

$$\omega_{\hat{T}}(a_i) > |E(\tilde{B}_1)| \geq \omega_{\hat{T}}(w_1).$$

Thus $C_d(\tilde{T}) = C_d(\hat{T})$ and hence $d_{\hat{T}}(C_d, S_c) \geq d_{\tilde{T}}(C_d, S_c)$.

If $\hat{T}$ is a path-star tree, then we are done. Otherwise, let $C_1, C_2, \ldots, C_s$ be the branches at w_1 in $\hat{T}$, where C_1 is the branch containing v_2. Note that, by applying continuously the process of detaching a path from a vertex of degree at least three and attaching it at a pendant vertex, we may convert any given tree into a path.

Now transform each of the branches C_i, $2 \leq i \leq s$ into paths at w_1 by continuously using the graph operation as in Lemma 3.1 (taking suitable pendant vertices for z) to obtain a new tree $\bar{T}$. The weight of w_j, $j \in \{1, 2\}$ in $\bar{T}$ is equal to that in $\hat{T}$. Applying similar arguments as before, it can be seen that the weight of any other vertex in $\bar{T}$ is greater than $\omega_{\bar{T}}(w_1)$. So $C_d(\bar{T}) = C_d(\hat{T})$. Also, Lemma 3.1 implies that $v_2 \in S_c(\bar{T})$. Therefore,

$$d_{\bar{T}}(C_d, S_c) \geq d_{\hat{T}}(C_d, S_c).$$

Let $\bar{C}_1, \bar{C}_2, \ldots, \bar{C}_s$ be the branches at w_1 in $\bar{T}$, where $\bar{C}_1$ contains v_2. Each $\bar{C}_i$, $i \neq 1$ is a path attached to w_1. If $s = 2$, then $\bar{T}$ is a path-star tree and we are done. Otherwise, again apply Lemma 3.1 continuously (taking $y = w_1$ and z the pendant vertices in $\bar{C}_i$, $i \neq 1$) to transform $\bar{T}$ to a new tree T' such that there are exactly two branches in T' at w_1, one is $\bar{C}_1$ containing v_2 and the other one is a path. Then T' is a path-star tree and $v_2 \in S_c(T')$ by Lemma 3.1. Note that $C_d(T')$ remains the same or moves away from $S_c(\bar{T})$. Therefore, $d_{T'}(C_d, S_c) \geq d_{\bar{T}}(C_d, S_c)$. This completes the proof. □

The position of the centroid of a path-star tree $P_{n-g,g}$ can be expressed in terms of g. The following result is straight-forward.

Theorem 3.3. *The centroid of the path-star tree $P_{n-g,g}$ is given by*

$$C_d(P_{n-g,g}) = \begin{cases} \begin{cases} \left\{\frac{n+1}{2}\right\}, & \text{if } g \leq \frac{n-1}{2} \\ \{n-g\}, & \text{if } g > \frac{n-1}{2} \end{cases}, & \text{if } n \text{ is odd,} \\ \begin{cases} \left\{\frac{n}{2}, \frac{n}{2}+1\right\}, & \text{if } g \leq \frac{n}{2}-1 \\ \{n-g\}, & \text{if } g > \frac{n}{2}-1 \end{cases}, & \text{if } n \text{ is even.} \end{cases}$$

In Example 1.6, the tree T is the path-star tree $P_{6,3}$ and the vertex v of T is identified with the vertex 5 of $P_{6,3}$. In this case, $3 = g \leq \frac{n-1}{2} = \frac{9-1}{2}$. Hence $C_d(P_{6,3}) = \{5\}$. When g increases while comparing to $n-g$, the centroid $C_d(P_{n-g,g})$ moves towards the vertex $n-g$. For example, in $P_{6,5}$, $C_d(P_{6,3}) = \{6\}$, which is the vertex $n-g$.

Using Theorems 2.4 and 3.3, we prove the following.

Theorem 3.4. *The distance between the centroid and the subtree core of the path-star tree $P_{n-g,g}$ is given by the following*: *If n is odd, then*

$$d_{P_{n-g,g}}(C_d, S_c) = \begin{cases} \begin{cases} \frac{2^g-g-1}{2}, & \textit{if } n-g \textit{ is even} \\ \frac{2^g-g-2}{2}, & \textit{if } n-g \textit{ is odd} \end{cases}, & \textit{if } 2^g+1 \le n-g \\ \begin{cases} \frac{n-1}{2}-g, & \textit{if } g \le \frac{n-1}{2} \\ 0, & \textit{if } g > \frac{n-1}{2} \end{cases}, & \textit{if } 2^g+1 > n-g. \end{cases}$$

If n is even, then

$$d_{P_{n-g,g}}(C_d, S_c) = \begin{cases} \begin{cases} \frac{2^g-g-2}{2}, & \textit{if } n-g \textit{ is even} \\ \frac{2^g-g-3}{2}, & \textit{if } n-g \textit{ is odd} \end{cases}, & \textit{if } 2^g+1 \le n-g \\ \begin{cases} \frac{n}{2}-1-g, & \textit{if } g \le \frac{n}{2}-1 \\ 0, & \textit{if } g > \frac{n}{2}-1 \end{cases}, & \textit{if } 2^g+1 > n-g. \end{cases}$$

Proof. First, assume that $2^g+1 \le n-g$. Then $2g < 2^g+g \le n-1$ and so $g < \frac{n-1}{2}$. This gives $g \le \frac{n}{2}-1$ if n is even. Thus, if n is odd, then

$$d_{P_{n-g,g}}(C_d, S_c) = \begin{cases} \frac{n-g+2^g}{2} - \left(\frac{n+1}{2}\right), & \text{if } n-g \text{ is even} \\ \frac{n-g-1+2^g}{2} - \left(\frac{n+1}{2}\right), & \text{if } n-g \text{ is odd.} \end{cases}$$

and if n is even, then

$$d_{P_{n-g,g}}(C_d, S_c) = \begin{cases} \frac{n-g+2^g}{2} - \left(\frac{n}{2}+1\right), & \text{if } n-g \text{ is even} \\ \frac{n-g-1+2^g}{2} - \left(\frac{n}{2}+1\right), & \text{if } n-g \text{ is odd.} \end{cases}$$

Now assume that $2^g+1 > n-g$. In this case, if n is odd, then

$$d_{P_{n-g,g}}(C_d, S_c) = \begin{cases} n-g-\left(\frac{n+1}{2}\right), & \text{if } g \le \frac{n-1}{2} \\ 0, & \text{if } g > \frac{n-1}{2}. \end{cases}$$

and if n is even, then

$$d_{P_{n-g,g}}(C_d, S_c) = \begin{cases} n-g-\left(\frac{n}{2}+1\right), & \text{if } g \le \frac{n}{2}-1 \\ 0, & \text{if } g > \frac{n}{2}-1. \end{cases}$$

Now the theorem follows from the above. □

For a given n, we now try to find a $g_0 \in \{2, 3, \ldots, n-3\}$ for which P_{n-g_0,g_0} will maximize the distance between the centroid and the subtree core among all trees on n vertices.

PROPOSITION 3.5

For a given $n \ge 5$, let g_0 be the smallest integer in $\{2, 3, \ldots, n-3\}$ satisfying $2^{g_0}+g_0 > n-1$. Then the path-star tree P_{n-g_0,g_0} maximizes the distance between the centroid and the subtree core among all trees on n vertices.

Proof. By Theorem 3.2, we need to show the following two inequalities:

$$d_{P_{n-g_0,g_0}}(C_d, S_c) \geq d_{P_{n-g_0+k,g_0-k}}(C_d, S_c) \text{ for } k \in \{1, 2, \ldots, g_0 - 2\}$$

and

$$d_{P_{n-g_0,g_0}}(C_d, S_c) \geq d_{P_{n-g_0-l,g_0+l}}(C_d, S_c) \text{ for } l \in \{1, 2, \ldots, n - g_0 - 3\}.$$

For $g \in \{2, 3, \ldots, n-3\}$, if $2^g + g \leq n - 1$, then $2g < 2^g + g$ implies $g < \frac{n-1}{2}$, that is, $g < \lfloor \frac{n}{2} \rfloor$. Thus if $g \geq \lfloor \frac{n}{2} \rfloor$, then $2^g + g > n - 1$. So $g_0 \leq \lfloor \frac{n}{2} \rfloor$ by the definition of g_0. Suppose that $g_0 = \lfloor \frac{n}{2} \rfloor$. If n even, then $g_0 = \frac{n}{2}$ and so

$$2^{\frac{n}{2}-1} + \frac{n}{2} - 1 = 2^{g_0-1} + g_0 - 1 \leq n - 1,$$

which gives $2^{\frac{n}{2}} \leq n$. This is possible only when n is 2 or 4. But our assumption is that $n \geq 5$. If n is odd, then $g_0 = \frac{n-1}{2}$, and so

$$2^{\frac{n-3}{2}} + \frac{n-3}{2} = 2^{g_0-1} + g_0 - 1 \leq n - 1$$

which gives $2^{\frac{n-3}{2}} \leq \frac{n+1}{2}$. This is possible only when n is 5 or 7. For these two values of n, it can be verified that $d_{P_{n-g_0,g_0}}(C_d, S_c) \geq d_{P_{n-g,g}}(C_d, S_c)$ for all possible values of g.

So assume that $g_0 < \lfloor \frac{n}{2} \rfloor$. Note that for $k \in \{1, 2, \ldots, g_0 - 2\}$,

$$2^{g_0-k} - (g_0 - k) - [2^{g_0-k-1} - (g_0 - k - 1)] = 2^{g_0-k-1} - 1 > 0.$$

This implies that

$$2^{g_0-1} - g_0 + 1 \geq 2^{g_0-k} - g_0 + k, \tag{16}$$

for all $k \in \{1, 2, \ldots, g_0 - 2\}$. If n is odd, then using (16), we get

$$\begin{aligned} d_{P_{n-g_0,g_0}}(C_d, S_c) - d_{P_{n-g_0+k,g_0-k}}(C_d, S_c) &\geq \frac{n-1}{2} - g_0 \\ &\quad - \left[\frac{2^{g_0-k} - g_0 + k - 1}{2}\right] \\ &\geq \frac{n-1}{2} - g_0 \\ &\quad - \left[\frac{2^{g_0-1} - g_0 + 1 - 1}{2}\right] \\ &= \frac{n - 1 - g_0 - 2^{g_0-1}}{2}. \end{aligned}$$

Similarly, if n is even, then

$$\begin{aligned}d_{P_{n-g_0,g_0}}(C_d, S_c) - d_{P_{n-g_0+k,g_0-k}}(C_d, S_c) &\geq \frac{n}{2} - 1 - g_0 \\ &\quad - \left\lceil \frac{2^{g_0-k} - g_0 + k - 2}{2} \right\rceil \\ &\geq \frac{n}{2} - 1 - g_0 \\ &\quad - \left\lceil \frac{2^{g_0-1} - g_0 + 1 - 2}{2} \right\rceil \\ &= \frac{n - 1 - g_0 - 2^{g_0-1}}{2}.\end{aligned}$$

Thus, in both the cases,

$$\begin{aligned}d_{P_{n-g_0,g_0}}(C_d, S_c) - d_{P_{n-g_0+k,g_0-k}}(C_d, S_c) &\geq \frac{n - 1 - g_0 - 2^{g_0-1}}{2} \\ &= \frac{n - 1 - [2^{g_0-1} + g_0 - 1 + 1]}{2} \\ &\geq \frac{n - 1 - (n - 1 + 1)}{2} \\ &= \frac{-1}{2}.\end{aligned}$$

Since the left-hand side is an integer, we must have

$$d_{P_{n-g_0,g_0}}(C_d, S_c) - d_{P_{n-g_0+k,g_0-k}}(C_d, S_c) \geq 0.$$

We now show that $d_{P_{n-g_0,g_0}}(C_d, S_c) \geq d_{P_{n-g_0-l,g_0+l}}(C_d, S_c)$ for $l \in \{1, 2, \ldots, n - g_0 - 3\}$. If n is odd, then

$$d_{P_{n-g_0,g_0}}(C_d, S_c) - d_{P_{n-g_0-l,g_0+l}}(C_d, S_c) \geq \frac{n-1}{2} - g_0 - \frac{n-1}{2} + g_0 + l = l.$$

Similar argument holds if n is even. This completes the proof. □

As a consequence of Proposition 3.5, we have the following.

Theorem 3.6. *Let T be a tree on $n \geq 5$ vertices and let g_0 be the smallest positive integer such that $2^{g_0} + g_0 > n - 1$. Then $d_T(C_d, S_c) \leq \lfloor \frac{n-1}{2} \rfloor - g_o$.*

4. Position of the centroid

In this section, we study the position of the centroid with respect to the center and the subtree core of a tree. We prove that the centroid always lies in the path connecting the center and the subtree core in any path-star tree. However, this statement need not be true for a general tree.

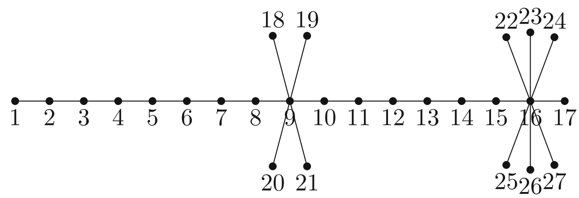

Figure 3. Centroid outside the path connecting the center and the subtree core.

PROPOSITION 4.1

In any path-star tree $P_{n-g,g}$, the centroid lies in the path connecting the center and the subtree core.

Proof. By Theorem 2.5, the vertex $\lceil\frac{n-g+1}{2}\rceil \in C(P_{n-g,g})$. Again, by Theorem 3.3, $C_d(P_{n-g,g}) = \{n-g\}$ or the vertex $\lceil\frac{n}{2}\rceil \in C_d(P_{n-g,g})$. First, assume that $C_d(P_{n-g,g}) = \{n-g\}$. This happens only when $g > \lfloor\frac{n-1}{2}\rfloor$. Then

$$2^g + 1 > 2^{\lfloor\frac{n-1}{2}\rfloor} + 1 > n - \lfloor\frac{n-1}{2}\rfloor > n-g.$$

So, by Theorem 2.4, $S_c(P_{n-g,g}) = \{n-g\}$. As $n-g > \lceil\frac{n-g+1}{2}\rceil$, the statement follows.

Now assume that $\lceil\frac{n}{2}\rceil \in C_d(P_{n-g,g})$. This happens only when $g \le \lfloor\frac{n}{2}\rfloor$. Since $\lceil\frac{n-g+1}{2}\rceil \le \lceil\frac{n}{2}\rceil$ for $g \in \{2, 3, \ldots, \lfloor\frac{n}{2}\rfloor\}$, the center $C(P_{n-g,g})$ is contained in the branch at $\lceil\frac{n}{2}\rceil$ which contains the vertex $\lceil\frac{n}{2}\rceil - 1$. If $2^g + 1 > n-g$, then $S_c(P_{n-g,g}) = \{n-g\}$ and the statement follows. If $2^g + 1 \le n-g$, then $\lceil\frac{n-g+2^g}{2}\rceil \in S_c(P_{n-g,g})$. Since $\lceil\frac{n-g+2^g}{2}\rceil > \lceil\frac{n}{2}\rceil$, $S_c(P_{n-g,g})$ is contained in the branch at $\lceil\frac{n}{2}\rceil$ which contains the vertex $n-g$ and the statement follows. □

We now give an example of a tree in which the centroid does not lie in the path connecting the center and the subtree core.

Example 4.2. Consider the tree T in figure 3. Observe that $C(T) = \{9\}$ and $C_d(T) = \{10\}$. We will show that $S_c(T) = \{9\}$. Let B_1, B_2 be the two components of $T - \{8, 9\}$ (deleting the edge $\{8, 9\}$ from T) containing vertices 8 and 9, respectively. Then

$$\begin{aligned} f_T(9) - f_T(8) &= f_{B_2}(9) + f_{B_2}(9) f_{B_1}(8) - f_{B_1}(8) - f_{B_2}(9) f_{B_1}(8) \\ &= f_{B_2}(9) - f_{B_1}(8) > 0. \end{aligned} \tag{17}$$

The last inequality holds since B_2 contains a copy of B_1 (by identifying the vertex 8 of B_1 with 9 of B_2) and B_2 has more vertices than B_1. So $f_T(9) > f_T(8)$. Let M and N be the two components of $T - \{9, 10\}$ containing vertices 9 and 10, respectively. Then

$$\begin{aligned} f_T(9) - f_T(10) &= f_M(9) + f_M(9) f_N(10) - f_N(10) - f_M(9) f_N(10) \\ &= f_M(9) - f_N(10) \\ &= (9 \times 2^4) - (6 + 2^7) \\ &= 144 - 134 > 0. \end{aligned} \tag{18}$$

So $f_T(9) > f_T(10)$ and hence $S_c(T) = \{9\}$. Thus $C_d(T)$ does not lie on the path connecting $C(T)$ and $S_c(T)$.

References

[1] Harary F, Graph Theory (1969) (Addison-Wesley Publishing Co.)
[2] Mitchell S L, Another characterization of the centroid of a tree, *Discrete Math.* **24** (1978) 277–280
[3] Patra K L, Maximizing the distance between center, centroid and characteristic set of a tree, *Linear Multilinear Algebra* **55** (2007) 381–397
[4] Székely L A and Wang H, On subtrees of trees, *Adv. Appl. Math.* **34** (2005) 138–155
[5] Zelinka B, Median and peripherians of trees, *Arch. Math.* **4** (1968) 87–95

COMMUNICATING EDITOR: B V Rajarama Bhat