A Conjecture on Different Central Parts of Binary Trees

Dinesh Pandey1,2 · Kamal Lochan Patra1,2

Received: 8 October 2020 / Revised: 1 August 2022 / Accepted: 6 November 2022
© The Author(s), under exclusive licence to Springer Japan KK, part of Springer Nature 2022

Abstract
Let \( \Omega_n \) be the family of binary trees on \( n \) vertices obtained by identifying the root of an rgood binary tree with a vertex of maximum eccentricity of a binary caterpillar. In the paper titled “On different middle parts of a tree” (The Electronic Journal of Combinatorics, 25 (2018), no. 3, paper 3.17, 32 pp), Smith et al. conjectured that among all binary trees on \( n \) vertices the pairwise distance between any two of center, centroid and subtree core is maximized by some member of the family \( \Omega_n \). We first obtain the rooted binary tree which minimizes the number of root-containing subtrees and then prove this conjecture. We also obtain the binary trees which maximize these distances.

Keywords  Binary tree · Center · Centroid · Subtree core · Distance

Mathematics Subject Classification  05C05 · 05C12 · 05C35

1 Introduction

All the graphs in this paper are simple, finite, connected and undirected. A tree \( T \) is a connected, acyclic graph. The vertex and edge sets of \( T \) are denoted by \( V(T) \) and \( E(T) \) respectively. A subtree of \( T \) is a connected subgraph of \( T \). For \( u, v \in V(T) \), the distance \( d_T(u, v) \) or simply \( d(u, v) \), is the number of edges in the path joining \( u \) and \( v \). The distance between two subsets \( X \) and \( Y \) of \( V(T) \) is denoted by \( d_T(X, Y) \) and defined as \( d_T(X, Y) = \min\{d(x, y) : x \in X, y \in Y\} \). The eccentricity of a vertex

\[ \epsilon(v) = \max\{d(v, u) : u \in V(T)\} \]

Kamal Lochan Patra
klpatra@niser.ac.in

Dinesh Pandey
dinesh.pandey@niser.ac.in

1 School of Mathematical Sciences, National Institute of Science Education and Research (NISER), P.O. Jatni, District-Khurda, Bhubaneswar, Odisha 752050, India

2 Homi Bhabha National Institute (HBNI), Training School Complex, Anushakti Nagar, Mumbai 400094, India

Published online: 19 November 2022
\( v \in V(T) \) is denoted by \( e(v) \) and defined as \( e(v) = \max\{d_T(u, v) : u \in V(T)\} \). The radius \( \text{rad}(T) \) of \( T \) is defined as \( \text{rad}(T) = \min\{e(v) : v \in V(T)\} \) and the diameter \( \text{diam}(T) \) of \( T \) is defined as \( \text{diam}(T) = \max\{e(v) : v \in V(T)\} \). It is clear that \( \text{diam}(T) = \max\{d_T(u, v) : u, v \in V(T)\} \). By degree of a vertex \( v \) in \( T \), we mean the number of edges incident with \( v \) and we denote it as \( \deg(v) \). A vertex \( v \) is called a pendant vertex if \( \deg(v) = 1 \). A set of vertices \( X \) of \( T \) lies on a path \( P : v_1v_2\cdots v_k \) (from \( v_1 \) to \( v_k \)) means \( X \subseteq \{v_1, v_2, \ldots, v_k\} \).

By specifying a vertex \( r \in V(T) \), we call \( T \) a rooted tree with root \( r \). A binary tree is a tree in which every non-pendant vertex has degree 3. A rooted binary tree is a tree in which the root has degree 2 and any other vertex is either a pendant vertex or a vertex of degree 3. Note that the number of vertices in a binary tree is always even and every binary tree on \( n \) vertices has \( \frac{n+2}{2} \) pendant vertices. The number of vertices in a rooted binary tree is always odd.

Let \( T \) be a rooted binary tree with root \( r \). For a vertex \( v \in V(T) \), the height \( h_T(v) \) of \( v \) is defined by \( h_T(v) = d(v, r) \). Let \( P(r, v) \) denote the path joining \( r \) and \( v \). For \( u, v \in V(T) \), \( v \) is called a successor of \( u \) if \( P(r, u) \subset P(r, v) \). If \( v \) is a successor of \( u \) and \( u \) is adjacent to \( v \), we call \( v \) a child of \( u \) and \( u \) the parent of \( v \). The height of \( T \) is denoted by \( h_T(T) \) and is defined as \( h_T(T) = \max\{h_T(v) : v \in V(T)\} \). We say a vertex \( v \in V(T) \) is at level \( l \) if its height is \( l \).

An ordering of the vertices of a rooted binary tree with height \( h \) consists of \( h + 1 \) linear orders \( \pi_0, \pi_1, \ldots, \pi_h \) such that \( \pi_\ell \) is a linear order of the vertices (from left to right) at level \( \ell \) and if \( u \leq v \) in \( \pi_{\ell+1} \) with \( p(u) \) and \( p(v) \) being the parents of \( u \) and \( v \) respectively, then \( p(u) \leq p(v) \) in \( \pi_\ell \), where \( 0 \leq \ell \leq h - 1 \).

**Definition 1.1** [12] A rooted binary tree with height \( h \) is called an rgood binary tree if

(i) the heights of any two of its pendant vertices differ by at most 1 and

(ii) there is an ordering of the vertices of the tree such that if \( u \) and \( v \) are on level \( h - 1 \) and \( u \) is the parent of a pendant vertex while \( v \) is not, then \( v \leq u \) in \( \pi_{h-1} \).

An rgood binary tree on 9 vertices with levels is shown in Fig. 1. A single vertex rooted binary tree is also rgood. All rgood binary trees on \( n \) vertices are isomorphic and we denote any such tree by \( T_{rg}^n \). A caterpillar is a tree which has a path such that every vertex not on the path is adjacent to some vertex on the path. A binary caterpillar is a caterpillar which is also a binary tree. Note that a binary caterpillar on \( n \) vertices has diameter \( \frac{n}{2} \).

In this paper we are interested in the following central parts of binary trees: center, centroid and subtree core. We recall the definitions of these central parts which can be found in some books and papers (see [5, 12]). The center of a tree \( T \) is the set of
vertices having minimum eccentricity. We denote the center of a tree $T$ by $C(T)$. An element of $C(T)$ is called a central vertex.

For $v \in V(T)$, a branch at $v$ is a maximal subtree of $T$ containing $v$ as a pendant vertex. The weight of $v$ is the maximal number of edges in any branch of $T$ at $v$. We denote the weight of a vertex $v \in V(T)$ by $W_T(v)$ or simply by $W(v)$. The weight of a branch $B$ at $v$ is the number of edges in $B$ and we denote it as $W_B(v)$. A vertex of minimal weight is called a centroid vertex of $T$ and the set of all centroid vertices is called the centroid of $T$ (see [5]). We denote the centroid of $T$ by $Cd(T)$. The following result is due to Jordan [6].

**Proposition 1.2** ([5], Theorem 4.2, Theorem 4.3)

1. The center of a tree consists of either a single vertex or two adjacent vertices.
2. The centroid of a tree consists of either a single vertex or two adjacent vertices.

It is straightforward that $C(T)$ intersects with the center of every longest path in $T$. If $|V(T)| = n$ and $|Cd(T)| = 2$ with $Cd(T) = \{u, v\}$, then $n$ must be even and $W(u) = W(v) = n/2$. Also, among the branches at $u$ (respectively, at $v$), the branch containing $v$ (respectively, $u$) has the maximum number of edges.

For $v \in V(T)$, $f_T(v)$ is the number of subtrees of $T$ containing $v$. The subtree core of $T$ is the set of vertices $v$ for which $f_T(v)$ attains its maximum. We denote the subtree core of a tree $T$ by $Sc(T)$. The following result is due to Székely and Wang.

**Proposition 1.3** ([12], Theorem 9.1) The subtree core of a tree consists of either a single vertex or two adjacent vertices.

The subtree core is the most recently defined central part of a tree. To prove Proposition 1.3, the authors used the fact that the function $f_T$ is strictly concave in the following sense.

**Lemma 1.4** ([12], Proof of Theorem 9.1) If $u, v, w$ are three vertices of a tree $T$ with $\{u, v\}, \{v, w\} \in E(T)$, then $2f_T(v) - f_T(u) - f_T(w) > 0$.

The concept of central parts in trees were started by Jordan ([6]) in 1869 with the definitions of center and centroid. Later many researchers contributed to it by giving definitions of median ([13]), telephone center ([9]), distance center ([7]), characteristic set ([8]) and subtree core ([12]) (also see [4]). But the idea of studying the pairwise distances between them is comparatively new. In the last 15 years, the distances between these middle parts in various classes of trees have been studied by many researchers (see [1, 3, 10, 11]). Here we consider the class of binary trees on $n$ vertices and the pairwise distances between the central parts center, centroid and subtree core.

### 1.1 Crg tree $T_{rg}^{n,l}$

Let $n \geq 4$ be a positive even integer and let $l \geq 3$ be a positive odd integer such that $l < n$. Let $T_{rg}^{n,l}$ denote the tree on $n$ vertices which is obtained by identifying the root of $T_{rg}^{l}$ with a vertex of maximum eccentricity of a binary caterpillar tree on $n - l + 1$ vertices (see Fig. 2). Such a tree $T_{rg}^{n,l}$ is called a crg tree. The binary caterpillar is the crg tree $T_{rg}^{n,3}$. 
Fig. 2 The crg tree \( T_{18}^{11} \)

We label the vertices of a longest path of the caterpillar part of \( T_{n}^{l} \) by \( 1, 2, \ldots, \left\lfloor \frac{n-1}{2} \right\rfloor = v \), where \( v \) is the root of the rgood part of it. We denote by \( \Omega_{n} \) the class of all crg trees on \( n \) vertices. Any binary tree on \( n \leq 8 \) vertices is isomorphic to a binary caterpillar. Due to the symmetry in the binary caterpillar trees, we observe the following:

The center, centroid and subtree core coincide in binary caterpillar trees.

This observation shows that among all binary trees on \( n \) vertices the minimum distance between any two of the above three central parts is zero. So, it is interesting to see which trees maximize these distances among all binary trees on \( n \) vertices. There are two non-isomorphic binary trees on 10 vertices and both are crg trees. Also in any crg tree on 10 vertices, the center, centroid and subtree core coincide. So throughout this paper, we consider binary trees on \( n \geq 12 \) vertices. Smith et al. conjectured the following result in [11] (see Conjecture 3.10).

Among all binary trees on \( n \) vertices, the pairwise distance between any two of center, centroid and subtree core is maximized by some trees of the family \( \Omega_{n} \).

In this paper, we prove this conjecture. The paper is organised in the following way: In Sects. 2 and 3, we develop some results on binary and rooted binary trees which are useful to prove our main results. In Sect. 4, we prove the above conjecture and obtain the trees which achieve the maximum distances.

2 Preliminaries

Let the height of \( T_{n}^{l} \) be \( h \geq 1 \). Then \( 2^{h} + 1 \leq n \leq 2^{h+1} - 1 \). There exists a positive integer \( \alpha \) such that \( n = 2^{h} + \alpha \), which gives \( h = \log_{2}(n - \alpha) \). There are two branches at the root of an rgood binary tree. A branch having maximum weight between the two, is called a heavier branch. If both the branches have the same weight then we say the rgood binary tree is complete. In this case any branch can be considered as heavier.

Let \( T \) be a rooted binary tree. Suppose the pendant vertices of \( T \) are in at least three different levels. Form a new rooted binary tree \( T' \) from \( T \) by moving a pair of pendant vertices with the same parent from the largest level to the smallest level which contains pendant vertices. Then \( \text{ht}(T') \leq \text{ht}(T) \). This leads to the next result which tells about the rooted binary trees with minimum height.

Lemma 2.1 Among all rooted binary trees on \( n \) vertices, \( \text{ht}(T_{n}^{l}) \leq \text{ht}(T) \) and equality holds when \( T \) is a rooted binary tree in which the heights of any two pendant vertices differ by at most one.
In the following result we determine the binary tree on \( n \) vertices which has maximum diameter.

**Lemma 2.2** Among all binary trees on \( n \) vertices, the binary caterpillar has the maximum diameter.

**Proof** Let \( \mathcal{B}_n \) be the set of all binary trees on \( n \) vertices and let \( k = \max \{ \text{diam}(T) | T \in \mathcal{B}_n \} \). Let \( T' \in \mathcal{B}_n \) such that \( \text{diam}(T') = k \). Let \( P : u_0u_1 \ldots u_k \) be a path of maximum length in \( T' \). Suppose \( T' \) is not a caterpillar. Then there exist two pendant vertices \( v_1, v_2 \in V(T') - V(P) \) adjacent to \( v \) such that \( v \) is not on the path \( P \). Delete the vertices \( v_1, v_2 \) and add them as pendant vertices at \( u_0 \) to get a new tree \( T'' \in \mathcal{B}_n \). Then \( \text{diam}(T'') > \text{diam}(T') \), which is a contradiction. This completes the proof. \( \square \)

For \( e = \{u, v\} \in E(T) \), let \( T(e, u) \) denote the component of \( T - e \) containing \( u \). The following lemma is very useful.

**Lemma 2.3** Let \( e = \{u, v\} \in E(T) \). Then \( |V(T(e, u))| > |V(T(e, v))| \) if and only if \( C_d(T) \subseteq V(T(e, u)) \).

**Proof** Let \( |V(T(e, u))| = k \) and \( |V(T(e, v))| = k' \). First suppose \( |V(T(e, u))| > |V(T(e, v))| \). Since \( k > k' \), it follows that \( W_t(v) = k \) (as the branch at \( v \) containing \( u \) has \( k \) edges and any other branch at \( v \) has at most \( k - 1 \) edges) and for any \( w' \in V(T(e, v)), w' \neq v, W_t(w') > W_t(v) \). So the only possible vertex of \( T(e, v) \) which may belong to \( C_d(T) \) is \( v \).

Let \( B \) be the branch at \( u \) containing \( v \). If \( W_t(u) = W_t(B) \), then \( W_t(u) = k' < k = W_t(v) \). If \( W_t(u) \neq W_t(B) \), then \( W_t(u) \) is the weight of a branch of \( T(e, u) \) at \( u \) and so \( W_t(u) \leq k - 1 < k = W_t(v) \). Hence \( \min \{ W_t(z) : z \in V(T) \} \leq W_t(u) < W_t(v) \). This implies that \( C_d(T) \subseteq V(T(e, u)) \).

Now suppose \( C_d(T) \subseteq V(T(e, u)) \). Let \( w \in C_d(T) \). Then \( W_t(w) \geq k' \) as the branch at \( w \) containing \( v \) has weight at least \( k' \). Since \( v \notin C_d(T) \), we have \( W_t(v) > W_t(w) \geq k' \). This implies that \( W_t(v) \) is the weight of the branch at \( v \) containing \( u \). So \( W_t(v) = k \), hence \( k > k' \). \( \square \)

**Corollary 2.4** Let \( v \) be the root of the rgood part of \( T_{rg}^{n,l} \) and let \( v' \) be the vertex in a heavier branch of the rgood part at \( v \) such that \( \{v, v'\} \in E(T_{rg}^{n,l}) \). If \( l \geq \frac{n}{2} + 1 \) then \( C_d(T_{rg}^{n,l}) \subseteq \{v, v'\} \). Moreover, if the rgood part is complete then \( C_d(T_{rg}^{n,l}) = \{v\} \).

**Proof** Let \( T \) be a crg tree with \( l \geq \frac{n}{2} + 1 \) and let \( T' \) be the rgood part of \( T \). Since \( l \geq \frac{n}{2} + 1 \), by Lemma 2.3, \( C_d(T) \subseteq V(T') \). Let \( e = \{w, v\} \in E(T') \) where \( w \neq v' \). Then a copy of \( T'(e, w) \) is properly contained in \( T'(e, v) \) and so \( |V(T(e, v))| > |V(T(e, w))| \). Hence by Lemma 2.3, \( C_d(T) \) is contained in the branch of \( T' \) at \( v \) containing \( v' \).

Suppose \( e' = \{u, v'\} \in E(T') \) where ht\( (u) = 2 \) in \( T' \). Then a copy of \( T'(e', u) \) is properly contained in \( T'(e', v') \) and so \( |V(T(e', v'))| > |V(T(e', u))| \). Hence by Lemma 2.3, \( C_d(T) \subseteq \{v, v'\} \).

If \( T' \) is complete then \( T' \) has two heavier branches at \( v \). So we have \( C_d(T) \subseteq \{v, v'\} \) and \( C_d(T) \subseteq \{v, w\} \) as both the branches of \( T' \) at \( v \) are heavier and hence \( C_d(T) = \{v\} \). \( \square \)
For \( l \geq \frac{n}{2} + 1 \), in Corollary 2.4, we proved that \( C_d(T_{rg}^{n,l}) \subseteq \{v, v'\} \). We also showed that \( C_d(T_{rg}^{n,l}) = \{v\} \) if the rgood part is complete. For many values of \( n \) and \( l \), the other two cases will also happen. For example, it can be checked that \( C_d(T_{rg}^{12,11}) = \{v'\} \) and \( C_d(T_{rg}^{14,13}) = \{v, v'\} \).

**Corollary 2.5** Let \( v \) be the root of the rgood part of \( T_{rg}^{n,l} \) and let \( v' \) be the vertex in a heavier branch of the rgood part at \( v \) such that \( \{v, v'\} \subseteq E(T_{rg}^{n,l}) \). If \( n = 4k \) and \( l \geq 2k + 1 \) then \( C_d(T_{rg}^{n,l}) \subseteq \{\{v\}, \{v'\}\} \).

**Proof** Let \( T \) be a crg tree with \( n = 4k \) and \( l \geq 2k + 1 \). By Corollary 2.4, \( C_d(T) \subseteq \{v, v'\} \). Let \( e = \{v, v'\} \subseteq E(T) \). If \( C_d(T) = \{v, v'\} \) then \( |V(T(e,v))| = |V(T(v,v'))| = 2k \). But both \( T(e,v) \) and \( T(e,v') \) are rooted binary trees with roots \( v \) and \( v' \), respectively and hence both must have an odd number of vertices. Thus a contradiction arises, so \( C_d(T) \subseteq \{\{v\}, \{v'\}\} \). \( \square \)

We will now prove a result similar to Lemma 2.3 related to the subtree core of trees.

**Lemma 2.6** Let \( e = \{u, v\} \subseteq E(T) \). Then \( S_c(T) \subseteq V(T(e,u)) \) if and only if \( f_{T(e,u)}(u) > f_{T(e,v)}(v) \).

**Proof** We have

\[
f_T(u) = f_{T(e,u)}(u) + f_{T(e,u)}(u)f_{T(e,v)}(v)
\]

and

\[
f_T(v) = f_{T(e,v)}(v) + f_{T(e,u)}(u)f_{T(e,v)}(v).
\]

So,

\[
f_T(u) - f_T(v) = f_{T(e,u)}(u) - f_{T(e,v)}(v).
\]

If \( f_{T(e,u)}(u) > f_{T(e,v)}(v) \) then \( f_T(u) > f_T(v) \). By Lemma 1.4, we have \( f_T(u) > f_T(w) \) for any \( w \in V(T(e,v)) \). So \( S_c(T) \subseteq V(T(e,u)) \). Conversely, if \( S_c(T) \subseteq V(T(e,u)) \) then by Lemma 1.4, \( f_T(u) > f_T(v) \) and so \( f_{T(e,u)}(u) > f_{T(e,v)}(v) \). This proves the result. \( \square \)

**Corollary 2.7** ([11], Proposition 1.7) A vertex \( u \in S_c(T) \) if and only if for each neighbour \( v \) of \( u \), \( f_{T(e,u)}(u) \geq f_{T(e,v)}(v) \) where \( e = \{v, u\} \). Furthermore if \( u \in S_c(T) \) and \( f_{T(e,u)}(u) = f_{T(e,v)}(v) \) for a neighbour \( v \) of \( u \) then \( v \in S_c(T) \).

In the following we discuss the position of center, centroid and subtree core of \( T_{rg}^n \).

**Lemma 2.8** Let \( v \) be the root of \( T_{rg}^n \), \( n \geq 1 \) and let \( v' \) be the vertex in a heavier branch of \( T_{rg}^n \) such that \( e = \{v, v'\} \subseteq E(T_{rg}^n) \). Then center, centroid and subtree core of \( T_{rg}^n \) are contained in the set \( \{v, v'\} \). Moreover if \( T_{rg}^n \) is complete then \( C(T_{rg}^n) = C_d(T_{rg}^n) = S_c(T_{rg}^n) = \{v\} \).
Proof For $n = 1$ or $3$, $T_{rg}^n$ is a path on $n$ vertices and hence the result follows. Now consider $n \geq 5$. Let $P$ be a longest path of $T_{rg}^n$. Then it must go through $v$ and we get, either $C(P) = \{v\}$ or $C(P) = \{v, v'\}$ depending on whether the length of $P$ is even or odd, respectively. So, $C(T_{rg}^n) \in \{\{v\}, \{v, v'\}\}$. Let $w' \neq v'$ and $e' = \{v, w'\} \in E(T_{rg}^n)$. Since $v'$ is in a heavier branch, $|V(T_{rg}^n(e', v))| > |V(T_{rg}^n(e', w))|$. By Lemma 2.3, $C_d(T_{rg}^n) \subseteq V(T_{rg}^n(e', v))$. Since $n \geq 5$, we have $\deg(v') = 3$. Let $e^1 = \{v', v_1\}, e^2 = \{v', v_2\} \in E(T_{rg}^n)$. For $i = 1, 2$, $T_{rg}^n(e^i, v')$ contains a copy of $T_{rg}^n(e^i, v_i)$. By Lemma 2.3, $C_d(T_{rg}^n) \subseteq V(T_{rg}^n(e^i, v'))$. Hence $C_d(T_{rg}^n) \subseteq \{v, v'\}$.

Since $v'$ is in a heavier branch, the tree $T_{rg}^n(e', v)$ properly contains a copy of the tree $T_{rg}^n(e', w')$. So $f_{T_{rg}^n(e', v)}(v) > f_{T_{rg}^n(e', w')}(w)$ and hence by Lemma 2.6, $S_c(T_{rg}^n) \subseteq V(T_{rg}^n(e', v))$. Also for $i = 1, 2$, the tree $T_{rg}^n(e^i, v')$ properly contains a copy of the tree $T_{rg}^n(e^i, v_i)$. So by Lemma 2.6, $S_c(T_{rg}^n) \subseteq V(T_{rg}^n(e^i, v'))$ for $i = 1, 2$. Hence $S_c(T_{rg}^n) \subseteq \{v, v'\}$.

If $T_{rg}^n$ is complete then $T_{rg}^n$ has two heavier branches at $v$ and in this case $C(T_{rg}^n) = C_d(T_{rg}^n) = S_c(T_{rg}^n) = \{v\}$.

Lemma 2.9 Let $v$ be the root of the rgood part $X$ of $T_{rg}^{n,l}$. Let $v'$ be the vertex on a heavier branch at $v$ with $\{v, v'\} \in E(T_{rg}^{n,l})$. Then center, centroid and subtree core of $T_{rg}^{l,l}$ lie on the path from $v$ to $v'$.

Proof Let $B_1$ and $B_2$ be the two branches at $v$ lying on the rgood part of the tree $T_{rg}^{n,l}$ and let $B_1$ be heavier. Let $h$ be the height of the rgood part of $T_{rg}^{n,l}$ and let $l = d_{T_{rg}^{l}}(1, v)$.

Suppose $b_1$ is a pendant vertex of $T_{rg}^{n,l}$ lying on $B_1$ such that $d(v, b_1) = h$ and $b_2$ is a pendant vertex of $T_{rg}^{n,l}$ lying on $B_2$ whose distance from $v$ is maximum over all vertices of $B_2$. If $h > l$ then the path from $b_1$ to $b_2$ is a longest path in $T_{rg}^{n,l}$ and hence $C(T_{rg}^{n,l}) \subseteq \{v, v'\}$. If $h \leq l$ then the path joining $B_1$ and $B_2$ is a longest path in $l$ which implies that $C(T_{rg}^{n,l})$ lies on the path joining $v$ and $v'$. Hence $C(T_{rg}^{n,l})$ lies on the path joining $v$ and $v'$. Let $e^1 = \{v, w\}$ where $w$ is a vertex in $B_2$. Then $|V(T_{rg}^{n,l}(e^1, v))| > |V(T_{rg}^{n,l}(e^1, w))|$. As $T_{rg}^{n,l}(e^1, v)$ properly contains a copy of $T_{rg}^{n,l}(e^1, v)$. So by Lemma 2.3, $C_d(T_{rg}^{n,l}) \subseteq V(T_{rg}^{n,l}(e^1, v))$. Let $e^2 = \{v', u\}$ where $u$ is an arbitrary child of $v'$. Then $T_{rg}^{n,l}(e^2, v')$ properly contains a copy of $T_{rg}^{n,l}(e^2, u)$. Therefore, $|V(T_{rg}^{n,l}(e^2, v'))| > |V(T_{rg}^{n,l}(e^2, u))|$. And so by Lemma 2.3, $C_d(T_{rg}^{n,l}) \subseteq V(T_{rg}^{n,l}(e^2, v'))$. Thus we get, $C_d(T_{rg}^{n,l}) \subseteq V(T_{rg}^{n,l}(e^1, v)) \cap V(T_{rg}^{n,l}(e^2, v'))$. Since for any tree on $n \geq 3$ vertices, the pendant vertices cannot be centroid vertices, $C_d(T_{rg}^{n,l})$ lies on the path joining $1$ and $v'$.

Using Lemma 2.6, the proof for the subtree core is similar to the proof for the centroid.

Corollary 2.10 Let $v$ be the root of the rgood part of $T_{rg}^{n,l}$ and let $v'$ be the vertex on a heavier branch at $v$ with $\{v, v'\} \in E(T_{rg}^{n,l})$. Then $C(T_{rg}^{n,l}) \neq \{v'\}$.

Proof Let $T'$ be the rgood part of $T_{rg}^{n,l}$ and also let $ht(T') = h$. Suppose $C(T_{rg}^{n,l}) = \{v'\}$. Then $\text{diam}(T_{rg}^{n,l}) = 2(h - 1)$, which is a contradiction as $\text{diam}(T_{rg}^{n,l}) \geq 2h - 1$. □
3 Root-Containing Subtrees

To prove our main results, it is important to know the rooted binary trees which minimize or maximize the number of root-containing subtrees. In [11], the authors have obtained the rooted binary tree which maximizes the number of root-containing subtrees. Here we obtain the rooted binary tree which minimizes the number of root-containing subtrees.

Proposition 3.1 ([11], Corollary 3.9) Among all rooted binary trees on n vertices, \( T^n_{rg} \) uniquely maximizes the number of root-containing subtrees.

For a tree \( T \) with \( u, v \in V(T) \), we denote the number of subtrees of \( T \) containing \( u \) and \( v \) by \( f_T(u, v) \).

Lemma 3.2 Let \( T \) be a rooted binary tree with root \( r \) and \( x \) be a pendant vertex in \( T \). Let \( y \) be a vertex other than \( x \) in the path joining \( r \) and \( x \) (\( y \) may be equal to \( r \)). Then, \( f_T(r, y) \geq 2f_T(r, x) \) and equality holds if and only if \( y \) is adjacent to \( x \).

Proof Suppose \( x_0 \) is the vertex adjacent to \( x \) in \( T \) and let \( T_0 \) be the tree \( T - x \). Then, \( f_T(r, x) = f_{T_0}(r, x_0) \) and

\[
f_T(r, y) = f_{T_0}(r, y) + f_{T_0}(r, x_0) \geq 2f_{T_0}(r, x_0) = 2f_T(r, x).
\]

The inequality holds, as any tree containing \( r \) and \( x_0 \) must contain \( r \) and \( y \) and equality holds if and only if \( y = x_0 \).

We denote the rooted binary tree on \( n \) vertices with exactly two vertices at every level (except the zero level) by \( T^n_{r,2} \).

Proposition 3.3 Among all rooted binary trees on \( n \) vertices, the tree \( T^n_{r,2} \) uniquely minimizes the number of root-containing subtrees.

Proof Let \( T \) be a rooted binary tree with root \( r \) in which there are more than two vertices at some levels. Let \( x \) be a pendant vertex of \( T \) such that \( d(r, x) = dT(r) \). Let \( y \) be the vertex nearest to \( r \) (\( y \) may be the same as \( r \)) such that every branch at \( y \) contains more than two vertices (let \( l \) be the smallest level in which there are more than two vertices. Then there are exactly four vertices at level \( l \). The non-vertex at level \( l - 2 \) is the vertex \( y \)). Then the path joining \( r \) and \( x \) must contain \( y \). Let \( y_0 \), \( y_1 \) and \( y_2 \) be the vertices adjacent to \( y \). Let the branch at \( y \) containing \( y_0 \) be the branch which contains \( r \). If \( y = r \), then \( y_1 \) and \( y_2 \) are the only two vertices adjacent to \( y \).

Let \( X \) and \( Y \) be the branches at \( y \) containing \( y_1 \) and \( y_2 \), respectively and let \( x \) be in the branch \( Y \). Then \( X' = X - y \) is a rooted binary tree with root \( y' = y_1 \). Let \( T' \) be the branch of \( T \) at \( y_1 \) containing \( y \). Then \( T \) can be obtained from \( T' \) and \( X' \) by identifying \( y_1 \) of \( T' \) with \( y' \) of \( X' \). So

\[
f_T(r) = f_{T'}(r) + f_{T'}(r, y_1)(f_{X'}(y') - 1).
\]
Construct a new tree $\hat{T}$ from $T'$ and $X'$ by identifying $x$ of $T'$ with $y'$ of $X'$. Then $\hat{T}$ is a rooted binary tree with root $r$ and $|V(T)| = |V(\hat{T})|$. This gives

$$f_\hat{T}(r) = f_{T'}(r) + f_{T'}(r, x)(f_{X'}(y') - 1).$$

So we have

$$f_T(r) - f_\hat{T}(r) = (f_{X'}(y') - 1)(f_{T'}(r, y_1) - f_{T'}(r, x)).$$

We have $f_{X'}(y') > 1$ as $|V(X')| \geq 3$. Since $y_1$ is a pendant vertex in $T'$ and $y$ is adjacent to $y_1$, by the equality condition of Lemma 3.2, we have $2f_{T'}(r, y_1) = f_{T'}(r, y)$. Then by Lemma 3.2, $f_{T'}(r, y_1) = \frac{1}{2}f_{T'}(r, y) > f_{T'}(r, x)$ as $y$ is not adjacent to $x$. Hence $f_T(r) - f_\hat{T}(r) > 0$. If there are exactly two vertices at every level of $\hat{T}$ then we are done. Otherwise, take $\hat{T}$ as $T$ and repeat the above process till we get the rooted binary tree with exactly two vertices at every level (except level zero). This completes the proof. □

**Corollary 3.4** Let $T$ be a rooted binary tree on $n$ vertices with root $r$. Then $f_T(r) \geq 3 \times 2^{\frac{n-1}{2}} - 2$ and equality holds if and only if $T \cong T_{r, 2}^n$.

**Proof** Let $T$ be a rooted binary tree on $n$ vertices with root $r$. By Proposition 3.3, $f_T(r) \geq f_{T_{r, 2}^n}(r')$ where $r'$ is the root of $T_{r, 2}^n$ and equality holds if and only if $T \cong T_{r, 2}^n$. Suppose $u$ and $v$ are the two vertices of $T_{r, 2}^n$ adjacent to $r'$ among which $u$ is pendant. Let $S_n$ be the number of subtrees of $T_{r, 2}^n$ containing $r'$. We have $S_1 = 1$ and for $n \geq 3$,

$$S_n = 2S_{n-2} + 2$$

where the number of subtrees containing $r'$ but not $v$ is 2 and the number of subtrees containing both $r$ and $v$ is $2S_{n-2}$. By solving this recurrence relation, we get $S_n = 3 \times 2^{\frac{n-1}{2}} - 2$. This proves the result. □

Let $r$ be the root of $T_{r, 2}^n$. It seems difficult to find the value of $f_{T_{r, 2}^n}(r)$. We will only be able to give a bound for $f_{T_{r, 2}^n}(r)$ which is a solution of a non-linear recurrence relation. Let $h$ be the height of $T_{r, 2}^n$ and let $m = 2^{h+1} - 1$. For $n \geq 3$, $2^h - 1 < n \leq m$ and the rooted binary tree $T_{r, 2}^m$ is complete.

Let $A_h$ be the number of subtrees of $T_{r, 2}^m$ containing the root $r$. We have $A_0 = 1$. For $h \geq 1$, let $u$ and $v$ be the vertices adjacent to $r$. Then

$$A_h = 1 + A_{h-1} + A_{h-1}(1 + A_{h-1}) = (A_{h-1} + 1)^2$$

where the first 1 is for the subtree containing only the single vertex $r$, the second term $A_{h-1}$ counts the number of subtrees containing $r$ and $v$ but not $u$ and the third term $A_{h-1}(1 + A_{h-1})$ counts the number of subtrees containing $r$ and $u$. The solution of the non-linear recurrence relation $A_h = (A_{h-1} + 1)^2$ for $h \geq 1$, $A_0 = 1$ is known.
(see [2] and sequence A004019 in the On-Line Encyclopedia of Integer Sequences) and the value is \( A_h = \lceil k^{2^h} \rceil - 1 \) where \( k \approx 2.25851845 \). Then for \( h \geq 1 \), we have
\[
A_{h-1} < f_{T'_{rg}}(r) \leq A_h.
\]
It would be nice to know the exact value of \( f_{T'_{rg}}(r) \).

## 4 Center, Centroid and Subtree Core

For a tree \( T \), by \( d_T(C, C_d) \), \( d_T(C, S_c) \) and \( d_T(C_d, S_c) \) we mean \( d_T(C(T), C_d(T)) \), \( d_T(C(T), S_c(T)) \) and \( d_T(C_d(T), S_c(T)) \), respectively. In this section we obtain the binary trees which maximize the pairwise distances between the central parts center, centroid and subtree core over all binary trees on \( n \) vertices. We first consider the pair center and centroid.

### 4.1 Center and Centroid

**Theorem 4.1** Among all binary trees on \( n \) vertices, the distance between center and centroid is maximized by a crg tree.

**Proof** Let \( T \) be a binary tree on \( n \) vertices with \( d_T(C, C_d) \geq 1 \). Let \( u \in C(T) \) and \( v \in C_d(T) \) such that \( d_T(C, C_d) = d(u, v) \). Let \( e = \{v, w\} \in E(T) \) such that \( w \) lies on the path joining \( u \) and \( v \) (note that \( w \) may be the same as \( u \)). Let \( |V(T(e, v))| = k \). The component \( T(e, v) \) is a rooted binary tree with root \( v \). Since \( C_d(T) \subseteq V(T(e, v)) \), by Lemma 2.3 we have \( |V(T(e, v))| > |V(T(e, w))| \).

If \( T(e, v) \) is an irrelevant binary tree then rename the tree \( T \) as \( T' \). Otherwise, form a new tree \( T' \) from \( T \) by replacing the component \( T(e, v) \) with \( T_{rg}^k \) rooted at \( v \). Since \( |V(T'(e, w))| < k = |V(T'(e, v))| = |V(T_{rg}^k)| \), by Lemma 2.3 we have \( C_d(T') \subseteq V(T'(e, v)) \). By Lemma 2.1, we have \( \text{ht}(T(e, v)) \geq \text{ht}(T_{rg}^k) \). As \( C(T) \) is contained in \( T(e, w) \), any longest path of \( T' \) cannot entirely lie in \( T'(e, v) \) and so \( \text{diam}(T') \leq \text{diam}(T) \). If \( \text{diam}(T') = \text{diam}(T) \) then there is a path \( P \) which is a longest path in both \( T \) and \( T' \) and so \( C(T') = C(T) \). If \( \text{diam}(T') < \text{diam}(T) \) then all the longest paths of \( T' \) must contain \( v \). So, while moving to \( T' \) from \( T \), \( C(T') \) is either same as \( C(T) \) or moves away from the vertex \( v \) as compared to \( C(T) \). Hence, \( d_{T'}(C, C_d) \geq d_T(C, C_d) \).

If \( T' \subseteq \Omega_n \) then the result follows. Otherwise let \( |V(T'(e, w))| = l \). Construct a new tree \( T'' \) from \( T' \) by replacing \( T'(e, w) \) with \( T_{rg}^l \) rooted at \( w \). Observe that \( T'' \in \Omega_n \) and by Lemma 2.2, \( \text{diam}(T'') > \text{diam}(T') \). Since the increment in the length of the longest paths of \( T'' \) occurs in a branch at \( w \) not containing \( v \), while moving to \( T'' \) from \( T' \), \( C(T'') \) moves away from \( v \) as compared to \( C(T') \). Also, \( |V(T''(e, v))| > |V(T''(e, w))| \) and \( T''(e, v) \) is the same as \( T'(e, v) \). So \( C_d(T'') = C_d(T') \). Hence \( d_{T''}(C, C_d) \geq d_{T'}(C, C_d) \geq d_T(C, C_d) \). This proves the result. \( \square \)

**Theorem 4.2** Among all crg trees on \( n \geq 12 \) vertices, the distance between center and centroid is maximized by the tree \( T_{rg}^{n, l} \), where \( l = 2\lceil \frac{n}{4} \rceil + 1 \).
Let $v$ be the root of the right good part of $T_{rg}^{n,l}$ and let $d_{Trg}^{n,l}(C, C _d) = \alpha$. Since $n$ is even, we consider two cases depending on whether $n$ is of the form $4k$ or $4k + 2$. The proofs for $n = 4k$ and $n = 4k + 2$ are very similar and we will only include the proof for $n = 4k$.

Let $n = 4k$ for some $k \geq 3$. Then $l = 2\left\lceil \frac{n}{4} \right\rceil + 1 = 2k + 1$. In $T_{rg}^{n,l}$, the right good part has $2k + 1$ vertices and the caterpillar part has $2k$ vertices. The weight $W_t(v) = 2k - 1$ and the weight of any other vertex of $T_{rg}^{n,l}$ is greater than $2k - 1$. Following the numbering of vertices mentioned in Section 1.1, we have $C_d(T_{rg}^{n,l}) = \{v\} = \{k + 1\}$. The diameter of the caterpillar part is $k$ and the height of the right good part is less than $k$. So, $C(T_{rg}^{n,l})$ lies on the path from $1$ to $k + 1$.

First consider the trees $T_{rg}^{n,l}, T_{rg}^{n,l-2}, \ldots, T_{rg}^{n,5}$. Note that $T_{rg}^{n,5}$ is a binary caterpillar. Then $C_d(T_{rg}^{n,l-i}) = \{k + 1\}$ for $0 \leq i \leq l - 5$. In the above sequence of trees, the center lies on the path from $1$ to $k + 1$. In $T_{rg}^{n,l-i}$, if the vertex numbered $u$ is the central vertex nearest to $k + 1$, then the central vertex in $T_{rg}^{n,l-i-2}$ nearest to $k + 1$ is either $u$ or $u + 1$. So, $d_{Trg}^{n,l-i}(C, C _d) \geq d_{Trg}^{n,l-i-2}(C, C _d)$ for $0 \leq i \leq l - 7$.

Now consider the sequence of trees $T_{rg}^{n,l}, T_{rg}^{n,l+2}, \ldots, T_{rg}^{n,n-1}$. For $0 \leq j \leq n - l - 1$, let $v_j$ be the root of the right good part of $T_{rg}^{n,l+j}$ and $v_j'$ be the vertex in a heavier branch of the right good part of $T_{rg}^{n,l+j}$ adjacent to $v_j$. From Lemma 2.3, it follows that $C_d(T_{rg}^{n,l+j}) = \{v_j, v_j'\}$. If $d_{Trg}^{n,l}(C, C _d) = \alpha = 0$ then $C(T_{rg}^{n,l}) \in \{\{k + 1\}, \{k, k + 1\}\}$. The height of the right good part of $T_{rg}^{n,l}$ is $k$ or $k - 1$ depending on $C(T_{rg}^{n,l}) = \{k + 1\}$ or $C(T_{rg}^{n,l}) = \{k, k + 1\}$, respectively. The right good part of $T_{rg}^{n,l}$ has at least $2^{k-1} + 1$ vertices and hence $2^{k-1} + 1 \leq l = 2k + 1$. So $k \leq 4$. Thus $\alpha = 0$ implies that $n \leq 16$ and in these cases it can be checked that $d_{Trg}^{n,l+j}(C, C _d) = 0$ for $0 \leq j \leq n - l - 1$.

If $\alpha \geq 1$ then $n \geq 20$. If $t$ is the smallest positive integer such that $C_d(T_{rg}^{n,l+t}) = \{v_j'\}$ then $C_d(T_{rg}^{n,l+t+p}) = \{v_{j+p}\}$ for $0 \leq p \leq n - l - t - 1$. In the tree $T_{rg}^{n,l+2}$, let $e = \{v_2, v_3\}$. Let $B_1$ and $B_2$ be the two components of $T_{rg}^{n,l+2} - e$ containing $v_2$ and $v_3$, respectively. Since $n \geq 20$, $|V(B_1)| > 2k$ and $|V(B_2)| < 2k$. Thus by Lemma 2.3, $C_d(T_{rg}^{n,l+2}) \subseteq B_1$ and as $C_d(T_{rg}^{n,l+2}) = \{v_2, v_3\}$, we get $C_d(T_{rg}^{n,l+2}) = \{v_2\}$. Therefore $d_{Trg}^{n,l+2}(C, C _d) = \{\alpha, \alpha - 1\}$. If $d_{Trg}^{n,l+4}(C, C _d) = \alpha$ then $d_{Trg}^{n,l+4}(C, C _d) = \{\alpha, \alpha - 1\}$ for $0 \leq j \leq n - l - 1$. Hence the distance between center and centroid of all the trees in the above sequence is at most $\alpha$. This proves the result.

We will now find the distance between center and centroid of $T_{rg}^{n,l}$ for $l = 2\left\lceil \frac{n}{4} \right\rceil + 1$. Let $h$ be the height of the right good part of $T_{rg}^{n,l}$. Then $h$ is the smallest positive integer such that $l \leq 2^{h+1} - 1$. This implies that $\left\lceil \frac{n}{4} \right\rceil \leq 2^h - 1$.

The caterpillar part of $T_{rg}^{n,l}$ contains $2\left\lfloor \frac{n}{4} \right\rfloor$ vertices. So, the diameter of the caterpillar part is $\left\lfloor \frac{n}{4} \right\rfloor$ and hence the root of the right good part is numbered by $\left\lfloor \frac{n}{4} \right\rfloor + 1$. The central vertex which is nearest to the root of the right good part is numbered by $\left\lfloor \frac{1}{4} + \frac{1 + h}{2} \right\rfloor + 1$. Thus the distance between center and centroid of $T_{rg}^{n,l}$ is $\left\lfloor \frac{n}{4} \right\rfloor - \left\lfloor \frac{1}{4} + \frac{1 + h}{2} \right\rfloor$, where $h$
is the smallest positive integer such that $\lceil \frac{n}{4} \rceil \leq 2^h - 1$. This leads to the following corollary.

**Corollary 4.3** Let $T$ be a binary tree on $n$ vertices and let $h$ be the smallest positive integer such that $\lceil \frac{n}{4} \rceil \leq 2^h - 1$. Then

$$d_{T}(C, C_d) \leq \left\lceil \frac{n}{4} \right\rceil - \left\lfloor \frac{\lceil \frac{n}{4} \rceil + 1 + h}{2} \right\rfloor$$

and equality happens if $T \cong T_{rg}^{n,1}$ where $l = 2\lceil \frac{n}{4} \rceil + 1$.

### 4.2 Center and Subtree Core

**Theorem 4.4** Among all binary trees on $n$ vertices, the distance between center and subtree core is maximized by a crg tree.

**Proof** Let $T$ be a binary tree on $n$ vertices with $d_{T}(C, S_c) \geq 1$. Let $u \in C(T)$ and $v \in S_c(T)$ such that $d_{T}(C, S_c) = d(u, v)$. Let $e = \{v, w\} \in E(T)$ such that $w$ lies on the path joining $u$ and $v$. Let $|V(T(e, v))| = k$. The component $T(e, v)$ is a rooted binary tree with root $v$. Since $S_c(T) \subseteq V(T(e, v))$, by Lemma 2.6 we have $f_{T(e, v)}(v) > f_{T(e, v)}(w)$.

If $T(e, v)$ is an rgood binary tree then rename the tree $T$ by $T'$. Otherwise, form a new tree $T'$ from $T$ by replacing the component $T(e, v)$ with $T_{rg}^{k}$ rooted at $v$. By Proposition 3.1, $f_{T_{rg}^{k}}(v) \geq f_{T(e, v)}(v) > f_{T(e, v)}(w)$ and hence by Lemma 2.6, $S_c(T') \subseteq V(T'(e, v))$. By Lemma 2.1, we have $ht(T(e, v)) \geq ht(T_{rg}^{k})$. So, while moving to $T'$ from $T$, $C(T')$ is either same as $C(T)$ or moves away from the vertex $v$ as compared to $C(T)$. Hence, $d_{T'}(C, S_c) \geq d_{T}(C, S_c)$.

If $T' \in \Omega_n$ then the result follows. Otherwise let $|V(T'(e, w))| = l$. Construct a new tree $T''$ from $T'$ by replacing $T'(e, w)$ with $T_{r,2}^{l}$ rooted at $w$. Observe that $T'' \in \Omega_n$. In $T''$ the length of the longest path is more than the length of the longest path of $T'$ and the increment occurs in a branch at $w$ containing the center. So, while moving to $T''$ from $T'$, $C(T'')$ moves away from $v$ as compared to $C(T')$. By Proposition 3.3, $f_{T_{r,2}^{l}}(w) < f_{T'(e, w)}(w)$. So $f_{T''(e, v)}(v) > f_{T''(e, w)}(w)$ and hence by Lemma 2.6, $S_c(T'') \subseteq V(T''(e, v))$. Thus $d_{T''}(C, S_c) \geq d_{T}(C, S_c)$. This proves the result. \hfill $\square$

**Theorem 4.5** In any crg tree $T_{rg}^{n,1}$, the centroid lies on the path connecting the center and the subtree core.

**Proof** In a binary caterpillar on $n$ vertices the center, centroid and subtree core are the same. So we can consider crg trees which are not caterpillars. Let $T$ be a crg non-caterpillar tree, and let $T'$ be the rgood part of $T$. Let $v$ be the root of $T'$ and let $v'$ be the vertex in a heavier branch of $T'$ such that $\{v, v'\} \in E(T')$. By Corollary 2.9, the center, centroid and subtree core of $T$ lie on the path from 1 to $v'$.

Let $u$ be the central vertex of $T$ nearest to the vertex 1 and let $u'$ be the vertex adjacent to $u$ which lies on the path from 1 to $u$ ($u'$ may be the same as 1). Let $e = \{u', u\} \in E(T)$. By Corollary 2.9 and Corollary 2.10, $u$ lies on the path from 1 to
v and the component $T(e, u')$ is a rooted binary tree with root $u'$ in which every level has exactly two vertices (except the zero level). Hence $|V(T(e, u'))| < |V(T(e, u))|$ and by Lemma 2.3, $C_d(T) \subseteq V(T(e, u))$. Thus $C_d(T)$ lies on the path from $C(T)$ to $v'$.

Let $w$ be the centroid vertex of $T$ nearest to $v'$ and let $w'$ be the vertex adjacent to $w$ which lies on the path from 1 to $w$. Let $e' = \{w', w\}$ be an edge in $T$. If $w' \in C_d(T)$ then $|V(T(e', w))| = |V(T(e', w'))|$, otherwise by Lemma 2.3, we have $|V(T(e', w))| > |V(T(e', w'))|$. We consider the following two cases.

**Case I:** $w$ lies on the path from 1 to $v$

In this case, $T(e', w')$ is a rooted binary tree in which there are exactly two vertices at every level (except the zero level). As $|V(T(e', w'))| \geq |V(T(e', w'))|$, by Proposition 3.3 it follows that $f_T(e', w)(w) > f_T(e', w')(w')$ and hence by Lemma 2.6, $S_c(T) \subseteq V(T(e', w))$.

**Case II:** $w = v'$

In this case $v'$ is a centroid vertex. Let $|V(T(e', w'))| = s$. Note that $T(e', w)$ is an rgood binary tree which contains a copy of $T_{rg}^s$. So by Proposition 3.1, $f_T(e', w)(w) \geq f_T(e', w')(w')$. If $f_T(e', w)(w) > f_T(e', w')(w')$ then by Lemma 2.6 and Corollary 2.9, $S_c(T) = \{v'\}$. Otherwise, $f_T(e', w)(w) = f_T(e', w')(w')$. In this case, both $T(e', w)$ and $T(e', w')$ are rooted binary trees with roots $w$ and $w'$, respectively and $|V(T(e', w))| \geq |V(T(e', w'))|$. Since $T(e', w)$ is an rgood binary tree, by Proposition 3.1, $T(e', w) \cong T(e', w')$ and hence $C_d(T) = \{v, v'\} = S_c(T)$. This completes the proof. □

**Corollary 4.6** In a crg tree $T_{rg}^{n,l}$, among center, centroid and subtree core, the center is nearest to the vertex 1.

**Proof** We rename the crg tree $T_{rg}^{n,l}$ as $T$. Let $u \in C(T)$ and $v \in S_c(T)$ such that $d(1, C(T)) = d(1, u)$ and $d(1, S_c) = d(1, v)$. We show that $d(1, v) \geq d(1, u)$. Suppose $d(1, v) < d(1, u)$. Let $w$ be the vertex adjacent to $u$ in the path joining 1 and $u$. Consider the edge $e = \{w, u\}$. Then $v \in V(T(e, u))$. Let $k = |V(T(e, u))|$. As $u \in C(T)$ is the central vertex nearest to 1, by Corollary 2.10, $|V(T(e, u))| \geq k + 2$. Note that the tree $T(e, u)$ is $T_{rg}^{k,2}$. So by Proposition 3.3, $f_T(e, u)(w) < f_T(e, u)(u)$. Hence by Lemma 2.6, $v \in V(T(e, u))$, which is a contradiction. □

Let $l$ be an odd integer and let $r$ be the root of $T_{rg}^{l}$. We denote the number $f_{T_{rg}^l}(r)$ by $R_l$. Note that for the tree $T_{rg}^{n,l}$, if $l > n - l$ then by Corollary 3.4, $R_l > 3 \times 2^{\frac{n-l-1}{2}} - 2$.

**Lemma 4.7** Let $n \geq 12$ be even and let $l$ be the smallest positive odd number such that $R_l > 3 \times 2^{\frac{n-l-1}{2}} - 2$. For an even integer $i$ with $2 \leq i \leq l - 3$, let $e = \{\frac{n-l+3}{2}, \frac{n-l+5}{2}\} \in E(T_l)$ where $T_l = T_{rg}^{n,l-i}$. Then $S_c(T_l) \subseteq V(T_l(e, \frac{n-l+3}{2}))$ or $S_c(T_l) = \{\frac{n-l+3}{2}, \frac{n-l+5}{2}\}.$

**Proof** Since $n \geq 12, l \geq 5$ and $T_l$ is defined for every even integer $i$ with $2 \leq i \leq l - 3$. Also since $l$ is the smallest positive odd integer such that $R_l > 3 \times 2^{\frac{n-l-1}{2}} - 2$, we have $R_{l-i} \leq R_{l-2} \leq 3 \times 2^{\frac{n-l+1}{2}} - 2$.

Let $B^1_l$ and $B^2_l$ be the components of $T_l - e$ containing the vertices $\frac{n-l+3}{2}$ and $\frac{n-l+5}{2}$, respectively. The component $B^2_l$ is a rooted binary tree on $l - 2$ vertices with
Theorem 4.9. Suppose \( R_l \) is an rgood tree if \( i = 2 \). By Proposition 3.1, \( f_{B_l^2}(n-i+5) \leq R_{l-2} \) for \( 2 \leq i \leq l - 3 \). The component \( B_l^1 \) of \( T_i - e \) is a rooted binary tree on \( n - l + 2 \) vertices with root \( n-i+3 \). By Corollary 3.4, \( f_{B_l^1}(n-i+3) = 3 \times 2^{n-i+1} - 2 \).

Thus we have \( f_{B_1^2}(n-l+5) = 3 \times 2^{n-l+1} - 2 \geq R_{l-2} \geq f_{B_l^2}(n-l+5) \) for \( 2 \leq i \leq l - 3 \). Suppose \( S_c(T_i) \neq \{n-l+3, \frac{n-l+5}{2}\} \) then \( f_{B_l^1}(n-l+3) > f_{B_l^2}(n-l+5) \). By Lemma 2.6, we have \( S_c(T_i) \subseteq V(B_i^1) \). This completes the proof. \( \square \)

**Lemma 4.8** Let \( n \geq 12 \) be even and let \( l \) be the smallest positive odd number such that \( R_l > 3 \times 2^{n-l-1} - 2 \). Then \( S_c(T^l) = \{v_j\} \) where \( v_j \) is the root of the rgood part of \( T^l \) for \( 0 \leq j \leq 2 \).

**Proof** Since \( T^l \) is a binary caterpillar and \( n \geq 12 \), we have \( l \geq 7 \). For \( 0 \leq j \leq 2 \), let \( v_j' \) be the vertex in a heavier branch of the rgood part of \( T^l \) with \( j \in \{v_j, v_j'\} \in E(T^l) \). We denote \( T^l \) and \( T^l \) by \( T_j \) and \( T_j' \), respectively. Clearly \( T_j \) and \( T_j' \) are two rooted binary trees with roots \( v_j \) and \( v_j' \), respectively. Since \( R_l > 3 \times 2^{n-l-1} - 2 \), we have \( S_c(T^l) \subseteq \{v_0, v_0'\} \) and hence \( S_c(T^l) \subseteq \{v_j, v_j'\} \). The component \( T_j \) of \( T^l \) is a rooted binary tree with at most \( l - 2 \) vertices. The component \( T_j \) of \( T^l \) is rooted with at least \( n - l + 2 \) vertices. Also \( T_j \) has more than two vertices at level 2. Thus we have,

\[
f_{T_j}(v_j') \leq R_{l-2} \leq 3 \times 2^{n-l+1} - 2 < f_{T_j}(v_j)
\]

for \( 0 \leq j \leq 2 \). Hence by Lemma 2.6, \( S_c(T^l) = \{v_j\} \) for \( 0 \leq j \leq 2 \). \( \square \)

**Theorem 4.9** Let \( n \geq 12 \) be even and let \( l \) be the smallest positive odd number such that \( R_l > 3 \times 2^{n-l-1} - 2 \). Among all crg trees on \( n \) vertices, the distance between center and subtree core is maximized by the tree \( T^l \).

**Proof** Let \( v \) be the root of the rgood part of \( T^l \). Then by Lemma 4.8, \( S_c(T^l) = \{v\} \). Following the numbering of vertices mentioned in Section 1.1, \( v \) is numbered as \( \frac{n-l+3}{2} \) in \( T^l \). Let the vertex numbered as \( u \) be the central vertex of \( T^l \) nearest to the vertex \( \frac{n-l+3}{2} \). Then by Corollary 4.6, \( u \) lies on the path joining 1 and \( \frac{n-l+3}{2} \) and

\[
d_{T^l}(u, \frac{n-l+3}{2}) = d_{T^l}(C, S_c).
\]

Let \( i \) be an even integer with \( 2 \leq i \leq l - 3 \). Consider the crg tree \( T^l \). Since the center of a tree is the same as the center of every longest path in it, the central vertex of \( T^l \) nearest to \( \frac{n-l+3}{2} \) is \( u + k \) for some \( k \geq 0 \). Also by Lemma 4.7, \( S_c(T^l) = \{\frac{n-l+3}{2}, \frac{n-l+5}{2}\} \) or lies on the path from 1 to \( \frac{n-l+3}{2} \). Hence, we have

\[
d_{T^l}(C, S_c) \leq d_{T^l}(u, S_c) \leq d_{T^l}(u, \frac{n-l+3}{2}) = d_{T^l}(C, S_c)
\]

\( \square \) Springer
for $2 \leq i \leq l - 3$.

Consider the sequence of trees $T_{rg}^{n,l+j}$ for $0 \leq j \leq n - l - 1$ with $j$ even. Let $z$ be the vertex in the caterpillar part of $T_{rg}^{n,l}$ such that $e = \{v, z\} \in E(T_{rg}^{n,l})$ (here $v$ is the root of the rgood part of $T_{rg}^{n,l}$). Let $B_1$ and $B_2$ be the components of $T_{rg}^{n,l} - e$ containing $z$ and $v$, respectively. Note that $B_1 \cong T_{rg}^{n-l-j} \leq T_2$ and $B_2 \cong T_{rg}^{l}$. It is given that $f_{T_{rg}}(v) > 3 \times 2^{\frac{n-l-1}{2}} - 2 = f_{B_1}(z)$ and so $f_{T_{rg}}^{n,l+j}(w) > f_{T_{rg}}^{l}(v)$ where $w$ is the root of the rgood part of $T_{rg}^{n,l+j}$. So, by Lemma 2.6 and Lemma 2.9 it follows that $S_C(T_{rg}^{n,l+j}) \leq \{w, w'\}$ where $w'$ is the vertex in a heavier branch with $\{w, w'\} \in E(T_{rg}^{n,l+j})$. In $T_{rg}^{n,l+j}$, $w$ is numbered as $\frac{n-\alpha - j + 3}{2}$. Let $d_T^{n,l}(C, S_C) = \alpha$. We have two cases:

**Case I: $\alpha \geq 1$**

By Lemma 4.8, $d_{T_{rg}}^{n,l+2}(C, S_C) \in \{\alpha, \alpha - 1\}$. Note that there exists a $j$ for which $C(T_{rg}^{n,l+j}) = \{w, w'\}$. Let $j'$ be the smallest positive even integer such that $d_{T_{rg}}^{n,l+j'}(C, S_C) = 0$. Then $d_{T_{rg}}^{n,l+k}(C, S_C) = \{0, 1\}$ for $j' \leq k \leq n - l - 1$ and $d_{T_{rg}}^{n,l+k}(C, S_C) \leq \alpha$ for $0 \leq k \leq j' - 2$. Hence

$$d_{T_{rg}}^{n,l+j}(C, S_C) \leq d_{T_{rg}}^{n,l}(C, S_C)$$

for $2 \leq j \leq n - l - 1$.

**Case II: $\alpha = 0$**

Since $n$ is even, $n = 4k$ or $4k + 2$ for some $k$. So $l \leq 2k + 1$ as $l$ is the smallest positive odd number such that $R_l > 3 \times 2^{\frac{n-l-1}{2}} - 2$. It can be checked that $C(T_{rg}^{14,7}) = \{4\}$ and $S_C(T_{rg}^{14,7}) = \{5\}$. So $d_{T_{rg}}^{14,7}(C, S_C) = 1$ and hence for $n = 4k + 2$ for some $k \geq 3$, $d_{T_{rg}}^{n,l}(C, S_C) \geq 1$. It can also be checked that $C(T_{rg}^{20,11}) = \{5\}$ and $S_C(T_{rg}^{20,11}) = \{6\}$. So $d_{T_{rg}}^{20,11}(C, S_C) = 1$ and hence for $n = 4k$ for some $k \geq 5$, $d_{T_{rg}}^{n,l}(C, S_C) \geq 1$.

If $n = 12$ then $l = 7$ and it can be easily checked that $d_{T_{rg}}^{12,7}(C, S_C) = 0$. If $n = 16$ then $l = 9$ and it can be checked that $d_{T_{rg}}^{16,9}(C, S_C) = d_{T_{rg}}^{16,11}(C, S_C) = d_{T_{rg}}^{16,13}(C, S_C) = d_{T_{rg}}^{16,15}(C, S_C) = 0$. Hence if $d_{T_{rg}}^{n,l}(C, S_C) = 0$ then $d_{T_{rg}}^{n,l+j}(C, S_C) = 0$ for $2 \leq j \leq n - l - 1$. This completes the proof.

**Corollary 4.10** Let $T$ be a binary tree on $n \geq 12$ vertices. Let $r$ be the root of the rooted binary tree $T_{rg}^{l}$, and let $l$ be the smallest positive integer such that $f_{T_{rg}}(r) > 3 \times 2^{\frac{n-l-1}{2}} - 2$. Then

$$d_T(C, S_C) \leq d_{T_{rg}}^{n,l}(C, S_C).$$

**4.3 Centroid and Subtree Core**

**Theorem 4.11** Among all binary trees on $n$ vertices, the distance between centroid and subtree core is maximized by a crg tree.
Proof Consider a binary tree $T$ on $n$ vertices with $d_T(C_d, S_c) \geq 1$. Our aim is to construct a crg tree $\tilde{T} \in \Omega_n$ such that $d_T(C_d, S_c) \geq d_T(C_d, S_c)$. Let $u \in C_d(T)$ and $v \in S_c(T)$ such that $d_T(C_d, S_c) = d(u, v)$. Let $u'$ and $v'$ be the vertices adjacent to $u$ and $v$, respectively and lie on the path joining $u$ and $v$. Let $e^1 = \{u', u''\}$, $e^2 = \{v', v''\} \in E(T)$.

Let $|V(T(e^2, v))| = k$. The component $T(e^2, v)$ is a rooted binary tree with root $v$. Since $S_c(T) \subseteq V(T(e^2, v))$, by Lemma 2.6 we have $f_{T(e^2, v)}(v) > f_{T(e^2, v)}(v')$. If $T(e^2, v)$ is an rgd binary tree then rename the tree $T$ by $T'$. Otherwise, form a new tree $T'$ from $T$ by replacing the component $T(e^2, v)$ with $T^k$ rooted at $v$. By Proposition 3.1, $f_{T^k}(v) \geq f_{T(e^2, v)}(v) > f_{T(e^2, v)}(v')$ and hence by Lemma 2.6, $S_c(T') \subseteq V(T'(e^2, v))$. Since $C_d(T') \subseteq V(T(e^1, u))$, by Lemma 2.3 we have $|V(T(e^1, u))| > |V(T(e^1, u'))|$. As $d_T(C_d, S_c) \geq 1$ and $|V(T(e^2, v))| = |V(T'(e^2, v))|$, $C_d(T') = C_d(T')$. Hence, $d_{T'}(C_d, S_c) \geq d_T(C_d, S_c)$.

If $T' \in \Omega_n$ then the result follows. Otherwise let $|V(T'(e^2, v'))| = l$. Construct a new tree $T''$ from $T'$ by replacing $T'(e^2, v')$ with $T^k_{r,2}$ rooted at $v'$. Observe that $T'' \in \Omega_n$. By Proposition 3.3, $f_{T^k_{r,2}}(v') < f_{T'(e^2, v')}(v')$. So $f_{T''(e^2, v')}(v') > f_{T'(e^2, v')}(v')$ and hence by Lemma 2.6, $S_c(T'') \subseteq V(T''(e^2, v'))$. Also we can construct $T''$ from $T'$ stepwise such that in each step the centroid is the same as the centroid of $T'$ or moves away from $v$. For that choose a longest path $P$ starting from $v'$ containing the centroid of $T'$ in the rooted binary tree $T'(e^2, v')$ with root $v'$. Let $x$ be the end point of the path $P$. Delete two pendant vertices with the same parent, where the parent is not on the path $P$ and add them as pendant vertices at $x$ to get a new tree $T'_1$. If $T'_1$ is not isomorphic to $T''$ then take $T'_1$ as $T'$ and continue this process till $T''(e^2, v')$ becomes the tree $T^k_{r,2}$ and we reach the tree $T''$. In each step of the process, the centroid is either same as the centroid of the tree in the previous step or moves away from $v$. Hence, $d_{T''}(C_d, S_c) \geq d_{T'}(C_d, S_c) \geq d_T(C_d, S_c)$. This completes the proof. \hfill \Box

Theorem 4.12 Let $T$ be a binary tree on $n \geq 12$ vertices. Let $r$ be the root of the rooted binary tree $T^l_{rg}$ and let $l$ be the smallest positive integer such that $f_{T^l_{rg}}(r) > 3 \times 2^{n-l-1} - 2$. Then

$$d_T(C_d, S_c) \leq \begin{cases} d_{T^l_{rg}}(C_d, S_c) & \text{if } d_{T^l_{rg}}(C_d, S_c) \geq 1 \\ 1 & \text{otherwise.} \end{cases}$$

Proof Since $n$ is even, $n = 4k$ or $4k + 2$ for some $k$. So $l \leq 2k + 1$ as $l$ is the smallest positive odd number such that $R_l > 3 \times 2^{n-l-1} - 2$. Then $C_d(T^{n,l}_{rg})$ lies on the path from $1$ to $\frac{n-l+3}{2}$. Let the vertex numbered $u$ be the centroid vertex of $T^{n,l}_{rg}$ nearest to the vertex $\frac{n-l+3}{2}$. Then by Lemma 4.8,

$$d_{T^{n,l}_{rg}}(u, \frac{n-l+3}{2}) = d_{T^{n,l}_{rg}}(C_d, S_c).$$

Let $i$ be an even integer with $2 \leq i \leq l-3$. Consider the crg tree $T^{n,l-i}_{rg}$. Then the centroid vertex of $T^{n,l-i}_{rg}$ nearest to $\frac{n-l+3}{2}$ is $u$ for $2 \leq i \leq l-3$. Also by Lemma 4.7,
$S_c(T_{rg}^{n,l+i}) = \{ \frac{n-l+3}{2}, \frac{n-l+5}{2} \}$ or lies on the path from 1 to $\frac{n-l+3}{2}$. So we have

$$d_{T_{rg}^{n,l+i}}(C_d, S_c) = d_{T_{rg}^{n,l+i}}(u, S_c) \leq d_{T_{rg}^{n,l+i}}(u, \frac{n-l+3}{2}) = d_{T_{rg}^{n,l}}(C_d, S_c)$$

for $2 \leq i \leq l-3$.

Consider the sequence of trees $T_{rg}^{n,l+j}$ for $0 \leq j \leq n-l-1$ with $j$ even. Then by Lemma 2.6 and Lemma 2.9, $S_c(T_{rg}^{n,l+j}) \subseteq \{w, w'\}$ for $0 \leq j \leq n-l-1$ where $w$ is the root of the rgood part of $T_{rg}^{n,l+j}$ and $w'$ is the vertex in a heavier branch with $e = \{w, w'\} \in E(T_{rg}^{n,l+j})$ (as discussed in the proof of Theorem 4.9). In $T_{rg}^{n,l+j}$, $w$ is numbered as $\frac{n-l+j}{2}$.

Let $\alpha$. We have two cases:

Case I: $\alpha \geq 1$

Let $j'$ be the smallest positive even integer such that $d_{T_{rg}^{n,l+j'}}(C_d, S_c) = 0$. Then $d_{T_{rg}^{n,l+k}}(C_d, S_c) \in \{0, 1\}$ for $j' \leq k \leq n-l-1$ and $d_{T_{rg}^{n,l+k}}(C_d, S_c) \leq \alpha$ for $0 \leq k \leq j'-2$. Hence

$$d_{T_{rg}^{n,l+j}}(C_d, S_c) \leq d_{T_{rg}^{n,l}}(C_d, S_c)$$

for $2 \leq j \leq n-l-1$.

Case II: $\alpha = 0$

In this case $d_{T_{rg}^{n,l+j}}(C_d, S_c) \in \{0, 1\}$, for $2 \leq j \leq n-l-1$.

Hence the result follows from Theorem 4.11.  

Acknowledgements The authors wish to sincerely thank the referees for their detailed comments which significantly improved the submitted version of the paper.

Funding The author Dinesh Pandey is supported by UGC Fellowship scheme (Sr. No. 2061641145), Government of India.

Declarations

Conflict of interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

References

1. Abrew, N., Fritscher, E., Justel, C., Kirkland, S.: On the characteristic set, centroid and center for a tree. Linear Multilinear Algebra 65(10), 2046–2063 (2017)
2. Aho, A.V., Sloane, N.J.A.: Some doubly exponential sequences. Fibonacci Quart. 11(4), 429–437 (1973)
3. Desai, D.N.S., Patra, K.L.: Maximizing distance between center, centroid and subtree core of trees. Proc. Indian Acad. Sci. Math. Sci. 129, 7–8 (2019)
4. Goldman, A.J.: Optimal center location in simple networks. Transp. Sci. 5, 212–221 (1971)
5. Harary, F.: Graph theory. Addison-Wesley Publishing Co, Reading (1969)
6. Jordan, C.: Sur les assemblages de lignes. J. Reine Angew. Math. 70, 185–190 (1869)
7. Kang, A., Ault, D.: Some properties of a centroid of a free tree. Inf. Process. Lett. 4(1), 18–20 (1975)
8. Merris, R.: Characteristic vertices of trees. Linear Multilinear Algebra 22(2), 115–131 (1987)
9. Mitchel, S.L.: Another characterization of the centroid of a tree. Discrete Math. 24(3), 277–280 (1978)
10. Patra, K.L.: Maximizing the distance between center, centroid and characteristic set of a tree. Linear Multilinear Algebra 55(4), 381–397 (2007)
11. Smith, H., Székely, L., Wang, H., Yuan, S.: On different middle parts of a tree. Electron. J. Combin. 25(3), 32 (2018)
12. Székely, L.A., Wang, H.: On subtrees of trees. Adv. in Appl. Math. 34(1), 138–155 (2005)
13. Zelinka, B.: Median and Peripherian of trees. Arch. Math. 4, 87–95 (1968)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.