Algorithmic Operations Research Vol.5 (2010) 105–118

Some Structural Properties of a Least Central Subtree of a Tree

Martti Hamina and Matti Peltola

Mathematics Division, Faculty of Technology, University of Oulu, PB 4500, 90014 Oulu, Finland

Abstract

We consider the graph center problem in the joinsemilattice $L(T)$ of all subtrees of a tree T. A subtree S of a tree T is a central subtree of T if S has the minimum eccentricity in the joinsemilattice. The graph center of the joinsemilattice is the set of all central subtrees. A central subtree with the minimum number of points is a least central subtree of a tree T. Thus least central subtrees of T are, in some sense, the best possible connected substructures of T among all connected substructures. We show that every tree is a unique least central subtree of some larger tree. Our main result points out the importance of the cardinality of the nodes of degree two. Low cardinality guarantees uniqueness and explicit construction for the least central subtree.

Key words: Joinsemilattice of subtrees, least central subtree, center of tree.

1. Introduction

The middle part of a graph has important applications in transportation, facility planning and location problems. Much research has been devoted to define that middle part of a tree. The most common centrality concepts are the center (points with minimum eccentricity), the centroid (points where maximum branchweight attains minimum value) and the path center (path with minimum eccentricity). Here we consider another centrality concept, the subtree center of a tree. This concept does not restrict the structure of the middle part of a tree. It can be a point or a path or some other kind of subtree such that the subtree is the most central when compared with all subtrees of the tree.

For every tree T there is a joinsemilattice $L(T)$ of subtrees of T, where the meet $S_1 \wedge S_2$ of subtrees S_1 and S_2 equals the subtree induced by the intersection of the point sets of S_1 and S_2 whenever the intersection is nonempty. The join $S_1 \vee S_2$ of subtrees S_1 and S_2 is the least subtree of T containing the subtrees S_1 and S_2. Note that the empty graph is not a subtree of T, and thus, in general, there is no least element in $L(T)$. The distance in the joinsemilattice $L(T)$ is the same as the distance in the undirected Hasse diagram graph G_L of $L(T)$. A subtree S of a tree T is a central subtree if S has the minimum eccentricity in the joinsemilattice $L(T)$. The graph center of the joinsemilattice is the set of all central subtrees. A central subtree with the minimum number of points is a least central subtree of a tree T. A more detailed discussion with appropriate references is given in our article [3].

A tree may have several least central subtrees. We are able to prove uniqueness of the least central subtree for certain tree classes e.g. for caterpillars and homeomorphically irreducible trees. Furthermore, we present estimates for the size and $L-$eccentricity of least central subtrees.

In the previous work [8] Nieminen and Peltola proved that the leaves of the tree cannot be points of any least central subtree. In addition they proved that the intersection of any two least central subtrees is nonempty. Recently Hamina and Peltola [3] improved the latter result by proving that any least central subtree contains the center of the tree and at least one point of the centroid of the tree. Thus any least central subtree divides the tree into the middle part and the peripheral part. In the case of multiple least central subtrees there exists a transition region between the middle part and the peripheral part.

There are many applications which can be expressed as a graph optimization problem of the type

$$W = \sum_{e \in E(T)} w(e) = \min!$$

where T is a spanning tree of the graph G and $w(e)$ is

Email: Martti Hamina [martti.hamina@ee.oulu.fi], Matti Peltola [matti.peltola@ee.oulu.fi].

the cost for edge e. The understanding of the behaviour of the subtree center may be useful for the choice of appropriate spanning trees.

2. Some results on least central subtrees

In this section we collect some basic results for least central subtrees from articles [8] and [3]. For any subtree S we denote by $|S|$ the number of points in the subtree. Figure 1 shows the construction of the joinsemilattice and the distance in the Hasse diagram graph G_L. We have drawn all subtrees of a tree with six points and the corresponding joinsemilattice of subtrees. The center (two peripheries), the centroid (lightgray shading) and the least central subtree (octagonal, gray shading) are marked in the joinsemilattice.

The graph center of the joinsemilattice graph is the set of all central subtrees of the tree T. Our example in Figure 1 is very simple. The joinsemilattice center consists of one subtree. This subtree is the least central subtree too. In general, the structure of the joinsemilattice and the joinsemilattice center is not that simple.

It is well known that for all trees T of size $|T| = n$ paths have the least number of subtrees and stars have the largest number of subtrees. Thus $\frac{1}{2}n(n+1) \leq |G_L| \leq 2^{n-1} + n - 1$; see [11]. In the case of stars almost all points of the graph G_L are points of the graph center $C(G_L)$. Only leaves v_i and complements of leaves $T \setminus v_i$ are excluded from the joinsemilattice center. Even the star itself is a point of the joinsemilattice center. Thus among all trees of size $|T| = n$, the cardinality of the underlying joinsemilattice graph center varies considerably: $1 \leq |C(G_L)| \leq 2^{n-1} - n + 1$. For all paths the repeated procedure of stripping away leaves gives all central subtrees. Again the path itself is included into the joinsemilattice center. Joinsemilattices of paths show that the graph center $C(G_L)$ need not be a connected subgraph of the joinsemilattice graph G_L.

Thus the graph center of G_L (the set of central subtrees) is too large for effective treatment of our problem. Therefore we have adopted another optimization criterion. Among all subtrees lying in the joinsemilattice center, the best is the one with minimal size. That is our least central subtree. For paths and stars this modification resolves the problem. In both cases the least central subtree is unique and coincides with the center of the tree.

Let S_1 and S_2 be subtrees of a tree T. Let $d_L(S_1, S_2)$ denote the distance between subtrees S_1 and S_2 in the joinsemilattice graph G_L. The $L-$eccentricity of the subtree S_1 is $e_L(S_1) = \max\{d_L(S_1, S) \mid S \text{ is a subtree of } T\}$. Clearly $e_L(S_1)$ equals the eccentricity of the subtree S_1 in the joinsemilattice graph G_L. The subtree S_1 is a central subtree of a tree T if it has the minimum $L-$eccentricity. Least central subtrees C_L of a tree T are solutions of the discrete optimization problem

$$e_L(C_L) = \min_{S_1 \subset T} \max\{d_L(S_1, S) \mid S \text{ is a subtree of } T\}$$

subject to the additional constraint that among all subtrees satisfying the minimax criteria, only those subtrees which are minimal in size, are selected. Thus least central subtrees of T are, in some sense, the best possible connected substructures of T among all connected substructures. Note that the definition of least central subtrees in terms of the given optimization problem can be interpreted as a property of the tree itself. The construction of the underlying undirected Hasse diagram graph is based purely on the neigbourhood relation between subtrees of a tree.

Moreover, the following lemma ([8], Lemma 1) shows how to define the distance $d_L(S_1, S_2)$ by using the properties of the tree T. In particular, there is no need to construct the joinsemilattice which is of much higher cardinality than the underlying tree.

Lemma 1 *Let G_L be the semilattice graph of all subtrees of a tree T, and S_1 and S_2 be two subtrees of T. Then the distance between S_1 and S_2 in G_L is*

$$d_L(S_1, S_2) = 2|S_1 \vee S_2| - |S_1| - |S_2|$$

$$= \begin{cases} |S_1| + |S_2| + 2(d_T(S_1, S_2) - 1), & \text{if } S_1 \cap S_2 = \emptyset \\ |S_1| + |S_2| - 2|S_1 \cap S_2|, & \text{if } S_1 \cap S_2 \neq \emptyset. \end{cases}$$

Proof. The basic observation is that the existence of a line (S_1, S_2) in G_L implies that S_1 is obtained from S_2 by adding/removing a point. Thus $d_L(S_1 \vee S_2, S_i) = |S_1 \vee S_2| - |S_i|$, $i = 1, 2$. Because of the median algebra property of $L(T)$, a shortest path $S_1 - S_2$ goes through $S_1 \vee S_2$. This implies that $d_L(S_1, S_2) = 2|S_1 \vee S_2| - |S_1| - |S_2|$.

If $S_1 \cap S_2 = \emptyset$ then $d_T(S_1, S_2) \geq 1$ and the number of points on the geodesic $S_1 - S_2$ is $d_T(S_1, S_2) - 1$. Thus $|S_1 \vee S_2| = |S_1| + |S_2| + d_T(S_1, S_2) - 1$.

If $S_1 \cap S_2 \neq \emptyset$ then $S_1 \wedge S_2$ exists and $|S_1 \vee S_2| = |S_1| + |S_2| - |S_1 \cap S_2|$. The proof follows by combining these facts. □

Nieminen and Peltola proved the following theorem in the paper [8].

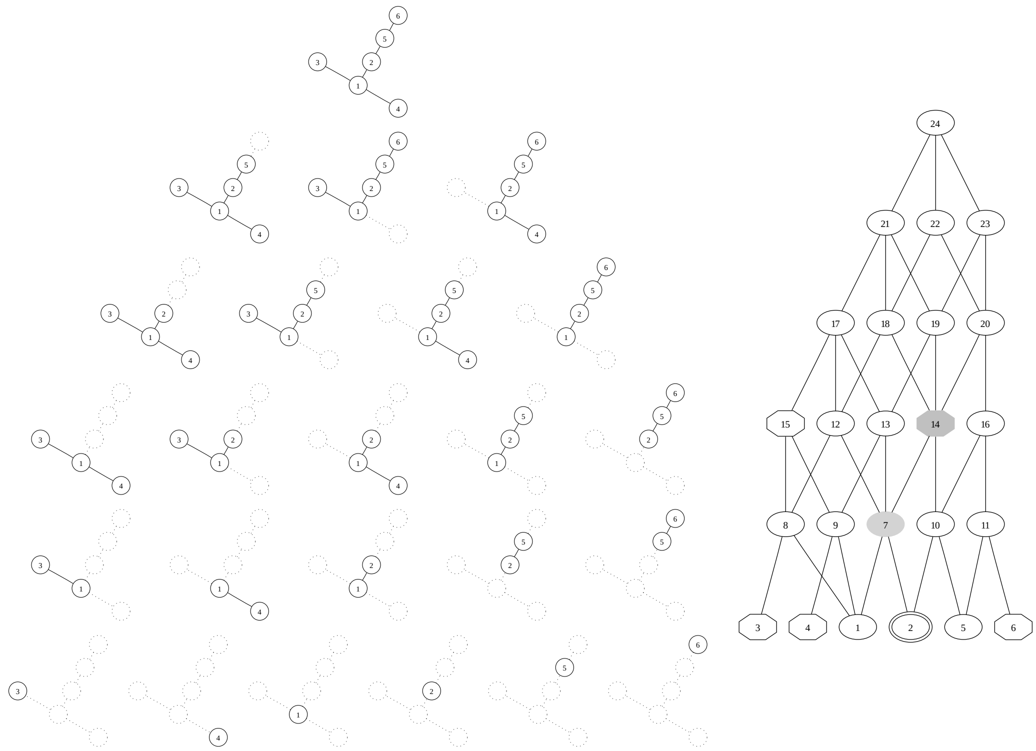

Fig. 1. The set of all subtrees of a tree and the corresponding joinsemilattice.

Theorem 1 *If C_L is any least central subtree of tree T, then the subtree C_L does not contain any endpoint of T. Furthermore, any two least central subtrees have a nonempty intersection.*

We remark that Theorem 1 implies that the intersection of an arbitrary number of least central subtrees is nonempty. This is a consequence of the acyclicity of a tree.

Recently Hamina and Peltola [3] improved the result given in Theorem 1.

Theorem 2 *The center of a tree is a subtree of every least central subtree. Any least central subtree of a tree T contains a point of the centroid. Any least central subtree contains the center, at least one point of the centroid, and the path from center to centroid.*

The proof is given in our article [3]. We remark that these two theorems give a good starting point for writing a practical algorithm. An outline of the algorithm is given in section 5.

In the following Theorem we prove that every tree T is a unique least central subtree of some larger tree.

Theorem 3 *For any tree T_0 there exists a tree T such that T_0 is a unique least central subtree of T with $L-$eccentricity $e_L(T_0) = |T_0| + 1$.*

Proof. We may assume that $|T_0| \geq 2$. Let T be a tree obtained from T_0 by inserting one leaf on every point of T_0. Thus $|T| = 2|T_0|$. Since the least central subtree cannot contain leaves then every least central subtree of T is a subtree of T_0. For every subtree T_s of T_0 such that $|T_s| \leq |T_0| - 2$, we have $e_L(T_s) \geq d_L(T_s, T) = |T| - |T_s| \geq |T| - (|T_0| - 2) = |T_0| + 2$. If T_s is a subtree of T_0 such that $|T_s| = |T_0| - 1$, then $T_s = T_0 \setminus \{v_1\}$, where the point v_1 is a leaf of T_0. Let u_1 be a leaf of a tree T such that u_1 is a neighbour of v_1 and $u_1 \notin T_0$. Then $d_T(T_s, u_1) = 2$ and

$$\begin{aligned} d_L(T_s, \{u_1\}) &= |T_s| + 1 + 2(d_T(T_s, u_1) - 1) \\ &= |T_s| + 3 = |T_0| + 2. \end{aligned}$$

Thus for every subtree T_s of T_0 we have $e_L(T_s) \geq |T_0| + 2$. Since $d_L(T_0, T) = |T_0|$, it suffices to prove that $d_L(T_0, S) \leq |T_0| + 1$ for every subtree S of T, $S \neq T$. Two cases arise (i) $S \cap T_0 = \emptyset$ and (ii) $S \cap T_0 \neq \emptyset$.
(i) Since $T \setminus T_0$ contains only leaves, S is a tree of one point, $d_T(T_0, S) = 1$ and

$$d_L(T_0, S) = |T_0| + |S| + 2(d_T(T_0, S) - 1) = |T_0| + 1.$$

(ii) Assume first that there exists a point $w_1 \in T_0 \setminus S$ (otherwise $S \cap T_0 = T_0$). Since $S \cap T_0 \neq \emptyset$, we may assume that w_1 is a neighbour of S. Let v_1 be a leaf of T such that v_1 is a neighbour of w_1. Let S_1 be a subtree induced by S and points w_1 and v_1. Clearly $S_1 \cap T_0 = (S \cap T_0) \cup \{w_1\} \neq \emptyset$ and $|S_1| = |S| + 2$. Furthermore,

$$d_L(T_0, S_1) = |S_1| + |T_0| - 2|S_1 \cap T_0| = |S| + 2 + |T_0| - 2|S \cap T_0| - 2 = d_L(T_0, S).$$

By repeating the process a finite number of times we obtain a sequence of subtrees $S, S_1, S_2, \ldots, S_n$ such that $d_L(T_0, S_i) = d_L(T_0, S)$, $i = 1, \ldots, n$. The process ends, since finally $S_n \cap T_0 = T_0$. In the case $S_n \cap T_0 = T_0$, we have

$$d_L(T_0, S_n) = |T_0| + |S_n| - 2|T_0 \cap S_n| = |S_n| - |T_0| \leq |T| - |T_0| = |T_0|.$$

The proof is complete. □

3. Some useful tree classes

In this section we recall some special tree classes and present useful general estimates for trees. The reader may find the references concerning tree enumeration [5], [**?**], [6], [9] and [10] interesting. Let $n_k(T)$ be the number of nodes of degree k. In particular $n_1(T)$ is the number of leaves of T. A **caterpillar** is a tree for which the points that are not leaves induce a path. See [1], [6] for more information. Let T' be the subtree obtained by removing all leaves of T. For caterpillars the tree T' is a path. Let $\deg v$ be the **degree** of the point v and let $\Delta = \max_{v \in T} \deg v$ be the maximum degree. It is well known that the degree sum satisfies $\sum_{v \in T} \deg v = 2(|T| - 1)$.

A tree is called **starlike** (c.f. spider) if exactly one point of the tree has degree greater than two. Let P_n denote the path on n points. By $T_{k_1,k_2,\ldots,k_s}$ we denote the starlike tree which has a point v_0 of degree s and which has the property that the graph $T_{k_1,k_2,\ldots,k_s} \setminus \{v_0\}$ is a forest of paths $P_{k_1}, P_{k_2}, \ldots, P_{k_s}$. Thus $|T_{k_1,k_2,\ldots,k_s}| = k_1 + k_2 + \ldots + k_s + 1$.

A tree T is a **Cayley tree of degree** n if each non-leaf point has a constant number n of branches. A tree is called **homeomorphically irreducible** if there are no points of degree 2. Every Cayley tree of degree n, with $n \geq 3$ is homeomorphically irreducible. We say that a tree is **almost homeomorphically irreducible** if there is exactly one point of degree 2. In the following lemma we have collected some useful, probably well known, estimates for trees. See [1] p. 106 for some references. Note the importance of the cardinality of nodes of degree 2. For the convenience of the reader we give the proofs.

Lemma 2 *For all trees* $|T'| = |T| - n_1(T)$ *and* $\operatorname{diam} T - 1 \leq |T'| \leq n_1(T) + n_2(T) - 2$. *Moreover, the estimates*

$$\begin{aligned}
\Delta + \operatorname{diam} T &\leq n_1(T) + \operatorname{diam} T \leq |T| + 1,\\
|T| + 1 &\leq 2n_1(T) + n_2(T) - 1,\\
\operatorname{diam} T &\leq n_1(T) + n_2(T) - 1,\\
2\operatorname{diam} T &\leq |T| + n_2(T),
\end{aligned}$$

are true.

Proof. Let P be any diametral path. Then P contains exactly two leaves and $|P| = \operatorname{diam} T + 1 \leq |T| - n_1(T) + 2$. This implies the claim $n_1(T) + \operatorname{diam} T \leq |T| + 1$. The general formula for the number of leaves of a tree is

$$n_1(T) = 2 + \sum_{k=3}^{\Delta} (k-2) n_k(T).$$

Furthermore, there exists at least one node of maximum degree. Thus we obtain the lower bound

$$n_1(T) = 2 + \sum_{k=3}^{\Delta} (k-2) n_k(T) \geq 2 + (\Delta - 2) n_\Delta(T) \geq 2 + (\Delta - 2) = \Delta.$$

The proof is complete for the first estimate. We have for the size of the subtree T' the following formulas

$$\begin{aligned}
|T'| &= n_2(T) + \sum_{k=3}^{\Delta} n_k(T),\\
|T'| &= |T| - n_1(T) = |T| - 2 - \sum_{k=3}^{\Delta} (k-2) n_k(T).
\end{aligned}$$

These imply the equality

$$n_2(T) + \sum_{k=3}^{\Delta} n_k(T) = |T| - 2 - \sum_{k=3}^{\Delta} (k-2) n_k(T)$$

which yields by solving with respect to $|T|$ the following result

$$|T| = 2 + n_2(T) + \sum_{k=3}^{\Delta} (k-1) n_k(T).$$

Here we get easily the lower bound

$$\begin{aligned}
|T| &\geq 2 + n_2(T) + \sum_{k=3}^{\Delta} 2\, n_k(T)\\
&= 2 + n_2(T) + 2 \sum_{k=3}^{\Delta} n_k(T)\\
&= 2 + n_2(T) + 2(|T| - n_1(T) - n_2(T))\\
&= 2 + 2|T| - 2n_1(T) - n_2(T).
\end{aligned}$$

This gives $|T| + 2 \leq 2\,n_1(T) + n_2(T)$ proving the second inequality. The first and second estimate imply

$$n_1(T) + \operatorname{diam} T \leq |T| + 1 \leq 2\,n_1(T) + n_2(T) - 1$$

which proves the third result. Finally, we obtain by using third and first inequality

$$2 \operatorname{diam} T - n_2(T) \leq \operatorname{diam} T + n_1(T) - 1 \leq |T|$$

proving the last estimate. The bounds for the size of the subtree T' follow from the previous estimates. □

The following estimates define a method for describing the starlike property and caterpillar property of a tree

$$\begin{aligned} &0 \leq |T| - n_1(T) - n_2(T) \leq n_1(T) - 2, \\ &0 \leq |T| - n_1(T) - \operatorname{diam} T + 1 + \min(1, n_1(T) - 2). \end{aligned}$$

Starlike trees satisfy $|T| - n_1(T) - n_2(T) - 1 = 0$. Caterpillar trees (including paths) are extremal in the sense that $|T| + 1 - n_1(T) - \operatorname{diam} T = 0$.

We can give more specific results for trees with low cardinality of nodes of degree two.

Lemma 3 *Let T be a homeomorphically irreducible tree and let $v \in T'$. Then $|T'| \leq n_1(T) - 2$ and for any branch B_v at v we have $|B_v \cap T'| \leq |B_v \setminus T'|$. Moreover, we have the estimates*

$$\begin{aligned} n_1(T) &\geq \operatorname{diam} T + 1, \\ n_1(T) &\geq \tfrac{|T|}{2} + 1, \\ \operatorname{diam} T &\leq \tfrac{|T|}{2}. \end{aligned}$$

Let T be any almost homeomorphically irreducible tree. Let $x \in T'$ be the point of degree 2 and let $v \in T'$, $v \neq x$. Then $|T'| \leq n_1(T) - 1$ and for any branch B_v at v not containing x we have $|B_v \cap T'| \leq |B_v \setminus T'|$. For branches B_v at v containing x we have $|B_v \cap T'| \leq |B_v \setminus T'| + 1$. Finally for the branches at x we have $|B_x \cap T'| \leq |B_x \setminus T'|$.

Proof. The estimates follow from the results of Lemma 2 by substituting $n_2(T) = 0$. In particular, we have for any irreducible tree

$$|T'| \leq n_1(T) - 2. \qquad (*)$$

Let $v \in T'$ and let B_v be any branch at v. Clearly, the subtree B_v is homeomorphically irreducible and $|B_v| = |B_v \cap T'| + |B_v \setminus T'|$. Furthermore we have

$$n_1(B_v) = |B_v \setminus T'| + 1.$$

Denoting by B'_v the subtree of the branch B_v obtained by removing the leaves (of B_v) we have by (*)

$$|B_v \cap T'| - 1 = |B'_v| \leq n_1(B_v) - 2 = |B_v \setminus T'| - 1.$$

This yields the estimate $|B_v \cap T'| \leq |B_v \setminus T'|$. The proof for almost irreducible trees is similar. □

Theorem 4 *For any caterpillar tree the least central subtree is unique.*

Proof. Assume to the contrary that there exists a caterpillar T which contains two least central subtrees C_L and C'_L such that $C_L \cap C'_L \neq \emptyset$. Moreover $C_L \cup C'_L$ is a subpath of any diametral path of T. Clearly $|C_L \setminus C'_L| = |C'_L \setminus C_L|$ and we have $d_L(C_L, C'_L) = 2|C_L \setminus C'_L| = 2|C'_L \setminus C_L| \geq 2$. We consider first the case $|C'_L \setminus C_L| \geq 2$ and prove the existence of a least central subtree $C''_L = (C_L \setminus \{u\}) \cup \{v\}$.

Assume that for any $u \in C_L \setminus C'_L$ and for any $v \in C'_L \setminus C_L$ there exists a subtree S of T such that $d_L(C''_L, S) > e_L(C_L)$. We may assume that $u \in C_L \setminus C'_L$ is a leaf of C_L and since $C_L \cap C'_L \neq \emptyset$, we may assume that v is a neighbour of C_L. Now $|C''_L| = |C_L| = |C'_L|$ and thus $d_L(C''_L, T) = d_L(C_L, T) \leq e_L(C_L)$ implying $S \neq T$. For a contradiction it suffices to prove that $d_L(C_L, S) \geq d_L(C''_L, S) = e_L(C''_L)$ or $d_L(C'_L, S) \geq d_L(C''_L, S) = e_L(C''_L)$. Two cases arise, either (1) $S \cap C''_L = \emptyset$ or (2) $S \cap C''_L \neq \emptyset$.

(1) If $S \cap C''_L = \emptyset$, then $d_L(C''_L, S) = |C''_L| + |S| + 2(d_T(C''_L, S) - 1)$. If u is on the $S - C''_L$ geodesic, then $C_L \cap S \subseteq \{u\}$ and since $u \notin C'_L$, and $S \cap C''_L = \emptyset$ we have $S \cap C'_L = \emptyset$ and $d_T(C'_L, S) \geq d_T(C''_L, S)$. Then

$$\begin{aligned} d_L(C'_L, S) &= |C'_L| + |S| + 2(d_T(C'_L, S) - 1) \\ &\geq |C''_L| + |S| + 2(d_T(C''_L, S) - 1) \\ &= d_L(C''_L, S), \end{aligned}$$

which is a contradiction. If $u \notin S$ and u is not on the $C''_L - S$ geodesic, then $S \cap C_L = \emptyset$, $d_T(C_L, S) \geq d_T(C''_L, S)$ and thus

$$\begin{aligned} d_L(C_L, S) &= |C_L| + |S| + 2(d_T(C_L, S) - 1) \\ &\geq |C''_L| + |S| + 2(d_T(C''_L, S) - 1) \\ &= d_L(C''_L, S), \end{aligned}$$

which is a contradiction.

(2) Assume that $S \cap C''_L \neq \emptyset$. Then $e_L(C''_L) = d_L(C''_L, S) = |C''_L| + |S| - 2|C''_L \cap S|$. Four subcases arise.

(2.1) If $u, v \in S$, then $|C_L \cap S| = |((C''_L \cap S) \cup \{u\}) \setminus \{v\}| = |C''_L \cap S| > 0$. Thus

$$\begin{aligned} d_L(C_L, S) &= |C_L| + |S| - 2|C_L \cap S| \\ &= |C''_L| + |S| - 2|C''_L \cap S| = d_L(C''_L, S), \end{aligned}$$

which is a contradiction.

(2.2) If $u \notin S$ and $v \notin S$, then $C_L \cap S = C''_L \cap S \neq \emptyset$ and

$$\begin{aligned} d_L(C_L,S) &= |C_L|+|S|-2|C_L\cap S| \\ &= |C''_L|+|S|-2|C''_L\cap S| = d_L(C''_L,S), \end{aligned}$$

which is again a contradiction.

(2.3) If $u\notin S$ and $v\in S$, then two cases arise. If $S\cap C_L=\emptyset$, then, since v is a neighbour of C_L, we have $d_T(C_L,S)=1$ and

$$\begin{aligned} d_L(C_L,S) &= |C_L|+|S|+2(d_T(C_L,S)-1) \\ &= |C''_L|+|S| \geq |C''_L|+|S|-2|C''_L\cap S| \\ &= d_L(C''_L,S) \end{aligned}$$

which is a contradiction. If $S\cap C_L\neq\emptyset$ then $C_L\cap S=(C''_L\cap S)\setminus\{v\}$, $|C_L\cap S|=|C''_L\cap S|-1$ and thus

$$\begin{aligned} d_L(C_L,S) &= |C_L|+|S|-2|C_L\cap S| \\ &= |C''_L|+|S|-2(|C''_L\cap S|-1) \\ &= |C''_L|+|S|-2|C''_L\cap S|+2 \\ &= d_L(C''_L,S)+2, \end{aligned}$$

which is a contradiction.

(2.4) Assume that $u\in S$ and $v\notin S$. If $C'_L\cap S=\emptyset$, then $d_T(C'_L,S)\geq 1$ and

$$\begin{aligned} d_L(C'_L,S) &= |C'_L|+|S|+2(d_T(C'_L,S)-1) \\ &\geq |C''_L|+|S| \\ &\geq |C''_L|+|S|-2|C''_L\cap S| = d_L(C''_L,S), \end{aligned}$$

which is a contradiction. Thus we may assume that $C'_L\cap S\neq\emptyset$. Note that $|C''_L\cap S|=|C_L\cap S|-1$. Since $C_L\cup C'_L$ is a path and $u\in S$, we have $(C_L\setminus C'_L)\cap S=C_L\setminus C'_L$. Now we obtain

$$\begin{aligned} d_L(C'_L,S) &= |C'_L|+|S|-2|C'_L\cap S| \\ &= |C''_L|+|S|-2|C''_L\cap S|+2(|C''_L\cap S| \\ &\quad -|C'_L\cap S|) \\ &= d_L(C''_L,S)+2(|C_L\cap S|-|C'_L\cap S|-1) \\ &= d_L(C''_L,S)+2(|(C_L\setminus C'_L)\cap S|-1) \\ &= d_L(C''_L,S)+2(|C_L\setminus C'_L|-1). \end{aligned}$$

This is a contradiction for $|C_L\setminus C'_L|\geq 2$. By (1) and (2) we may assume that C_L and C'_L differ only one point. Let u and u' be the only point of $C_L\setminus C'_L$ and $C'_L\setminus C_L$ respectively. Thus $C_L\cup C'_L=C_L\cup\{u'\}=C'_L\cup\{u\}$ and $C_L\cup C'_L$ is a path with two leaves u and u'. Two cases arise (3) $e_L(C_L)=e_L(C'_L)=|T|-|C_L|$ and (4) $e_L(C_L)=e_L(C'_L)>|T|-|C_L|$.

(3) Let $e_L(C_L)=|T|-|C_L|$. Since $d_L(C_L\cup C'_L,T)=|T|-|C_L\cup C'_L|=|T|-|C_L|-1$ there exists a subtree $S\neq T$ such that $e_L(C_L\cup C'_L)=d_L(C_L\cup C'_L,S)$. We may assume that S is the maximal of such subtree i.e. for any subtree S' such that $S\subseteq S', S'\neq S$ we have $d_L(C_L\cup C'_L,S')<d_L(C_L\cup C'_L,S)$. Two cases arise (3.1) $S\cap(C_L\cup C'_L)=\emptyset$ and (3.2) $S\cap(C_L\cup C'_L)\neq\emptyset$.

(3.1) We consider first the case, where S is a leaf of T such that $d_T(C_L\cup C'_L,S)=1$. Since u is not a leaf of T there exists a node w on the diametral path of T such that $d_T(w,u)=1$ and $w\notin C_L\cup C'_L$. Furthermore, $d_L(C_L\cup C'_L,S)=d_L(C_L\cup C'_L,\{w\})$.

Hence we may assume that S contains at least one point of a diametral path of T. Since $S\cap(C_L\cup C'_L)=\emptyset$, and since $C_L\cup C'_L$ is a subpath of a diametral path, either u or u' is on the geodesic from S to $C_L\cup C'_L$. Assume that $d_T(u',S)=d_T(C_L\cup C'_L,S)$ (another case is similar). Now $d_T(C_L,S)=d_T(C_L\cup C'_L,S)+1$ and

$$\begin{aligned} d_L(C_L,S) &= |C_L|+|S|+2(d_T(C_L,S)-1) \\ &= |C_L\cup C'_L|-1+|S| \\ &\quad +2(d_T(C_L\cup C'_L,S)+1-1) \\ &= e_L(C_L\cup C'_L)+1>e_L(C_L), \end{aligned}$$

which is a contradiction.

(3.2) If $u,u'\in S$, then $C_L\cup C'_L\subseteq S$, thus $S=T$, which is a contradiction. Two subcases arise (a) $u\in S$ and $u'\notin S$ (the case $u'\in S$ and $u\notin S$ is similar) and (b) $u,u'\notin S$.

(a) If $(C_L\cup C'_L)\cap S=\{u\}$, then $d_T(C'_L,S)=1$ and we obtain a contradiction

$$\begin{aligned} d_L(C'_L,S) &= |C'_L|+|S|+2(d_T(C'_L,S)-1) \\ &= |C'_L\cup C_L|+|S|-2|(C'_L\cup C_L)\cap S|+1 \\ &= e_L(C_L\cup C'_L)+1. \end{aligned}$$

Thus $C'_L\cap S\neq\emptyset$. Since $u'\notin S$ and $u\in S$, we have $|C'_L\cap S|=|(C_L\cup C'_L)\cap S|-1$ and

$$\begin{aligned} d_L(C'_L,S) &= |C'_L|+|S|-2|C'_L\cap S| \\ &= |C_L\cup C'_L|+|S|-2|(C_L\cup C'_L)\cap S|+1 \\ &= e_L(C_L\cup C'_L)+1, \end{aligned}$$

a contradiction.

(b) Assume $u,u'\notin S$. Since $S\cap(C_L\cup C'_L)\neq\emptyset$, there exists a point u'' of $C_L\cup C'_L$ such that u'' is a neighbour of S and u'' is on the geodesic from S to u'. Since S is a maximal subtree such that $d_L(C_L\cup C'_L,S)=e_L(C_L\cup C'_L)$, u'' cannot have a point $v\notin C_L\cup C'_L$ as a neighbour. Otherwise, we have

$$\begin{aligned} & d_L(C_L\cup C'_L,S\cup\{u'',v\}) \\ &= |C_L\cup C'_L|+|S|+2-2(|(C_L\cup C'_L)\cap S|+1) \\ &= |C_L\cup C'_L|+|S|-2|(C_L\cup C'_L)\cap S| \\ &= e_L(C_L\cup C'_L) \end{aligned}$$

contradicting the maximality of S. Then clearly $u'' \neq u'$. Let S'' be a branch of u'' such that $u' \in S''$. Clearly $S \cap S'' = \emptyset$ and $|S \cup S''| = |S| + |S''|$. Since $u \notin S$, $u \notin S''$, we have $|S'' \setminus (C_L \cup C'_L)| = |S'' \setminus C'_L|$ and $|S'' \cap (C_L \cup C'_L)| = |S'' \cap C'_L|$. Combining these facts, we obtain

$$\begin{aligned} & d_L(C_L \cup C'_L, S \cup S'') - e_L(C_L \cup C'_L) \\ & = |S''| - 2|(C_L \cup C'_L) \cap S''| \\ & = |S'' \setminus (C_L \cup C'_L)| - |(C_L \cup C'_L) \cap S''| \\ & = |S'' \setminus C'_L| - |S'' \cap C'_L|. \end{aligned}$$

The property that u'' cannot contain a leaf as a neighbour implies that $T \setminus S''$ induces a subtree of T and $T = (T \setminus S'') \cup S''$. Then

$$\begin{aligned} & e_L(C'_L) - d_L(C'_L, T \setminus S'') \\ & = |T| - |C'_L| - (|C'_L| + |T \setminus S''| - 2|C'_L \cap (T \setminus S'')|) \\ & = |T| - 2|C'_L| - |T| + |S''| + 2(|C'_L \cap T| \\ & \quad -|C'_L \cap S''|) \\ & = |S'' \setminus C'_L| - |C'_L \cap S''| \\ & = d_L(C_L \cup C'_L, S \cup S'') - e_L(C_L \cup C'_L). \end{aligned}$$

By the maximality of S we have $d_L(C_L \cup C'_L, S \cup S'') - e_L(C_L \cup C'_L) < 0$. Thus the identity above implies a contradiction $e_L(C'_L) - d_L(C'_L, T \setminus S'') < 0$.

(4) Assume $e_L(C_L) > |T| - |C_L|$, thus $e_L(C_L) \geq |T| - |C_L| + 1$. Then since $|C_L \cap C'_L| = |C_L| - 1$, we have $d_L(C_L \cap C'_L, T) = |T| - |C_L \cap C'_L| = |T| - |C_L| + 1$ and since $C_L \cap C'_L$ is not a least central subtree of T, there exists $S \neq T$ such that $e_L(C_L \cap C'_L) = d_L(C_L \cap C'_L, S) > |T| - |C_L| + 1$. Two cases arise, (4.1) $S \cap (C_L \cap C'_L) = \emptyset$ and (4.2) $S \cap (C_L \cap C'_L) \neq \emptyset$.

(4.1) Clearly we may assume that either u or u' is on the geodesic from $C_L \cap C'_L$ to S. Otherwise S is a leaf of T and a neighbour of a point of $C_L \cap C'_L$, which implies that $d_L(C_L, S) > d_L(C_L \cap C'_L, S)$. Assume that u' is on the geodesic (another case is analogous). Then clearly $d_T(C_L, S) = d_T(C_L \cap C'_L, S)$ and since $C_L \cap S = \emptyset$, we have

$$\begin{aligned} & d_L(C_L, S) \\ & = |C_L| + |S| + 2(d_T(C_L, S) - 1) \\ & = |C_L \cap C'_L| + 1 + |S| + 2(d_T(C_L \cap C'_L, S) - 1) \\ & = e_L(C_L \cap C'_L) + 1, \end{aligned}$$

a contradiction.

(4.2) If $(C_L \cap C'_L) \cap S \neq \emptyset$, then since $S \neq T$ either u or u' is not a point of S. We may assume that $u \notin S$. Then clearly $|C_L \cap S| = |(C_L \cap C'_L) \cap S|$ and

$$\begin{aligned} d_L(C_L, S) & = |C_L| + |S| - 2|C_L \cap S| \\ & = |C_L \cap C'_L| + 1 + |S| - 2|(C_L \cap C'_L) \cap S| \\ & = e_L(C_L \cap C'_L) + 1, \end{aligned}$$

a contradiction. By subcases (3) and (4) the Theorem follows. □

Recently, it turned out that the cardinality of nodes of degree two has a connection to the uniqueness of least central subtrees. In some cases we are able to prove that the unique C_L equals T', the subtree of T obtained by removing all leaves of T.

Theorem 5 *For any homeomorphically irreducible tree T the unique least central subtree is T'. For any almost homeomorphically irreducible tree T the unique least central subtree is T'. For any Cayley tree of degree n the unique least central subtree is T'. In all cases above the minimum $L-$eccentricity is $e_L(T') = |T| - |T'| = n_1(T)$.*

Proof. Cayley trees are a subclass of homeomorphically irreducible trees. We prove first the result for homeomorphically irreducible trees. We may assume $|T| \geq 4$. Since C_L cannot contain leaves, every least central subtree is a subtree of T'. Let $n_1(T)$ be the number of leaves of $|T|$. For every subtree T_s of T' such that $|T_s| \leq |T'| - 1$, we have

$$\begin{aligned} e_L(T_s) & \geq d_L(T_s, T) = |T| - |T_s| \geq |T| - (|T'| - 1) \\ & = |T| - |T| + n_1(T) + 1 = n_1(T) + 1. \end{aligned}$$

Since $d_L(T', T) = |T| - |T'| = n_1(T)$ it suffices to prove that $d_L(T', S) \leq n_1(T)$ for every subtree S of T, $S \neq T$. Two cases arise (i) $S \cap T' = \emptyset$ (ii) $S \cap T' \neq \emptyset$.

(i) Since $T \setminus T'$ contains only leaves of T we have S is a tree of one point and $d_T(T', S) = 1$. Then, by Lemma 1 and Lemma 3

$$\begin{aligned} d_L(T', S) & = |T'| + |S| + 2(d_T(T', S) - 1) \\ & = |T'| + 1 \leq n_1(T) - 1. \end{aligned}$$

(ii) If $S \cap T' = T'$, we have

$$\begin{aligned} d_L(T', S) & = |T'| + |S| - 2|T' \cap S| = |S| - |T'| \\ & \leq |T| - |T'| = n_1(T). \end{aligned}$$

Assume that there exists a point $v \in T' \setminus S$. Since $S \cap T' \neq \emptyset$, we may assume that v is a neighbour of S. We can choose the point $w \in S$ such that w is a neighbour of v. Let B_v be any branch at v not containing S. For all trees $|B_v| = |B_v \cap T'| + |B_v \setminus T'|$. In particular, for branches of homeomorphically irreducible trees rooted at v we have $|B_v \cap T'| \leq |B_v \setminus T'|$. Let $\tilde{S}_v$ be a maximal subtree of B_v such that $|\tilde{S}_v \cap T'| = |\tilde{S}_v \setminus T'|$. We obtain $\tilde{S}_v$ from the branch B_v by omitting some leaves in order to obtain the balance. Clearly $v \in \tilde{S}_v$. Let $S_1 = S \cup \tilde{S}_v$. Now $|S_1| = |S| + |\tilde{S}_v|$, $S_1 \cap T' = (S \cap T') \cup (\tilde{S}_v \cap T') \neq$

$\emptyset$ and $|S_1 \cap T'| = |S \cap T'| + |\tilde{S}_v \cap T'|$. Then

$$\begin{aligned} d_L(T', S_1) &= |S_1| + |T'| - 2|S_1 \cap T'| \\ &= |S| + |\tilde{S}_v| + |T'| \\ &\quad -2(|S \cap T'| + |\tilde{S}_v \cap T'|) \\ &= d_L(T', S) + |\tilde{S}_v| - 2|\tilde{S}_v \cap T'| \\ &= d_L(T', S). \end{aligned}$$

If $S_1 \cap T' = T'$, then as before $d_L(T', S_1) = |S_1| - |T'| \le |T| - |T'| \le n_1(T)$. If $S_1 \cap T' \neq T'$ then we repeat the process described above. Hence we obtain a sequence of subtrees $S = S_0 \subseteq S_1 \subseteq S_2 \subseteq \ldots \subseteq S_n$ such that $d_L(T', S_{i+1}) = d_L(T', S_i)$ for each index $i = 0, \ldots, n-1$. At final stage $S_n \cap T' = T'$. Moreover, $d_L(T', S_n) = |S_n| - |T'| \le |T| - |T'| \le n_1(T)$. Thus $C_L = T'$ is the unique least central subtree.

Here we consider the case of almost homeomorphically irreducible trees. The case $|T| = 3$ is clear (one tree), there are no such trees with $|T| = 4$. Thus we may assume that $|T| \ge 5$. Let $x \in T'$ be the unique point of degree 2. For every subtree T_s of T' such that $|T_s| \le |T'| - 1$, we have $e_L(T_s) \ge n_1(T) + 1$. Since $d_L(T', T) = |T| - |T'| = n_1(T)$ it suffices to prove that $d_L(T', S) \le n_1(T)$ for every subtree S of T, $S \neq T$. Two cases arise (i) $S \cap T' = \emptyset$ (ii) $S \cap T' \neq \emptyset$.

(i) Since $T \setminus T'$ contains only leaves of T we have S is a tree of one point and $d_T(T', S) = 1$. Then, by Lemma 3 $d_L(T', S) = |T'| + |S| + 2(d_T(T', S) - 1) = |T'| + 1 \le n_1(T)$.

(ii) If $S \cap T' = T'$, we have $d_L(T', S) = |T'| + |S| - 2|T' \cap S| = |S| - |T'| \le |T| - |T'| = n_1(T)$. If $x \in S$ then the proof of homeomorphically irreducible tree is applicable. Thus we may assume that $x \in T' \setminus S$. Let $v \in T'$ be a neighbour of S and let $w \in S$ be a neighbour of v. We consider the branches B_v not containing the point w. By Lemma 3 we have three cases for branches at v.

(1) If $v \neq x$ and $x \notin B_v$ then $|B_v \cap T'| \le |B_v \setminus T'|$.

(2) If $v \neq x$ and $x \in B_v$ then $|B_v \cap T'| \le |B_v \setminus T'| + 1$.

(3) If $v = x$ then $B_v = B_x$ and $|B_v \cap T'| \le |B_v \setminus T'|$.

Case (1) is similar to the case of a homeomorphically irreducible tree. We may assume that either one of the branches B_v contains the point x or $v = x$. If $v \neq x$ then we repeatedly process the branches at v not containing the point x. These branches can be treated by the method of homeomorphically irreducible trees. Let $\tilde{S}_v$ be a maximal subtree of B_v such that $|\tilde{S}_v \cap T'| = |\tilde{S}_v \setminus T'|$. Let $S_1 = S \cup \tilde{S}_v$. Clearly $|S_1| = |S| + |\tilde{S}_v|$ and $S_1 \cap T' = (S \cap T') \cup (\tilde{S}_v \cap T') \neq \emptyset$, $|S_1 \cap T'| = |S \cap T'| + |\tilde{S}_v \cap T'|$. Then

$$\begin{aligned} d_L(T', S_1) &= |S| + |\tilde{S}_v| + |T'| \\ &\quad -2(|S \cap T'| + |\tilde{S}_v \cap T'|) \\ &= d_L(T', S) + |\tilde{S}_v| - 2|\tilde{S}_v \cap T'| \\ &= d_L(T', S). \end{aligned}$$

If $S_1 \cap T' = T'$, then as before $d_L(T', S_1) = |S_1| - |T'| \le |T| - |T'| \le n_1(T)$. If $S_1 \cap T' \neq T'$ then we repeat the process described above. Hence we obtain a sequence of subtrees $S = S_0 \subseteq S_1 \subseteq S_2 \subseteq \ldots \subseteq S_n$ such that $d_L(T', S_{i+1}) = d_L(T', S_i)$ for each index $i = 0, \ldots, n-1$. Moreover the point $x \in T'$ is a neighbour of S_n. We can choose the point $w \in S_n$ such that w is a neighbour of x. We consider the branch B_x not containing the point w. By Lemma 3 $|B_x \cap T'| \le |B_x \setminus T'|$. Let $\tilde{S}_x$ be a maximal subtree of B_x such that $|\tilde{S}_x \cap T'| = |\tilde{S}_x \setminus T'|$. At final stage the subtree $S_{n+1} = S_n \cup \tilde{S}_x$ satisfies $S_{n+1} \cap T' = T'$. Moreover, $d_L(T', S_{n+1}) = |S_{n+1}| - |T'| \le |T| - |T'| \le n_1(T)$. Thus $C_L = T'$ is the unique least central subtree. $\square$

4. Bounds for $L-$eccentricity and size of C_L

The following Theorem gives sharp estimates for the $L-$eccentricity of least central subtrees. The upper bound is attained by stars and paths. Moreover, there exist certain caterpillars which attain the lower bound.

Theorem 6 *We have the bounds for the eccentricity of the least central subtree*

(1) $\frac{1}{2}(|T| + 1) \le e_L(C_L) \le |T| - 1$,
(2) $e_L(C_L) \ge n_1(T)$,
(3) $e_L(C_L) \ge \operatorname{diam} T$.

Proof. The upper bound follows from the elementary property (see [8], Theorem 5) $e_L(C_L) \le e_L(x) \le |T| - 1$, provided that T is not a path with even number of points. A direct calculation shows that $e_L(C_L) = |T| - 1$ for paths with even number of points. We use the estimates

$$\begin{aligned} e_L(C_L) &\ge d_L(C_L, T) = |T| - |C_L|, \\ e_L(C_L) &\ge d_L(C_L, v_l) \\ &= |C_L| + 1 + 2(d_T(C_L, v_l) - 1), \\ &\quad \text{for leaves } v_l \end{aligned}$$

and obtain the lower bound $2\, e_L(C_L) \ge |T| + 1 + 2(d_T(C_L, v_l) - 1) \ge |T| + 1$ by adding inequalities. Estimate (2) is clear. For the proof of (3), let v_1 and v_2 be two different leaves on any diametral path of T. Let $c_1 \in C_L$ and $c_2 \in C_L$ such that $d_T(C_L, v_i) = d_T(c_i, v_i)$, $i = 1, 2$. Now $e_L(C_L) \ge d_L(C_L, v_i)$, $i = 1, 2$ and we have

$$\begin{aligned}2e_L(C_L) &\geq |C_L| + |v_1| + 2[d_T(C_L, v_1) - 1]\\ &\quad + |C_L| + |v_2| + 2[d_T(C_L, v_2) - 1]\\ &= 2|C_L| + 2[d_T(C_L, v_1) + d_T(C_L, v_2) - 1]\\ &= 2|C_L| + 2[d_T(c_1, v_1) + d_T(c_2, v_2) - 1].\end{aligned}$$

Thus

$$\begin{aligned}e_L(C_L) &\geq |C_L| + d_T(c_1, v_1) + d_T(c_2, v_2) - 1\\ &= |C_L| - d_T(c_1, c_2) + d_T(c_1, v_1)\\ &\quad + d_T(c_1, c_2) + d_T(c_2, v_2) - 1\\ &= |C_L| - d_T(c_1, c_2) + \operatorname{diam} T - 1 \geq \operatorname{diam} T,\end{aligned}$$

since $|C_L| - d_T(c_1, c_2) - 1 \geq 0$ for the subtree C_L.

Table 1 and Table 2 show that the lower bounds in Theorem 6 agree well with computational results. Boldface numbers represent the count of cases where the least central subtree is not unique.

Recently we obtained a result concerning least central subtrees with minimal values of $L-$eccentricity. It gives sufficient conditions for a least central subtree of a tree to attain maximum size or maximum diameter. These results are formulated in Theorem 7.

Theorem 7 *Let C_L be a least central subtree of a tree T.*

(1) If $e_L(C_L) = n_1(T)$ then $C_L = T'$ and C_L is unique.
(2) If $e_L(C_L) = \operatorname{diam} T$ then C_L is a path.
(3) If $2e_L(C_L) = |T| + 1$ then $\operatorname{diam} T = \operatorname{diam} C_L + 2$.

Proof. The proof of (1) is clear. For the proof of (2) we use the result

$$e_L(C_L) \geq |C_L| - d_T(c_1, c_2) + \operatorname{diam} T - 1 \geq \operatorname{diam} T$$

obtained in the proof of Theorem 6. If $e_L(C_L) = \operatorname{diam} T$ then we have by previous estimate $|C_L| - d_T(c_1, c_2) - 1 = 0$. Thus C_L is a path.

In the case (3) the tree size is an odd number. We have for all trees

$$\begin{aligned}2e_L(C_L) &\geq (|T| - |C_L|) + (|C_L| - d_T(c_1, c_2)\\ &\quad + \operatorname{diam} T - 1)\\ &= |T| - 1 + \operatorname{diam} T - d_T(c_1, c_2)\\ &= |T| + 1 + \operatorname{diam} T - 2 - d_T(c_1, c_2)\\ &\geq |T| + 1.\end{aligned}$$

If $2e_L(C_L) = |T|+1$ then $\operatorname{diam} T - 2 - d_T(c_1, c_2) = 0$. Thus $\operatorname{diam} C_L = \operatorname{diam} T - 2$. □

Remark 1. Note that case (1) in Theorem 7 implies the following claim. If $e_L(T') = n_1(T)$ then $e_L(C_L) = n_1(T)$ and according to Theorem 7 $C_L = T'$ and C_L is unique. This follows from $n_1(T) = e_L(T') \geq e_L(C_L) \geq n_1(T)$ and Theorem 7 case (1) is in use.

Note that Table 3 agrees with Remark 1. In the case $e_L(C_L) = n_1(T) = 7$ all 72 trees are such that $C_L = T'$, in all cases the joinsemilattice center consists of just one subtree, there are no cases with nonunique least central subtree, and $L-$eccentricity is $e_L(C_L) = d_L(C_L, T) = |T| - |T'|$.

The numbers in Table 3 are interpreted as follows. The middle number is the count for all trees. The upper left number is the count for trees with $e_L(C_L) > d_L(C_L, T)$. The lower left number is the count for trees with $C_L = T'$. The upper right number is the count for trees with unique joinsemilattice center. The lower right number is the count for trees with nonunique least central subtree. Only nonzero values are printed.

In Theorem 8 we bound the size of any least central subtree. We are able to prove the Theorem under the assumption that $e_L(C_L) \leq |T| - |C_L| + 1$. Table 4 and Table 5 illustrate the Theorem from the numerical point of view.

Theorem 8 *If $e_L(C_L) = |T| - |C_L|$ or $e_L(C_L) = |T| - |C_L| + 1$ then we have the following bounds for the size of any least central subtree*

(1) $2|C_L| \leq |T|$,
(2) $|C_L| \leq |T| - n_1(T) + 1$,
(3) $|C_L| \leq |T| - \operatorname{diam} T + 1$.

Proof. We assume that $e_L(C_L) = |T| - |C_L| + k$, with $k \geq 0$. Thus

$$\begin{aligned}|T| - |C_L| + k \geq d_L(C_L, v_l) &= |C_L| + 1\\ &\quad + 2(d_T(C_L, v_l) - 1), \text{ for leaves } v_l\end{aligned}$$

Solving this with respect to $|C_L|$ gives

$$2|C_L| \leq |T| + k - 1 - 2(d_T(C_L, v_l) - 1) \leq |T| + k - 1$$

and estimate (1) follows for $k = 0$ and $k = 1$. Our assumption together with estimate (2) in Theorem 6 gives

$$|T| - |C_L| + k \geq n_1(T)$$

and estimate (2) follows by solving the inequality with respect to C_L. The proof for estimate (3) is analogous. □

Remark 2. The distributions in Table 4 and Table 5 are in agreement with the estimate (1) given in Theorem 8. Estimates (2) and (3) are sharp only for trees with even size. We conjecture that either $e_L(C_L) = |T| - |C_L|$ or $e_L(C_L) = |T| - |C_L| + 1$. We are working in order to prove the conjecture.

The results in all tables are computed by constructing all free trees, finding the subtrees and constructing the

Table 1

	Diameter										
$e_L(C_L)$	2	3	4	5	6	7	8	9	10	11	Sum
11	1									1	2
10		5	1					1	5		12
9			20,**1**	9	2	1	13	22			67,**1**
8			16,**1**	58,**14**	94,**42**	101,**4**	53				322,**61**
7			8	43	71	26					148
Sum	1	5	45,**2**	110,**14**	167,**42**	128,**4**	66	23	5	1	551,**62**

Distribution of 551 trees of size 12 with respect to diameter and $L-$eccentricity.

Table 2

	Diameter											
$e_L(C_L)$	2	3	4	5	6	7	8	9	10	11	12	Sum
12	1										1	2
11		5	1						1	5		12
10			25,**1**	10	2		1	12	28			78,**1**
9			23,**1**	72,**10**	53,**3**	28,**10**	90,**25**	76				342,**49**
8			13,**2**	89,**24**	275,**129**	290,**124**	128					795,**279**
7			3	16	37	16						72
Sum	1	5	64,**4**	187,**34**	367,**132**	334,**134**	219,**25**	88	29	5	1	1301,**329**

Distribution of 1301 trees of size 13 with respect to diameter and $L-$eccentricity.

Table 3

	Number of leaves											
$e_L(C_L)$	2	3	4	5	6	7	8	9	10	11	12	Sum
12	1										${}_{1}1$	${}_{1}2$
11		5	1						1	${}_{5}5$		${}_{5}12$
10		4	26	10	1		2	11_{1}	${}_{24}24$			${}_{24}78_{1}$
9		3^{3}	40_{15}	83_{14}	${}^{1}53_{6}$	25^{1}_{3}	66_{11}	${}_{72}72$				${}^{1}_{72}342^{4}_{49}$
8			11^{11}	126^{81}_{45}	296^{55}_{170}	${}^{4}_{4}228^{57}_{64}$	${}_{134}134^{1}$					${}^{4}_{138}795^{205}_{279}$
7						${}_{72}72^{72}$						${}_{72}72^{72}$
Sum	1	12^{3}	78^{11}_{15}	219^{81}_{59}	${}^{1}350^{55}_{176}$	${}^{4}_{76}325^{130}_{67}$	${}_{134}202^{1}_{11}$	${}_{72}83_{1}$	${}_{24}25$	${}_{5}5$	${}_{1}1$	${}^{5}_{312}1301^{281}_{329}$

Distribution of trees of size 13 with respect to number of leaves and $e_L(C_L)$.

Table 4

	Diameter											
$\|C_L\|$	2	3	4	5	6	7	8	9	10	11	12	Sum
6			3	16	37	20						76
5			13,**2**	89,**24**	275,**129**	286,**124**	129					792,**279**
4			23,**1**	72,**10**	53,**3**	28,**10**	89,**25**	76				322,**49**
3			25,**1**	10	2		1	12	28			68,**1**
2		5	1						1	5		12
1	1										1	2
Sum	1	5	64,**4**	187,**34**	367,**132**	334,**134**	219,**25**	88	29	5	1	1301,**329**

Distribution of 1301 trees of size 13 with respect to diameter and C_L-size.

Table 5

	Diameter										
$\lvert C_L \rvert$	2	3	4	5	6	7	8	9	10	11	Sum
6			2	15	41	26					84
5			6	28	30	19,**4**	53				136,**4**
4			16,**1**	58,**14**	94,**42**	82	1	22			273,**57**
3			20,**1**	9	2	1	12		5		49,**1**
2		5	1					1		1	8
1	1										1
Sum	1	5	45,**2**	110,**14**	167,**42**	128,**4**	66	23	5	1	551,**62**

Distribution of 551 trees of size 12 with respect to diameter and C_L-size.

joinsemilattice graph of subtrees. We have computed complete distributions up to tree size $|T| = 15$ with similar results. The problem of generating free trees of n nodes was very efficiently solved by Li and Ruskey in [7]. The implementation of their algorithm in C-language is available via Combinatorial Object Server (COS) web-page http://theory.cs.uvic.ca/.

5. An algorithmic approach

This section is directed to a reader with background from algorithmic/computational complexity/optimization issues. Here we give a high level algorithmic approach.

The joinsemilattice $L(T)$ contains a complete information of the structure of a tree T. The nodes in the corresponding graph G_L of $L(T)$ are all subtrees of the given tree T Let S_1 and S_2 be two different subtrees of T. There exists a line from S_1 to S_2 in graph G_L if and only if S_1 is obtained from S_2 either by adding a neighbouring point or by removing a leaf of S_2. The basic tree operations concerning any subtree S of T are

- The operation of adding to S a neighbouring point $\{v\}$ of S: $\mathrm{EXPAND}\,(S, v) := S \mapsto S \cup \{v\}$.
- The operation of removing from S a leaf $\{u\}$ of S: $\mathrm{SHRINK}\,(S, u) := S \mapsto S \setminus \{u\}$.
- The exchange operation (swapping) i.e. removing a leaf of S and simultaneously adding a neighbour of $S \setminus \{u\} : \mathrm{SWAP}\,(S, u, v) := S \mapsto (S \setminus \{u\}) \cup \{v\}$.

Let $\{v_i, 1 = 1, \ldots, n(S)\}$ be the set of different neighbours of S. Let $\{u_i, i = 1, \ldots, n_1(S)\}$ be the set of leafs of S. It is clear that the set

$$N(S) := \Big(\bigcup_{i=1}^{n(S)} \mathrm{EXPAND}\,(S, v_i)\Big) \bigcup \Big(\bigcup_{i=1}^{n_1(S)} \mathrm{SHRINK}\,(S, u_i)\Big)$$

contains all neighbouring subtrees of S in the joinsemilattice. Furthermore

$$\begin{aligned}
&d_L(S, \mathrm{EXPAND}\,(S, v_i)) = 1,\\
&|\,\mathrm{EXPAND}\,(S, v_i)| = |S| + 1\\
&\text{for all } \ i = 1, \ldots, n(S),\\
&d_L(S, \mathrm{SHRINK}\,(S, u_i)) = 1,\\
&|\,\mathrm{SHRINK}\,(S, u_i)| = |S| - 1\\
&\text{for all } \ i = 1, \ldots, n_1(S).
\end{aligned}$$

The swapping operation produces subtrees with distance $d_L(S, \mathrm{SWAP}\,(S, u_i, v_j)) = 2$ and size $|\,\mathrm{SWAP}\,(S, u_i, v_j)| = |S|$. It is clear that for a given subtree S the repeated use of operation SWAP produces all subtrees of size $|S|$. In other words the swapping operation is closed among all subtrees of given size.

Let U be a subtree of T with eccentricity $e_L(U)$. By general properties of the eccentricity sequence we have for all neighbouring subtrees $V \in N(U)$ that $e_L(U) - 1 \le e_L(V) \le e_L(U) + 1$. Let $P(C, C_R)$ denote the least path containing the center and the point of centroid which is nearest to the center. There exists a geodesic in the joinsemilattice that connects $P(C, C_R)$, a least central subtree C_L, and the subtree T'. In extreme cases this geodesic may be reduced to a point in the joinsemilattice. In general, there are several geodesics even though the least central subtree is unique. In the general case there may be many least central subtrees. We have been analyzing subtree perturbations of type EXPAND, SHRINK and SWAP concerning least central subtrees. Unfortunately we cannot say much when the subtree under perturbation process is not a least central subtree.

In the following we give an outline of the algorithm for constructing least central subtrees.

Let $T = < \{v_1, \ldots, v_n\} >$.
Compute degree sequence of T.
Identify leaves of T.
$LCSup = T'$.
/* Every LCS is subtree of T'.
If $n_2(T) = 0$ then
/* Homeomorphically irreducible tree.
  $LCS = LCSup$
  $E_L(LCS) = |T| - |T'|$
else if $n_2(T) = 1$ then
/* Almost homeomorphically
  $LCS = LCSup$
/* irreducible tree.
  $E_L(LCS) = |T| - |T'|$
else if $n_2(T) \geq 2$ then
  Compute eccentricity $E_L(LCSup)$.
  if $E_L(LCSup) = n_1(T)$ then
/* See Remark 1 after Theorem 7.
    $LCS = LCSup$
    $E_L(LCS) = |T| - |T'|$
  else
    Find center C.
    Find centroid C_R.
    Let $LCSlow = P(C, C_R)$.
/* Every LCS contains $P(C, C_R)$.
    If $LCSlow = LCSup$ then
/* The geodesic reduces to a point.
      $LCS = LCSup$
      $E_L(LCS) = E_L(LCSup)$
    else
      ANALYZE all subtrees on the geodesics between $LCSlow$ and $LCSup$ by computing L-eccentricities and finding least subtrees in size with minimum $L-$eccentricity.
    end if
  end if
end if

Here we give a brief interpretation of least central subtrees. Assume that we have a fixed tree structure (eg. organization hierarchy, image layout hierarchy). Let S_1 and S_2 be two subtrees within this hierarchical structure. We assume that there is a unit cost for operations EXPAND (S_1, v) and SHRINK (S_1, u). According to Lemma 1, the cost of deformation from substructure S_1 to substructure S_2 is given by $L-$distance $d_L(S_1, S_2)$. Thus least central subtrees are the smallest possible substructures that are deformable into any other substructure within the underlying tree hierarchy with least cost.

6. Some conclusions and examples

Our results give a new point of view into trees. There are two tree classes: trees with unique least central subtrees and trees with several least central subtrees. Homeomorphically irreducible trees are the smallest trees which contain complete information of the tree branching structure. For this class of trees we have uniqueness and explicit a priori construction of the least central subtree. The unique least central subtree for any homeomorphically irreducible tree is obtained by stripping away leaves. One edge subdivision of any homeomorphically irreducible tree preserves this property. For the class of trees with exactly one node of degree two the unique least central subtree can be constructed by the same method. Swapping between different least central subtrees is impossible because there is no room for swapping. These results have several practical special cases e.g. Cayley trees are homeomorphically irreducible. A full binary tree with root node of degree two is another practical example. Furthermore, caterpillars are a tree class with unique least central subtrees. Again swapping is forbidden but the reason is different; there are no free leaves for swapping between least central subtrees.

However, we believe that trees with multiple least central subtrees are of practical interest as spanning trees. Our results show that flexible spanning trees should have a sufficient deviation away from a caterpillar tree. On the other hand the tree branching should not be too strong. The number of nodes of degree 2 must be large enough in order to guarantee some deviation away from the class of irreducible trees.

There are several subjects for further research. We feel that the most important open problem is our conjecture concerning possible values of $L-$eccentricity. We conjecture that either $e_L(C_L) = |T| - |C_L|$ or $e_L(C_L) = |T| - |C_L| + 1$. We claim that this property is true for all joinsemilattices generated by subtrees of a tree. This property is needed for C_L-size bounds given in Theorem 8. Furthermore, we believe that this result would be helpful for a more detailed description of the ANALYZE-stage of our high level algorithm given in section 5.

For some problems it might be useful to use the largest central subtree instead of the least central subtree. What can be said about the case when the least central subtree and largest central subtree coincide? Is it possible to characterize trees with unique joinsemilattice center?

Finally, we present some examples concerning our results. There are trees such that the tree has a bicentroid and the other point of the centroid is not a point of the least central subtree. There are trees such that the center, the centroid, the least central subtree and the path center are different subtrees. For more information on center, centroid and path center see [1]. Examples 3 and 4 illustrate the swapping mechanism.

Example 1. Let p be the point of degree 5 of the starlike caterpillar tree in Figure 2. The largest branch at p is the path $<\{p, q, r, s, t\}>$. Then the center of the tree $T_{1,1,1,1,4}$ is $C =<\{q, r\}>$, the centroid is $C_R =<\{p\}>$, the least central subtree is $C_L =<\{p, q, r\}>$ and the path center is $C_P =<\{p, q, r, s\}>$.

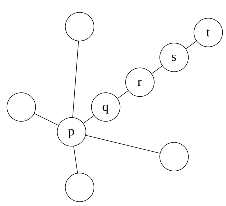

Fig. 2. Centrality for a caterpillar tree $T_{1,1,1,1,4}$.

Example 2. The largest branch at p of the starlike caterpillar tree in Figure 3 is the path $<\{p, q, r, s, t\}>$. For the tree $T_{1,1,1,4}$ the center is $C =<\{q, r\}>$, the centroid is $C_R =<\{p, q\}>$, the least central subtree is $C_L =<\{q, r\}>$ and the path center is $C_P =<\{p, q, r, s\}>$. Here p, the other point of the centroid is not a point of the least central subtree.

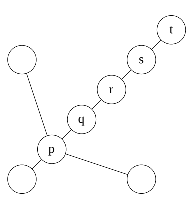

Fig. 3. Centrality for a caterpillar tree $T_{1,1,1,4}$.

Example 3. Subdivision trees of stars are ideal symmetric swappers. The smallest member of the family has 3 least central subtrees of size two. The second tree with $|T_{2,2,2,2}| = 9$, $|C_L| = 3$, $e_L(C_L) = 6$ has 6 least central subtrees. The third tree with $|T_{2,2,2,2,2}| = 11$, $|C_L| = 4$, $e_L(C_L) = 7$ has 10 least central subtrees.

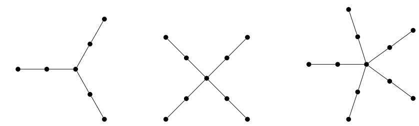

Fig. 4. Centrality of trees $T_{2,2,2}$, $T_{2,2,2,2}$ and $T_{2,2,2,2,2}$.

These trees are obtained by subdividing the star graphs $K_{1,n}$ with n leaves. We have in general $|K_{1,n}| = n + 1$. We prefer the complete bipartite graph notation for stars. The tree size is $|T_{2,2,\ldots,2}| = 2n + 1$ for $n \geq 3$. Least central subtree size is $|C_L| = n - 1$ and $e_L(C_L) = d_L(C_L, T) = n + 2$. Center node v_0 is contained in all least central subtrees and leaves cannot be in least central subtrees. There are $\binom{n}{n-2} = \binom{n}{2}$ least central subtrees, since this binomial coefficient is the number of ways to select $n-2$ nodes from n candidates. For these trees we have always $d_L(C_L, C'_L) \leq 4$.

Example 4. All subtrees obtained by swapping between points of least central subtrees are not necessarily least central subtrees. In the following example $|T| = 11$, $|C_L| = 4$. The intersection and union of least central subtrees are $<\{p, q\}>$ and $<\{p, q, r, s, t, u\}>$ respectively. There are five least central subtrees $<\{p, q, r, s\}>$, $<\{p, q, r, t\}>$, $<\{p, q, r, u\}>$, $<\{p, q, s, t\}>$, $<\{p, q, s, u\}>$, with $L-$eccentricity $e_L(C_L) = 7$. The subtree $<\{p, q, t, u\}>$ is not a least central subtree, since $e_L(<\{p, q, t, u\}>) = 8$.

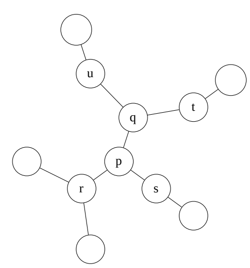

Fig. 5. An example on forbidden swapping.

All graphs in this article are drawn by using the dot graph drawing system, [2].

References

[1] Buckley, F. and Harary, F., Distance in graphs, Addison Wesley, Reading Mass, 1990.

[2] Gansher E., Koutsofios E. and North S., Drawing graphs with dot, 2006. http://www.graphviz.org/Documentation/dotguide.pdf.

[3] Hamina, M. and Peltola, M., Least Central Subtrees, Center and Centroid of a Tree, Networks, 2010, DOI: 10.1002/net.20402.

[4] Harary, F. and Palmer, E.M., Graphical enumeration, Academic Press, New York-London, 1973.

[5] Harary, F. and Prins, G., The number of homeomorphically irreducible trees, and other species, Acta Math., 1959, (**101**), 141–162.

[6] Harary, F. and Schwenk, A.J., The number of caterpillars, Discrete Math., 1973, (**6**), 359–365.

[7] Li, G.,and Ruskey, F., The Advantages of Forward Thinking in Generating Rooted and Free Trees, 10th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA), 1999, 939–940.

[8] Nieminen, J. and Peltola, M., The subtree center of a tree, Networks, 1999, (**34**), 272–278.

[9] Otter R., The number of trees, Ann. of Math., 1948, (**49**), 583–599.

[10] Riordan J., The enumeration of trees by height and diameter, IBM J. Res. Develop., 1960, (**4**), 473–478.

[11] Szekely, L.A. and Wang, H., On subtrees of trees, Advances in Applied Mathematics 34 (2005), 138–155.

Received 20-4-2009; revised 15-9-2010; accepted 22-9-2010