Available online at www.sciencedirect.com

SCIENCE DIRECT®

Advances in Applied Mathematics 34 (2005) 138–155

ADVANCES IN Applied Mathematics

www.elsevier.com/locate/yaama

On subtrees of trees ☆

L.A. Székely *, Hua Wang

Department of Mathematics, University of South Carolina, Columbia, SC 29208, USA

Received 11 February 2004; accepted 12 July 2004

Available online 15 September 2004

Abstract

We study that over a certain type of trees (e.g., all trees or all binary trees) with a given number of vertices, which trees minimize or maximize the total number of subtrees (or subtrees with at least one leaf). Trees minimizing the total number of subtrees (or subtrees with at least one leaf) usually maximize the Wiener index, and vice versa. In [L.A. Székely, H. Wang, Binary trees with the largest number of subtrees, submitted for publication], we described the structure of binary trees maximizing the total number of subtrees, here we provide a formula for this maximum value. We extend here the results from [L.A. Székely, H. Wang, Binary trees with the largest number of subtrees, submitted for publication] to binary trees maximizing the total number of subtrees with at least one leaf—this was first investigated by Knudsen [Lecture Notes in Bioinformatics, vol. 2812, Springer-Verlag, 2003, 433–446] to provide upper bound for the time complexity of his multiple parsimony alignment with affine gap cost using a phylogenetic tree.

Also, we show that the techniques of [L.A. Székely, H. Wang, Binary trees with the largest number of subtrees, submitted for publication] can be adapted to the minimization of Wiener index among binary trees, first solved in [M. Fischermann, A. Hoffmann, D. Rautenbach, L.A. Székely, L. Volkmann, Discrete Appl. Math. 122 (1–3) (2002) 127–137] and [F. Jelen, E. Triesch, Discrete Appl. Math. 125 (2-3) (2003) 225–233].

Using the number of subtrees containing a particular vertex, we define the *subtree core* of the tree, a new concept analogous to, but different from the concepts of *center* and *centroid*.

☆ This research was supported in part by the NSF contracts DMS 0072187 and 0302307.
* Corresponding author.
E-mail addresses: szekely@math.sc.edu (L.A. Székely), hwang0@math.sc.edu (H. Wang).

0196-8858/$ – see front matter
doi:10.1016/j.aam.2004.07.002

Keywords: Center; Centroid; Subtree core; Number of subtrees; Wiener index; Multiple parsimony alignment with affine gap cost; Caterpillar; Binary tree; Tree

1. Terminology and notation

All graphs in this paper will be finite, simple and undirected. A *tree* $T = (V, E)$ is a connected, acyclic graph. We refer to vertices of degree 1 of T as *leaves*. The unique path connecting two vertices v, u in T will be denoted by $P_T(v, u)$. For a tree T and two vertices v, u of T, the *distance* $d_T(v, u)$ between them is the number of edges on the connecting path $P_T(v, u)$. For a vertex v of T, define the *distance of the vertex* as

$$g_T(v) = \sum_{u \in V(T)} d_T(v, u),$$

the sum of distances from v to all other vertices. Let

$$\sigma(T) = \frac{1}{2} \sum_{v \in V(T)} g_T(v)$$

denote the *Wiener index* of T, which is the sum of distances for all unordered pairs of vertices.

We call a tree (T, r) *rooted at the vertex* r (or denote just by T if it is clear what the root is) by specifying a vertex $r \in V(T)$. For any two different vertices u, v in a rooted tree (T, r), we say that v is a *successor* of u, if $P_T(r, u) \subset P_T(r, v)$. Furthermore, if u and v are adjacent to each other and $d_T(r, u) = d_T(r, v) - 1$, we say that u is a *parent* of v and v is a *child* of u.

If v is any vertex of a rooted tree (T, r), let $T(v)$, *the subtree induced by* v, denote the rooted subtree of T that is induced by v and all its successors in T, and is rooted at v.

The *height* of a vertex v of a rooted tree T with root r is $h_T(v) = d_T(r, v)$, and the *height* of a rooted tree T is $h(T) = \max_{v \in T} h_T(v)$, the maximum height of vertices.

A *binary* tree is a tree T such that every vertex of T has degree 1 or 3. A *rooted binary tree* is a tree T with root r, which has exactly two children, while every other vertex of T has degree 1 or 3. A rooted binary tree T is *complete*, if it has height h and 2^h leaves for some $h \geqslant 0$. In addition, we also take a single vertex to be a rooted binary tree of height 0.

A *caterpillar tree* 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 tree* is a caterpillar tree, which is also a binary tree.

For a tree T and a vertex v of T, let $f_T(v)$ denote the number of subtrees of T that contain v, let $F(T)$ denote the number of non-empty subtrees of T.

2. Introduction

In [10] we characterized the binary tree on n vertices with the largest number of subtrees. We observe that the very same tree minimizes the Wiener index among binary trees