DIMACS Technical Report 2000-41
December 2000

The Center Function on Trees

by

F. R. McMorris[1]
Department of Applied Mathematics
Illinois Institute of Technology
Chicago, IL 60610

Fred S. Roberts[2]
Department of Mathematics
Rutgers University
New Brunswick, NJ 08903

Chi Wang
New Media Department
J. Walter Thompson, USA
Detroit, MI 48226

[1]Research supported by the Office of Naval Research grant N00014-00-1-0004.
[2]Research supported by NSF grant NSF-SBR97-09134 to Rutgers University.

ABSTRACT

When (X, d) is a finite metric space and $\pi = (x_1, \ldots, x_k) \in X^k$, a central element for π is an element x of X for which $\max\{d(x, x_i) : i = 1, \ldots, k\}$ is minimum. The function that returns the set of all central elements for any tuple π is called the center function on X. In this note, the center function on finite trees is characterized.

1 Introduction

Let (X, d) be a finite metric space and $X^* = \bigcup_{k \geq 1} X^k$, where X^k is the k-fold cartesian product of X.. Elements of X^* are called *profiles* and denoted $\pi = (x_1, \ldots, x_k)$, $\pi' = (y_1, \ldots, y_m)$, etc. A central problem in location theory and consensus theory is to find those points in X that are "closest" to any given profile π. Most of the work done in this area focuses on developing algorithms to find these points [7,17]. In recent years, there have been axiomatic studies of the procedures themselves and these have resulted in a much better understanding the process of location and consensus [4,5,9,10,19]. Without any conditions imposed, a *location function* (*consensus function*) *on* X is simply a mapping $L_X : X^* \to 2^X \backslash \{\emptyset\}$, where 2^X denotes the set of all subsets of X. When the set X is clear from the context, we will write L instead of L_X.

Let $D : X \times X^* \to \mathbb{R}$ be a function such that $D(x, \pi)$ represents a measure of "remoteness" of x to the profile π. An attractive class of location functions on (X, d) is defined by letting $L(\pi) = \{x \in X : D(x, \pi) \text{ is minimum}\}$. Two important location functions in this class are the *median function*, Med (defined by letting $D(x, \pi) = \sum_{i=1}^{k} d(x, x_i)$, where $\pi = (x_1, \ldots, x_k)$) and the *center function*, Cen (defined by letting $D(x, \pi) = \max\{d(x, x_1), \ldots, d(x, x_k)\}$). If $G = (X, E)$ is a finite connected graph and d is the usual geodesic metric on X, we will refer to a location function on X as being a *location function on the graph G*. Thus a profile in a graph can be thought of as a sequence of vertices, repetitions allowed. To date, there has been some success in characterizing Med on various types of graphs [3,5,11,12,13,14,15,16]. However, there is not much known for Cen on graphs, although Foster and Vohra [8] study the center function on edge-weighted trees allowing elements of $L(\pi)$ to be introduced between vertices if necessary. Of course, if π consists of each vertex of a graph appearing exactly once in π, then $Med(\pi)$ is the usual median of the graph and $Cen(\pi)$ is the usual center of the graph. The goal of this note is to characterize the center function on trees, where a *tree* is a finite connected graph without cycles. Our work was stimulated by the manuscript of Foster and Vohra. Graph theory terminology not defined here can be found in [6].

2 Axioms

We will now recall some axioms for a location function L defined on a tree $T = (X, E)$ which were used in [14,15] for studies of Med. We will also introduce axioms used to characterize Cen on trees and then compare and contrast Med and Cen with respect to these properties.

Anonymity (A): For every profile $\pi = (x_1, \ldots x_k) \in X^*$ and permutation σ of $\{1, \ldots k\}$, $L(\pi) = L(\pi^\sigma)$, where $\pi^\sigma = (x_{\sigma(1)}, \ldots, x_{\sigma(k)})$.

Betweenness (B): $L((x, y)) = \{z : d(x, y) = d(x, z) + d(z, y)\}$ for any $x, y \in X$, not necessarily distinct.

Consistency (C): If $L(\pi) \cap L(\pi') \neq \emptyset$ for profiles π and π', then $L(\pi\pi') = L(\pi) \cap L(\pi')$, where

$\pi\pi'$ is the concatenation of π and π', i.e., if $\pi = (x_1, \ldots, x_k)$ and $\pi' = (y_1, \ldots, y_m)$ then $\pi\pi' = (x_1, \ldots, x_k, y_1, \ldots, y_m)$.

As a consequence of results holding for graphs more general than trees, the following result was established in [15].

Theorem 1. Let L be a location function on a tree T. Then L is the median function Med on T if and only if L satisfies properties (A), (B) and (C).

It is not hard to see that Med satisfies (A), (B) and (C) on any finite metric space and that Cen also satisfies (A) in this general setting. However, Cen does not satisfy (B) or (C), even on trees. To see that Cen does not satisfy (B) a small example can easily be constructed, or one could use the following result of Slater [18]. Let $\{\pi\}$ be the set of vertices making up the profile π.

Lemma 1. For a profile π in a tree T, $Cen(\pi)$ is precisely the center of the smallest subtree of T that contains $\{\pi\}$.

Let a and b be vertices in a tree and $\pi = (a, b)$. Then, using Lemma 1, $Cen((a, b))$ is the usual center of entire unique path joining a and b. Hence $Cen(a, b)$ is a K_1 or K_2 and can be obtained by repeatedly deleting the pendant vertices of the path until a K_1 or K_2 remains. We refer to this K_1 or K_2 as the *middle* of the path between a and b. Using this procedure and Lemma 1, the next result follows easily.

Lemma 2. For a profile π in a tree T, $Cen(\pi) = Cen((u, v))$ where $d(u, v) = \max\{d(x, y) : x, y \in \{\pi\}\}$.

The following axiom is easily seen to be satisfied by Cen on a tree.

Middleness (M): Let L be a location function on a tree T, and $x, y \in T$ not necessarily distinct. Then $L((x, y))$ is the middle of the unique path joining x and y in T.

Now consider the path $x_1x_2x_3x_4$, and let $\pi = (x_1, x_4)$ and $\pi' = (x_1, x_3)$. Then $Cen(\pi) = \{x_2, x_3\}$ and $Cen(\pi') = \{x_2\}$ so that $Cen(\pi) \cap Cen(\pi') = \{x_2\}$. However, $Cen((x_1, x_4, x_1, x_3)) = \{x_2, x_3\}$ so Cen does not satisfy the Consistency axiom. We will show that it does satisfy the following slightly weaker axiom.

Quasi-Consistency (QC): If $L(\pi) = L(\pi')$ for profiles π and π', then $L(\pi\pi') = L(\pi)$.

Clearly if a location function L satisfies (C) then it satisfies (QC).

Lemma 3. The center function satisfies (QC) on any finite metric space.

Proof: For a profile $\pi = (x_1, \ldots, x_k)$ let $D(x, \pi) = \max\{d(x, x_1), \ldots, d(x, x_k)\}$. Assume that $Cen(\pi) = Cen(\pi')$, $a \in Cen(\pi\pi')$, and $b \in Cen(\pi) = Cen(\pi')$. ¿From the definition of Cen we have $D(a, \pi\pi') \leq D(b, \pi\pi')$, $D(b, \pi) \leq D(a, \pi)$, and $D(b, \pi') \leq D(a, \pi')$. Clearly

$$\begin{aligned} D(b, \pi\pi') = \max\{D(b, \pi), D(b, \pi')\} &\leq \max\{D(a, \pi), D(a, \pi')\} \\ &= D(a, \pi\pi') \leq D(b, \pi\pi'). \end{aligned} \tag{1}$$

Thus $D(a,\pi\pi') = D(b,\pi\pi')$ which implies $b \in Cen(\pi\pi')$ and therefore $Cen(\pi) \subseteq Cen(\pi\pi')$. We now claim that $a \in Cen(\pi) = Cen(\pi')$. For this, it suffices to show that $D(a,\pi) = D(b,\pi)$ or $D(a,\pi') = D(b,\pi')$. From (1), $\max\{D(b,\pi), D(b,\pi')\} = \max\{D(a,\pi), D(a,\pi')\}$. If $D(b,\pi) < D(a,\pi)$ and $D(b,\pi') < D(a,\pi')$, then
$\max\{D(b,\pi), D(b,\pi')\} < \max\{D(a,\pi), D(a,\pi')\}$, a contradiction. □

If $T = (X,E)$ is a tree and $S \subseteq X$, we let $T(S)$ denote the smallest subtree of T containing S. For the profile π and $x \in X$, let $\pi - x$ denote the profile obtained from π by removing x everywhere in π and reducing the length accordingly. Thus if $x \notin \{\pi\}$ we have $\pi - x = \pi$. $\{\pi\} - x$ will be the set of vertices appearing in π not equal to x. The following axiom, like (M), is formulated only for trees.

Redundancy (R): Let L be a location function on a tree T. If $x \in T(\{\pi\}-x)$, then $L(\pi-x) = L(\pi)$.

One final axiom is needed before our characterization.

Population Invariant (PI): If $\{\pi\} = \{\pi'\}$, then $L(\pi) = L(\pi')$.

Clearly Cen satisfies (R) and (PI) on any tree T. Also, if L satisfies (PI), then L satisfies (A). Note that Med satisfies neither (R) nor (PI).

Theorem 2. Let L be a location function on a tree T. Then L is the center function Cen on T if and only if L satisfies properties (M), (PI), (QC) and (R).

Proof: We have seen above that Cen satisfies (M), (PI), (QC) and (R) on T. For the converse, let L be a location function on the tree T and suppose L satisfies (M), (PI), (QC) and (R).

Let π be a profile such that $T(\{\pi\})$ is a path in T. If $|\{\pi\}| \leq 2$, then by (M), (PI) and Lemma 1, we have $L(\pi) = Cen(\pi)$. If $|\{\pi\}| > 2$, let $T(\{\pi\}) = I[a,b]$ where $I[a,b]$ denotes the path in T joining the vertices a and b. For $x \in I[a,b]$ with $x \notin \{a,b\}$, (R) gives $L(\pi - x) = L(\pi)$. Repeating this process, starting with $\pi - x$, and using (R) and (PI), we get $L(\pi) = L((a,b))$. From Lemma 1 and (M) we have $L(\pi) = Cen(\pi)$.

Now assume $T(\{\pi\})$ is not a path. We will show $L(\pi) = Cen(\pi)$ by using induction on $|\{\pi\}|$. Suppose first that $|\{\pi\}| = 3$ and that $T(\{\pi\})$ forms a, possibly subdivided, star $K_{1,3}$ with pendant vertices x, y, z and central vertex c. If $d(x,y) = d(y,z) = d(x,z)$, then from (M), we get $L((x,y)) = \{c\} = L((y,z))$. By (PI) and (QC) we have $L(\pi) = \{c\}$. Lemma 1 and (PI) yield $Cen(\pi) = \{c\}$, and we are done with this case. Without loss of generality we now suppose that either $d(x,y) \geq d(y,z) > d(x,z)$, or $d(x,y) > d(y,z) \geq d(x,z)$. These two cases are similar so we will only do the last one. Let $C = L((x,y))$, so we know that C is either a single vertex or two adjacent vertices. Using (M) and the relationships of the distances we get that C is on the path $I[c,y]$. Since $d(y,c) > d(z,c)$, we can pick $w \in I[c,y]$ so that the distance from w to the closest vertex in C is equal to the distance from z to the closest vertex in C. Then by (M), $L((w,z)) = C$. From (QC) this implies that $L((x,y,w,z)) = C$ and using (R), or (PI) if $w = y$, we have $L((x,y,z)) = C$. By (PI), $L(\pi) = C$. Finally, by Lemma 1, Lemma 2, and (PI) we know that $Cen(\pi) = C$.

For the induction step, assume $|\{\pi\}| = n+1$ with $n \geq 3$ and that $L(\pi') = Cen(\pi')$ for all profiles π' where $|\{\pi'\}| \leq n$. Let $\{x, y\} \subset \{\pi\}$ where x and y are selected so that $d(x, y) = \max\{d(r, s) : r, s \in \{\pi\}\}$. If K is the middle of $I[x, y]$, we know that $K = Cen((x, y))$ and by Lemma 2, we then have that we have $K = Cen(\pi)$. Now let $a, b \in \{\pi\}$ with $\{a, b\} \cap \{x, y\} = \emptyset$. Set $\pi_1 = \pi - a$ and $\pi_2 = \pi - b$. Since $x, y \in T(\{\pi_1\}) \cap T(\{\pi_2\})$, Lemmas 1 and 2 give $Cen(\pi_1) = K = Cen(\pi_2)$. Using the induction hypothesis, $L(\pi_1) = Cen(\pi) = L(\pi_2)$. Now (QC) gives $L(\pi_1\pi_2) = Cen(\pi)$ and since $\{\pi\} = \{\pi_1\pi_2\}$, (PI) implies that $L(\pi_1\pi_2) = L(\pi)$. Therefore $L(\pi) = Cen(\pi)$ and the proof is complete. $\square$

3 Examples

We now give four examples showing the independence of the axioms in Theorem 2. The verifications are left to the reader.

Example 1. Let T be any tree with at least three vertices and define L by $L(\pi) = T(\{\pi\})$ for any profile π. Then L satisfies $(PI), (R), (QC)$ on T but does not satisfy (M).

Example 2. Let $T = K_{1,3}$ and define L as follows: For any profile π that contains all three pendant vertices, set $L(\pi) = V(T)$ and let $L(\pi) = Cen(\pi)$ otherwise. Then L satisfies $(PI), (M)$ and (R) on T, but does not satisfy (QC).

Example 3. Let T be the path consisting of two vertices a and b. Define $L(\pi) = \{a, b\}$ if $\{\pi\} = \{a, b\}$ and with a and b occurring an equal number of times in π. Otherwise, let $L(\pi) = \{z\}$ where z occurs most frequently in π. Then L satisfies $(M), (R)$ and (QC), but does not satisfy (PI).

Example 4. Let T be the path *abcd*. For distinct $x, y \in \{a, b, c, d\}$ define $L((x)) = \{x\}$ and $L((x, y))$ to be the middle of the path between x and y. For profiles π where $|\{\pi\}| = 3$ define $L(\pi)$ as follows: $L((a, b, c)) = \{b\}, L((b, c, d)) = \{c\}, L((a, b, d)) = \{a, b, d\}, L((a, c, d)) = \{a, c, d\}, L((a, b, c, d)) = \{b, c\}$. For any profile π require $L(\pi) = L(\pi')$ for every π' with $\{\pi\} = \{\pi'\}$. Then L satisfies $(M), (PI)$, and (QC), but does not satisfy (R).

4 References

[1] J. P. Barthélemy and M. F. Janowitz, A formal theory of consensus, SIAM J. Discrete Math. 4 (1991), 305-322.

[2] J. P. Barthélemy, B. Leclerc and B. Monjardet, On the use of ordered sets in problems of comparison and consensus of classifications, J. Classification 3 (1986), 187-224.

[3] J. P. Barthélemy and F. R. McMorris, The median procedure for n-trees, J. Classification 3 (1986), 329-334.

[4] J. P. Barthélemy and B. Monjardet, The median procedure in cluster analysis and social choice theory, Math. Social. Sci. 1 (1981), 235-268.

[5] J. P. Barthélemy and B. Monjardet, The median procedure in data analysis: New results and open problems, in Classification and Related Methods of Data Analysis, H. H. Bock, ed., Elsevier Science, Amsterdam, 1988, 309-316.

[6] G. Chartrand and L. Lesniak, Graphs & Digraphs, Chapman & Hall, New York, 1996.

[7] M. S. Daskin, Network and Discrete Location: Models, Algorithms, and Applications, Wiley-Interscience, New York, 1995.

[8] D. P. Foster and R. V. Vohra, An axiomatic characterization of the absolute centre of a tree, unpublished manuscript, 1991.

[9] P. Hansen and F. S. Roberts, An impossibility result in axiomatic location theory, Math. Oper. Res. 2 (1996), 195-208.

[10] R. Holzman, An axiomatic approach to location on networks, Math. Oper. Res. 15 (1990), 553-563.

[11] B. Leclerc, Lattice valuations, medians and majorities, Discrete Math. 111 (1993), 345-356.

[12] B. Leclerc, Median for weight metrics in the covering graphs of semilattices, Discrete Appl. Math. 49 (1994), 281-297.

[13] F. R. McMorris, The median procedure for n-trees as a maximum likelihood method, J. Classification 7 (1990), 77-80.

[14] F. R. McMorris, H. M. Mulder and R. C. Powers, The median function on median graphs and semilattices, Discrete Appl. Math. 101 (2000), 221-230.

[15] F. R. McMorris, H. M. Mulder and F. S. Roberts, The median procedure on median graphs, Discrete Appl. Math. 84 (1998), 165-181.

[16] F. R. McMorris and R. C. Powers, The median procedure in a formal model of consensus, SIAM J. Discrete Math. 14 (1995), 507-516.

[17] P. B. Mirchandani and R. L. Francis, eds., Discrete Location Theory, John Wiley and Sons, New York, 1990.

[18] P. J. Slater, Centers to centroids in graphs, J. Graph Th. 2 (1978), 209-222.

[19] R. Vohra, An axiomatic characterization of some locations in trees, European J. Operational Res. 90 (1996), 78-84.