Bohdan Zelinka
Medians and peripherians of trees

Archivum Mathematicum, Vol. 4 (1968), No. 2, 87--95

Persistent URL: `http://dml.cz/dmlcz/104654`

MEDIANS AND PERIPHERIANS OF TREES

Bohdan Zelinka, Liberec

Received October 11, 1967

In [1], page 30, the concepts of the vertex median, the edge median and the center of gravity are introduced. On the page 66 the problem is posed, to study the properties of these concepts in the case of trees. Here some results concerning the vertex median and the edge median are given. Further, the concept of the peripherian is defined and some of its properties are studied.

Let G be a graph with the vertex set V and edge set E, let ν_v be its number of vertices, ν_e its number of edges. By the symbol $d(a, x)$ the distance of two vertices is denoted. If we have a subgraph G' of G, we shall denote the vertex set of G' by $V(G')$, its cardinality by $\nu_v(G')$.

The *vertex deviation* of a vertex a is by definition the number

$$m_1(a) = \frac{1}{\nu_v} \sum_{x \in V} (a, x).$$

A vertex a_1 with the minimal vertex deviation is called the *vertex median* of the graph G, its vertex deviation is called the *mean vertex deviation* of G and denoted by $m_1(G)$.

Analogously to the concept of the vertex median we can define also the edge median. If h is an edge joining the vertices x and y, and u a vertex, we define the distance

$$d(u, h) = \frac{1}{2} [d(u, x) + d(u, y)].$$

Then we define the edge median as the vertex b_1, for which the sum

$$M_1(b) = \frac{1}{\nu_e} \sum_{h \in E} d(b, h),$$

called the edge deviation of b, is minimal. The value of this sum for $b = b_1$ is called the mean edge deviation of G and denoted by $M_1(G)$.

Now let G be a finite tree. Study the properties of the above defined concepts. We consider always only finite trees.

Theorem 1. *Between the vertex deviation and the edge deviation of a vertex a of a tree G the following relation holds:*

$$M_1(a) = \frac{\nu_e + 1}{\nu_e} m_1(a) - \frac{1}{2}$$

Proof. Let a be a vertex of a tree G. For every non-negative integer n let $\Delta_a(n)$ be the set of exactly all vertices of the tree G whose distance from a is equal to n. As G is a tree, these sets have following properties:

a) The set $\Delta_a(0) = \{a\}$ and all edges incident with a join a with vertices of $\Delta_a(1)$.

b) For $n \geqq 1$ let $v \in \Delta_a(n)$. Then there exists exactly one edge joining v with a vertex of $\Delta_a(n-1)$ and all other edges incident with v join v with vertices of $\Delta_a(n+1)$.

Let $\varrho(v)$ be the degree of a vertex v. From the properties a) and b) it follows that

$$\text{card } \Delta_a(n+1) = \sum_{v \in \Delta_a(u)} [\varrho(v) - 1]. \tag{1}$$

for every positive integer n and

$$\text{card } \Delta_a(1) = \varrho(a).$$

Now the vertex deviation of a can be expressed as follows:

$$m_1(a) = \frac{1}{\nu_v} \sum_{n=1}^{\infty} n \text{ card } \Delta_a(n) \tag{2}$$

We can use formally the infinite sum, because only for finitely many numbers n the set $\Delta_a(n)$ is non-empty.

For the edge deviation of a, following [1], page 30, we have

$$M_1(a) = \frac{1}{2\nu_e} \sum_{v \in V} \varrho(v)\, d(a, v)$$

Analogously we may use the decomposition into the sets $\Delta_a(n)$ and write the equality

$$M_1(a) = \frac{1}{2\nu_e} \sum_{n=1}^{\infty} \left(n \sum_{v \in \Delta_a(n)} \varrho(v) \right)$$

If $\Delta_a(n) \neq 0$, all sums taken over $v \in \Delta_a(n)$ will be assumed to be zero.

We shall adapt the expression on the right-hand side.

$$M_1(a) = \frac{1}{2\nu_e} \sum_{n=1}^{\infty} \left[n \left(\sum_{v \in \Delta_a(n)} [\varrho(v) - 1] + \text{card } \Delta_a(n) \right) \right]$$

Applying (1), we get

$$M_1(a) = \frac{1}{2\nu_e} \sum_{n=1}^{\infty} n \, [\text{card } \Delta_a(n+1) + \text{card } \Delta_a(n)].$$

We can write this so:

$$M_1(a) = \frac{1}{2\nu_e} \sum_{n=1}^{\infty} n \text{ card } \Delta_a(n+1) + \frac{1}{2\nu_e} \sum_{n=1}^{\infty} n \text{ card } \Delta_a(n).$$

The first term on the right-hand side we adapt as follows:

$$\frac{1}{2\nu_e}\sum_{n=1}^{\infty} n \text{ card } \Delta_a(n+1) =$$

$$= \frac{1}{2\nu_e}\sum_{n=1}^{\infty} (n+1) \text{ card } \Delta_a(n+1) - \frac{1}{2\nu_e}\sum_{n=1}^{\infty} \text{card } \Delta_a(n+1)$$

Evidently

$$\sum_{n=1}^{\infty} \text{card } \Delta_a(n+1) = \nu_v - 1 - \text{card } \Delta_a(1) = \nu_e - \text{card } \Delta_a(1)$$

Using (2), the first term on the right-hand side can be adapted:

$$\frac{1}{2\nu_e}\sum_{n=1}^{\infty} (n+1) \text{ card } \Delta_a(n+1) = \frac{1}{2\nu_e}\sum_{n=2}^{\infty} n \text{ card } \Delta_a(n) =$$

$$= \frac{1}{2\nu_e}\left(\sum_{n=1}^{\infty} n \text{ card } \Delta_a(n) - \text{card } \Delta_a(1)\right) = \frac{\nu_v}{2\nu_e} m_1(a) - \frac{1}{2\nu_e} \text{card } \Delta_a(1)$$

Thus we have

$$M_1(a) = \frac{\nu_v}{\nu_e} m_1(a) - \frac{1}{2\nu_e} \text{card } \Delta_a(1) - \frac{1}{2} + \frac{1}{2\nu_e} \text{card } \Delta_a(1) =$$

$$= \frac{\nu_e + 1}{\nu_e} m_1(a) - \frac{1}{2},$$

which was to be proved.

Corollary 1. *For the mean vertex deviation* $m_1(G)$ *and the mean edge deviation* $M_1(G)$ *of a tree* G *the equality*

$$M_1(G) = \frac{\nu_e + 1}{\nu_e} m_1(G) - \frac{1}{2}$$

holds.

Corollary 2. *The vertex median and the edge median of a tree coincide with one another.*

Evidently, between vertex and edge deviations there is a linear dependence with the positive coefficient, so they must attain their minima in the same vertex.

Before expressing another theorem, we shall prove a lemma. According to Corollary 2 we can omit the epitheta "vertex" and "edge" and speak only about medians.

Lemma 1. *Let* a_1 *be the median of the tree* G. *Let* $B_1(a_1), \ldots, B_l(a_1)$ *be the branches of* G *from* a_1 (see [1], page 64). *Then for each* i, $1 \leqq i \leqq l$, *the inequality*

$$\nu_v[B_i(a_1)] \leqq \nu_v/2 + 1 \tag{3}$$

holds.

Proof. Let $1 \leqq i \leqq l$, let u_i be the vertex of the branch $B_i(a_1)$ which is joined by an edge with a_1. As a_1 is the median of G, there must be $m_1(a_1) \leqq m_1(u_i)$. Now if x is a vertex of the branch $B_i(a_1)$ different from a_1, the equality

$$d(u_i, x) = d(a_1, x) - 1$$

holds. If y is a vertex of G not belonging to $B_i(a_1)$ or equal to a_1, the equality

$$d(u_i, y) = d(a_1, y) + 1$$

holds. We have

$$m_1(u_i) = \frac{1}{\nu_v}\sum_{z \in V} d(u_i, z) = \frac{1}{\nu_v}\left(\sum_{x \in V(B_i)-\{a_1\}} d(u_i, x) + \sum_{y \in [V \dot{-} V(B_i)] \cup \{a_1\}} d(u_i, y)\right) =$$

$$= \frac{1}{\nu_v}\left(\sum_{x \in V(B_i)-\{a_1\}} (d(a_1, x) - 1) + \sum_{y \in [V - V(B_i)] \cup \{a_1\}} (d(a_1, y) + 1)\right) =$$

$$= \frac{1}{\nu_v}\left(\sum_{y \in V(B_i)-\{a_1\}} d(a_1, x) + \sum_{y \in [V \dot{-} V(B_i)] \cup \{a_1\}} d(a_1, y) -\right.$$

$$\left. - \nu_v(B_i) + 1 + \nu_v - \nu_v(B_i) + 1\right) = \frac{1}{\nu_v}\left(\sum_{z \in V} d(a_1, z) + \nu_v - 2\nu_v(B_i) + 2\right) =$$

$$= m_1(a_1) + \frac{1}{\nu_v}[\nu_v - 2\nu_v(B_i) + 2].$$

Thus from $m_1(a_1) \leqq m_1(u_i)$ we get

$$0 \leqq \frac{1}{\nu_v}[\nu - 2\nu(B_i) + 2].$$

from which it follows that

$$\nu_v(B_i) \leqq \nu_v/2 + 1$$

for all i, $1 \leqq i \leqq l$.

Now we can prove a theorem.

Theorem 2. *A tree has either exactly one median, or exactly two medians joined by an edge.*

Proof. Assume that there exists a tree G which has two medians $a_1^{(1)}, a_1^{(2)}$ such that $d(a_1^{(1)}, a_1^{(2)}) \geqq 2$. Let $B_1(a_1^{(1)}), \ldots, B_p(a_1^{(1)})$ be the branches of G from $a_1^{(1)}$ and $B_1(a_1^{(2)}), \ldots, B_2(a_1^{(2)})$ be the branches of G from $a_1^{(2)}$. Assume without the loss of generality that $a_1^{(1)}$ belongs to $B_1(a_1^{(2)})$ and $a_1^{(2)}$ belongs to $B_1(a_1^{(1)})$. Evidently all the branches $B_2(a_1^{(1)}), \ldots, B_p(a)_1^{(1)}$ are subgraphs of the branch $B_1(a_1^{(2)})$ and all the branches $B_2(a_1^{(2)}), \ldots, B_2(a_1^{(2)})$ are subgraphs of the branch $B_1(a_1^{(1)})$. According to Lemma 1 we have

$$\nu_v(B_1(a_1^{(1)}) \leqq \nu_v/2 + 1.$$

Thus the sum

$$\sum_{i=2}^{p} [\nu_v(B_i(a_1^{(1)}))-1] = \nu_v - \nu_v(B_1(a_1^{(1)}) \geqq \nu_v/2-1.$$

This is the number of all vertices belonging to any of the branches $B_2(a_1^{(1)}), \ldots, B_p(a_1^{(1)})$ except $a_1^{(1)}$. As written above, all these vertices belong to $B_1(a_1^{(2)})$ .Beside them whole the path P from $a_1^{(1)}$ to $a_1^{(2)}$ belongs to $B_1(a_1^{(2)})$. The vertices of the path P belong to the branch $B_1(a_1^{(1)})$, so except $a_1^{(1)}$ they do not belong to any of the branches $B_2(a_1^{(1)}), \ldots, B_p(a_1^{(1)})$. The number of vertices of the path P is at least three, because the distance between $a_1^{(1)}$ and $a_1^{(2)}$ is greater than or equal to two. The branch $B_1(a_1^{(1)})$ must therefore contain at least $\nu_v/2+2$ vertices, which leads to a contradiction.

Thus the distance between two medians of a tree is at most one, this means that those medians are joined by an edge. Therefore there cannot exist three different medians. (If they existed, they would form a triangle, which is impossible in a tree.)

On Fig. 1a we see a tree with one median, on Fig. 1b a tree with two medians.

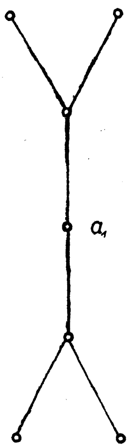

Fig. 1a.

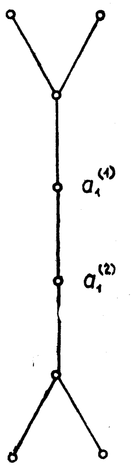

Fig. 1b.

Further in [1] the concept of the mass center of a tree is defined. If v is a vertex of the tree G, the maximal number of vertices of a branch of G from v is called the weight at v. The vertex of G with the minimal weight is called the mass center of G and its weight is called the weight of G.

Theorem 3. *The median of a tree coincides with its mass center.*

Proof. We have shown that for the median a_1 the inequality (3) holds and that it holds only for the median. Thus for each other vertex there exists some branch of G from that vertex such that its number of vertices is greater than $\nu_v/2+1$. From that it follows that every mass center of a tree must be its median. It remains to prove that every median of a tree is its mass center. For that it suffices to consider only the case of a tree with two medians (if there is only one median, the mass center must be identical with it as above proved). Let the tree G have two medians $a_1^{(1)}$ and $a_1^{(2)}$. Let $B_1(a_1^{(1)}), \ldots, B_p(a_1^{(1)})$ be the branches of G from $a_1^{(1)}$ and $B_1(a_1^{(2)}), \ldots, B_q(a_1^{(2)})$ the branches of G from $a_1^{(2)}$. We may assume without the loss of generality that $a_1^{(1)}$ belongs to $B_1(a_1^{(2)})$ and $a_1^{(1)}$ belongs to $B_1(a_1^{(2)})$. The branch $B_1(a_1^{(1)})$ contains therefore all branches $B_2(a_1^{(2)}), \ldots, B_q(a_1^{(2)})$ and the branch $B_1(a_1^{(2)})$ contains all branches $B_2(a_1^{(1)}), \ldots, B_p(a_1^{(1)})$ as its subgraphs. So if there were

$$\nu_v[B_1(a_1^{(1)})] < \nu_v/2 + 1,$$

the total number of vertices of the branches $B_2(a_1^{(2)}), \ldots, B\ (a_1^{(2)})$ would be less than $\nu_v/2$ because the branch $B_1(a_1^{(1)})$contains except these branches also the vertex $a_1^{(1)}$ which does not belong to any of them. The union of the branches $B_2(a_1^{(2)}), \ldots, B_q(a_1^{(2)})$ has a unique common vertex $a_1^{(2)}$ with the branch $B_1(a_1^{(2)})$. Thus the number of vertices of $B_1(a_1^{(2)})$ would be greater than $\nu_v/2 + 1$, which would be a contradiction. So if the tree G has two medians $a_1^{(1)}, a_1^{(2)}$, the branches $B_1(a_1^{(1)}), B_1(a_1^{(2)})$ have the same number of vertices equal to $\nu_v/2 + 1$. This is evidently the weight of the tree G, so both the medians of G are also its mass centers.

From Theorem 3 it follows that in the case of a tree it suffices to use only one of the terms "median" and "mass center". According to my personal opinion it is better to use the term "median" for avoiding misunderstandings due to the similarity of the terms "mass center" and "gravity center".

Theorem 4. *For a tree G with ν_v vertices the mean vertex deviation $m_1(G)$ satisfies the inequalities*

$$1 - \frac{1}{\nu_v} \leqq m_1(G) \leqq \frac{\nu_v}{4} - \frac{1}{4\nu_v}$$

in the case when ν_v is odd and

$$1 - \frac{1}{\nu_v} \leqq m_1(G) \leqq \frac{\nu_v}{4}$$

in the case when ν_v is even. The equality $m_1(G) = 1 - \dfrac{1}{\nu_v}$ occurs in the case of a star graph, the equality $m_1(G) = \dfrac{\nu_v}{4} - \dfrac{1}{4\nu_v}$ or $m_1(G) = \dfrac{\nu_v}{4}$ in the case of a path.

Proof. Let a_1 be the median of the tree G. The distance between a_1 and an arbitrary vertex different from a_1 is at least one. As the tree G has ν_v vertices, the sum of these distances is at least $\nu_v - 1$. Now we see that the distance between the median and an arbitrary vertex of G is at most $\nu_v/2$, as the path connecting these two vertices must be contained in one of the branches of G from a_1. Assume at first that ν_v

is odd. In every branch B of G from a_1 there can be at most n vertices with the distance greater than or equal to $\nu_v/2 - n$ because to each of those vertices a path from a_1 to it must exist in G containing $\nu_v/2 - n + 1$ vertices (including a_1). These paths together must contain at least $\nu_v/2 - n$ vertices with the distance from a_1 less than $\nu_v/2 - n$ (the case of exactly $\nu_v/2 - n$ vertices occurs if for all those paths such vertices are common). Thus the branch B contains at least $\nu_v/2 - n$ vertices with the distance from a_1 less than $\nu_v/2 - n$ and the number of the remaining vertices must be at most n. So the maximal mean vertex deviation can occur, if the number of vertices in each branch of G from a_1 with the distance greater than or equal to $\nu_v/2 - n$ is equal to n for every n. Such a case occurs, if G is a path with n vertices and a_1 is its center. Then G has two branches from a_1, each of them is a path with $(\nu_v + 1)/2$ vertices (evidently there cannot be more than two branches of G from a_1 containing such a number of vertices). Then using the well-known formula for the sum of a finite arithmetic sequence we obtain the upper bound for $m_1(G)$.

For ν_v even the proof is analogous.

Corollary 3. *For a tree G with ν_e edges the mean edge deviation satisfies the inequalities.*

$$\frac{1}{2} \leqq M_1(G) \leqq \frac{\nu_e}{4} + \frac{1}{2}$$

in the case when ν_e is odd and

$$\frac{1}{2} \leqq M_1(G) \leqq \frac{\nu_e}{4} + \frac{1}{4\nu_e}$$

in the case when ν_e is even. The equality $M_1(G) = \frac{1}{2}$ occurs in the case of a star graph, the equality $M_1(G) = \frac{\nu_e}{4} + \frac{1}{2}$ or $M_1(G) = \frac{\nu_e}{4} + \frac{1}{4\nu_e}$ in the case of a path.

This follows immediately from Theorem 4 using the equality from Theorem 1.

So as the minimal value of the vertex deviation, also its maximal value may be studied. We shall call it the periphery vertex deviation and the vertex in which this value is attained will be called the vertex peripherian. Analogously we may define the periphery edge deviation and the edge peripherian. The periphery vertex deviation will be denoted by $\overline{m}_1(G)$, the periphery edge deviation will be denoted by $\overline{M}_1(G)$.

Corollary 4. *The vertex peripherian and the edge peripherian of a tree coincide with one another.*

This follows from Theorem 1.

In the following we shall speak only about a "peripherian", omitting the epitheta "vertex" and "edge".

Theorem 5. *The peripherians of a tree are always its end vertices.*

Proof. We shall make analogous considerations to the proof of the Lemma 1. Let $\bar{a}_1$ be the peripherian of the tree G, let $B_1(\bar{a}_1), \ldots, B_l(\bar{a}_1)$ be the branches of G from $\bar{a}_1$. For $1 \leqq i \leqq l$, let u_i be the vertex of the branch $B_i(\bar{a}_1)$, which is joined by an edge with $\bar{a}_1$. As $\bar{a}_1$ is the peripherian of G, there must be $m_1(\bar{a}_1) \geqq m_1(u_i)$. Now

$$m_1(u_i) = m_1(\bar{a}_1) + \frac{1}{\nu_v}[\nu_v - 2\nu_v(B_i(\bar{a}_1)) + 2]$$

holds (see the proof of Lemma 1). Thus from $m_1(\bar{a}_1) \geqq m_1(u_i)$ we get

$$0 \geqq \frac{1}{\nu_v}(\nu_v - 2\nu_v(B_i(\bar{a}_1)) + 2),$$

from which it follows that

$$v_v[B_i(\bar{a}_1)] \geqq \mathrm{v}_v/2 + 1.$$

So each branch of G from $\bar{a}_1$ must have at least $\nu_v/2 + 1$ vertices. Two different branches from $\bar{a}_1$ could have exactly one common vertex, so if there existed any two ones, the total number of their vertices would be at least $\nu_v + 1$, which is impossible, because G has only ν_v vertices. Therefore there exists only one branch of G from $\bar{a}_1$, which means that $\bar{a}_1$ is an end vertex of G.

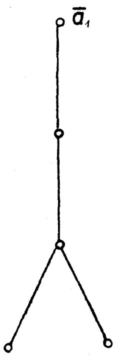

Fig. 2.

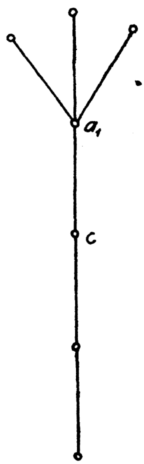

Fig. 3.

The number of peripherians of a tree is not so bounded as the number of medians.

Theorem 6. *For each positive integer n there exists a tree with exactly n peripherians.*

Proof. For $n = 1$ such a tree is on Fig. 2. This example shows us also that an end vertex of a tree need not be its peripherian. For $n = 2$ it is an arbitrary path, because it contains two end vertices and one of them can be mapped on the other by an automorphism. For $n \geqq 3$ it is the star graph with n edges.

Evidently every tree must have at least one peripherian.

Theorem 7. *The peripherian of a tree never coincides with its median,* *with the exception of the trivial cases of a tree consisting of one isolated vertex and of a tree consisting of one edge with its end vertices.*

Proof follows immediately from Lemma 1 and Theorem 5.

Theorem 8. *For the periphery vertex deviation of a tree G the inequalities*

$$2 - \frac{3}{v_v} \leqq \overline{m}_1(G) \leqq \frac{v_v - 1}{2}$$

hold. The equality $\overline{m}_1(G) = 2 - \frac{3}{v_v}$ *occurs in the case of a star graph, the equality* $m_1(G) = \frac{v_v - 1}{2}$ *occurs in the case of a path.*

Proof is analogous to the proof of Theorem 4.

Corollary 5. *For the periphery edge deviation of a tree G the inequalities*

$$\frac{3}{2} - \frac{1}{v_e} \leqq \overline{M}_1(G) \leqq \frac{v_e}{2}$$

hold. The equality $\overline{M}_1(G) = \frac{3}{2} - \frac{1}{v_e}$ *occurs in the case of a star graph, the equality* $M_1(G) = \frac{v_e}{2}$ *occurs in the case of a path.*

Proof follows immediately from Theorem 8 and Theorem 1. We see that upper bounds are the same in Theorem 8 and in this corollary. Applying Theorem 1 we may easily show that the vertex deviation in a vertex of a tree can be equal to the edge deviation if and only if it is equal to $v_e/2$.

At the end we note that the median of a tree need not coincide with its center. It is shown on Fig. 3, where a_1 is the median and c is the center of the tree.

REFERENCE

[1] Ore O., *Theory of Graphs*, Providence 1962.

Katedra matematiky
Vysoká škola strojní a textilní, Liberec
Czechoslovakia
Department of Mathematics