

Notes on betweenness centrality of a graph

S. Gago ^{a*}, J. Hurajová and T. Madaras ^{b †}

May 19, 2009

^a Dept. Matemàtica Aplicada IV, EPSC, Universitat Politècnica de Catalunya
Av. Canal Olímpic s/n, 08860 Castelldefels, Barcelona (Spain)
email: `sgago@ma4.upc.edu`

^b Institute of Mathematics,
Jesenná 5, 04001 Košice, Slovakia
email: `jana.hurajova@student.upjs.sk`
`tomas.madaras@upjs.sk`

Abstract

The betweenness centrality of a vertex of a graph is the portion of shortest paths between all pairs of vertices passing through that vertex. We study selected general properties of this invariant and its relations to distance parameters (diameter, mean distance); also, there are studied properties of graphs whose vertices have the same value of betweenness centrality.

1 Introduction

Throughout this paper, we consider connected finite graphs without loops or multiple edges; we will use standard graph theory terminology by [5].

*Research supported by the Ministry of Science and Technology (Spain) under project TEC2005-03575 and MTM2008-06620-C03-01/MTM

†This work was supported by Science and Technology Assistance Agency under the contract No. APVV-0007-07.

Among the applications of graph theory the analysis of complex networks is the one which has an important place, as it describes relations between real-world objects of various nature. As a part of such an analysis, there are investigated objects that play the key role within a network. The measure of importance of objects is usually expressed in terms of so called centrality indices. In [9], Freeman introduced several such indices; the one with the most frequent use is the betweenness centrality which can be described, using the graph terminology, as the portion of shortest paths of a graph that pass through a selected vertex. Apart of the use in real network analysis, the attention is recently focused also on its mathematical properties. In [4], there are studied the Laplacian spectral bounds of betweenness and its relation to other invariants of graphs like diameter or mean distance. The paper [7] deals with estimations on betweenness of vertices in bipartite graphs. In [8], the betweenness of vertices within their ego networks (that is, within graphs induced on their neighbourhoods) is discussed as well as its correlation with betweenness of vertices in random graphs and real-world networks. The concept of betweenness was also defined for edges and applied in clustering algorithms for determining the communities within social networks [13].

The aim of these notes is to continue the study of selected properties of betweenness centrality and its relations with other graph invariants. The Section 2 contains basic definitions and properties of betweenness. In Section 3, we study estimations of the average vertex betweenness of general and triangle-free graphs. We also explore the structure of triangle-free graph having a vertex with the betweenness reaching the theoretical minimum. The influence of selected graph operations (like edge or vertex addition) on the average betweenness is discussed in Section 4. The final Section 5 deals with betweenness-selfcentric graphs (that is, the graphs having the same betweenness for each vertex). We give several constructions for such graphs; besides we show that each strongly regular graph possesses this property (which partially answers the question in [14] on the existence of betweenness-selfcentric graphs with trivial automorphism group). Furthermore, we show that the only betweenness-selfcentric graphs with universal vertex are complete graphs.

2 Basic definitions and properties

Given a connected graph $G = (V, E)$, n and m denote the number of its vertices and edges, respectively. The set of all neighbors of a vertex x is $N(x)$, and the degree of x is $d(x) = |N(x)|$. The average degree of G is the ratio $\bar{d} = 2m/n$. We denote by δ and Δ the minimum and the maximum degree of the vertices of the graph, respectively. Given two vertices $u, v \in V$, $d(u, v)$ stands for their distance (that is, the length of a shortest u - v -path) in G . The complement of the graph G is denoted by \bar{G} and the graph constructed by taking t disjoint copies of G is denoted by tG .

For any two vertices $u, v \in V$, $d(u, v)$ denotes their distance (that is, the length of a shortest u - v -path) in G ; the mean distance \bar{l} of G is equal to $\frac{1}{n(n-1)} \sum_{(u,v) \in V^2} d(u, v)$. By $\sigma_{u,v}$ we denote the number of all shortest u - v -paths, and by $\sigma_{u,v}(x)$ the number of all shortest u - v -paths passing through a vertex $x \neq u, v$. Let $b_{u,v}(x) = \sigma_{u,v}(x)/\sigma_{u,v}$, then the *betweenness* $B(x)$ of a vertex x is defined as $\sum_{(u,v) \in V^2} b_{u,v}(x)$ (note that, if u and v belong to different components of G , then the corresponding term is not included in the sum).

The betweenness of the graph G is $\bar{B}(G) = \frac{1}{n} \sum_{x \in V} B(x)$.

We note that other authors (cf. [9]) consider the above defined sum for unordered pairs of vertices. Our definition matches the one given in [4] and allows to transfer the concept of betweenness to directed graphs (see also [?]). Throughout this paper, all sums involving betweenness are considered as being summed over ordered pairs of vertices from a given set.

Using the definition of betweenness, we obtain that $B(x) \leq (n-1)(n-2)$ and the equality holds if and only if $G \cong K_{1,n-1}$ ([9]). Another useful formula that describes the connection of betweenness and mean distance is contained in the following

Theorem 1 ([4]). $\bar{B}(G) = (n-1)(\bar{l}-1)$.

Proof. Given two vertices u, v of G at a distance $d = d(u, v)$, the sets (called layers, [4]) $P_{u,v}^h = \{w \in V \mid d(u, w) = h, d(w, v) = d - h\}$, for any $0 \leq h \leq d$ (where $P_{u,v}^0 = \{u\}$ and $P_{u,v}^d = \{v\}$). Let $P_{u,v} = \bigcup_{h=0}^d P_{u,v}^h$. Considering the layers $P_{u,v}^h$, $1 \leq h \leq d-1$, all the shortest paths from u to v cross through all the vertices of each layer, thus $\sum_{w \in P_{u,v}^h} b_{u,v}(w) = 1$, $1 \leq h \leq d-1$. For any other vertex of the graph $w \notin P_{u,v}^h$, there is no shortest path from u to

v crossing through w , so $b_{u,v}(w) = 0$. Therefore

$$\sum_{w \in V} b_{u,v}(w) = \sum_{w \in P_{u,v}} b_{u,v}(w) = \sum_{h=0}^d \sum_{w \in P_{u,v}^h} b_{u,v}(w) = \sum_{h=1}^{d-1} 1 = d - 1. \quad (1)$$

Thus we obtain

$$\bar{B}(G) = \frac{1}{n} \sum_{w \in V} B(w) = \frac{1}{n} \sum_{w \in V} \sum_{u,v \in V} b_{u,v}(w) = \frac{1}{n} \sum_{u,v \in V} \sum_{w \in V} b_{u,v}(w),$$

and applying (1) we get $\bar{B}(G) = \frac{1}{n} \sum_{u,v \in V} (d(u,v) - 1) = (n-1)(\bar{l} - 1)$. \square

Another useful tools in determining the betweenness of vertices are graph symmetries. It is not hard to see (details are left to reader) that every two vertices which belong to the same orbit under the action of the automorphism group $Aut(G)$ on a graph G have the same value of betweenness centrality index; hence, vertex-transitive graphs possess the same value for each vertex.

3 Estimations based on local parameters

In this section, we will study bounds on betweenness of a graph which are based on parameters of neighbourhoods of its vertices. The following estimations of particular sums will be useful in the next:

Proposition 1.

$$a) \quad \sum_{x \in V(G)} \frac{1}{n - d(x)} \geq \frac{n}{n - \bar{d}}$$

$$b) \quad \sum_{x \in V(G)} d^2(x) \geq 2m\bar{d}.$$

Proof. a) By the inequality on the harmonic and arithmetic means,

$$\frac{n}{\sum_{x \in V(G)} \frac{1}{n - d(x)}} \leq \frac{\sum_{x \in V(G)} (n - d(x))}{n} = \frac{n^2 - 2m}{n},$$

which yields

$$\frac{\sum_{x \in V(G)} \frac{1}{n-d(x)}}{n} \geq \frac{n}{n^2 - 2m} = \frac{1}{n - \bar{d}},$$

and the result follows.

b) By the Cauchy-Schwartz inequality,

$$\left(\sum_{x \in V(G)} d(x) \cdot 1 \right)^2 \leq \left(\sum_{x \in V(G)} 1^2 \right) \cdot \left(\sum_{x \in V(G)} d^2(x) \right),$$

hence

$$(2m)^2 \leq n \sum_{x \in V(G)} d^2(x),$$

which implies the result. \square

In the following, we present several local estimations of betweenness of a vertex based on its neighbourhood. Let u, v be two nonadjacent vertices of G having a common neighbour x . By $M_{u,v}$ we denote the set of vertices $y \in V(G)$ such that $uy \notin E(G)$ or $vy \notin E(G)$ (note that $u, v \in M_{u,v}$), and by $N_{u,v}(x)$ the set of the vertices of $M_{u,v}$ which are neighbours of x ; we put $\mu_{u,v} = |M_{u,v}|, \nu_{u,v}(x) = |N_{u,v}(x)|$.

Lemma 1. $B(x) \geq \sum_{u,v \in N(x)} \frac{1}{n - \nu_{u,v}(x)}$.

Proof. Let $u, v \in N(x), uv \notin E(G)$. Then $\sigma_{u,v}(x) = 1$ and $b_{u,v}(x) = 1/\sigma_{u,v}$. Moreover, $\sigma_{u,v} \leq n - \mu_{u,v} \leq n - \nu_{u,v}(x)$, so $b_{u,v}(x) \geq 1/(n - \nu_{u,v}(x))$ and $B(x) = \sum_{\substack{(u,v) \in V^2 \\ u,v \neq x}} b_{u,v}(x) \geq \sum_{\substack{u,v \in N(x) \\ uv \notin E(G)}} b_{u,v}(x) \geq \sum_{\substack{u,v \in N(x) \\ uv \notin E(G)}} \frac{1}{n - \nu_{u,v}(x)}$. \square

As a consequence, given a vertex x of a triangle-free graph G and its neighbours $u, v, \nu_{u,v}(x) = d(x)$, we obtain

Corollary 1. *If G is a triangle-free graph, then, for every vertex $x \in V(G)$,*

$$\frac{d(x)(d(x) - 1)}{n - d(x)} \leq B(x).$$

Based on these results, we obtain several estimation for the betweenness of a triangle-free graph:

Theorem 2. *If G is a triangle-free graph, then $\overline{B}(G) \geq \frac{\delta(\delta-1)}{n-\bar{d}}$.*

Proof. In each inequality of Proposition 1, we can use the lower bound $\delta(\delta-1)$ for the numerator. Now, summing the obtained inequalities for all vertices of G and multiplying this sum by $\frac{1}{n}$, we have

$$\frac{\delta(\delta-1)}{n} \sum_{x \in V(G)} \frac{1}{n-d(x)} \leq \frac{1}{n} \sum_{x \in V(G)} B(x) = \overline{B}_G.$$

By Proposition 1a)

$$\frac{\delta(\delta-1)}{n} \cdot \frac{n}{n-\bar{d}} \leq \frac{\delta(\delta-1)}{n} \sum_{x \in V(G)} \frac{1}{n-d(x)}.$$

and the result follows. \square

Theorem 3. *If G is a triangle-free graph, then $\overline{B}(G) \geq \frac{1}{n-\delta} \left(\bar{d}^2 - \frac{n}{2} \right)$.*

Proof. Using the lower bound on $B(x)$ from Proposition 1 and the estimation on the denominator, we obtain $\frac{d(x)(d(x)-1)}{n-\delta} \leq B(x)$. Now,

$$\sum_{x \in V(G)} d(x)(d(x)-1) = \sum_{x \in V(G)} d^2(x) - \sum_{x \in V(G)} d(x) = \sum_{x \in V(G)} d^2(x) - 2m.$$

As G is triangle-free, $m \leq \frac{n^2}{4}$ by Turán theorem. Finally, using Proposition 1 b), we obtain

$$\overline{B}(G) \geq \frac{1}{n(n-\delta)} \sum_{x \in V(G)} d^2(x) \geq \frac{1}{n(n-\delta)} \left(2m\bar{d} - \frac{n^2}{2} \right) = \frac{1}{n-\delta} \left(\bar{d}^2 - \frac{n}{2} \right).$$

\square

Theorem 4. *If G is a triangle-free graph, then $\overline{B}(G) \geq \frac{n(n-1)}{n-\bar{d}} - n - \bar{d} + 1$.*

Proof. Summing the inequalities from Proposition 1 and dividing with n , we obtain

$$\begin{aligned} \overline{B}(G) &\geq \frac{1}{n} \sum_{x \in V(G)} \frac{d(x)(d(x) - 1)}{n - d(x)} = \frac{1}{n} \sum_{x \in V(G)} \left(-d(x) + 1 - n + \frac{n(n - 1)}{n - d(x)} \right) = \\ &\frac{1}{n} \left(- \sum_{x \in V(G)} d(x) + (1 - n) \sum_{x \in V(G)} 1 + n(n - 1) \sum_{x \in V(G)} \frac{1}{n - d(x)} \right) \geq \\ &\frac{1}{n} \left(-2m + (1 - n)n + n(n - 1) \frac{n}{n - \overline{d}} \right) = -\overline{d} + 1 - n + \frac{n(n - 1)}{n - \overline{d}}. \end{aligned}$$

□

In the next, we consider graphs such that $B(x)$ attains the lower bound from Proposition 1 for some vertex x of G . Note that such a graph is not necessarily triangle-free: the graph on Fig. 3 is vertex-transitive, hence, for each vertex x , $B(x)$ is the same and equal to $2 = 3(3 - 1)/(6 - 3)$.

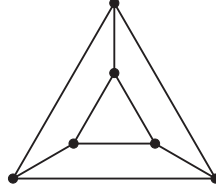


Fig.1

Nevertheless, if we restrict our study to triangle-free graphs, we obtain the complete characterization:

Lemma 2. *Let G be a triangle-free graph. Then the following statements are equivalent:*

- a) *there exists a vertex $x \in V(G)$ such that $B(x) = \frac{d(x)(d(x) - 1)}{n - d(x)}$,*
- b) *G is isomorphic to a complete bipartite graph.*

Proof. Suppose first that $G \cong K_{r,s}$. Let R, S be bipartitions of G such that $|R| = r$, $|S| = s$ and let $x \in R$. Then

$$B(x) = \sum_{\substack{(u,v) \in V^2 \\ u,v \neq x}} b_{u,v}(x) = \sum_{(u,v) \in S^2} b_{u,v} + \sum_{\substack{(u,v) \in R^2 \\ u,v \neq x}} b_{u,v}(x) + \sum_{\substack{u \in R, v \in S \\ u \neq x}} b_{u,v}(x).$$

If $(u, v) \in S^2$, then $d(u, v) = 2$, $\sigma_{u,v} = r$ and $\sigma_{u,v}(x) = 1$; hence, $\sum_{(u,v) \in S^2} b_{u,v}(x) = \sum_{(u,v) \in S^2} \frac{1}{r} = \frac{s(s-1)}{r}$. If $u \in R$ and $v \in S$, then $d(u, v) = 1$, $\sigma_{u,v} = 1$ and $\sigma_{u,v}(x) = 0$ which yields $\sum_{\substack{u \in R, v \in S \\ u \neq x}} b_{u,v}(x) = 0$. Finally, if $(u, v) \in R^2$, $u, v \neq x$, then $d(u, v) = 2$ and $\sigma_{u,v}(x) = 0$, thus $\sum_{\substack{(u,v) \in R^2 \\ u,v \neq x}} b_{u,v}(x) = 0$. So, we obtain

$$B(x) = \frac{s(s-1)}{r} = \frac{d(x)(d(x)-1)}{n-d(x)}.$$

Conversely, suppose that $B(x) = \frac{d(x)(d(x)-1)}{n-d(x)}$ for some $x \in V(G)$. As $B(x) = \sum_{u,v \in N(x)} b_{u,v}(x) + \sum_{u,v \notin N(x)} b_{u,v}(x) + \sum_{u \in N(x), v \notin N(x)} b_{u,v}(x)$, this implies that $\sum_{u,v \in N(x)} b_{u,v}(x) \leq B(x)$. Since G is triangle-free, $\sigma_{u,v}(x) = 1$ for $u, v \in N(x)$, thus $\sum_{u,v \in N(x)} b_{u,v}(x) = \sum_{u,v \in N(x)} \frac{1}{\sigma_{u,v}}$. Besides for $u, v \in N(x)$ it also

holds $\sigma_{u,v} \leq n-d(x)$. Therefore $\sum_{u,v \in N(x)} \frac{1}{\sigma_{u,v}} \geq \sum_{u,v \in N(x)} \frac{1}{n-d} = \frac{d(d-1)}{n-d}$.

Thus $\frac{d(x)(d(x)-1)}{n-d(x)} = B(x) = \sum_{u,v \in N(x)} \frac{1}{\sigma_{u,v}}$. This equality holds if and only

if $\sigma_{u,v} = n-d(x)$ for every $u, v \in N(x)$, which means that u and v are adjacent to every other vertex $y \notin N(x)$. In addition, both sets $N(x)$ and $V(G) \setminus N(x)$ are independent, which together implies that G is a complete bipartite graph. \square

Note that from this proof, it follows that, if there is a single vertex of a triangle-free graph that reaches theoretical minimum betweenness bound, then all vertices do so.

4 Betweenness and local graph operations

In this section we will study the influence of local graph operations on the betweenness of a graph.

Proposition 1. *Let G be an n -vertex graph and G' be the graph obtained from G by connecting two vertices $u, v \in V(G)$ at distance $d = d(u, v) > 1$ with a new edge. Then*

$$\bar{B}(G') \leq \bar{B}(G) - \frac{2(d-1)}{n}.$$

Proof. Since $d_G(u, v) = d$, $d_{G'}(u, v) = 1$ and $d_{G'}(x, y) \leq d_G(x, y)$ for $(x, y) \neq (u, v)$, we obtain that $\sum_{(x,y) \in V^2} d_{G'}(x, y) \leq \sum_{(x,y) \in V^2} d_G(x, y) - 2(d-1)$. From

this we obtain

$$\frac{\sum_{(x,y) \in V^2} d_{G'}(x, y)}{n(n-1)} \leq \frac{\sum_{(x,y) \in V^2} d_G(x, y)}{n(n-1)} - \frac{2(d-1)}{n(n-1)}$$

$$\bar{l}(G') \leq \bar{l}(G) - \frac{2(d-1)}{n(n-1)}$$

$$\bar{l}(G') - 1 \leq (\bar{l}(G) - 1) - \frac{2(d-1)}{n(n-1)}$$

$$(n-1)(\bar{l}(G') - 1) \leq (n-1)(\bar{l}(G) - 1) - \frac{2(d-1)}{n}$$

which yields the desired statement by Lemma 1. □

Corollary 1. *Let $G' = (V', E')$ be a factor subgraph of a graph $G = (V, E)$ and let $r = |E \setminus E'|$. Then*

$$\bar{B}(G) \leq \bar{B}(G') - \frac{2r}{n}.$$

Proof. Adding the edges of $E \setminus E'$ one by one to G , using Proposition 1 in each step, the betweenness of obtained graph is decreased by $\frac{2}{n}$ or more, thus, the total drop is at least $\frac{2r}{n}$. □

Theorem 1. Let $G = (V, E)$ be a graph of order n and size e , and let T be one of its spanning trees. Then

$$\overline{B}(G) \leq \overline{B}(T) - \frac{2(e - n + 1)}{n}.$$

Proof. Any spanning tree is a factor subgraph of G with $n - 1$ edges, so, applying Proposition 1 with $r = e - n + 1$, we get the result. \square

Corollary 2. Let G be a hamiltonian graph of order n and size e . Then

$$\overline{B}(G) \leq \begin{cases} \frac{(n^2 - 4n + 4)}{4} - \frac{2(e - n)}{n} & \text{for } n \text{ even,} \\ \frac{(n^2 - 4n + 3)}{4} - \frac{2(e - n)}{n} & \text{for } n \text{ odd.} \end{cases}$$

Proof. Let C_n be the hamiltonian cycle of G . We calculate the betweenness of C_n and then apply Corollary 1 to find the bounds. As C_n is vertex-transitive, $B(w) = \overline{B}(G)$ for all w of C_n , so we just need to calculate $B(w)$ for only one vertex.

If n is odd ($n = 2k + 1$), let u_i be the vertex at a distance i on the left of w and let v_j be the vertex at a distance j on the right of w (in a clockwise orientation of C_n). If $i + j \leq k$ we will have to cross through w if we go from u_i to v_j , so $b_{u_i v_j}(w) = b_{v_j u_i}(w) = 1$. Then $B(w) = \sum_{i=1}^{k-1} \sum_{j=1}^{k-i} 2 = 2 \sum_{i=1}^{k-1} (k - i) = k(k - 1)$.

Otherwise, if n is even ($n = 2k$), each vertex of the graph has one vertex at a maximum distance k , and two possible paths to get to it. The contribution of the vertices that are not at maximum distance from w will be equal to the ones of the odd case and we just have to subtract $\frac{1}{2}$ for the vertex at the maximum distance, so $B(w) = k(k - 1) - 2(k - 1)/2 = (k - 1)^2$. \square

We note that these bounds are tight for graphs with many edges and few vertices (dense graphs). Using the relation between the betweenness centrality and the mean distance of a graph given by the Theorem 1, the bounds are also useful for the mean distance of a graph. They also improve the bounds for the mean distance of Doyle in [6].

The influence of adding a vertex to a graph on the betweenness is more complicated. We discuss first the addition of a pendant vertex:

Proposition 2. *Let G be a graph of order n , G' the graph obtained by connecting a new vertex v to a vertex $w \in G$ with degree $d(w) = d$. If $\overline{B}(G) \leq 4(n-1) - 2d$ then $\overline{B}(G) \leq \overline{B}(G')$.*

Proof. Let $B(w)$ and $B'(w)$ be the betweenness of a vertex w in G and G' , respectively. Put $B'_{n+1} = B'_v$ and $B'_n = B'_w$. The vertex v has degree 1, then $B'_{n+1} = B'_v = 0$. How much B'_w is affected? All the shortest paths between w and the $n-1$ vertices of the graph cross through w , so $B'_w = B'_n = B_n + 2(n-1)$. By the way, considering all the δ_w neighbors of w , the new B'_n will be increased in $2(n-1-\delta_w)$ as all the shortest paths between w and the rest of the vertices of G will cross through all these vertices. We can not know how the rest of B_w will increase, but we can insure that $\sum_{w=0}^{n-1-\delta_w} B'_w \geq \sum_{w=0}^{n-1-\delta_w} B_w$. Then $\sum_{w=0}^n B'_w = \sum_{w=0}^{n-1-\delta_w} B'_w + \sum_{w=n-1-\delta_w}^{n-1} B'_w + B'_n \geq \sum_{w=0}^n B_w + 2(n-1-\delta_w) + 2(n-1)$, and so

$$\overline{B}_{G'} = \frac{\sum_{w=0}^n B'_w}{n+1} + \frac{B'_{n+1}}{n+1} \geq \frac{n\overline{B}_G + 4(n-1) - 2\delta_w}{n+1}.$$

Observe that the larger the degree of the vertex to which we connect the new vertex, the lower the bound. The increment of the betweenness will be

$$\overline{B}_{G'} - \overline{B}_G \geq [-\overline{B}_G + 4(n-1) - 2\delta_w]/(n+1) \geq 0.$$

Therefore the right side of the inequality is positive if $\overline{B}_G \leq 4(n-1) - 2\delta_w$. \square

Proposition 3. *Let G be a graph of order n , let G' be the graph obtained connecting a new vertex w to two vertices u, v of G at a distance $d(u, v) = 1$ or $d(u, v) = 2$, then*

$$\frac{1}{n+1}[n\overline{B}_G + 2(n-2)] \leq \overline{B}_{G'}.$$

Proof. The proof is similar to the one of the previous proposition:

If $d(u, v) = 1$, the three vertices form a triangle, so $B_w = 0$. The sum of the new betweenness of u and v will be $B'_u + B'_v = B_u + B_v + 2(n-1)$, since all the shortest paths connecting w to the rest of the graph pass through u and v . Also the betweenness of the neighbors of u and v will increase, although we do not know how much, so $\frac{1}{n+1}[n\overline{B}_G + 2(n-2)] \leq \overline{B}_{G'}$.

If $d(u, v) = 2$, we call i_1, \dots, i_s the s intermediate vertices connecting u and v . In G we have $b_{i_i}(u, v) = 1/s$ and $\sum_{i=1}^s b_{i_i}(u, v) = 1$, then in G'

$w = i_{s+1} \Rightarrow b_{i_l}(u, v) = 1/(s+1)$ and $\sum_{l=1}^{s+1} b_{i_l}(u, v) = \sum_{l=1}^{s+1} 1/(s+1) = 1$, but $B'_u + B'_v = B_u + B_v + 2(n-2)$ as before, therefore with the same reasoning as the other case $\frac{1}{n+1}[n\bar{B}_G + 2(n-2)] \leq \bar{B}_{G'}$. \square

For $d = 3$ the \bar{B}_G can increase or decrease, depending on the graph.

Theorem 2. *Let T_n be a tree of order n , w a vertex of degree $d = d(w) > 1$, Δ the maximum degree of the graph, and m_1, \dots, m_d be the sizes of the branches of T_n (with respect to w). Then*

1. $B(w) = \sum_{i,j=1, i \neq j}^d m_i m_j$ for every $w \in V(T_n)$,
2. $B_T^{max} \leq (n - \Delta)(\Delta - 1)(n - 1)^2/n\Delta$ where the upper bound is reached for a tree with a root vertex of degree Δ and with all its branches of the same size.

Proof. In a tree, every pair of vertices is connected by unique path. Thus, if u, v belong to different components of $T_n - w$, then $b_{u,v}(w) = 1$. We obtain then $B(w) = \sum_{i=1}^d m_i(n - m_i) = \sum_{i,j=1, i \neq j}^d m_i m_j = f(m_1, \dots, m_d) = f(m_1, \dots, m_d)$.

Now, we calculate the maximum of this function under the condition $m_1 + \dots + m_d = n - 1$. Assume first that there exist $i, j \in \{1, \dots, d\}$ such that $m_i \geq m_j + 2$; without loss of generality, let $i = 1, j = 2$. Set $m'_1 = m_1 - 1, m'_2 = m_2 + 1, m'_i = m_i$ for $i \geq 3$. Now $f(m'_1, \dots, m'_d) = \sum_{\substack{i,j=3 \\ i \neq j}}^d m_i m_j + \sum_{j=3}^d m'_1 m_j + \sum_{i=3}^d m_i m'_2 + m'_1 m'_2 = \sum_{\substack{i,j=1 \\ i \neq j}}^d m_i m_j - \sum_{j=3}^d m_j + \sum_{i=3}^d m_i - m_2 + m_1 - 1 = f(m_1, \dots, m_d) + m_1 - m_2 - 1 > f(m_1, \dots, m_d)$. Thus, $f(m_1, \dots, m_d)$ is maximized in the case when $|m_i - m_j| \leq 1$ for all $i, j = 1, \dots, d$.

Next, consider the above maximization problem for positive real numbers x_1, \dots, x_d . In this case, we can use the Lagrange multipliers formula with the auxiliary function $F(x_1, \dots, x_d, \lambda) = \sum_{i=1}^d m_i \left(\sum_{j \neq i} x_j \right) - \lambda \left(\sum_{i=1}^d x_i - n + 1 \right)$. Solving the system formed by its partial derivatives equal to 0, and isolating λ , we obtain $\lambda = 2 \sum_{j \neq i} m_j = n - 1 - m_i \forall i = 1, \dots, d$. As λ is the same for all equations, all x_i are equal too, so the maximum is reached when $x_i = (n - 1)/d$ and its value is $f^*\left(\frac{n-1}{d}, \dots, \frac{n-1}{d}\right) = (n - 1)^2(d - 1)/d$. This value is then also an upper bound for $f(m_1, \dots, m_d)$, so,

$$B(w) \leq (n - 1)^2(d - 1)/d \leq (n - 1)^2(\Delta - 1)/\Delta.$$

Now recall that T_n contains at least Δ pendant vertices which have zero betweenness. Thus, for the betweenness of T , we obtain

$$\overline{B}_{T_n} = \frac{1}{n} \sum_{w \in V(T_n)} B(w) \leq \frac{1}{n} \sum_{\substack{w \in V(T_n) \\ d(w) \neq 1}} (n-1)^2 \frac{d(w) - 1}{d(w)} \leq \frac{1}{n} (n-\Delta)(n-1)^2 \frac{\Delta - 1}{\Delta}.$$

and this bound is reached for a tree with a root vertex of degree Δ and with all its branches of the same size. \square

Corollary 3. *Let G be a graph of order n , size e and maximum degree of vertices Δ , then*

$$\overline{B}_G \leq B_G^{max} \leq (n-1)^2 \frac{(\Delta-1)}{\Delta} - \frac{2(e-n+1)}{n}.$$

Proof. Considering any spanning tree containing as a root vertex the vertex of maximum degree, and applying Proposition 1 and (1), we get the result. \square

Now we found a bound for the maximum betweenness of a tree and we generalize it for a general graph G .

Theorem 3. *Let G be a graph of order n , $w \in V$ of eccentricity e_w and $\Gamma_k(w)$ the set of vertices at a distance k from w . If we connect a new vertex u to w , then*

$$\overline{B}_{G'} = \frac{1}{n+1} \left[n \overline{B}_G + 2 \sum_{k=1}^{e_w} k |\Gamma_k(w)| \right].$$

note 1. *We note that the second summand depends on the eccentricity of the vertex w and its number of extremal vertices, that is, the more vertices at extremal distance the vertex w has, the more \overline{B}_G will be increased.*

Proof. Given a vertex $w \in V$, we denote by $n_k = |\Gamma_k(w)|$, and we denote by B_i and B'_i the betweenness of the vertex i in G and G' respectively. Connecting a new vertex u to w , all the shortest paths that go from u to the vertices at a distance $l > k$, will cross through the vertices of $\Gamma_k(w)$. For this reason their betweenness will be increased as $\sum_{v \in \Gamma(k)} B'_v = \sum_{v \in \Gamma(k)} B_v + 2 \sum_{i > k} n_i$, $1 \leq k \leq e_k - 1$. The betweenness of the extremal vertices are not affected. Adding the betweenness of all these sets

$$\sum_{k=0}^{e_w} \sum_{v \in \Gamma(k)} B'_v = \sum_{k=0}^{e_w} \sum_{v \in \Gamma(k)} B_v + 2 \sum_{k=0}^{e_w-1} \sum_{i > k} n_i.$$

Finally we divide all by $n + 1$ and get the result. \square

Theorem 4. *Let T_n be a tree of n vertices and diameter D , then*

- *If $n - D$ is odd, $\bar{B}_{T_n} \leq 1 + \frac{(n-4)D}{2} - \frac{D^3 - 6D^2 - D + 6}{6n}$.*
- *If $n - D$ is even, $\bar{B}_{T_n} \leq 1 + \frac{(n-4)D}{2} - \frac{D^3 - 6D^2 + 2D}{6n}$.*

Proof. Using Theorem 3 we are going to construct a tree with maximum \bar{B} . We start from a path P_{D+1} of diameter D and connect the $n - D - 1$ other vertices without incrementing the diameter, in such a way that the total betweenness would be the maximum. For constructing the tree, we consider the two vertices with maximal eccentricity (apart from the ends): u_2 and u_D . Both have the same number of vertices at a maximum distance (1 in this moment). We connect a new vertex to one of them. Then the second vertex to be added must be connected to the other one, because that vertex would have the maximum eccentricity and also the maximum number of vertices at extremal distance (2 at the moment). The third vertex can be connected again to one of these vertices u_2 or u_D , but the fourth would have to be connected to the other one, for the same reason.

Following this procedure we just have to add the $n - D - 1$ vertices to the second and the D vertex of the path. The betweenness of the tree depends on two cases:

$$\text{If } n - D = 2k + 1, \quad \bar{B}_{T_n} = \frac{1}{n} \sum_{i=2}^D 2(k-1+i)(n-k-i) + \frac{2k(k+1)}{n}.$$

$$\text{If } n - D = 2k + 2, \quad \bar{B}_{T_n} = \frac{1}{n} \sum_{i=2}^D 2(k+i)(n-k-1-i) + \frac{2(k+1)^2}{n}.$$

Simplifying these sums we get the result. \square

Corollary 4. *Let G be a graph with n vertices, e edges and diameter $D > 1$, then*

$$\text{If } n - D \text{ is odd, } \bar{l} \leq \frac{n}{(n-1)} + \frac{(n-4)D}{2(n-1)} - \frac{D^3 - 6D^2 - D + 6}{6n(n-1)} - \frac{2(e-n+1)}{n(n-1)}.$$

$$\text{If } n - D \text{ is even, } \bar{l} \leq \frac{n}{(n-1)} + \frac{(n-4)D}{2(n-1)} - \frac{D^3 - 6D^2 + 2D}{6n(n-1)} - \frac{2(e-n+1)}{n(n-1)}.$$

We notice that these bounds are reached for star graphs.

Example 1. If we consider the graph of the Figure 3, $n = 9$, $e = 15$, $D = 3$ and its mean distance is $\bar{l} = 1.75$.

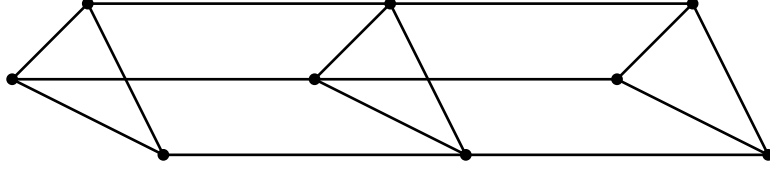


Figure 1: $K_3 \times P_3$

The known bounds for the mean distance and the bound of the Corollary 4 give:

$$\text{Mohar [12]} \quad \bar{l} \leq \frac{n}{n-1} \left\lceil \frac{\Delta + \theta_2}{4\theta_2} \ln(n-1) \right\rceil = 3.375,$$

$$\text{Kouider-Winkler [17]} \quad \bar{l} \leq \frac{n}{\delta+1} + 2 = 4.25,$$

$$\text{Beezer et al. [1]} \quad \bar{l} \leq \frac{1}{n(n-1)} \left\lceil \frac{(n+1)n(n-1) - 2e}{\delta+1} \right\rceil = 2.38,$$

$$\text{Yebra et al. [18]} \quad \bar{l} \leq \frac{1}{n-1} \sum_{Q_k(0) \leq n-1} \left\lceil \frac{n(n-1)}{Q_k^2(0) + n-1} \right\rceil = 2.25,$$

$$\text{Corollary 4} \quad \bar{l} \leq 1.91.$$

5 Betweenness-selfcentric graphs

In this section, we will study graphs having the same value of betweenness for all their vertices (that is, *betweenness-selfcentric* graphs). Clearly, any vertex-transitive graph is betweenness-selfcentric; on the other hand, we show that there are many betweenness-selfcentric graphs which are not transitive. Namely, we show that this property holds for wide class of strongly regular graphs (recall that a graph G is *strongly regular* with parameters (n, k, λ, μ)

if G is k -regular of order n , every pair of adjacent vertices has λ common neighbours and every pair of nonadjacent vertices has μ common neighbours).

We first prove a lemma for graphs of diameter 2:

Lemma 3. *Let G be a graph of diameter 2. Then, for every vertex $x \in V(G)$,*

$$B(x) = \sum_{\substack{u,v \in N(x) \\ uv \notin E(G)}} \frac{1}{\sigma_{u,v}}.$$

Proof. By definition,

$$B(x) = \sum_{\substack{u,v \in V(G) \\ u,v \neq x}} \frac{\sigma_{u,v}(x)}{\sigma_{u,v}} = \sum_{u,v \in N(x)} \frac{\sigma_{u,v}(x)}{\sigma_{u,v}} + \sum_{u,v \notin N(x)} \frac{\sigma_{u,v}(x)}{\sigma_{u,v}} + \sum_{\substack{u \in N(x), \\ v \notin N(x)}} \frac{\sigma_{u,v}(x)}{\sigma_{u,v}}.$$

Now, if both $u, v \notin N(x)$, then each u - v -path containing x has length at least 4; since $d(u, v) \leq 2$, we get $\sigma_{u,v}(x) = 0$. If $u \in N(x), v \notin N(x)$ and $uv \in E(G)$, then trivially $\sigma_{u,v}(x) = 0$; if $uv \notin E(G)$, then $d(u, v) = 2$ but there is no u - v -path passing through x , so again $\sigma_{u,v}(x) = 0$. Hence, in the expression for $B(x)$, the second and third sums are equal to zero, leaving

$$B(x) = \sum_{u,v \in N(x)} \frac{\sigma_{u,v}(x)}{\sigma_{u,v}}.$$

Since $\sigma_{u,v}(x) = 1$ for nonadjacent $u, v \in N(x)$ and 0 for adjacent ones, we are done. \square

Corollary 2. *Every strongly regular graph is betweenness-selfcentric.*

Proof. Let G be a strongly regular graph with parameters (n, k, λ, μ) and $x \in V(G)$. Every edge xy belongs to exactly λ triangles; this implies that there is $\frac{k\lambda}{2}$ edges between vertices of $N(x)$. Note that each strongly regular graph has diameter 2. Then, by previous lemma, we obtain that the sum

$$B(x) = \sum_{\substack{u,v \in N(x) \\ uv \notin E(G)}} \frac{1}{\sigma_{u,v}}$$

consists of $k(k-1) - k\lambda = k(k-\lambda-1)$ summands through ordered pairs of vertices; now, for each nonadjacent $u, v \in N(x)$, $\sigma_{u,v} = \mu$, thus $B(x) = \frac{k}{\mu}(k-\lambda-1)$. Hence, G is betweenness-selfcentric. \square

By [?], each finite group may serve as the automorphism group of some strongly regular graph, so, we obtain

Corollary 3. *For each finite group A , there exists a betweenness-selfcentric graph G such that $\text{Aut}(G) \cong A$.*

Little is known about betweenness-selfcentric nonregular graphs; we present one class of such graphs:

Lemma 4. *For all integers $t \geq 2, r \geq 1$, the graph $\overline{tK_{1,r}}$ is betweenness-selfcentric.*

Proof. It is easy to see that $\overline{tK_{1,r}}$ has diameter 2. Let x be a vertex of $\overline{tK_{1,r}}$. By Lemma 3, $B(x) = \sum_{\substack{u,v \in N(x) \\ uv \notin E(G)}} \frac{1}{\sigma_{u,v}}$. In this sum, each pair u, v originates from the same copy of $K_{1,r}$ (which does not contain x) in $tK_{1,r}$ before complementation in the way that one of u, v is the central vertex; we obtain that $\sigma_{u,v}(x) = 1$ and $\sigma_{u,v} = (t-1)(r+1)$, hence, $B(x) = (t-1)r \cdot \frac{1}{(t-1)(r+1)} = \frac{r}{r+1}$. \square

As all these graphs have diameter 2, it would be interesting to construct, for each integer $d > 2$, an infinite family of betweenness-selfcentric nontransitive graphs of diameter d .

References

- [1] R.A. Beezer, J.E. Riegsecker, B.A. Smith, Using minimum degree to bound average distance, *Discrete Math.* **226** (2001), 365-371.
- [2] F. Buckley, L. Superville, Mean distance in line graphs, *Congr. Numer.* **32** (1981) 153-162.
- [3] F. Buckley, L. Superville, Distance distributions and mean distance problems, *Proceedings of the Third Caribbean Conference on Combinatorics and Computing*, Bridgetown, 1981, pp. 67-76.
- [4] F. Comellas, S. Gago, Spectral bounds for the betweenness of a graph, *Linear Algebra Appl.* **423** (2007) 7480.
- [5] R. Diestel, *Graph Theory*, Springer, New York, 1997
- [6] J.K. Doyle, J.E. Graver, Mean distance in a graph, *Discrete Math.* **17** (1977), 147-154.

- [7] M.G. Everett, P. Sinclair, P.A. Dankelmann Some centrality results new and old *Journal of. Mathematical Sociology*, **28**(4) (2004), 215-227
- [8] M.G. Everett, S. P. Borgatti, Ego networks, *Social Networks* **27** (2005), 3138.
- [9] L. C. Freeman, A set of measures of centrality based on betweenness, *Sociometry* **40** (1977), 35.
- [10] K.I. Goh, B. Kahng, D. Kim, Universal behavior of load distribution in scale-free networks, *Phys. Rev. Lett.* **87** (2001), 278701.
- [11] H. Hong, B.J. Kim, M.Y. Choi, H. Park, Factors that predict better synchronizability on complex networks, *Phys. Rev. E* **69** (2004), 067105.
- [12] B. Mohar, Eigenvalues, diameter and mean distance in graphs, *Graphs. Combin.* **7** (1991), 53-64.
- [13] M.E.J. Newman, M. Girvan, Finding and evaluating community structure in networks, *Phys.Rev.E.* **69** (2004), 026113.
- [14] K. Phelps, Latin square graphs and their automorphism groups, *Ars Combinatoria* **7** (1979), 273-299.
- [15] J. Plesnik, On the sum of all distances in a graph or digraph, *J. Graph Theory* **8** (1984), 1-21.
- [16] D.J. Watts, S.H. Strogatz, Collective dynamics of 'small-world' networks, *Nature* **393** (1998), 440-442.
- [17] M. Kouider, P. Winkler, Mean distance and minimum degree, *J. Graph Theory* **25** (1997), 95-99.
- [18] J.A. Rodriguez, J.L.A. Yebra, Bounding the diameter and the mean distance of a graph from its eigenvalues: Laplacian versus adjacency matrix methods, *Discrete Math.* **196** (1999), 267-275.