The betweenness centrality of a graph [*]

S. Gago

Dept. Matemàtica Aplicada IV, EPSC, Universitat Politècnica de Catalunya

Av. Canal Olímpic s/n, 08860 Castelldefels, Barcelona (Spain)

email: sgago@ma4.upc.edu

Abstract

A measure of the centrality of a vertex of a graph is the portion of shortest paths crossing through it between other vertices of the graph. This is called betweenness centrality and here we study some of its general properties, relations with distance parameters (diameter, mean distance), local parameters, symmetries, etc. Some bounds for this parameter are obtained, using them to improve all the known bounds for the mean distance of the graph.

1 Introduction

Freeman introduced a set of centrality indices for social networks [1]. One of them is the betweenness centrality of a vertex, that gives us an idea of the importance of the vertex in a social network. Recently other authors have introduced the same concept for edges [2], studying its distribution in complex networks and using it for finding communities in them. Other authors showed that is strongly related with the distribution of load in an interconnection network, and that it is an important factor to predict synchronizability of dynamic networks. The strong relation of the parameter for vertices with the mean distance and other known parameters of the graph is the object of our study. Laplacian spectral bounds of the betweenness centrality for vertices and for edges and relations with other indices of networks have been studied by the author et al. in [3].

2 Definitions and first properties

Let $u, v \in V(G)$, if $\sigma_{uv}(w)$ denotes the number of shortest paths (geodetic paths) from vertex u to vertex v that go through w, and σ_{uv} is the total number of geodetic paths from u to v, then

---

and so the betweenness of vertex w is $B_w = \sum_{u,v \neq w} b_{uv}(w)$. The (vertex) betweenness of a graph $G = (V, E)$ of order n is $\overline{B}_G = (\sum_{w \in V} B_w)/n$ and the maximum (vertex) betweenness of G is $B_G^{max} = \max\{B_w \mid w \in V\}$.

We can find some properties of this parameter for a graph G of order n:

1. $0 \leq b_{uv}(w) \leq 1 \quad \forall \, u, v, w \in V$,

2. $0 \leq B_G^{min} \leq \overline{B}_G \leq B_G^{max} \leq (n-1)(n-2) \quad \forall \, w \in V$, reaching the upper bound for the central vertex of a star graph,

3. $b_{uv}(w) = b_{vu}(w) \quad \forall \, u, v, w \in V$,

4. if $\delta_w = 1$ then $B_w = 0$.

Symmetries in graphs are also useful to simplify the calculation of these parameters. For instance, if π is an automorphism of G and $u, v \in G$ are two vertices such that $\pi(u) = v$, then all the shortest paths crossing for u will be the same of the ones $B_u = B_v$. Therefore, in a vertex-transitive graph G the group of automorphisms of the graph $Aut(G)$ acts transitively on $V(G)$, so $B_u = B_G^{max} = \overline{B}_G, \quad \forall u \in V$. For two isomorphic graphs, G_1 and G_2 there exists a bijective mapping Φ between the vertices of G_1 and G_2, such that $\Phi(G_1) = G_2$. Then all the shortest paths of one graph will be equal to the ones of the other, so $B_u = B_{\Phi(u)}$ for all $u \in V_1$, and $\overline{B}_{G_1} = \overline{B}_{G_2}$.

3   Relations with distance parameters

The idea of centrality in a graph is related with the idea of distance. A vertex will be more central if we have to cross through it going to other vertices of the graph. In this section we study a first approximation of the betweenness to some known distance parameters of the graph as the mean distance, the diameter, the radius...

Given two vertices u and v of G at a distance $d = d(u, v)$, we consider the following sets called lawyers $P_{uv}^h = \{w \in V \mid \ d(u, w) = h, \ \ d(w, v) = d - h\}, \quad 0 \leq h \leq d$, introduced in [3], and their union as $P_{uv} = \bigcup_{h=0}^d P_{uv}^h$. Of course $P_{uv}^0 = \{u\}$ and $P_{uv}^d = \{v\}$. Considering the lawyers P_{uv}^h, $1 \leq h \leq d-1$, all the shortest paths from u to v cross through all the vertices of each lawyer, thus $\sum_{w \in P_{uv}^h} b_{uv}(w) = 1, \quad 1 \leq h \leq d-1$. For any other vertex of the graph $w \notin P_{uv}^h$, there is no shortest path from u to v crossing through w, so $b_{uv}(w) = 0$. Then

$$\sum_{w \in V} b_{uv}(w) = \sum_{w \in P_{uv}} b_{uv}(w) = \sum_{h=0}^{d} \sum_{w \in P_{uv}^h} b_{uv}(w) = \sum_{h=1}^{d-1} 1 = d - 1. \tag{1}$$

We can use this to prove the following theorem:

2

$$\overline{B}_G = (n-1)(\bar{l}-1).$$

Proof.

$$\overline{B}_G = \frac{1}{n}\sum_{w\in V} B_w = \frac{1}{n}\sum_{w\in V}\sum_{u,v\in V} b_{uv}(w) = \frac{1}{n}\sum_{u,v\in V}\sum_{w\in V} b_{uv}(w),$$

and applying (1)

$$\overline{B}_G = \frac{1}{n}\sum_{u,v\in V}(d(u,v)-1) = (n-1)(\bar{l}-1).$$

$\square$

The diameter of the graph also give us two lower bounds for these parameters:

Theorem 2. *Let G be a graph of order n, $w \in V(G)$ with eccentricity $e(w)$, then*

$$\frac{e(w)(e(w)-1)(e(w)-2)}{3n} \leq \overline{B}_G, \qquad \frac{e(w)(e(w)-2)}{2} \leq B_G^{max}.$$

Proof. Since the eccentricity of w is $e(w)$, there is at least one vertex of the graph at a distance e_w from w, and so there is at least a path of $e(w)$ vertices contained into the graph $P_{e(w)} \subseteq G$. The idea of the proof is based into calculate the betweenness of $P_{e(w)}$ and compare it with the betweenness of the graph G. Let $V(P_{e(w)}) = \{w_1, \dots, w_{e(w)}\}$ be the set of vertices of the path. First we note that $B_{w_1} = B_{w_{e(w)}} = 0$. For any other vertex w_k we suppose that there are $k-1$ vertices on its left and $e(w) - k$ on its right. Thus $B_{w_k} = 2(k-1)(e(w)-k)$ and

$$\overline{B}_{P_{e(w)}} = \frac{1}{e(w)}\sum_{k=2}^{e(w)-1} 2(k-1)(e(w)-k) = (e(w)-1)(e(w)-2)/3.$$

Therefore the betweenness of the whole graph G will be bigger than this value

$$\overline{B}_G = \frac{1}{n}\sum_{w\in V} B_w \geq \frac{1}{n}\sum_{w_k\in P_{e(w)}} B_{w_k} = e(w)(e(w)-1)(e(w)-2)/3n.$$

By the same way, we can deduce the bound for B_G^{max} from the $B_{P_{e_w}}^{max}$. To calculate the second, we can consider $B_{w_k} = 2(k-1)(e(w)-k) = -2k^2 + 2(e(w)+1)k - 2e(w) = f(k)$ as a function of k, and look for its maximum. The first derivative is $f'(k) = -4k + 2(e(w)+1) = 0$, then $k_0 = \frac{e(w)+1}{2}$. The second derivative when $k = k_0$ is $f''(k_0) = -4 < 0$, so k_0 is a maximum. If $e(w)$ is odd, $k = (e(w)+1)/2 \in \mathbb{Z}$, then $B_G^{max} \geq B_{P_{e(w)}}^{max} = (e(w)-1)^2/2$. And if $e(w)$ is even, $k = (e(w)+1)/2 \notin \mathbb{Z}$, the maximum is reached when $k_0 = e(w)/2$ or $k_0 = e(w)/2 + 1$ and its value is $B_G^{max} \geq B_{P_{e(w)}}^{max} = e(w)(e(w)/2 - 1)$. $\square$

Corollary 1. *Let G be a graph of order n, diameter D and mean distance $\bar{l}$, then*

$$1 + \frac{D(D-1)(D-2)}{3n(n-1)} \leq \bar{l}.$$

Now we are going to study the relation of the betweenness with a local parameter of the vertices of the graph called the clustering coefficient. This parameter was introduced by Watts and Strogatz [4] as a measure of the connectivity of the neighbourhood of the vertices of the graph. The clustering coefficient of a vertex C_u, is the fraction of the number of edges connecting the neighbors of a vertex n_u among the total number of possible edges between them, that is, if u has degree δ_u the clustering parameter of the vertex u is $C_u = \dfrac{2n_u}{\delta_u(\delta_u - 1)}$. We can also define the clustering parameter of the graph G as the average of the clustering parameters of its vertices $\overline{C} = \dfrac{1}{n}\sum_{u \in V} C_u$. The last one also give us an idea about the number of triangles of the graph. For instance, the unique graph with $\overline{C} = 1$ is the complete graph K_n, and a graph with $\overline{C} = 0$ is a triangle-free graph.

We can find some results relating this parameter to the betweenness.

Lemma 1. *Let G be a graph of order n, $w \in G$, then $B_w = 0 \Leftrightarrow C_w = 1$.*

Proof. If $B_w = 0$ there are no shortest paths connecting two vertices of the graph containing w, thus w has degree 1 or all the neighbors of w must be connected, and therefore if and only if $C_w = 1$. $\square$

Proposition 1. *Let $G = (V, E)$ be a graph of order n and size e, $w \in V$ a vertex of degree $\delta_w > 1$ and $C_w = 0$, then*
$$\frac{\delta_w(\delta_w - 1)}{n - \delta_w} < B_w.$$

Proof. Since $C_w = 0$ the vertex w has δ_w neighbors no connected between them. Let u and v be two of these neighbors, that is the edges $e_{uw} \in E$, $e_{vw} \in E$, then $b_{uv}(w) = \sigma_{uv}(w)/\sigma_{uv} \geq 1/\sigma_{uv}$. On the other hand $\sigma_{uv} \leq n - \delta_w$, since the worse thing that could happen for going from u to v is that we would have to cross through the rest of the vertices of the graph. Hence $b_{uv}(w) \geq 1/\sigma_{uv} \geq 1/(n - \delta_w)$ and $\sum_{e_{uw} \in E, e_{vw} \in E} b_{uv}(w) \geq \delta_w(\delta_w - 1)/(n - \delta_w)$. Finally
$$B_w \geq \sum_{e_{uw} \in E, e_{vw} \in E} b_{u,v}(w) + \sum_{e_{uw} \notin E, e_{vw} \notin E} b_{u,v}(w) \geq \frac{\delta_w(\delta_w - 1)}{n - \delta_w}.$$
$\square$

Now applying this result we obtain two lower bounds for the betweenness of triangle-free graphs:

Corollary 2. *Let G be a triangle-free graph of order n, let Δ be the maximum degree of the vertices of the graph and δ the minimum degree, then*
$$\frac{\delta(\delta - 1)}{n - \Delta} \leq \overline{B}_G \leq B_G^{max}.$$

4

vertices,

$$\frac{r(r-1)}{n-r} \le \overline{B}_G \le B_G^{max}.$$

We can get another result for vertices w such that $C_w \ne 0$:

Proposition 2. *Let G be a graph of order n, $w \in V(G)$ a vertex of degree δ_w, then*

- *if $\delta_w > 1$,* $\quad 2(n - \delta_w - 1) \le \sum_{i=1}^{\delta_w} B_{u_i}.$

- *if $\delta_w = 1$,* $\quad 2(n-2) \le B_{u_1}.$

Proof. All the shortest paths connecting w with the rest of the $n - \delta_w - 1$ vertices of the graph cross through its neighbors $u_1, \ldots, u_{\delta_w}$. Then $B_{u_i} \ge \sum_{j=\delta_w}^{n-1} b_{wu_j}(u_i)$ and

$$\sum_{i=1}^{\delta_w} B_{u_i} \ge \sum_{i=1}^{\delta_w} \sum_{j=\delta_w}^{n-1} b_{wu_j}(u_i) = 2(n - \delta_w - 1).$$

$\square$

Corollary 4. *Let n_1 be the number of vertices of degree 1 of the graph G,*

$$\frac{2(n-2)n_1}{n} \le \overline{B}_G.$$

Corollary 5. *If the graph G is vertex-transitive of degree r,*

$$\frac{2(n-r-1)}{r} \le B_i = B_G^{max} = \overline{B}_G.$$

4   Some Bounds

In this section we are going to study what happens with the betweenness and the maximum betweenness of the graph when we make some operations like connecting two vertices with an new edge or connecting a new vertex with one or more vertices of the graph. This will help us to find some bounds for the betweenness in some particular cases of graphs like trees, hamiltonian, etc.

4.1   Adding a new edge

Proposition 3. *Let G be a graph of order n and G' the graph obtained connecting two vertices $u, v \in V(G)$ at distance $d = d(u, v) > 1$ with a new edge, then*

$$\overline{B}_{G'} \le \overline{B}_G - 2(d-1)/n.$$

5

path between u and v. We denote by $b_w(u,v)$ and $b'_w(u,v)$ the betweennes of w in G and G'.

By previous result (1) we know that $\sum_{i=1}^{h} b_{w_i}(u,v) = 2 - 1 = 1$. Connecting u and v with a new edge, all the shortest paths that cross through those vertices $w_1, \ldots, w_h$ are eliminated thus $\sum_{i=1}^{h} b'_{w_i}(u,v) = \sum_{i=1}^{h} b'_{w_i}(v,u) = 0$ and $\sum_{w \in V(G')} B'_w \leq \sum_{w \in V(G)} B_w - 2$. By the same reasoning, the rest of the shortest paths between two vertices of the graph containing u and v will not pass through $w_1, \ldots, w_h$, thus $\sum_u B'_u \leq \sum_u B_u - 2$ and $\overline{B}_{G'} \leq \overline{B}_G - 2/n$.

If the vertices u, v are at a distance $d(u,v) = d > 2$, following a similar reasoning with the intermediate vertices we get that $\sum_u B'_u \leq \sum_u B_u - 2(d-1)$, and so on. $\qquad \square$

Corollary 6. *Let $G' = (V', E')$ be a maximal connected subgraph of a graph $G = (V, E)$ such that $V' = V$, $E' \subset E$, and let $m = | E \setminus E' |$, then*

$$\overline{B}_G \leq \overline{B}_{G'} - 2m/n.$$

Proof. The result can be proved by applying Proposition 3 in m steps. $\qquad \square$

Theorem 3. *Let $G = (V, E)$ be a graph of order n and size $e > n$, and let T be one of its spanning trees, then*

$$\overline{B}_G \leq \overline{B}_T - 2(e - n)/n.$$

Proof. Any spanning tree is a maximal connected subgraph of G with n edges, so applying Proposition 3 with $m = e - n$ we get the result. $\qquad \square$

Corollary 7. *Let G be a hamiltonian graph of order n and size $e > n$, then*

- $\overline{B}_G \leq (n^2 - 4n)/4 - 2(e - n)/n$ *if n is even.*

- $\overline{B}_G \leq (n^2 - 4n + 3)/4 - 2(e - n)/n$ *if n is odd.*

Proof. Let C_n be the hamiltonian cycle containing all the vertices of G. We can calculate the betweenness of C_n and then apply the last result to find the bounds. As C_n is vertex-transitive, $B_G^{max} = \overline{B}_G = B_w$ for all $w \in V$, 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. 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 $B_w = \sum_{i=1}^{k-1} \sum_{j=1}^{k-i} 2 = 2 \sum_{i=1}^{k-1} (k-i) = k(k-1)$.

- 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

add $\frac{1}{2}$ for the vertex at the maximum distance, so $B_w = k(k-1) + 2(k-1)/2 = k^2 - 1$.

Finally we just have to apply Proposition 3 again to get the result. $\qquad\square$

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 [5].

4.2   Adding a new vertex

Proposition 4. *Let G be a graph of order n, G' the graph obtained connecting a new vertex v to a vertex $w \in G$ with degree δ_w, then $\overline{B}_G \leq \overline{B}_{G'}$.*

Proof. We denote as B_w the betweenness of a vertex before adding the new vertex and B'_w the betweenness of a vertex afterwards, and we define $B'_{n+1} = B'_v$ 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

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

Supposing that $2n - 5/2 \leq \delta_w$ and noting that $\delta_w \leq n-1$, $2n - 5/2 \leq n-1 \Leftrightarrow n \leq 3/2$. Therefore, for $n \geq 1$ the betweenness increases. $\qquad\square$

Proposition 5. *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*

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

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 $B'_u + B'_w = B_u + B_w + 2(n-1)$, 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 $i_1, \ldots, i_s$ the s intermediate vertices connecting u and v. In G we have $b_{i_l}(u, v) = 1/s$ and $\sum_{l=1}^{s} b_{i_l}(u, v) = 1$, 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 $\frac{1}{n+1}[n\overline{B}_G + 2(n-2)] \leq \overline{B}_{G'}$. $\qquad \square$

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

Theorem 4. *Let T_n be a tree of order n, w a vertex of degree $\delta_w > 1$, Δ the maximum degree of the graph, and $m_1, \ldots, m_{\delta_w}$ the size of the branches of T_n (with respect to w), then*

1. $B_w = \sum_{i,j=1, i \neq j}^{\delta_w} m_i m_j$ *for every $w \in V(T_n)$,*

2. $B_T^{max} \leq (n-1)^2(\Delta - 1)/\Delta$, *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. To go from the m_i vertices of one branch of the three to the others $n - m_i$ vertices we have to cross through w, so $B_w = \sum_{i=1}^{\delta_w} m_i \cdot (n - m_i) = \sum_{i,j=1, i \neq j}^{\delta_w} m_i m_j = f(m_1, \ldots, m_{\delta_w})$. Now we can apply the Lagrange multipliers formula to calculate the maximum of this function under the condition $m_1 + \cdots + m_{\delta_w} = n - 1$. The auxiliar function is $F(m_1, \ldots, m_{\delta_w}, \lambda) = \sum_{i=1}^{\delta_w} m_i \left(\sum_{j \neq i} m_j \right) - \lambda \left(\sum_{i=1}^{\delta_w} m_i - n + 1 \right)$. So we have to solve the system formed by its partial derivatives equal to 0, and isolate λ

$$\lambda = 2 \sum_{j \neq i} m_j = n - 1 - m_i \qquad \forall i = 1, \ldots, \delta_w.$$

As λ is the same for all the equations, the m_i must be all equal too, so the maximum will be reached when $m_i = (n-1)/\delta_w$ and its value will be $f\left(\frac{n-1}{\delta_w}, \ldots, \frac{n-1}{\delta_w}\right) = (n-1)^2(\delta_w - 1)/\delta_w$. Finally considering the value of this maximum as a function of δ_w, we note that the function is increasing, thus $B_T^{max} \leq (n-1)^2(\Delta - 1)/\Delta$. The bound will be reached for a tree with a root vertex of degree Δ and with all its branches of the same size. $\qquad \square$

Corollary 8. *Let G be a graph of order n, size e and maximum degree of vertices Δ, then*

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

Proof. Considering any spanning tree containing as a root vertex the vertex of maximum degree, and applying Proposition 3 and (1), we get the result. $\qquad \square$

graph G.

Theorem 5. *Let G be a graph of order n, $w \in V$ of excentricity 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*

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

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 $\sum_{v\in\Gamma(k)} B'_v = \sum_{v\in\Gamma(k)} B_v + 2\sum_{i>k} n_i$, $\quad 1 \le k \le e_k - 1$. 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. $\qquad\square$

Theorem 6. *Let T_n be a tree of n vertices and diameter D, then*

- *If $n - D$ is odd,* $\quad \overline{B}_{T_n} \le 1 + \dfrac{(n-4)D}{2} - \dfrac{D^3 - 6D^2 - D + 6}{6n}.$

- *If $n - D$ is even,* $\quad \overline{B}_{T_n} \le 1 + \dfrac{(n-4)D}{2} - \dfrac{D^3 - 6D^2 + 2D}{6n}.$

Proof. Using Theorem 5 we are going to construct a tree with maximum $\overline{B}$. 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.

9

D vertex of the path. The betweenness of the tree depends on two cases:

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

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

Simplifying these sums we get the result. $\qquad\square$

Corollary 9. *Let G be a graph with n vertices, $e > n$ edges and diameter D, then*

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

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

Proof. $\qquad\square$

Example 1. *If we consider the graph of the Figure 1, $n = 9$, $e = 15$, $D = 3$ and its mean distance is $\bar{l} = 1.75$. As $n - D = 6 = 2k + 2$ we apply the second bound of the Corollary 9, obtaining $\bar{l} \leq 1.75$. Therefore this is an example of graph where the upper bound is reached.*
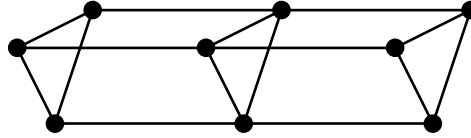

Figure 1: $K_3 \times P_3$

Example 2. *For the graph of Figure 2 , $n = 10$, $e = 11$, $D = 7$ and its mean distance is $\bar{l} = 3.022$. As $n - D = 3$ we apply the first bound of the Corollary 9, obtaining $\bar{l} \leq 3.33$. The bound of Mohar [11] gives $\bar{l} \leq 15$, and the bounds of Kouider et al. [8] and Beezer et al. [10] are $\bar{l} \leq 5.33$ and $\bar{l} \leq 3.57$ respectively.*
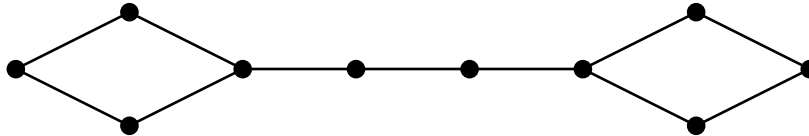

Figure 2: Example 2.

[1] L. C. Freeman, A set of measures of centrality based on betweenness, Sociometry **40** (1977), 35.

[2] M.E.J. Newman, M. Girvan, Finding and evaluating community structure in networks, Phys.Rev.E. **69** (2004), 026113.

[3] F. Comellas, S. Gago, Spectral bounds for the betweenness of a graph, Linear Algebra Appl. **423** (2007) 7480.

[4] D.J. Watts, S.H. Strogatz, Collective dynamics of 'small-world' networks, Nature **393** (1998), 440-442.

[5] J.K. Doyle, J.E. Graver, Mean distance in a graph, Discrete Math. **17** (1977), 147-154.

[6] K.I. Goh, B. Kahng, D. Kim, Universal behavior of load distribution in scale-free networks, Phys. Rev. Lett. **87** (2001), 278701.

[7] H. Hong, B.J. Kim, M.Y. Choi, H. Park, Factors that predict better synchronizability on complex networks, Phys. Rev. E **69** (2004), 067105.

[8] M. Kouider, P. Winkler, Mean distance and minimum degree, J. Graph Theory **25** (1997), 95-99.

[9] J. Plesnik, On the sum of all distances in a graph or digraph, J. Graph Theory **8** (1984), 1-21.

[10] R.A. Beezer, J.E. Riegsecker, B.A. Smith, Using minimum degree to bound average distance, Discrete Math. **226** (2001), 365-371.

[11] B. Mohar, Eigenvalues, diameter and mean distance in graphs, Graphs. Combin. **7** (1991), 53-64.

[12] 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.

[13] F. Buckley, L. Superville, Mean distance in line graphs, Congr. Numer. **32** (1981) 153-162.