



Also available at http://amc-journal.eu ISSN 1855-3966 (printed edn.), ISSN 1855-3974 (electronic edn.) ARS MATHEMATICA CONTEMPORANEA 12 (2017) 261–270

Relative edge betweenness centrality

Damir Vukičević

Department of Mathematics, University of Split, Teslina 12, HR-21000 Split, Croatia

Riste Škrekovski *

FMF, University of Ljubljana, Jadranska ulica 21, 1000 Ljubljana. & Faculty of Information Studies, Ljubljanska cesta 31A, p.p. 603, 8000 Novo mesto & FAMNIT, University of Primorska, Glagoljaška 8, 6000 Koper, Slovenia

Aleksandra Tepeh *

Faculty of Information Studies, Ljubljanska cesta 31A, p.p. 603, 8000 Novo mesto & Faculty of Electrical Engineering and Computer Science, University of Maribor, Smetanova 17, 2000 Maribor, Slovenia

Received 3 June 2015, accepted 29 August 2016, published online 9 December 2016

Abstract

We introduce a new edge centrality measure - relative edge betweenness $\gamma(uv) = b(uv)/\sqrt{c(u)c(v)}$, where b(uv) is the standard edge betweenness and c(u) is the adjusted vertex betweenness. In this alternative definition, the importance of an edge is normalized with respect to the importance of its end-vertices. This gives a better presentation of the "local" importance of an edge, i.e. its importance in the near neighborhood. We present sharp upper and lower bounds on this invariant together with the characterization of graphs attaining these bounds. In addition, we discuss the bounds for various interesting graph families, and state several open problems.

Keywords: Centrality measures, betweenness centrality, social networks.

Math. Subj. Class.: 05C35

^{*}This work is supported in part by Slovenian research agency ARRS, program no. P1–00383, project no. L1–4292, and Creative Core–FISNM–3330-13-500033.

E-mail addresses: vukicevic@pmfst.hr (Damir Vukičević), skrekovski@gmail.com (Riste Škrekovski), aleksandra.tepeh@gmail.com (Aleksandra Tepeh)

1 Introduction

One of the fundamental problems in network analysis is to determine the importance (or the *centrality*) of a particular vertex or an edge in a network. Since the 1950's many centrality indices have evolved, each with specific application and based on a different concept of what makes a vertex or an edge central to a network. One of the most common measures of the importance of an edge is its betweenness, i.e. the number of shortest paths passing through that edge (normalized in the case where multiple shortest paths between some vertices occur). More precisely, *betweenness of an edge e, b(e)*, is given by

$$b(e) = \sum_{\{k,l\} \in \binom{V}{2}} \frac{\sigma_{k,l}(e)}{\sigma_{k,l}},$$

where $\sigma_{k,l}$ denotes the number of shortest paths between vertices k and l, and $\sigma_{k,l}(e)$ denotes the number of shortest paths connecting k and l that pass through the edge e. Caporossi at. al [1] defined *adjusted vertex betweenness* of a vertex u, c(u), as the sum of the betweennesses of all edges incident to the vertex u, i.e.

$$c\left(u\right) = \sum_{v \in N(u)} b\left(uv\right),$$

where N(u) is the set of neighbors of the vertex u. In the original definition given by Freeman [2], *betweenness of a vertex* u, b(u), was defined as the number of the shortest paths that contain the vertex u as an interior vertex. It can be shown that it holds

$$c(u) = 2b(u) + n - 1.$$

Some centrality indices, e.g., degree centrality, reflect local properties of the underlying graph, while others, like betweenness centrality, give information about the global network structure, as they are based on shortest path computation and counting [7]. In [3] extremal graphs with respect to vertex betweenness for certain graph families were considered. Some recent applications of betweenness centrality include analyzing social and protein interaction networks [6, 4, 5] and traffic flow optimization [8, 9].

Note that betweenness of an edge measures only importance of a link to the entire network, and that link of the highest betweenness may be completely unimportant to some vertex such that no shortest (or even reasonably short) path from that vertex passes through this link. Hence, on the level of individuals the same link can be observed in completely different way. If we observe a social network, the existence of an edge depends on the level of importance attributed to this edge by its adjacent vertices. These vertices (actors) are the ones that create, sustain and destroy this edge (relationship). Namely, they decide how much they want to invest in their friendship. Obviously, edges with high betweenness should be valuable to both vertices considering that many information circulate through such edges.

In this paper, we are interested in measuring the relative importance of an edge to its end-vertices. To clarify the notion of the relative importance, one can use also analogy with a business venture. It is important to partners in this venture if the size of a deal is large comparing to the sizes of the companies involved. While a thousand dollars deal might be extremely important to (say) individual building contractor, it is a very small job to a large corporation. Hence, in order to estimate the value of the edge to its end vertices, one needs to normalize it using their adjusted betweennesses. We define the *relative betweenness* as

$$\gamma(uv) = \frac{b(uv)}{\sqrt{c(u)c(v)}}.$$
(1.1)

Note that we use the geometric mean between c(u) and c(v). The reason why we use the geometric mean and not the arithmetic is that the geometric mean is much lower than the arithmetic mean when there is a large difference between c(u) and c(v), say $c(u) \ll c(v)$. Hence in this case, $\gamma(uv)$ is significantly larger compared to the usage of the arithmetic mean. This corresponds to the relationship between persons with large difference in the influence. We may assume that this relationship is of great value to the actor u and hence he will put significant effort in sustaining this relationship. Therefore, we give to such edge high relative betweenness.

The aim of this paper is to measure how weak can be the weakest link and how strong may be the strongest link in the (social) network. In other words, we are interested in finding extremal values for relative edge betweenness and in analyzing distribution of relative betweenness values on edges of graphs. In Section 2 we give sharp upper and lower bound for relative betweenness in the case of general graphs, and characterize graphs for which these bounds are attained. In Section 3 the bounds for various graphs classes are discussed. We conclude the paper by stating some open problems.

2 Bounds for general graphs

Let B_n be the graph obtained from the complete bipartite graph $K_{2,n-2}$ by adding the edge that connects two vertices in the part of cardinality 2, see Fig. 1. Let this edge be denoted as e^* .



Figure 1: Graph B_n and the edge e^* .

Theorem 2.1. Let uv be an edge of a connected graph G with $n \ge 3$ vertices. Then,

$$\gamma(uv) \ge \frac{2}{n^2 - 3n + 4}.$$
(2.1)

Moreover, equality holds if and only if G is isomorphic to B_n and uv corresponds to e^* .

Proof. Let us denote by $\alpha(u)$ the sum of all edge betweennesses of edges incident to u (but not to v), and by $\alpha(v)$ the sum of all edge betweennesses of edges incident to v (but not to u). In other words, $\alpha(u) = c(u) - b(uv)$ and $\alpha(v) = c(v) - b(uv)$. Note that $\alpha(u) + \alpha(v)$ consists of three types of contributions:

- (a) those from shortest paths not having neither u nor v as an end-vertex. They contribute in total at most $2\binom{n-2}{2} = n^2 5n + 6$.
- (b) those from the shortest path with an end-vertex u, but not v. They contribute all together at most n-2.
- (c) those from the shortest path with an end-vertex v, but not u. They contribute all together at most n-2.

By (a)-(c), it holds $\alpha(u) + \alpha(v) \le n^2 - 3n + 2$. Note that $c(u) + c(v) = \alpha(u) + \alpha(v) + 2b(uv)$. Now, as $b(uv) \ge 1$, we have

$$\gamma(uv) = \frac{b(uv)}{\sqrt{c(u)c(v)}} \ge \frac{2b(uv)}{c(u) + c(v)} = \frac{2b(uv)}{\alpha(u) + \alpha(v) + 2b(uv)}$$

$$\ge \frac{2b(uv)}{n^2 - 3n + 2 + 2b(uv)} \ge \frac{2}{n^2 - 3n + 4},$$
(2.2)

which establishes inequality (2.1). Note that equality in (2.1) holds if and only c(u) = c(v), b(uv) = 1 and $\alpha(u) + \alpha(v) = n^2 - 3n + 2$. The latter implies that the numbers given in (a), (b) and (c) are in fact the exact values of contributions to $\alpha(u) + \alpha(v)$ of the paths described in (a), (b) and (c), respectively.

To assure that the shortest paths having neither u nor v as an end-vertex contribute exactly $n^2 - 5n + 6$, the contribution has to be 2 for each pair of vertices from $V(G) \setminus \{u, v\}$. But this is not the case if there exists an edge xy for $x, y \in V(G) \setminus \{u, v\}$. Since G is connected we have $V(G) \setminus \{u, v\} = N(u) \cup N(v)$. Moreover, since b(uv) = 1, every vertex from $N(u) \cup N(v)$ is adjacent to both u and v. We infer that if for an edge uv of a graph G equality holds in (2.1), then G is isomorphic to B_n and uv corresponds to e^* . \Box

Now we establish the upper bound for $\gamma(uv)$. To prove its sharpness we will consider graphs containing an edge with certain properties. An edge uv of a graph is called a *handle* if u is a pendant vertex and the set $N(v) \setminus \{u\}$ induces a clique. We will denote the edge uv by h^* , see Fig. 2. Note that the path P_n contains two handles.

Theorem 2.2. Let uv be an edge of a connected graph G with $n \ge 3$ vertices. Then,

$$\gamma(uv) \le \sqrt{\frac{n-1}{3n-5}}.$$

Moreover, equality holds if and only if G contains a handle.

Proof. Let us introduce the following notation:

(a) p_u (resp. p_v) denotes the contribution to c(u) (resp. c(v)) of all shortest paths for which both end-vertices are different from u and v, and which do not pass through the edge uv;



Figure 2: A handle h^* in a graph.

- (b) p_{uv} is the contribution to b(uv) of all the shortest paths for which both end-vertices are different from u and v;
- (c) q_{uv} (resp. q_{vu}) is the contribution to b(uv) of all paths that start in u (resp. v) and pass through the edge uv, but do not finish in v (u, respectively).

Note that $q_{uv} + q_{vu} \le n-2$ and $c(u) = p_u + 2p_{uv} + 2q_{vu} + n-1$. In c(u), the second and the third summand appear with factor 2 since each contributing path passes through the edge uv and another edge incident with u, and the last summand corresponds to all paths that start in u. Analogously, $c(v) = p_v + 2p_{uv} + 2q_{uv} + n - 1$. Further, $b(uv) = p_{uv} + q_{uv} + q_{vu} + 1$. Hence,

$$\gamma(uv) = \frac{b(uv)}{\sqrt{c(u)c(v)}} \le \frac{p_{uv} + q_{uv} + q_{vu} + 1}{\sqrt{(2p_{uv} + 2q_{vu} + n - 1)(2p_{uv} + 2q_{uv} + n - 1)}}.$$
 (2.3)

We need to prove that

$$\frac{p_{uv} + q_{uv} + q_{vu} + 1}{\sqrt{(2p_{uv} + 2q_{vu} + n - 1)(2p_{uv} + 2q_{uv} + n - 1)}} \le \sqrt{\frac{n - 1}{3n - 5}}$$

This is equivalent to

$$(p_{uv} + q_{uv} + q_{vu} + 1)^2 (3n - 5) - (2p_{uv} + 2q_{vu} + n - 1)(2p_{uv} + 2q_{uv} + n - 1)(n - 1) \le 0,$$

which is further equivalent to

$$(-1-n)p_{uv}^{2} + (-2+4n-2n^{2}+(6-2n)(n-2-q_{uv}-q_{vu}))p_{uv} + ((q_{uv}+q_{vu}+1)^{2}(3n-5) - (2q_{vu}+n-1)(2q_{uv}+n-1)(n-1)) \le 0.$$
(2.4)

It is obvious that $-n-1 \le 0$, $-2+4n-2n^2 < 0$, $6-2n \le 0$ and $n-2-q_{uv}-q_{vu} \ge 0$. Hence, it is sufficient to prove that

$$(q_{uv} + q_{vu} + 1)^2 (3n - 5) - (2q_{vu} + n - 1)(2q_{uv} + n - 1)(n - 1) \le 0.$$
(2.5)

Without loss of generality (because of the symmetry of the last equation), we may assume that $q_{uv} \ge q_{vu}$. Let us denote $s = q_{uv} + q_{vu}$ and $d = q_{uv} - q_{vu}$. Obviously, $0 \le d \le s \le n-2$. Inequality (2.5) reduces to

$$(s+1)^2(3n-5) - (s-d+n-1)(s+d+n-1)(n-1) \le 0,$$

and further to

$$(s+1)^2(3n-5) - \left((s+n-1)^2 - d^2\right)(n-1) \le 0.$$

Since $0 \le d \le s$, it is sufficient to prove that:

$$(s+1)^2(3n-5) - \left((s+n-1)^2 - s^2\right)(n-1) \le 0.$$
(2.6)

On the left-hand side we have a quadratic function in s,

$$f(s) = s^{2}(3n-5) + (-2n^{2} + 10n - 12)s - n^{3} + 3n^{2} - 4,$$
(2.7)

with quadratic coefficient 3n - 5 > 0 and roots n - 2 and $\frac{-n^2 + n + 2}{3n - 5}$. Hence, in order to prove (2.6), it is sufficient to show that $\frac{-n^2 + n + 2}{3n - 5} \leq 0$. A simple check shows that the numerator is negative and the denominator is positive. This completes the proof that $\gamma(uv) \leq \sqrt{\frac{n-1}{3n-5}}$.

Let $h^* = uv$ be an edge in a graph L_n such that d(u) = 1 and $N(v) \setminus \{u\}$ induces a clique. Then clearly $p_u = p_v = p_{uv} = q_{vu} = 0$ and $q_{uv} = n - 2$, hence the upper bound for γ is attained, i.e. $\gamma(h^*) = \sqrt{\frac{n-1}{3n-5}}$.

To prove the converse, assume G is a connected graph of order at least 3 with an edge uv such that $\gamma(uv) = \sqrt{\frac{n-1}{3n-5}}$. This implies equality in (2.3) and (2.4). From equality in (2.3) follows $p_u = p_v = 0$, and form equality in (2.4) we obtain that $p_{uv} = 0$ and that equality holds in (2.5). The latter is equivalent to the fact that $(s + 1)^2(3n - 5) - ((s + n - 1)^2 - d^2)(n - 1) = 0$. As a consequence (since $0 \le d \le s$) we have also equality in (2.6). Now, it follows that s and d coincide, and hence $s = q_{uv}$ and $q_{vu} = 0$. Equality in (2.6) implies also that s is a (positive) root of the quadratic function in (2.7), so $s = n - 2 = q_{uv}$.

To summarize, for $uv \in E(G)$ we have $p_u = p_v = p_{uv} = q_{vu} = 0$ and $q_{uv} = n - 2$. From this, observe that there is no vertex $w \in V(G) \setminus \{u, v\}$ such that $uw \in E(G)$ and $vw \notin E(G)$, otherwise we obtain a contradiction with $q_{vu} = 0$. The fact that $q_{uv} = n - 2$ implies that every shortest path from u to any other vertex in $V(G) \setminus \{u, v\}$ passes through the edge uv. Thus v is the only neighbor of u. Since G is connected and of order at least 3, $N(v) \setminus \{u\}$ is nonempty and induces a clique, otherwise we obtain a contradiction with $p_v = 0$. Thus, we infer that uv is a handle in G.

Corollary 2.3. For any edge e of a graph with $n \ge 3$ vertices, it holds that

$$\frac{2}{n^2-3n+4} \leq \gamma(e) \leq \sqrt{\frac{n-1}{3n-5}} \,.$$

3 Bounds for some graph classes

In this section, the bounds of Corollary 2.3 for various graph classes are considered.

Graphs with higher connectivity. The graphs containing handles, for which the upper bound in Corollary 2.3 is attained, belong to the class of graphs with bridges. Thus, one might wonder whether this bound can be improved if we forbid them. But it turns out that even in the case of k-connected graphs the leading term in upper bound remains essentially $\frac{1}{\sqrt{3}}$. To illustrate this, consider the graph $C_{n,k}$, constructed as follows: take a complete graph on n - k vertices, $k \ge 2$, n > 2k, choose k of its vertices, make a copy of each chosen vertex and join it with the original one, and finally add edges so that copied vertices induce a clique (see Fig. 3 where general situation is presented on the left, while the right graph is isomorphic to $C_{10,3}$). Let v be one of k chosen vertices and u its copy in the above construction of $C_{n,k}$. Then we have b(uv) = n - k, c(u) = n + k - 2, c(v) = 3n - 3k - 2, and thus $\gamma(uv) = \frac{n-k}{\sqrt{(n+k-2)(3n-3k-2)}}$.



Figure 3: A k-connected graph $C_{n,k}$ (left) and the graph $C_{10,3}$ (right).

In what follows, we discuss the bounds of the above corollary in a various interesting graph classes.

Bipartite graphs. The upper bound in Corollary 3.2 is clearly attained also when restricted to two-mode data networks (bipartite graphs). We now give an example of a bipartite graph and an edge of it that achieves asymptotically the lower bound $\Theta(n^{-2})$.

Proposition 3.1. There exist bipartite graphs on n vertices containing an edge e with $\gamma(e) \in \Theta(n^{-2})$.

Proof. To prove the claim consider the graph G from Fig. 4 constructed in the following way. Take eight independent sets A_0, A_1, \ldots, A_7 , each of size k, connect each vertex of A_i with each vertex of A_{i+1} , index being taken modulo 8. Finally, take two new adjacent vertices a_0 and a_1 , and connect each vertex of A_i with $a_i \pmod{2}$, for $i = 0, \ldots, 7$. Now, we will show that

$$b(a_0a_1) \in \Theta(1)$$
 and $c(a_0), c(a_1) \in \Theta(n^2)$,

which immediately implies that $\gamma(a_0 a_1) \in \Theta(1/n^2)$.

First, we evaluate $b(a_0a_1)$. Consider the contribution to $b(a_0a_1)$ of shortest paths that contain a_0a_1 according to their length. The edge a_0a_1 is the only path of length 1 that

contains a_0a_1 , and it contributes 1 to $b(a_0a_1)$. There are 8k paths of length 2 containing a_0a_1 and each of them contributes 1/(2k+1) to $b(a_0a_1)$. Notice that there are $8k^2$ paths of length 3 that contain a_0a_1 , and each of them contributes $1/(k^2 + 4k + 1)$ to $b(a_0a_1)$. Observe that no shortest path of length 4 or more contains a_0a_1 as the diameter of this graph is 3. Summing up all together, we obtain that $b(a_0a_1)$ is slightly less than 13.

Now, we evaluate $c(a_0)$. Note that any shortest path from a vertex in A_0 to a vertex in A_4 is of length 2 and it contributes 2 to $c(a_0)$. As these paths are k^2 , it follows that $c(a_0) \in \Theta(k^2) = \Theta(n^2)$. Similarly, we evaluate $c(a_1)$.



Figure 4: A bipartite graph from the proof of Proposition 3.1.

Trees. As the trees are bipartite graphs the same upper bound holds but regarding the lower bound we get slightly different result.

Theorem 3.2. For any edge e of a tree with $n \ge 3$ vertices, it holds that

$$\frac{1}{\sqrt{n-1}} \le \gamma(e) \le \sqrt{\frac{n-1}{3n-5}},\tag{3.1}$$

and the lower bound is attained at an edge of the *n*-star unless n = 4 and *e* is the middle edge of a 4-path, in which case $\gamma(e) = \frac{4}{7}$.

Proof. Obviously, the upper bound holds by Corollary 2.3, and is attained at any edge incident to a leaf and to a vertex of degree 2, e.g. such an edge is an end-edge of a path on *n*-vertices.

Now, we argue the lower bound. In its proof we use the following notation: for an edge $f = w_1 w_2$, as T - f has two components, we name them by $T_f(w_1)$ and $T_f(w_2)$, where the first one contains w_1 and the second contains w_2 .

Let T be a tree and e = xy its edge with minimum γ . Suppose $T_e(x)$ has a vertices, and $T_e(y)$ has b vertices; hence n = a + b.

We claim that x is of degree a. Suppose to the contrary that it is of degree strictly less than a. Then there is a leaf u of T that belongs to $T_e(x)$ and is not adjacent to x. Let f = xv be the edge the removal of which from T, separates x and u. Then u belongs to the component $T_f(v)$, and let this component have s vertices. As u and v belong to $T_f(v)$, we have $s \ge 2$.

Let T^* be the tree obtained from T by first removing u from T and then reattaching it to x. Notice that

$$c_{T^*}(x) - c_T(x) = (s-1)(n-s+1) + 1 \cdot (n-1) - s(n-s) = 2s - 2 > 0.$$

So, we have $c_T(x) < c_{T^*}(x)$. Notice that $b_T(e) = b_{T^*}(e) = ab$, and $c_T(y) = c_{T^*}(y)$. Thus, $\gamma_{T^*}(e) < \gamma_T(e)$, which is a contradiction. This establishes the claim.

Similarly we prove that y is of degree b, and hence T is a double star. Now, notice that

$$\gamma(e) = \frac{ab}{\sqrt{ab + (a-1)(n-1)}\sqrt{ab + (b-1)(n-1)}}$$

We want to prove that

$$\gamma(e) \ge \frac{1}{\sqrt{n-1}}$$

(unless the exceptional case) and this is equivalent to

$$(n-1)a^{2}b^{2} \ge (ab + (a-1)(n-1))(ab + (b-1)(n-1)).$$

As n = a + b, this equality can be rewritten into

$$((ab - a - b)(a + b - 2) - 1)(a - 1)(b - 1) \ge 0.$$
(3.2)

Notice that this inequality does not hold only if a = b = 2, but then G is a 4-path and e is its middle edge, which is the exceptional case. In all other cases, it is easy to see that (3.2) holds. Also notice that equation holds in (3.2) if a = 1 or b = 1 but in that case G is a star.

In the classes of graphs with girth 3 and 4 the lower bound of γ is asymptotically $\Theta(n^{-2})$, for the trees which are class of graphs of girth infinity, it is $\Theta(1/\sqrt{n})$. We expect that at girth 5 the following change happens.

Conjecture 3.3. In graphs G on $n \ge 3$ vertices and with girth ≥ 5 it holds $\gamma(e) \in \Omega(n^{-1})$ for every edge e.

Regarding the lower bound for trees we wonder if some finite girth may occur.

Problem 3.4. Is there any finite number g such that for graphs G with girth at least g, every edge e has $\gamma(e) \in \Omega(1/\sqrt{n})$?

References

 G. Caporossi, M. Paiva, D. Vukičević and M. Segatto, Centrality and betweenness: vertex and edge decomposition of the Wiener index, *MATCH Commun. Math. Comput. Chem.* 68 (2012), 293–302.

- [2] L. C. Freeman, A set of measures of centrality based on betweenness, *Sociometry* 40 (1977), 35–41, doi:10.2307/3033543.
- [3] J. Govorčin, J. Coroničová Hurajová, T. Madaras and R. Škrekovski, Extremal graphs with respect to vertex betweenness for certain graph families, *FILOMAT* (accepted).
- [4] M. Hahn and A. Kern, Comparative genomics of centrality and essentiality in three eukaryotic protein-interaction networks, *Mol. Biol. Evol.* 22 (2005), 803–806, doi:10.1093/molbev/msi072.
- [5] M. P. Joy, A. Brock, D. E. Ingber and S. Huang, High-betweenness proteins in the yeast protein interaction network, J. Biomed Biotechnol. 2005 (2005), 96–103, doi:10.1155/JBB.2005.96.
- [6] N. Kourtellis, T. Alahakoon, R. Simha, A. Iamnitchi and R. Tripathi, Identifying high betweenness centrality nodes in large social networks, *Soc. Netw. Anal. Min.* 3 (2013), 899–914, doi: 10.1007/s13278-012-0076-6.
- [7] M. Newman, Networks: an introduction, Oxford university press, 2010, doi:10.1093/acprof: oso/9780199206650.001.0001.
- [8] R. Puzis, Y. Altshuler, Y. Elovici, S. Bekhor, Y. Shiftan and A. Pentland, Augmented betweenness centrality for environmentally aware traffic monitoring in transportation networks, *Journal* of *Intelligent Transportation Systems* 17 (2013), 91–105, doi:10.1080/15472450.2012.716663.
- [9] A. Tizghadam and A. Leon-Garcia, Robust network planning in nonuniform traffic scenarios, *Comput. Commun.* 34 (2011), 1436–1449, doi:http://dx.doi.org/10.1016/j.comcom.2010.12.013.