involve
a journal of mathematics

Intersecting geodesics and centrality in graphs

Emily Carter, Bryan Ek, Danielle Gonzalez,
Rigoberto Flórez and Darren A. Narayan

msp

2019 vol. 12, no. 1

INVOLVE 12:1 (2019)
dx.doi.org/10.2140/involve.2019.12.31

Intersecting geodesics and centrality in graphs

Emily Carter, Bryan Ek, Danielle Gonzalez,
Rigoberto Flórez and Darren A. Narayan

(Communicated by Kenneth S. Berenhaut)

In a graph, vertices that are more central are often placed at the intersection of geodesics between other pairs of vertices. This model can be applied to organizational networks, where we assume the flow of information follows shortest paths of communication and there is a required action (i.e., signature or approval) by each person located on these paths. The number of actions a person must perform is linked to both the topology of the network as well as their location within it. The number of expected actions that a person must perform can be quantified by *betweenness centrality*. The betweenness centrality of a vertex v is the ratio of shortest paths between all other pairs of vertices u and w in which v appears to the total number of shortest paths from u to w. We precisely compute the betweenness centrality for vertices in several families of graphs motivated by different organizational networks.

1. Introduction

In a graph, vertices with higher centrality are often placed at the intersection of geodesics between other pairs of vertices. This model can be applied to organizational and social networks, where we assume the flow of information follows shortest paths of communication and there is a required action (i.e., signature or approval) by each person located on these paths.

A simple organizational structure can be designed using a binary tree. An example is shown in Figure 1.

We consider this structure where the CEO oversees a "left wing" and a "right wing". We first note the employees (E1, E2, E3, and E4) do not have to perform any actions, as they are on the periphery. The CEO will have to perform actions on any correspondence between people in different wings, for a total of 18 possible actions. Vice presidents VP1 and VP2 will have to perform actions on the correspondence between the two employees under them as well as correspondences between their two

MSC2010: 05C12, 05C82.
Keywords: betweenness centrality, shortest paths, distance.

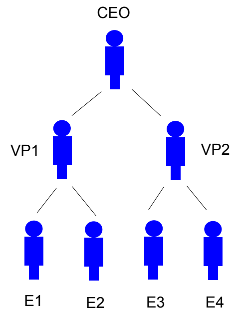

Figure 1. A binary tree organizational structure.

employees and anyone else in the company. This is a total of $2(1\times1)+2(2\times4)=18$ actions. Ironically, in this model, which would appear at first to distribute the work according to rank, the VPs actually have to perform as many actions as the CEO.

We next consider a ternary tree model where each person (except for the employees) oversees three people; see Figure 2. This slight change puts the most amount of work in the hands of the CEO. The CEO has to perform 96 actions, while each of the VPs perform 60 actions, and again the employees are not responsible for any actions.

The determination of the number actions is more complicated if the network contains cycles, since there can be multiple shortest paths, which can be used with equal probability. For our next example we will use the organizational network shown in Figure 3.

Consider the number of actions that B must make. Person B acts on communication between persons A and D and D and A. However A and D could choose to route their correspondence through person C rather than B. So B will only appear on half of the four shortest paths between A and D and D and A. (We will consider these paths to be equally likely to be followed.) Thus the total number of expected

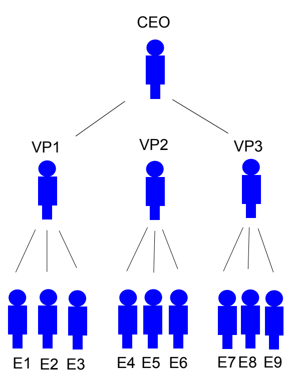

Figure 2. A ternary tree model.

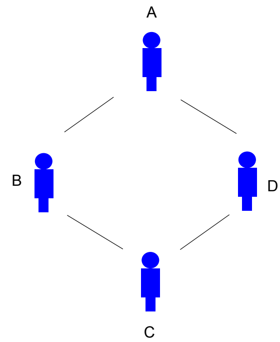

Figure 3. A network with a cycle.

actions by person B is 2, which is the same, by symmetry, for persons A, C, and D. This creates a balance of actions over every employee.

The expected number of actions can be quantified by *betweenness centrality*. This concept was introduced in [Freeman 1977] in the context of social networks. This concept has appeared frequently in both network and neuroscience literature [Brandes et al. 2016; Bullmore and Sporns 2009; Freeman et al. 1991; Guye et al. 2010; Pandit et al. 2013; White and Borgatti 1994]. The betweenness centrality of graphs was computed for various families of graphs including complete bipartite graphs, Cartesian products, wheel graphs, cocktail party graphs, ladder graphs, and cycles [Kumar and Balakrishnan 2016; Kumar et al. 2014].

In this paper, we determine the betweenness centrality for several other families of graphs motivated by organizational networks.

We first give some background with some elementary results.

Definition 1. The *betweenness centrality* of a vertex v, denoted $\mathrm{bc}(v)$, measures the frequency at which v appears on a shortest path between two other distinct vertices x and y. Let σ_{xy} be the number of shortest paths between distinct vertices x and y, and let $\sigma_{xy}(v)$ be the number of shortest paths between x and y that contain v. Then

$$\mathrm{bc}(v) = \sum_{x,y} \frac{\sigma_{xy}(v)}{\sigma_{xy}}$$

(for all distinct vertices x and y).

In our first lemma, we restate an elementary result on the lower and upper bounds of the betweenness centrality of a vertex. This was found by Gago et al. [2012] and Grassi et al. [2009].

Lemma 2. *For a given graph G with n vertices, $0 \leq \mathrm{bc}(v) \leq (n-1)(n-2)$ for all vertices v in G. Furthermore these bounds are tight.*

It is clear that if a vertex has a betweenness centrality of zero, it means that the vertex is likely to be less vital to the network than a vertex with a higher betweenness centrality. Gago et al. [2012] and Grassi et al. [2009] provided a classification for vertices to have a betweenness centrality of zero. We restate this as our next lemma. We recall that the *closed neighborhood of a vertex* is the subgraph induced by a vertex and its neighbors.

Lemma 3. *Given a vertex v, we have* $\mathrm{bc}(v) = 0$ *if and only if the closed neighborhood of v forms a complete subgraph.*

2. Betweenness centrality

We now investigate the betweenness centralities of vertices in several families of graphs, including star-like graphs, k-ary trees, complete multipartite graphs, and powers of paths and cycles. The following lemma can be implicitly found in [White and Borgatti 1994].

Lemma 4. *Let P_n be a path on vertices $v_1, v_2, \ldots, v_n$. Then* $\mathrm{bc}(v_i) = 2(i-1)(n-i)$.

Next we investigate complete multipartite graphs, making a small extension of known results for complete bipartite graphs [Kumar et al. 2014]. The complete multipartite graph $K_{n_1,n_2,\ldots,n_t}$ for $t \geq 2$ is the graph where the vertex set is partitioned into t partite sets $V_1, V_2, \ldots, V_t$ such that $|V_i| = n_i$ for each $1 \leq i \leq t$ and uv is an edge if and only if u and v belong to different partite sets.

In an application with personnel, people are divided into different groups where there are no direct connections among people in the same group, but there are direct connections between each pair of people in different groups. We give an example of a graph in Figure 4 where there are three vertices in one part, four in a second part, and five in a third part. This graph is denoted by $K_{3,4,5}$. The vertices with the highest betweenness centrality will be in the part of size 3 (since there will be the largest number of shortest paths routed through them) and the vertices with the

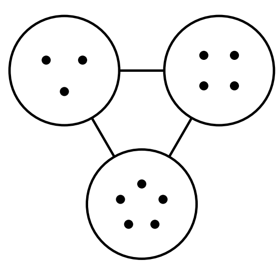

Figure 4. The complete multipartite graph $K_{3,4,5}$. The lines indicate that every vertex in one part is adjacent to every vertex in a different part.

lowest betweenness centrality will be in the part of size 5 (since they will have the smallest number of shortest paths routed through them). We explore this problem for the general class in our next lemma.

Lemma 5. *Let G be the complete multipartite graph $K_{n_1,n_2,\dots,n_t}$ where the vertices in part i are $v_{i,1}, v_{i,2}, \dots, v_{i,n_i}$ for all $1 \le i \le t$. Then for all $1 \le j \le n_i$.*

$$\mathrm{bc}(v_{i,j}) = \sum_{k=1,k\neq i}^{t} \frac{\binom{n_k}{2}}{\sum_{r=1,r\neq k}^{t} n_r}$$

Proof. We will compute $\mathrm{bc}(v_{i,j})$. Consider the shortest paths between vertices $v_{x,y}$ and $v_{x,z}$. We first determine the total number of shortest paths that contain $v_{i,j}$. Let $V_1, V_2, \dots, V_t$ represent the partite sets $K_{n_1,n_2,\dots,n_t}$. To determine the total number of shortest paths containing $v_{i,j}$ we count the number of pairs of distinct vertices in each part A_k where $k \neq i$, and divide by the number of vertices in $V(G) - A_i$. □

2.1. *Complete and balanced k-ary trees.* In a complete and balanced binary trees, there is a root vertex that is adjacent to exactly two other vertices. These vertices then have two "children" vertices. Let $k \ge 2$. A balanced k-ary with t levels will have k^i vertices at the i-th level for all $0 \le i \le t-1$. We generalize the class of trees found in the Introduction to include k-ary trees. Here there is a root vertex that has k neighbors and each of these k neighbors have k children. In balanced k-ary trees with t levels there will be k^i vertices at the i-th level for all $0 \le i \le t-1$.

We next determine the betweenness centrality of vertices in a k-ary tree.

Theorem 6. *Let G be a complete and balanced k-ary tree with levels $0, 1, \dots, t-1$. Let v_j be a vertex on level j. Then*

$$\mathrm{bc}(v_j) = -\frac{k^{t-j-1}-1}{(k-1)^2}(k - k^{t+1} - k^{t-j} + k^{t-j-1} + k^{t-j+1} - 1).$$

Proof. Consider a complete and balanced k-ary tree with levels $0, 1, \dots, t-1$. This tree will have $1 + k + k^2 + \cdots + k^{t-1} = (k^t - 1)/(k-1)$ vertices. We note that vertices in the same level will have the same betweenness centrality, so we will use v_i to denote a vertex on level i. Note that vertex v_j has k sets of $(k^{t-j} - 1)/(k-1)$ vertices beneath it. The paths that pass through v_j will either go between vertices beneath v_j in different subparts, or between any of these vertices and other vertices in the graph besides v_j. Hence

$$\begin{aligned}\mathrm{bc}(v_j) &= \Big(\frac{k^{t-(j+1)}-1}{k-1}\Big)(k-1)\Big(\frac{k^{t-(j+1)}-1}{k-1}\Big) \\ &\qquad + k\Big(\frac{k^{t-(j+1)}-1}{k-1}\Big)\Big(\frac{k^t-1}{k-1} - k\Big(\frac{k^{t-(j+1)}-1}{k-1}\Big) - 1\Big) \\ &= -\frac{k^{t-j-1}-1}{(k-1)^2}(k - k^{t+1} - k^{t-j} + k^{t-j-1} + k^{t-j+1} - 1).\end{aligned}$$

□

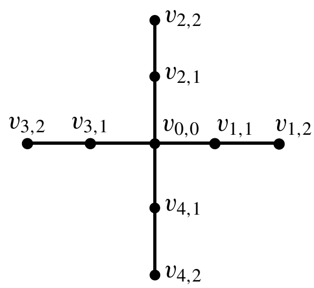

Figure 5. A star-like graph.

2.2. ***Star-like graphs.*** As shown earlier, star graphs include vertices with the highest and lowest possible betweenness centrality values. We next expand the investigation to include graphs that are obtained by subdividing the edges of the star graph $K_{1,n-1}$. These graphs will have a "center" and "pendant spokes". This model appears in a organization where communication moves along different lines that meet at a central processing person. This model is also found frequently in airports where concourses (with multiple gates along them) intersect at a common location. An example of a star-like graph is given in Figure 5.

It is clear that the center $v_{0,0}$ has the highest betweenness centrality, and the betweenness centrality of vertices will be less if they are located farther away from the center. We address the general problem in our next theorem.

Theorem 7. *Let G be a subdivided star graph with the center vertex $v_{0,0}$. Let the m paths pendant to the center have lengths $s_1, \dots, s_m$ and let $v_{l,k}$ be the k-th vertex from $v_{0,0}$ on the l-th pendant path. Then*

$$\mathrm{bc}(v_{0,0}) = 2\sum_{j=2}^{m} s_j \sum_{i=1}^{j-1} s_i \quad \textit{and} \quad \mathrm{bc}(v_{l,k}) = 2(s_l - k)\Bigl(\sum_{i \neq l} s_i + k\Bigr).$$

Proof. The center vertex will lie on optimal paths between two vertices if and only if the path connects vertices on different spokes. Thus it suffices to sum the number of pairs of vertices between spokes. Vertices on a spoke will lie on an optimal path if and only if the path is between a vertex further along the same spoke ($s_l - k$ vertices) and a vertex closer to the center or on a different spoke yielding $k + \sum_{i \neq l} s_i$ vertices. Finally we double the product to account for paths in either direction. □

Theorem 8. *Let G be a triangle graph with vertices $v_{0,0}$, $v_{1,0}$, and $v_{2,0}$ and pendant paths of lengths s_0, s_1, and s_2 incident to the three vertices, respectively. If $v_{l,k}$ is the k-th vertex on the l-th pendant path then*

$$\mathrm{bc}(v_{l,k}) = 2(s_l - k)\Bigl(k + 2 + \sum_{i \neq l} s_i\Bigr).$$

Proof. The vertex $v_{l,k}$ will be on an optimal path if and only if the path is between a vertex on the l-th pendant path further from the triangle ($s_l - k$ vertices) and a vertex closer to the triangle or on a different pendant path, giving a total of $k+2+\sum_{i\neq l} s_i$ vertices. Finally, we double this product to account for both directions. □

2.3. ***Powers of cycles and paths.*** We next consider cycles that have "redundant" connections. Consider a network of 20 people where there are direct links between adjacent people and also links between people that are spaced two apart. We will explore the betweenness centrality of this class of networks.

Recall that the k-th power of a graph G is denoted by G^k, which is defined as follows: $V(G^k) = V(G)$ and $v_i v_j \in E(G^k)$ if and only if the distance between v_i and v_j in G is less than or equal to k. Next, we investigate the betweenness centrality of vertices in powers of cycles.

Theorem 9. *Let $G = C_n^m$ with $n > 2m+1$ and let $d = \operatorname{diam}(C_n^m) = \lceil (n-1)/(2m)\rceil$. Then*

$$\begin{aligned}\operatorname{bc}(v) &= (d-1)\bigl(2\lceil \tfrac{1}{2}(n-1)\rceil - d\bigr) - (n-(n-1))\bigl(\lceil \tfrac{1}{2}(n-1)\rceil - 1\bigr)\\ &= d - \lceil \tfrac{1}{2}n - \tfrac{1}{2}\rceil - 2\lceil \tfrac{1}{2}n - \tfrac{1}{2}\rceil + 2d\lceil \tfrac{1}{2}n - \tfrac{1}{2}\rceil - d^2 + 1.\end{aligned}$$

Proof. Let $G = C_n^m$ with $n > 2m+1$. Let $r \equiv -\lceil \frac{1}{2}(n-1)\rceil \bmod m$ such that $m > r \geq 0$. Then $r = dm - \lceil \frac{1}{2}(n-1)\rceil$. The maximum number of intermediate vertices on any path is $d-1$. Let P_l be the set of shortest paths of length l where $m+1 \leq l \leq d$. The number of intermediate vertices in each shortest path is $\lceil l/m\rceil$. Let s be the number of internal vertices on a shortest path between two vertices where $1 \leq s \leq d-1$. For each path with length l, where $s = \lceil l/m\rceil - 1$ internal vertices are placed at particular locations, there exist s pair(s) of vertices where the path includes v. In the betweenness centrality this accounts for s terms equal to $1/|P_l|$. Since we can reverse any of these paths, this number is doubled. Then counting all paths of length l, the betweenness centrality for v will be

$$2|P_l| \cdot \frac{s}{|P_l|} = 2s.$$

Summing over all values of l gives

$$\sum_{l=m+1}^{\lceil (n-1)/2\rceil} 2\left(\left\lceil \frac{l}{m}\right\rceil - 1\right) = (d-1)\bigl(2\lceil \tfrac{1}{2}(n-1)\rceil - dm\bigr)$$

when n is not divisible by m. When n is divisible by m there will be two paths of the same distance between vertices that are diametrically opposite on the cycle. Hence the final term in the summation must be divided by 2, which yields

$$(d-1)\bigl(2\lceil \tfrac{1}{2}(n-1)\rceil - dm\bigr) - \left(\left\lceil \frac{\lceil \frac{1}{2}(n-1)\rceil}{m}\right\rceil - 1\right).$$

The betweenness centrality values for the two cases can be combined into a single function,

$$c(v) = (d-1)\big(2\lceil \tfrac{1}{2}(n-1)\rceil - dm\big) - \left(\left\lceil \frac{n}{m}\right\rceil - \left\lceil \frac{n-1}{m}\right\rceil\right)\left(\left\lceil \frac{\lceil \frac{1}{2}(n-1)\rceil}{m}\right\rceil - 1\right). \quad \square$$

2.3.1. *Powers of paths.* The problem of determining the betweenness centrality of vertices in a power of a path is considerably more difficult than powers of cycles. In the case of cycles, all of the vertices have the same betweenness centrality, but in a path the betweenness centrality of a vertex is dependent upon its location in the path. The m-th power of a path P_n is denoted by P_n^m with vertices $v_1, v_2, \dots, v_n$ and edges $v_i v_j$ whenever $|j-i| \le m$. For simplicity, an edge that joins two vertices where $j-i = t$ will be referred to as a t-hop.

We first consider P_n^2. We begin by defining a piecewise function, which will be used in the subsequent lemma:

$$f(i, j, k) = \begin{cases} \dfrac{j-i+1}{k-i+1} & \text{if } k-i \equiv 1 \bmod 2 \text{ and } j-i \equiv 1 \bmod 2, \\ \dfrac{k-j+1}{k-i+1} & \text{if } k-i \equiv 1 \bmod 2 \text{ and } j-i \equiv 0 \bmod 2, \\ 1 & \text{if } k-i \equiv 0 \bmod 2 \text{ and } j-i \equiv 0 \bmod 2, \\ 0 & \text{if } k-i \equiv 0 \bmod 2 \text{ and } j-i \equiv 1 \bmod 2. \end{cases}$$

Lemma 10. *If v_j is a vertex of P_n^2, then the between centrality of v_1 and v_n is zero and if $1 < j < n$, then the betweenness centrality is*

$$\mathrm{bc}(v_j) = \sum_{\substack{1<i<j \\ j<k<n}} f(i, j, k). \tag{1}$$

Proof. We prove this proposition for the case where n is even. The case where n is odd is similar. Clearly, $\mathrm{bc}(v_1) = \mathrm{bc}(v_n) = 0$. To calculate the betweenness centrality of any fixed v_j, where $1 < j < n$, we add all values given by the function $f(i, j, k)$ over all i and k, which gives (1).

For the first two cases in our piecewise function we note that since $k-i$ is odd, a shortest path between v_i and v_k must be composed of one 1-hop and $k-i$ 2-hops. In the first case we note that the 1-hop must be before the vertex v_j is reached. Since there are $j-i+1$ possible positions for the 1-hop, $\mathrm{bc}(v_j) = (j-i+1)/(k-i+1)$. In the second case the 1-hop must be after the vertex v_j is reached. Since there are $k-j+1$ possible positions for the 1-hop, $\mathrm{bc}(v_j) = (k-j+1)/(k-i+1)$. For the third and fourth cases since $k-i$ is even, any shortest path between v_i and v_k must be composed of 2-hops. When $k-j$ is even then all of these paths will contain v_j and when $k-j$ is odd then none of these paths contain v_j. $\square$

We extend this result for P_n^3 in our next lemma. We notice that there are $3^2 = 9$ cases for this step. First we define the following piecewise function. Let

$$g(i,j,k) = \begin{cases} 1 & \text{if } k-i \equiv 0 \bmod 3 \text{ and } j-i \equiv 0 \bmod 3, \\ \dfrac{(k-j+2)(k-j+5)}{(k-i+5)(k-i+2)} & \text{if } k-i \equiv 1 \bmod 3 \text{ and } j-i \equiv 0 \bmod 3, \\ \dfrac{(j-i+2)(j-i+5)}{(k-i+5)(k-i+2)} & \text{if } k-i \equiv 1 \bmod 3 \text{ and } j-i \equiv 1 \bmod 3, \\ \dfrac{2(j-i+1)(k-j+1)}{(k-i+5)(k-i+2)} & \text{if } k-i \equiv 1 \bmod 3 \text{ and } j-i \equiv 2 \bmod 3, \\ \dfrac{k-j+1}{k-i+1} & \text{if } k-i \equiv 2 \bmod 3 \text{ and } j-i \equiv 0 \bmod 3, \\ \dfrac{j-i+1}{k-i+1} & \text{if } k-i \equiv 2 \bmod 3 \text{ and } j-i \equiv 2 \bmod 3, \\ 0 & \text{otherwise.} \end{cases}$$

Lemma 11. *If v_j is a vertex of P_n^3, then the between centrality of v_1 and v_n is zero and if $1 < j < n$, then*

$$\text{bc}(v_j) = \sum_{\substack{1<i<j \\ j<k<n}} g(i,j,k).$$

Proof. Clearly, $\text{bc}(v_1) = \text{bc}(v_n) = 0$. To calculate the betweenness centrality of any fixed v_j, where $1 < j < n$, we add all values given by the function $f(i,j,k)$ over all i and k to obtain

$$\text{bc}(v_j) = \sum_{\substack{1<i<j \\ j<k<n}} f(i,j,k).$$

We consider a series of different cases.

When $k - i \equiv 0 \bmod 3$, the shortest path between v_i and v_k must consist of 3-hops. Hence these shortest paths will contain v_j if and only if $j - i \equiv 0 \bmod 3$.

When $k - i \equiv 2 \bmod 3$ the shortest path between v_i and v_k must consist of a single 2-hop and the rest 3-hops. If $j - i \equiv 1 \bmod 3$ then v_j will never appear on a shortest path between v_i and v_k. If $j - i \equiv 0 \bmod 3$ then there will only be 3-hops between v_i and v_j and a single 2-hop and the rest 3-hops between v_j and v_k. There are $\frac{1}{3}(k - j + 1)$ positions in which to place the 2-hop so that v_j lies on a shortest path between v_i and v_k. The total number of shortest paths between v_i and v_k is $\frac{1}{3}(k - i + 1)$. Hence the ratio is $(k - j + 1)/(k - i + 1)$. The case where $j - i \equiv 2 \bmod 3$ is done similarly.

The case where $k - i \equiv 1 \bmod 3$ is more complicated as a shortest path between v_i and v_k where $k - i \geq 4$ can have two different forms. The first is a composition of a single 1-hop and the rest 3-hops. The second is a composition of two 2-hops and the rest 3-hops. Hence from the $\frac{1}{3}(k-i+2)$ positions we must either choose a spot for the single 1-hop or choose two spaces for the two 2-hops. Hence the denominator will be

$$\binom{\frac{1}{3}(k-i+2)}{2}+\tfrac{1}{3}(k-i+2).$$

Simplifying we obtain that

$$\binom{\frac{1}{3}(k-i+2)}{2}+\tfrac{1}{3}(k-pi+2)=\tfrac{1}{18}(k-i+2)(k-i+5).$$

When $j-i\equiv 0 \bmod 3$, we know v_j will be on a shortest path between v_i and v_k if and only if there are only 3 -hops between v_i and v_j. Hence the numerator will be

$$\tfrac{1}{3}(k-j+2)+\binom{\frac{1}{3}(k-j+2)}{2}.$$

Simplifying we obtain that

$$\tfrac{1}{3}(k-j+2)+\binom{\frac{1}{3}(k-j+2)}{2}=\tfrac{1}{18}(k-j+2)(k-j+5).$$

The case where $j-i\equiv 1 \bmod 3$ is similar. When $j-i\equiv 2 \bmod 3$ then there will be a single 2-hop and the rest 3-hops between v_i and v_j, and the same for between v_j and v_k. Hence the numerator will be $\left(\frac{1}{3}(j-i+1)\right)\left(\frac{1}{3}(k-j+1)\right)$. □

We next investigate higher powers of paths and obtain a complete result for path powers with diameter 2.

We first give an example that shows a connection to the triangular numbers.

Example. Let $G=P_{15}^7$. We note that $\mathrm{bc}(v_j)=\mathrm{bc}(v_{16-j})$.

Clearly $\mathrm{bc}(v_1)=0$. We next compute $\mathrm{bc}(v_j)$ and consider all shortest paths containing v_j with the form $v_x-v_j-v_y$, where $1\le x<j<y\le 15$. We note that $d(G)=2$.

$\underline{\mathrm{bc}(v_2)}$: We first note that any shortest path containing v_2 must start with v_1 and end with v_9.

Of the paths of length 2 that connect v_1 and v_9, there are seven possible intermediate vertices $v_2, v_3, \dots, v_8$.

Since v_2 is one of these seven possibilities, $\mathrm{bc}(v_2)=\frac{1}{7}$.

$\underline{\mathrm{bc}(v_3)}$: We first note that any shortest path containing v_3 must have one of the following three forms:

v_1-v_9: Of the shortest paths connecting v_1 and v_9, there are six possible intermediate vertices $v_3, \dots, v_8$.

$v_2 - v_{10}$: Of the shortest paths connecting v_2 and v_{10}, there are seven possible intermediate vertices $v_3, \ldots, v_9$.

$v_1 - v_{10}$: Of the shortest paths connecting v_1 and v_{10}, there are seven possible intermediate vertices $v_3, \ldots, v_9$.

Hence $\mathrm{bc}(v_3) = 2\big(\frac{1}{7}\big) + \frac{1}{6} = \frac{19}{42}$.

$\underline{\mathrm{bc}(v_4)}$: We first note that any shortest path containing v_4 must have one of the following six forms:

$v_1 - v_9$: Of the shortest paths connecting v_1 and v_9, there are five possible intermediate vertices $v_4, \ldots, v_8$.

$v_2 - v_{10}$: Of the shortest paths connecting v_2 and v_{10}, there are six possible intermediate vertices $v_4, \ldots, v_9$.

$v_3 - v_{11}$: Of the shortest paths connecting v_3 and v_{11}, there are seven possible intermediate vertices $v_4, \ldots, v_{10}$.

$v_1 - v_{10}$: Of the shortest paths connecting v_1 and v_{10}, there are six possible intermediate vertices $v_4, \ldots, v_9$.

$v_2 - v_{11}$: Of the shortest paths connecting v_2 and v_{11}, there are seven possible intermediate vertices $v_4, \ldots, v_{10}$.

$v_1 - v_{11}$: Of the shortest paths connecting v_1 and v_{11}, there are seven possible intermediate vertices $v_4, \ldots, v_{10}$.

Hence $\mathrm{bc}(v_4) = 3\big(\frac{1}{7}\big) + 2\big(\frac{1}{6}\big) + \frac{1}{5} = \frac{101}{105}$.

For the sake of brevity we note that this pattern continues with the following observations.

$\underline{\mathrm{bc}(v_5)}$: Any shortest path containing v_5 is one of 10 forms where $d(v_x, v_y)$ are 8, 9, 10, or 11.

Hence $\mathrm{bc}(v_5) = 4\big(\frac{1}{7}\big) + 3\big(\frac{1}{6}\big) + 2\big(\frac{1}{5}\big) + \frac{1}{4} = \frac{241}{140}$.

$\underline{\mathrm{bc}(v_6)}$: Any shortest path containing v_6 is one of 15 forms where $d(v_x, v_y)$ are 8, 9, 10, 11, or 12.

Hence $\mathrm{bc}(v_6) = 5\big(\frac{1}{7}\big) + 4\big(\frac{1}{6}\big) + 3\big(\frac{1}{5}\big) + 2\big(\frac{1}{4}\big) + \frac{1}{3} = \frac{197}{70}$.

$\underline{\mathrm{bc}(v_7)}$: Any shortest path containing v_7 is one of 21 forms where $d(v_x, v_y)$ are 8, 9, 10, 11, 12, or 13.

Hence $\mathrm{bc}(v_7) = 6\big(\frac{1}{7}\big) + 5\big(\frac{1}{6}\big) + 4\big(\frac{1}{5}\big) + 3\big(\frac{1}{4}\big) + 2\big(\frac{1}{3}\big) + \frac{1}{2} = \frac{617}{140}$.

$\underline{\mathrm{bc}(v_8)}$: Any shortest path containing v_8 is one of 28 forms where $d(v_x, v_y)$ are 8, 9, 10, 11, 12, 13, or 14.

Hence $\mathrm{bc}(v_8) = 7\big(\frac{1}{7}\big) + 6\big(\frac{1}{6}\big) + 5\big(\frac{1}{5}\big) + 4\big(\frac{1}{4}\big) + 3\big(\frac{1}{3}\big) + 2\big(\frac{1}{2}\big) + 1 = 7$.

We note that the number of forms in each of these cases are triangular numbers. This pattern holds in general for $G = P_n^k$, where $n = 2k + 1$. We state this in our next theorem.

P_{15}^7	$\{0, \frac{1}{7}, \frac{19}{42}, \frac{101}{105}, \frac{241}{140}, \frac{197}{70}, \frac{617}{140}, 7, \frac{617}{140}, \frac{197}{70}, \frac{241}{140}, \frac{101}{105}, \frac{19}{42}, \frac{1}{7}, 0\}$
P_{14}^7	$\{0, \frac{1}{7}, \frac{19}{42}, \frac{101}{105}, \frac{241}{140}, \frac{197}{70}, \frac{617}{140}, \frac{617}{140}, \frac{197}{70}, \frac{241}{140}, \frac{101}{105}, \frac{19}{42}, \frac{1}{7}, 0\}$
P_{13}^7	$\{0, \frac{1}{7}, \frac{19}{42}, \frac{101}{105}, \frac{241}{140}, \frac{197}{70}, \frac{197}{70}, \frac{197}{70}, \frac{241}{140}, \frac{101}{105}, \frac{19}{42}, \frac{1}{7}, 0\}$
P_{12}^7	$\{0, \frac{1}{7}, \frac{19}{42}, \frac{101}{105}, \frac{241}{140}, \frac{241}{140}, \frac{241}{140}, \frac{241}{140}, \frac{101}{105}, \frac{19}{42}, \frac{1}{7}, 0\}$
P_{11}^7	$\{0, \frac{1}{7}, \frac{19}{42}, \frac{101}{105}, \frac{101}{105}, \frac{101}{105}, \frac{101}{105}, \frac{19}{42}, \frac{1}{7}, 0\}$
P_{10}^7	$\{0, \frac{1}{7}, \frac{19}{42}, \frac{19}{42}, \frac{19}{42}, \frac{19}{42}, \frac{19}{42}, \frac{19}{42}, \frac{1}{7}, 0\}$
P_{9}^7	$\{0, \frac{1}{7}, \frac{1}{7}, \frac{1}{7}, \frac{1}{7}, \frac{1}{7}, \frac{1}{7}, 0\}$
P_{8}^7	$\{0, 0, 0, 0, 0, 0, 0\}$

Table 1. Path powers with diameter 2. Note the nested nature of the prefixes ending with $0, \frac{1}{7}, \frac{19}{42}, \frac{101}{105}, \frac{241}{140}, \frac{197}{70}, \frac{617}{140}, 7$.

Theorem 12. *Let $G = P_n^k$, where $n = 2k+1$. Then*

$$\mathrm{bc}(v_j) = \sum_{i=1}^{j-1} \frac{j-i}{k+1-i}, \tag{2}$$

where $1 \le j \le k$ and $\mathrm{bc}(v_j) = \mathrm{bc}(v_{n+1-j})$.

Proof. When calculating $\mathrm{bc}(v_j)$, we note that v_j is contained in shortest paths between v_x and v_y, where $x < j < y \le n$ and $y - x = k + i$, where $1 \le i \le j-1$. This will account for $j-i$ pairs where the difference between indices is $k+1-i$.

For each pair of vertices v_x and v_y where $y - x = k+i$ there will be $j-i$ sets of paths. Each of these paths will have $k-i+1$ possible intermediaries. This contributes $(j-i)/(k-i+1)$ to $\mathrm{bc}(v_j)$. Summing these terms for all $1 \le i \le j-1$ will give the value of $\mathrm{bc}(v_j)$. Hence we have (2). □

Next we show how the previous theorem can be extended to cover all other path powers of diameter 2. We begin with an example.

Example 13. Let $G = P_{12}^7$. We first note that $\mathrm{bc}(v_j) = \mathrm{bc}(v_{13-j})$.

For v_j where $1 \le j \le 5$, the betweenness centrality values are identical to those in P_{15}^7 and can be computed using the exact same method. However the pattern used in the example with P_{15}^7 cannot be extended for $\mathrm{bc}(v_6)$ since $v_y \le 12$. As a result, the paths used in the computation of $\mathrm{bc}(v_5)$ and $\mathrm{bc}(v_6)$ are identical. Hence, $\mathrm{bc}(v_1) = 0$; $\mathrm{bc}(v_2) = \frac{1}{7}$, $\mathrm{bc}(v_3) = \frac{19}{42}$, $\mathrm{bc}(v_4) = \frac{101}{105}$, and $\mathrm{bc}(v_5) = \mathrm{bc}(v_6) = \frac{241}{140}$.

We observe that the betweenness centrality values in path powers with diameter 2 have a nested pattern (see Table 1).

We formalize this property in our next theorem.

Theorem 14. *Let $G = P_n^k$, where $n < 2k+1$ and $j < k$. Then*

$$\mathrm{bc}(v_j) = \sum_{i=1}^{j-1} \frac{j-i}{k-i+1} \tag{3}$$

for all $2 \le j \le n-k$ and $\mathrm{bc}(v_j) = \mathrm{bc}(v_{n+1-j})$. For P_n^k, the $\mathrm{bc}(v_j)$ are all equal for all $n-k \le j \le \lceil \frac{1}{2}n \rceil$.

Proof. When calculating $\mathrm{bc}(v_j)$, we note that v_j is contained in shortest paths between v_x and v_y where $x < j < y \le n$ and $y - x = k + i$ where $1 \le i \le j-1$. This will account for $j-i$ pairs where the difference between indices is $k+1-i$.

For each pair of vertices v_x and v_y where $y - x = k + i$ (where $k + i \le n$), there will be $j-i$ sets of paths. Each of these paths will have $k-i+1$ possible intermediaries. This contributes $(j-i)/(k-i+1)$ to $\mathrm{bc}(v_j)$. Summing these terms for all $1 \le i \le j-1$ will give the value of $\mathrm{bc}(v_j)$. Hence we have (3). For P_n^k, $\mathrm{bc}(v_j)$ is the same as in P_{2k+1}^k for the first $n-k$ terms. Then since there are the same number of pairs of vertices v_x and v_y where $y-x = k+i$ (where $k+i \le n$) in P_n^k, the $\mathrm{bc}(v_j)$ are all the same for $n-k \le j \le \lceil \frac{1}{2}n \rceil$. □

3. Conclusion

For path powers with larger diameter the problem becomes more complex. The case of P_n^m involves m^2 different cases, and as n increases the cases become more complicated. Hence the problem for general powers of paths is more difficult. We note that problem is tied to the number of integer partitions with a fixed upper bound on the size of each part [Ratsaby 2008]. The objective is to minimize the number of parts.

We pose the following problem.

Problem 15. *Determine the betweenness centrality for all vertices in P_n^m.*

Acknowledgements

The authors are grateful to an anonymous referee whose careful reading and comments improved the presentation of this paper. This research was supported by a National Science Foundation Research Experiences for Undergraduates Site Award Grant (#1062128) with cofunding from the Department of Defense. Darren Narayan was also supported by NSF Award #1019532. Rigoberto Flórez was partially supported by The Citadel Foundation.

References

[Brandes et al. 2016] U. Brandes, S. Borgatti, and L. Freeman, "Maintaining the duality of closeness and betweenness centrality", *Social Networks* **44** (2016), 153–159.

[Bullmore and Sporns 2009] E. Bullmore and O. Sporns, "Complex brain networks: graph theoretical analysis of structural and functional systems", *Nature Rev. Neuroscience* **10**:3 (2009), 186–198.

[Freeman 1977] L. C. Freeman, "A set of measures of centrality based upon betweenness", *Sociometry* **40**:1 (1977), 35–41.

[Freeman et al. 1991] L. C. Freeman, S. P. Borgatti, and D. R. White, "Centrality in valued graphs: a measure of betweenness based on network flow", *Social Networks* **13**:2 (1991), 141–154. MR

[Gago et al. 2012] S. Gago, J. Hurajová, and T. Madaras, "Notes on the betweenness centrality of a graph", *Math. Slovaca* **62**:1 (2012), 1–12. MR Zbl

[Grassi et al. 2009] R. Grassi, R. Scapellato, S. Stefani, and A. Torriero, "Betweenness centrality: extremal values and structural properties", pp. 161–175 in *Networks, topology and dynamics: theory and applications to economics and social systems*, edited by A. K. Naimzada et al., Lecture Notes in Economics and Mathematical Systems **613**, Springer, 2009. Zbl

[Guye et al. 2010] M. Guye, G. Bettus, F. Bartolomei, and P. J. Cozzone, "Graph theoretical analysis of structural and functional connectivity MRI in normal and pathological brain networks", *Magn. Reson. Mater. Phys.* **23**:5-6 (2010), 409–421.

[Kumar and Balakrishnan 2016] S. Kumar R. and K. Balakrishnan, "Betweenness centrality of Cartesian product of graphs", preprint, 2016. arXiv

[Kumar et al. 2014] S. Kumar Raghavan Unnithan, B. Kannan, and M. Jathavedan, "Betweenness centrality in some classes of graphs", *Int. J. Comb.* **2014** (2014), art. id. 241723. MR Zbl

[Pandit et al. 2013] A. S. Pandit, P. Expert, R. Lambiotte, V. Bonnelle, R. Leech, F. E. Turkheimer, and D. J. Sharp, "Traumatic brain injury impairs small-world topology", *Neurology* **80**:20 (2013), 1826–1833.

[Ratsaby 2008] J. Ratsaby, "Estimate of the number of restricted integer-partitions", *Appl. Anal. Discrete Math.* **2**:2 (2008), 222–233. MR Zbl

[White and Borgatti 1994] D. R. White and S. P. Borgatti, "Betweenness centrality measures for directed graphs", *Social Networks* **16**:4 (1994), 335–346.

Received: 2017-03-04 Revised: 2017-07-26 Accepted: 2018-01-20

emmacarter2014@gmail.com *School of Mathematical Sciences, Rochester Institute of Technology, Rochester, NY, United States*

bte14@math.rutgers.edu *Department of Mathematics, Rutgers University, Piscataway, NJ, United States*

dng2551@rit.edu *Department of Software Engineering, Rochester Institute of Technology, Rochester, NY, United States*

florezr1@citadel.edu *Department of Mathematics and Computer Science, The Citadel, Charleston, SC, United States*

dansma@rit.edu *School of Mathematical Sciences, Rochester Institute of Technology, Rochester, NY, United States*

involve

msp.org/involve

INVOLVE YOUR STUDENTS IN RESEARCH

Involve showcases and encourages high-quality mathematical research involving students from all academic levels. The editorial board consists of mathematical scientists committed to nurturing student participation in research. Bridging the gap between the extremes of purely undergraduate research journals and mainstream research journals, *Involve* provides a venue to mathematicians wishing to encourage the creative involvement of students.

See inside back cover or msp.org/involve for submission instructions. The subscription price for 2019 is US $/year for the electronic version, and $/year (+$, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues and changes of subscriber address should be sent to MSP.

Involve (ISSN 1944-4184 electronic, 1944-4176 printed) at Mathematical Sciences Publishers, 798 Evans Hall #3840, c/o University of California, Berkeley, CA 94720-3840, is published continuously online. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices.

Involve peer review and production are managed by EditFLOW® from Mathematical Sciences Publishers.

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

involve

2019 vol. 12 no. 1

1944-4176(2019)12:1;1-4