Social Networks 13 (1991) 155–168
North-Holland

Simultaneous group and individual centralities

Phillip Bonacich

Department of Sociology, University of California, Los Angeles, 405 Hilgard Ave., Los Angeles, CA 90024-1551, USA

A completely symmetric pair of measures of individual and group centrality is described. The centralities of groups are a function of the centralities of their members. The centralities of individuals are a function of the centralities of the groups to which they belong. It is shown that the standard approach to controlling for variations in group size is inadequate. It produces measures of centrality that are roughly proportional to the square roots of group sizes. A better way to control for variations in group size or in the numbers of individual memberships is presented. Comparisons are made with correspondence analysis.

Sociologists and others interested in the phenomenon of interlocking directorates among corporations believe that patterns of interlocking directorates will reflect relations between corporations. The centrality of corporations has been of particular interest. A firm is central because it is connected to firms that are not directly connected to each other. Such a firm is in advantageous position with respect to communication flows. Its information will be more complete, and it will tend to get information earlier than more peripheral or isolated firms. For example, Mintz and Schwartz (1985) have described the central, and presumably powerful, position of banks in the structure of interlocking directorships.

The kind of techniques I am going to describe need not, however, be confined to boards of directors. Someone studying voluntary organizations could use these techniques to describe the pattern of committee memberships or the pattern of attendance at group events. "Clique" events would be those attended only by a single category of organizational member, while "central" integrative events would be attended by a variety of members. The centrality of a group's events also should be related to communication within the group. If one wanted and announcement to diffuse widely and rapidly among those attending an annual meeting of the American Sociological Association, one would not necessarily choose the largest session. The largest session might contain sociologists whose attendance patterns at other sessions were similar. The announcement would diffuse more widely if it were made at a session (for example, opening Presidential Address) with a heterogeneous audience.

In any of these situations, one may wish to identify central individuals as well as central groups. Central directors of corporations share board memberships with directors who do not share board memberships with each other. These directors will receive more information and they will receive it earlier. The simultaneous mapping of individuals as well as groups can contribute to the richness of the picture that emerges. For example, one may know as much or more about some directors than about the firms they direct. A simultaneous mapping enables one to use both sets of information.

Centrality involves, using Breiger's term (Breiger 1974), the "duality" of groups and individuals. A central firm gets its central position from the board membership patterns of its members. They belong to the variety of boards that make that firm central. If its members belong to a constricted set of other boards, that firm would not be central. Dually a central individual should be one who belongs to a variety of important firms. One kind of centrality cannot be defined without reference to the other.

To illustrate, I shall use the following classic data on the attendance of eighteen women at a series of fourteen events (Homans 1950: 83). The rows and columns have been permuted to show a pattern.

Measures of centrality (Freeman 1979) take into account the indirect as well as the direct paths connecting positions. A connection to a peripheral position provides less information than does a connection to a central and information rich position. In the most commonly used measure of centrality in interlocking directorate research (Bonacich 1972a), central boards are those that overlap with other central boards.

Let s_{jk} be the number of common directors shared by boards j and k. The standard Bonacich measure of centrality assumes that the centrality of each board is a linear function of the centralities of the groups with which it overlaps and the amount of overlap.

Letting g_j be the centrality of group j:

$$\lambda g_j = \Sigma_k s_{jk} g_k . \tag{1}$$

The parameter λ is necessary for the system of equations to have the solution. These also can be described as a matrix equation:

$$Sg = \lambda g. \tag{2}$$

The vector of centrality scores g is an eigenvector of S and lambda is the largest eigenvalue of S. [1]

Let me suggest a dual approach to group and individual centrality. Let A be an n by m matrix showing the memberships of n individuals in m groups; $a_{ij} = 1$ if person i belongs to group j and $a_{ij} = 0$ otherwise. Let p and g be n and m dimensional vectors of individual and group centrality scores. Group centralities are a function of the centralities of their members and individual centralities are a function of the centralities of the groups to which they belong. This implies the following matrix equations, where lambda is a singular value of the matrix A (Greenacre 1984: Appendix A) and A^t is the transpose of A. [2]

$$Ag = \lambda p \tag{3}$$

$$A^t p = \lambda g. \tag{4}$$

These equations have an alternative form. Substituting Equation (1) into Equation (2) and vice versa produces the following two equations

$$AA^t = \lambda^2 p, \qquad A^t A g = \lambda^2 g. \tag{5}$$

The vector of individual centrality scores p is an eigenvector of AA^t and g is an eigenvector of A^tA. In both equations λ^2 is the eigenvalue. Both AA^t and A^tA are positive symmetric matrices. Hence, if they are irreducible they have a positive eigenvalue greater than or equal in

[1] This approach has been used to define centrality scores for any relation (Bonacich 1972b). Let R be a non-negative relational matrix. Then $Rc = \lambda c$ defines the centrality of a vertex in terms of the centralities of the points to which it is connected.

[2] The introduction of the singular value is necessary for the equations to have a non-zero solution.

magnitude to every other eigenvalue, the eigenvector corresponding to this eigenvalue is positive, and no other eigenvector is positive. [3] Since only positive centrality scores make sense, the centrality scores are the positive eigenvectors corresponding to this largest eigenvalue.

Since S in Equation (2) equals A^tA in Equation (5), the new group centrality measure g is the same as the standard measure. What is new is the interpretation of g; g_i is proportional to the sum of the individual centralities of the members of group i.

There is another connection between the new and old approaches. A group membership table, like Table 1, defines a "bipartite" graph (Busacker and Saaty 1965: 19), a graph whose points can be divided into two disjoint sets such that every edge connects points from different sets. The two sets are the individuals and the groups. The bipartite graph has $n+m$ vertices, where n is the number of individuals and m is the number of groups. A line connects individual i and group j in this graph if $a_{ij}=1$. The "degree" of a vertex is the number of other vertices to which it is connected. In this bipartite graph, the degree of a group is its size and the degree of a person is the number of groups to which he belongs. Corresponding to this graph is an $n+m$ by $n+m$ square matrix whose first m rows and columns correspond to groups and whose next n rows and columns represent individuals. Equations (3) and (4) are then equivalent to the following matrix equation:

$$\lambda\begin{pmatrix} g \\ p \end{pmatrix} = \begin{pmatrix} 0 & A^t \\ A & 0 \end{pmatrix}\begin{pmatrix} g \\ p \end{pmatrix}. \tag{6}$$

The new approach applies Equation (2) to the bipartite graph of relations among groups and individuals instead of the matrix of relations between groups only.

[3] A square matrix X is cogradient to a matrix Y if there exists a permutation matrix P such that $X=P^tYP$. A non-negative square matrix is reducible if it is cogradient to a matrix of the form

$$\begin{pmatrix} B & C \\ 0 & D \end{pmatrix}$$

where B and D are square submatrices and 0 is submatrix of zeros (Minc 1988: 5). An irreducible non-negative matrix has a real positive eigenvalue greater than or equal to the absolute value of every other eigenvalue, and there is a positive eigenvector corresponding to this eigenvalue (Minc 1988: 11). Moreover, because the eigenvectors of a symmetric matrix are orthogonal, g and p must be the only positive eigenvectors of A^tA and AA^t.

Table 1
Person (rows) by group (columns) membership matrix

		Groups													
		1	2	3	4	5	6	7	8	9	10	11	12	13	14
People	1	1	1	1	1	1	1	0	1	1	0	0	0	0	0
	2	1	1	1	0	1	1	1	1	0	0	0	0	0	0
	3	0	1	1	1	1	1	1	1	1	0	0	0	0	0
	4	1	0	1	1	1	1	1	1	0	0	0	0	0	0
	5	0	0	1	1	1	0	1	0	0	0	0	0	0	0
	6	0	0	1	0	1	1	0	1	0	0	0	0	0	0
	7	0	0	0	0	1	1	1	1	0	0	0	0	0	0
	8	0	0	0	0	0	1	0	1	1	0	0	0	0	0
	9	0	0	0	0	1	0	1	1	1	0	0	0	0	0
	10	0	0	0	0	0	0	1	1	1	0	0	1	0	0
	11	0	0	0	0	0	0	0	1	1	1	0	1	0	0
	12	0	0	0	0	0	0	0	1	1	1	0	1	1	1
	13	0	0	0	0	0	0	1	1	1	1	0	1	1	1
	14	0	0	0	0	0	1	1	0	1	1	1	1	1	1
	15	0	0	0	0	0	0	1	1	0	1	1	1	0	0
	16	0	0	0	0	0	0	0	1	1	0	0	0	0	0
	17	0	0	0	0	0	0	0	0	1	0	1	0	0	0
	18	0	0	0	0	0	0	0	0	1	0	1	0	0	0

The measures of individual centrality and group centrality are given in the columns labelled "∞" of Table 2 and 3.

In Tables 2 and 3 the vector of individual and group centrality scores of Equation (6) has been standardized so that its length $= m + n = 18 + 14 = 32$. If all groups and all individuals were equal in the centrality, all centrality scores would equal 1.00. The value 1.00 can serve as a base line for evaluating whether a particular score is high or low.

Controlling for size

One serious problem with these centrality measures is that they are strongly affected by the sheer size of groups and the number of groups to which an individual belongs. For example, the correlation between the event centrality scores in Table 2 and the sheer number attending that event is 0.97. This may or may not be a problem. One may not want to control for the size of groups or the number of individual memberships; the size of an annual family Christmas gathering is an

Table 2
Centrality scores for persons as a function of path length and correspondence analysis solution

Person	Path length						
	1	2	3	4	5	∞	CA
1	1.28	1.34	1.29	1.33	1.3	1.34	−0.80
2	1.12	1.19	1.16	1.2	1.18	1.24	−0.84
3	1.28	1.43	1.39	1.44	1.42	1.48	−0.65
4	1.12	1.2	1.17	1.21	1.2	1.25	−0.86
5	0.64	0.66	0.64	0.66	0.65	0.67	−0.97
6	0.64	0.75	0.75	0.78	0.78	0.84	−0.80
7	0.64	0.8	0.8	0.85	0.84	0.91	−0.51
8	0.48	0.65	0.65	0.68	0.68	0.72	−0.04
9	0.64	0.85	0.85	0.89	0.89	0.94	−0.17
10	0.64	0.83	0.82	0.86	0.84	0.87	0.43
11	0.64	0.77	0.74	0.77	0.75	0.75	0.83
12	0.96	1	0.95	0.96	0.93	0.88	1.05
13	1.12	1.2	1.15	1.18	1.15	1.11	0.87
14	1.28	1.26	1.18	1.18	1.14	1.06	0.87
15	0.8	0.87	0.83	0.85	0.83	0.8	0.81
16	0.32	0.47	0.47	0.5	0.5	0.53	0.30
17	0.32	0.35	0.32	0.32	0.31	0.28	1.10
18	0.32	0.35	0.32	0.32	0.31	0.28	1.10

Table 3
Centrality scores for groups as a function of path length and correspondence analysis solution

Group	Path length						
	1	2	3	4	5	∞	CA
1	0.48	0.51	0.53	0.53	0.54	0.57	−1.05
2	0.48	0.52	0.55	0.55	0.57	0.6	−0.97
3	0.96	0.94	0.97	0.96	0.98	1.01	−1.04
4	0.64	0.65	0.67	0.66	0.68	0.7	−1.04
5	1.28	1.2	1.25	1.22	1.25	1.29	−0.88
6	1.28	1.23	1.28	1.26	1.29	1.31	−0.57
7	1.6	1.5	1.56	1.52	1.55	1.53	−0.13
8	2.24	2.01	2.08	2.02	2.06	2.03	−0.03
9	1.92	1.66	1.69	1.62	1.63	1.52	0.51
10	0.8	0.76	0.77	0.74	0.75	0.68	1.12
11	0.64	0.53	0.5	0.46	0.45	0.36	1.22
12	0.96	0.89	0.91	0.87	0.88	0.81	1.02
13	0.48	0.49	0.5	0.49	0.49	0.45	1.17
14	0.48	0.49	0.5	0.49	0.49	0.45	1.17

important part of its integrative function. On the other hand, the size of a board of directors may reflect nothing about the position of the firm but simply be in conformity with some arbitrary rule. Similarly, variations in the number of memberships by an individual may be either meaningful or arbitrary. [4]

The conventional way of controlling for group size (Mariolis 1975; Mintz and Schwartz 1985; Mizruchi *et al.* 1986; Mizruchi 1989) is to modify A^tA before using Equation (5). Let n_{ij} be the overlap between groups i and j and let n_i be the size of group i. Let R be a matrix of standardized overlaps between groups in which the number of common members is divided by the geometric mean of their sizes.

$$r_{ij} = \frac{n_{ij}}{\sqrt{n_i}\sqrt{n_j}}. \tag{7}$$

The centrality measures g are contained in the eigenvector associated with the largest eigenvalue λ of S.

$$Rg = \lambda g. \tag{8}$$

Appendix A shows that these centrality scores do not correct adequately for group size. They tend to be correlated with the square root of group size, and hence with size itself. The same conclusions hold for individual centrality.

Let me suggest another approach based on the following result (the proof is contained in Appendix B).

Theorem. Assume that a symmetric non-negative matrix R has a single largest eigenvalue λ_1, and let v_1 be its eigenvector. [5] Let e be a column vector of ones. The infinite sum $R + \beta R^2 + \beta^2 R^3 + \beta^3 R^4 + \ldots$ converges if $|\beta| < 1/\lambda_1$. Moreover, as β approaches $1/\lambda_1$ from below, $(1 - \beta\lambda_1)(R + \beta R^2 + \beta^2 R^3 + \beta^3 R^4 + \ldots)e$ approaches v_1.

[4] This problem exists whenever an eigenvector is used as measure of centrality. Let R be any binary (0-1) matrix of relations between vertices. The matrix equation $Rc = \lambda c$ defines the centralities of points in terms of the total centralities of the points to which they are connected. This measure will almost invariably be highly related to degree.

[5] Because R is non-negative, it will have a positive eigenvalue at least as large in absolute value as any other eigenvalue (Minc 1988: 14). Any data matrix is extremely unlikely to have two equal largest eigenvalues.

In other words, the eigenvector of centrality scores can be thought of as proportional to an infinite sum, $Re + \beta R^2 e + \beta^2 R^3 e + \beta^3 R^4 e + \ldots$, where the first term, Re, gives the contribution of size, (graph degree), the sum of the first two terms gives the contribution to centrality of paths of length 1 and 2 in the bipartite graph of individuals and groups, and so on. The sum of the first k terms gives the centrality scores that would result if all paths longer than k were ignored. Let

$$c_k(\beta) = a_{k\beta} \Sigma_i \beta^{i-1} R^i, \qquad i = 1 \ldots k \tag{9}$$

where $a_{k\beta}$ is chosen so that the length of the vector $c_k(\beta)$ is one. I will define $c_{,}$ the centrality scores ignoring all paths longer than k, as the limit of $c_k(\beta)$ as β approaches $1/\lambda_1$.[6] In particular, c_1 is simply the degree-based centrality scores, and c_∞ is v_1, an eigenvector of R.

For any graph, not only the bipartite considered in this paper, the c_k form a sequence of centrality scores based on longer and longer paths whose limit is the eigenvector. Tables 2 and 3 show the sequence of centrality scores for persons and for groups. Figures 1 and 2 show the changes in centrality score $c_k - c_1$ for selected individuals and groups as paths longer than one are introduced. I suggest $c_\infty - c_1$ as a measure of centrality corrected for size.[7] The effect of size has been subtracted out.

Looking at the changes in group centrality scores, the three largest groups 7, 8, and 9 decline in relative centrality because, in a sense, they are larger than they are central. Group 11 declines because two of its members are the infrequent attenders 17 and 18. Looking at the changes in individual centrality, those increasing the most in centrality (7, 8, 9, 10, and 16) attended few events (so their degree-based centrality was small), but the events they attended were attended by many others. Person 14 becomes less central because she did not attend the most central event, number 8.

[6] The infinite sequence converges when $|\beta| < 1/\lambda_1$ and has a limit as β approaches $1/\lambda_1$, and each finite sequence converges when $\beta = 1/\lambda_1$. Thus c_k is defined whether k is finite or infinite.
[7] If there are no differences in either group size or in number of individual memberships, the approach described in this paper will give no differences in centrality. As shown by the theorem, the eigenvector is the limit of a weighted sum of the number of paths of various lengths emanating from each vertex. If all the vertices have the same number of paths of length one, then they all have the same number of paths of every length.

Correspondence analysis

The matrix A is a cross-classification of people by the groups to which they belong. The goal of correspondence analysis (Greenacre 1984) is to create multidimensional spaces for vectors corresponding to the row and column categories that describe the geometric "distances" between categories. The most important output from correspondence analysis is a row space and a column space of small dimension that provide a least-squared-error description of the similarities between the row categories and between the column categories.

Although the goals of correspondence analysis and centrality analysis are different, their mathematics are similar. Both depend on the singular value decomposition SVD (Greenacre 1984: 37). However, the different ways in which the cross-classifications are treated before the SVD is applied means that the two techniques are sensitive to different aspects of the matrix A. The most powerful and useful dimensions for correspondence analysis describe similarities in row or column pat-

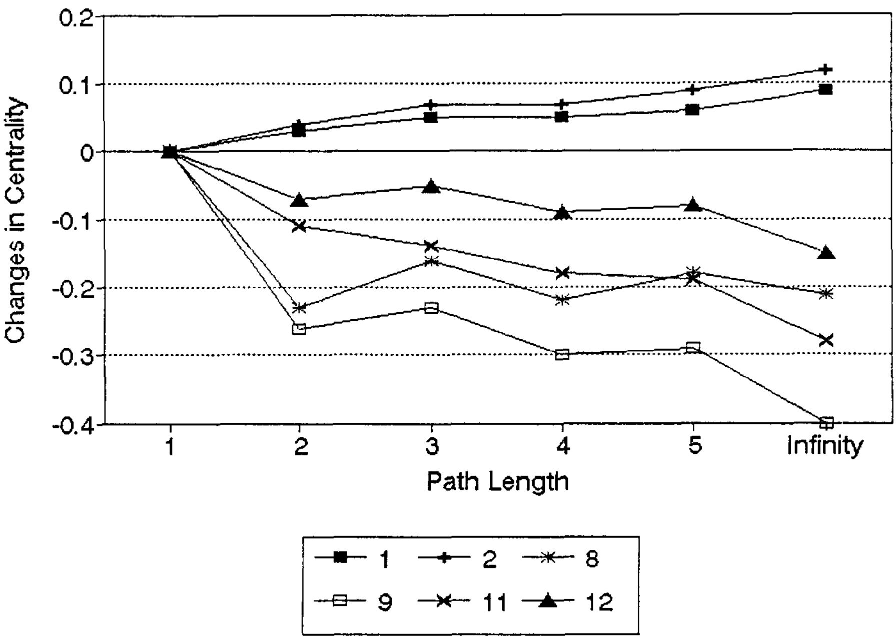

Fig. 1. Changes in centrality for selected groups by path length.

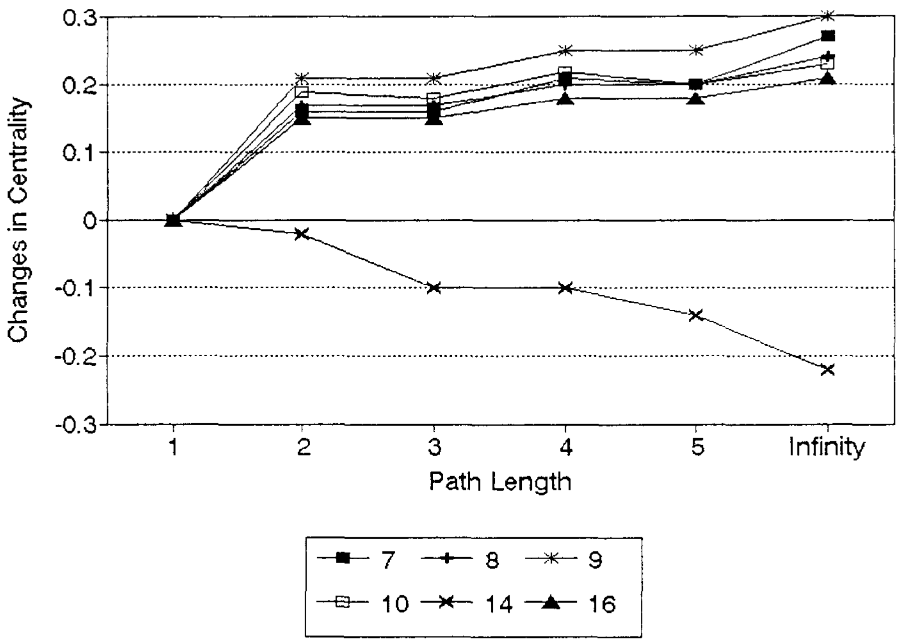

Fig. 2. Changes in centrality for selected groups by path length.

terns. Before the SVD is used, the cross-classification is modified so that these dimensions of similarity will be the most powerful. Among the less powerful dimensions that may be discarded in correspondence analysis there may be a dimension corresponding to centrality. However, correspondence analysis is not designed to highlight this dimension.

The last columns in Tables 2 and 3 give the scores of the rows and columns of Table 1 on the most powerful of the correspondence analysis dimensions. Rather than centrality, what they seem to capture is membership in the two cliques that attended two different sets of events. The highest scores for individuals, positive or negative, go to those whose attendance patterns were most characteristic of their cliques. The highest magnitudes for groups go to those that had the purest pattern of attendance by one clique or another. The third most powerful correspondence analysis dimension for both the rows and columns corresponded closely to the centrality analysis pattern. However, within correspondence analysis there is no way of detecting which dimensions, if any, correspond to centrality, whereas the techniques described in this paper are designed to isolate the centrality dimension.

Conclusions

The goal of this paper is to describe a related set of conceptually symmetric individual and group centrality measures for interlocking directorate data and to provide a new interpretation for an old measure of group centrality (Bonacich 1972a; Mizruchi *et al.* 1986). Based on this new approach, ways of controlling for variations in the sizes of boards and the numbers of boards individuals are members of are suggested.

Appendix A

This appendix will show why Equations (7) and (8) give centrality measures that tend to be highly correlated with the square root of group size, and hence with group size itself. Let n be a column vector of group sizes, n_{ij} the number of persons in both groups i and j, and N be the total number of persons.

We want a measure of group centrality that is uncorrelated with group size if there is no structure to the data. Assume, therefore, that group memberships are independent.

$$n_{ij} = \frac{n_i n_j}{N}. \tag{A1}$$

Then, Equation (7) becomes:

$$\begin{aligned} r_{ij} &= \frac{n_{ij}}{\sqrt{n_i}\sqrt{n_j}} \\ &= \frac{n_i n_j}{N\sqrt{n_i}\sqrt{n_j}} \\ &= \frac{\sqrt{n_i}\sqrt{n_j}}{N} \quad \text{for } i \neq j. \end{aligned} \tag{A2}$$

Ignoring the main diagonal (only a small proportion of the entire matrix), the following is (almost) true (where $n^{1/2}$ is a vector of the square roots of the elements of n).

$$R = \frac{n^{1/2}(n^{1/2})^{t}}{N}. \tag{A3}$$

From this it can be seen that $n^{1/2}$ is an eigenvector of R.

Appendix B

Let R be a non-negative symmetric matrix. R has a positive eigenvalue λ_1 whose absolute value is at least as great as any other eigenvalue (Minc 1988: 14). λ_1 has a non-negative eigenvector (Minc 1988: 14). For convenience, assume that λ_1 is strictly greater in magnitude than any other eigenvalue (which is almost certainly true for any data). Let the vector e consist entirely of ones, and let β be such that $|\beta| < 1/\lambda_1$. Then the following infinite series converges (Finkbeiner 1966: 242):

$$\begin{aligned} s(\beta) &= (I - \beta R)^{-1} Re \\ &= (I + \beta R + \beta^2 R^2 + \ldots) Re \\ &= (R + \beta R^2 + \beta^2 R^3 + \ldots) e. \end{aligned} \tag{B1}$$

Since R is symmetric, it has an orthonormal set of eigenvectors $\{v_i\}$, and,

$$R = \Sigma \lambda_i v_i v_i, \text{ where } v_i^t v_i = 1,\ v_i^t v_j = 0 \text{ for } i \neq j. \tag{B2}$$

It is easy to show that for positive k,

$$R^k = \Sigma \lambda_i^k v_i v_i^t. \tag{B3}$$

Therefore, because $\beta\lambda_1 < 1$,

$$s(\beta) = \Sigma_i(\lambda_i + \beta\lambda_i^2 + \beta^2\lambda_i^3 \ \ldots)v_i v_i^t e \tag{B4}$$

$$= \Sigma_i\left(\frac{\lambda_i}{1-\beta\lambda_i}\right)v_i v_i^t e.$$

It follows that:

$$s(\beta)(1-\beta\lambda_1) = \lambda_1 v_1 v_1^t e + \Sigma_i\left(\frac{\lambda_i(1-\beta\lambda_1)}{1-\beta\lambda_i}\right)v_i v_i^t e, \qquad i \geq 2. \tag{B5}$$

There is a finite number of terms in the summation in Equation (5), and each of them approaches zero as β approaches λ_1 from below. Therefore, as β approaches $1/\lambda_1$, $s(\beta)(1-\beta\lambda_1)$ approaches $\lambda_1 v_1 v_1^t e = (\lambda_1 v_{11}^t e)v_1$, and

$$\frac{s(\beta)(1-\beta\lambda_1)}{\lambda_1 v_1^t e} \text{ approaches } v_1. \tag{B6}$$

Note that when $\beta = 1/\lambda_1$, the left side of Equation (6) is undefined because $s(1/\lambda_1)$ does not converge and $1 - \beta\lambda_1 = 0$.

It might seem natural to control for size by partitioning $s(\beta)$ into a component due to degree (paths of length one in the bipartite graph) and a component due to longer paths.

$$s(\beta) = Re \qquad\qquad + (\beta R^2 + \beta^2 R^3 + \ \ldots)e$$

$$= \text{degree component} + \text{component due to longer paths.}$$

However, this is not a feasible solution. The relative sizes of the two components are a function of β. As β approaches $1/\lambda_1$, the second component increases without limit.

References

Bonacich, Phillip

1972a "Techniques for analyzing overlapping memberships," in: Herbert L. Costner (ed.) *Sociological Methodology* 1972, p. 176–185. San Francisco: Jossey-Bass.

Bonacich, Phillip

1972b "Factoring and weighting approaches to clique identification." *Journal of Mathematical Sociology 2*: 113–120.

Breiger, Ronald L.

1974 "The duality of persons and groups." *Social Forces 53*: 181–190.

Busacker, Robert and Thomas L. Saaty

1965 *Finite Graphs and Networks: An Introduction with Applications*. New York: McGraw-Hill.

Finkbeiner, Daniel T., II

1966 *Introduction to Matrices and Linear Transformations*. San Francisco: W.H. Freeman.

Freeman, Linton C.

1979 "Centrality in social networks: conceptual clarification." *Social Networks 1*: 215–239.

Greenacre, Michael J.

1984 *Theory and Applications of Correspondence Analysis*. New York: Academic Press.

Homans, George C.

1950 *The Human Group*. New York: Harcourt Brace Jovanovich.

Mariolis, Peter

1975 "Interlocking directorates and the control of corporations." *Social Science Quarterly 56*: 425–439.

Minc, Henryk

1988 *Nonnegative Matrices*. New York: John Wiley & Sons.

Mintz, Beth and Michael Schwartz

1985 *The Power Structure of American Business*. Chicago: University of Chicago Press.

Mizruchi, Mark S., Peter Mariolis, Michael Schwartz and Beth Mintz

1986 "Techniques for disaggregating centrality scores in social networks", in: Nancy B. Tuma (ed.) *Sociological Methodology* 1986, p. 26–48. San Francisco: Jossey-Bass.

Mizruchi, Mark S.

1989 "Similarity of political behavior among large American corporations". *American Journal of Sociology 95*: 401–24.