*Quality and Quantity*, 14 (1980) 585–592 585
Elsevier Scientific Publishing Company, Amsterdam – Printed in The Netherlands

# NOTE

# THE GATEKEEPER, PAIR-DEPENDENCY AND STRUCTURAL CENTRALITY *

LINTON C. FREEMAN

*School of Social Sciences, University of California, Irvine, CA 92717, U.S.A.*

This is a note to introduce a new measure of a kind of structural centrality called pair-dependency. Pair-dependency explicates the centrality-related notion of the gatekeeper. Moreover, it turns out to be a fundamental structural property of communication networks that provides the basis for the derivation of two standard measures of structural centrality.

## The Concept of Centrality and its Measurement

The concept of centrality as applied to human communication networks was introduced by Bavelas (1948). He reasoned that when a person was located between others in a network, he or she had the potential for control of their communication and was, therefore, somehow central. But when he went on to specify a measure of centrality, Bavelas (1950) lost sight of this intuition. Instead he borrowed from graph theory and based his measure on the communication distances between persons; a person who was close to others was central. It turned out, as Bavelas' student, Leavitt (1951) pointed out, that the closeness-based Bavelas measure was an index, not of the control potential of a point, but rather of its independence of such control.

In any case, discussions of the intuitive concept of centrality and of alternative procedures for its measurement have been maintained over

* Presented at the Seminar on Communication Network Analysis, East-West Communication Institute, January 2–9, 1979, East-West Center, Honolulu, Hawaii, U.S.A.

a 30 year period. This whole history was reviewed in a recent paper by Freeman (1979). It is relevant here only to note that the original Bavelas' intuition was not embodied in a formally defined measure until 1971 when Anthonisse (1971) created an index of what he called "rush." His measure, however, was not published and essentially the same index of centrality was derived independently by Freeman (1977).

The Anthonisse–Freeman measure is a property of points in a network. It may be calculated for any point in a symmetrical communication network. It is an index of the degree to which each point falls on the shortest paths (geodesics) between all reachable pairs of others and thus can facilitate, inhibit or distort their communication. Thus, it is a measure of betweenness that embodies the original Bavelas intuition.

Even before Bavelas wrote about centrality, his teacher, Lewin (1947), had proposed a related – but more limited – idea. Lewin suggested that a gatekeeper in a communication network was a person in a position to control "the travelling of a news item *through certain communication channels in a group*" (emphasis added).

On the face of it, this sounds very much like the idea introduced by Bavelas. If not identical, these are at least very similar ideas. One might suspect that Bavelas was simply repeating his teacher's insight and calling it centrality instead of gatekeeping. There is, however, one important difference between the two conceptions. While Lewin and subsequent users of the gatekeeper concept stress the idea of control of certain channels of communication, Bavelas, and others who refer to centrality, emphasize the potential of points for control of communication over the total network.

Rogers and Agarwala-Rogers (1976, p. 133), for example, characterized a gatekeeper as "an individual who is located in a communication structure so as to control messages *flowing through a communication channel*" (emphasis added). They went on to give an illustration (1976, p. 134): "If you have ever tried to get a rush memo to your boss, and his secretary told you he was 'in conference,' you know what a gatekeeper is." The reference to a communication channel in both of these statements suggests that a gatekeeper is not conceived as being in a general sort of position of control like a position high in centrality based on betweenness. Instead he or she is the keeper of the gate controlling communication to and from a particular other person vis-a-vis the rest of the network!

Bavelas (1948) gave an illustration of an Italian-American work group containing only one English speaking worker and a boss who spoke only English. Since she was the only translator in a two-way communication process, Bavelas reasoned that the bilingual worker was in a special position (gatekeeper) in terms of her ability to control the others in the group [1].

Thus, to construct an index that embodies the gatekeeper idea we need to look at the degree to which a given point must depend upon a specific other (the gatekeeper) to get information to and from still other points. This may be done by measuring the pair-dependency exhibited by dyads in a network.

The proposed measure of pair-dependency is an index of the degree to which a particular point must depend upon a specific other – as a relayer of messages – in communicating with all others in the network. It may be calculated for any pair of points in a network. It is itself an important property of point pairs, but it also provides a measure from which two of the three established kinds of indexes of global point centrality may be derived as functions.

## Measuring Pair-Dependency

Consider a graph representing the symmetrical relation, "communicates with" for a set of people. When a pair of points is linked by an edge so that they can communicate directly without intermediaries, they are said to be adjacent. A set of edges linking two or more points $(p_i, p_j, p_k)$ such that $p_i$ is adjacent to $p_j$ and $p_j$ is adjacent to $p_k$ constitute a path from $p_i$ to $p_k$. The shortest path linking a pair of points is called a geodesic. There can, of course, be more than one geodesic linking any pair of points.

Now let $g_{ik}$ = the number of geodesics linking a pair of points, $p_i$ and $p_k$, and $g_{ik}(p_j)$ = the number of such geodesics that contain point $p_j$ as an intermediary between $p_i$ and $p_k$, then:

$$b_{ik}(p_j) = \frac{g_{ik}(p_j)}{g_{ik}}$$

Thus, $b_{ik}(p_j)$ is the proportion of geodesics linking $p_i$ and $p_k$ that contain $p_j$; it is an index of the degree to which $p_i$ and $p_k$ need $p_j$ in order to communicate along the shortest path linking them together. Since it is a proportion, $0 \leqslant b_{ik}(p_j) \leqslant 1$. Moreover, when $b_{ik}(p_j) = 1$, $p_j$ is strictly between $p_i$ and $p_k$; they cannot communicate along the geodesic or geodesics linking them without its support in relaying messages. In such a situation the communication between $p_i$ and $p_k$ is completely at the whim of $p_j$; he can distort or falsify any information passing through him.

Now we can define pair-dependency as the degree to which a point, $p_i$, must depend upon another, $p_j$, to relay its messages along geodesics to and from all other reachable points in the network. Thus, for a network containing $n$ points,

$$d_{ij}^* = \sum_{k=1}^{n} b_{ik}(p_j)(i \neq j \neq k)$$

is the pair-dependency of $p_i$ on $p_j$.

We can calculate the pair-dependencies of each point on every other point in the network and arrange the results in a matrix,

$$\mathbf{D} = [d_{ij}^*]$$

Each entry in **D** is an index of the degree to which the point designated by the row of the matrix must depend on the point designated by the column to relay messages to and from others. Thus **D** captures the importance of each point as a gatekeeper with respect to each other point – facilitating or perhaps inhibiting its communication.

Whenever any person is in a position to be a gatekeeper for communications, others must depend on that person. A gatekeeper position, however, may be either rather wide or quite narrow in its impact. A given position may be minimal in its overall impact on a total network of communications, but have great importance to, say, one of its close neighbors. Consider, for example, the network shown in Fig. 1. As Table 1 shows, point $b$ exhibits relatively little betweenness based on its overall centrality. Its sum is 10, and compared with point $d$ - whose sum is 18 – the overall centrality of $b$ is relatively low. From the view of point $a$, however, $b$ is dominant, $a$ needs $b$ to reach all 5 other points while it needs $c$ only for 4. Point $a$ then is locally dependent upon $b$.

Obviously, such local pair effects are of great potential importance to the points affected. In large networks, where individuals may be at considerable distance from one another, global patterns may be submerged and pair effects may be the main factors determining information flows. Moreover, it is reasonable to suspect that the importance of pair-dependency may be increased when a person in a position of authority or power – as in the cases described above – must depend upon a person in a position of considerably less power or authority as

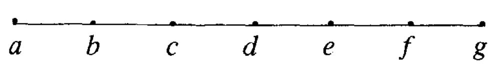


Fig. 1. A seven position chain network.

TABLE I

Local dependencies for the chain shown in Fig. 1.

| | *a* | *b* | *c* | *d* | *e* | *f* | *g* |
|---|---|---|---|---|---|---|---|
| *a* | 0 | 5 | 4 | 3 | 2 | 1 | 0 |
| *b* | 0 | 0 | 4 | 3 | 2 | 1 | 0 |
| *c* | 0 | 1 | 0 | 3 | 2 | 1 | 0 |
| *d* | 0 | 1 | 2 | 0 | 2 | 1 | 0 |
| *e* | 0 | 1 | 2 | 3 | 0 | 1 | 0 |
| *f* | 0 | 1 | 2 | 3 | 4 | 0 | 0 |
| *g* | 0 | 1 | 2 | 3 | 4 | 5 | 0 |
| Σ | 0 | 10 | 16 | 18 | 16 | 10 | 0 |

his or her sole channel of communication to a large set of others. In such cases the high status person virtually abdicates his or her advantage.

## Local Dependency and Standard Centrality Measures

In addition to its importance as a concept in its own right, pair-dependency provides a basis for clarifying two standard measures of centrality. The measure based on betweenness, $C_B(p_j)$, and that based on closeness, $C_C(p_i)^{-1}$, are both functions of **D**. Thus, **D** embodies both betweenness and closeness – it reflects both the degree to which others must depend on a given point and that to which it must depend on others.

The betweenness index was defined by Freeman (1977) as:

$$C_B(p_j) = \sum_{i<k}^{n} \sum^{n} b_{ik}(p_j)$$

If we sum down the columns of **D**, we get:

$$\sum_{i=1}^{n} d_{ij}^* = \sum_{i=1}^{n} \sum_{k=1}^{n} b_{ik}(p_j)$$

$$= 2 \sum_{i<k}^{n} \sum^{n} b_{ik}(p_j)$$

$$= 2C_B(p_j)$$

since, in an irreflexive symmetrical matrix, the sum of the entire matrix is twice the sum of the upper triangle. Or, put another way, the sum of ordered pairs,

$$\sum_{i=1}^{n} \sum_{k=1}^{n}$$

is twice the sum of unordered pairs,

$$\sum_{i}^{n} \sum_{< k}^{n} .$$

The closeness based measure of centrality was defined by Sabidussi (1966) as:

$$C_C(p_i)^{-1} = \sum_{k=1}^{n} d(p_i, p_k)$$

where $d(p_i, p_k)$ = the number of edges in the geodesic linking $p_i$ and $p_k$.

Consider a pair of points, $p_i$ and $p_k$, that are connected by a single geodesic of length $d$. Such a geodesic must pass through $d - 1$ intermediate points or, put another way, $d - 1$ points stand between $p_i$ and $p_k$.

The situation is somewhat more complicated if two or more geodesics connect $p_i$ and $p_k$. In this case intermediate points take fractional values of betweenness according to the proportion of geodesics linking $p_j$ and $p_k$ upon which they fall.

Recall that $g_{ik}$ = the number of geodesics linking $p_i$ and $p_k$, and $g_{ik}(p_j)$ = the number of such geodesics that contain point $p_j$. If we consider the set of geodesics linking $p_i$ and $p_k$ and sum for all points that are intermediate on any of the geodesics, $p_j$, we get:

$\sum_{j=1}^{n} g_{ik}(p_j)$ = the total number of points falling on all geodesics linking $p_i$ and $p_k$.

Then if we divide by the number of such geodesics, we get:

$$\frac{\sum_{j=1}^{n} g_{ik}(p_j)}{g_{ik}} = \sum_{j=1}^{n} \frac{g_{ik}(p_j)}{g_{ik}}$$

$$= \sum_{j=1}^{n} b_{ik}(p_j)$$

which is simply the number of points falling on any one of the geodesics linking $p_i$ with $p_k$. Such a geodesic, of course, must by definition have a length of one greater than the number of intermediate points it contains. Thus,

$$\sum_{j=1}^{n} b_{ik}(p_j) = d(p_i, p_k) - 1$$

Then if we sum for all the geodesics linking $p_i$ with all other points, $p_k$, in a connected graph we get:

$$\sum_{k=1}^{n} \sum_{j=1}^{n} b_{ik}(p_j) = \left[ \sum_{k=1}^{n} d(p_i, p_k) \right] - \left[ \sum_{k=1}^{n} 1 \right]$$

and since $i \neq k$, the sum:

$$\sum_{k=1}^{n} 1 = n - 1$$

and:

$$\sum_{k=1}^{n} \sum_{j=1}^{n} b_{ik}(p_j) = \left[ \sum_{k=1}^{n} d(p_i, p_j) \right] - (n - 1)$$

But since:

$$\sum_{k=1}^{n} \sum_{j=1}^{n} b_{ik}(p_j) = \sum_{j=1}^{n} d_{ij}^{*}$$

according to the definition above, then:

$$\sum_{j=1}^{n} d_{ij}^{*} = C_c(p_i)^{-1} - (n - 1) ,$$

and the closeness based measure of centrality of a point, $Cc(p_i)^{-1}$, is a function of the row sums of **D**.

From this perspective it is apparent that both the betweenness and closeness based measures of point centrality are determined by the same structural elements of a communication network. Both are functions of pair-dependency. While the betweenness based measure depends upon a point's potential for control of communications, the closeness based measure depends on its independence of such potential control by others.

## Acknowledgement

This report grew out of a computer conference on social networks supported by the National Science Foundation under Grant No. DSI77.16578. Any opinions, findings and conclusions or recommendations expressed in this publication are those of the author and do not necessarily reflect the views of the National Science Foundation.

The author is grateful to Paul W. Holland, Peter and Trudy Johnson-Lenz and Brian Foster for their helpful comments on an earlier draft of this paper.

## Note

1 Similar statements and illustrations have been given by Freeman et al. (1963), Deutscher (1967) and Allen (1969). Moreover, the same sort of reasoning seems to be embodied in the idea of the liaison (Jacobson and Seashore, 1951) and that of the cosmopolite (Rogers and Agarwala-Rogers, 1976, pp. 139–140).

## References

Allen, Thomas J. (1969). "The world; your company: a gate for information! Who guards the gate?" *Innovation* 3: 33–39.

Anthonisse, Jac M. (1971). *The Rush in a Graph.* Amsterdam: Mathematisch Centrum (mimeographed).

Bavelas, Alex (1948). "A mathematical model for group structure," *Applied Anthropology* 7: 16–39.

Bavelas, Alex (1950). "Communication patterns in task oriented groups," *Journal of the Acoustical Society of America* 57: 271–282.

Deutscher, Irwin (1967). "The gatekeeper in public housing," in I. Deutscher and E. Thompson). eds., *Among the People: Studies of the Urban Poor.* New York: Basic Books.

Freeman, Linton C. (1977). "A set of measures of centrality based on betweenness," *Sociometry* 40: 35–41.

Freeman, Linton C. (1979). "Centrality in social networks: I. Conceptual clarification," *Social Networks* 1: 215–239.

Freeman, Linton C., Thomas J. Fararo, Warner Bloomberg, Jr. and Morris H. Sunshine (1963). "Locating leaders in local communities: a comparison of some alternative approaches," *American Sociological Review* 28: 791–798.

Jacobson, Eugene and Stanley Seashore (1951). "Communication patterns in complex organizations," *Journal of Social Issues* 7: 28–40.

Leavitt, Harold (1951). "Some effects of certain communication patterns on group performance," *Journal of Abnormal and Social Psychology* 46: 38–50.

Lewin, Kurt (1947). "Group decision and social change," in T.M. Newcomb and E.L. Hartley, eds., *Readings in Social Psychology.* New York: Holt.

Rogers, Everett M. and Rekha Agarwala-Rogers (1976). *Communication in Organizations.* New York: Free Press.

Sabidussi, Gert (1966). "The centrality index of a graph," *Psychometrika* 31: 581–603.