SPATIAL MODELLING OF SOCIAL NETWORKS WITH APPLICATIONS TO THE DIFFUSION PROCESS: AN INITIAL ANALYSIS

George A. Barnett
Communication Research Laboratory
Rensselaer Polytechnic Institute
Troy, N. Y.

Paper Presented at Workshop
On
Metric Multidimensional Scaling

International Communication Association
Philadelphia, Pennsylvania

May, 1979

The author acknowledges the assistance of Joseph Woelfel, James Dinkelacker, Michael Miller and the staff of the Communication Institute at the East-West Center, Honolulu, Hawaii, without whom this paper would not have been possible.

SPATIAL MODELLING OF SOCIAL NETWORKS:
WITH APPLICATIONS TO THE DIFFUSION PROCESS

A social network may be precisely described by a N x N matris X where N equals the number of nodes or interacting units in the network.[1] The value in each cell is some measured attribute of the relationship or link among the nodes. While there exists a variety of techniques for analyzing this matrix, sociometry (Moreno, 1934), matrix manipulations (Forsyth & Katz, 1946; Festinger, 1949), network analysis (Richards, 1974; Pitts, 1979) and non-metric multidimensional scaling (Goldstein, et al., 1966; Lankford, 1974), none of these methods precisely meet the requirements of the real number system and therefore are incapable of precisely describing changes in the network over time.[2] Longitudinal metric multidimensional scaling (Woelfel & Fink, forthcoming) however, may be used to analyze over-time changes in social networks (Gillham & Woelfel, 1977).

One practical application of the analysis of social networks is in the study of the diffusion of innovations (Rogers, 1975; Danowski, 1976). Implicit in the study of this topic is the change in a social system over time as a new product, practice or idea spreads through the network. Indeed, the prediction of the future adoption of an innovation may be enhanced by the accurate description of the social network of the adopting system at earlier points in time. Among the variables which have been shown to predict diffusion are the boundary point or port where the innovation enters the network (Katz & Kahn, 1978), the interconnectedness of the system (Danowski, 1976), the degree of integration (Roberts & O'Reilly, 1978) and the strength of individual links.

Controlling for all mediating networks (such as telephones and the broadcast media) information and innovations will spread spatially in the same manner that a wave on a pond will grow from a point where a stone breaks the water's surface or the way in which energy is exchanged among particles in a thermodynamic system when the temperature (energy level) of a particular particle is altered (Dodd, 1955). In contemporary society, however, mediating networks are present and they have been shown to predict information dissemination and the rates of adoption of innovations (Rogers, 1971). Indeed as Korzenny (1977) points out, the electronic media alter the "communication distance" among networks by making people "electronically propinquitious."

While both orientations to the study of diffusion are well known, they have rarely been combined into a single theroretical model. Hägerstrand (1967) represents a notable exception. Additionally, while the mathematical transformations among the models are known (Barnett, 1975, 1978), they have not been empirically verified. This paper proposes such a test.

The diffusion process may be described by the formula

$$ds/dt = c/s^{k-1} \quad . \tag{1.0}$$

where c = a scaling constant,

s = distance and

k = some arbitrary power.

Furthermore, k may be specified such that it equals the dimensionality of the space in which the object or information is diffused. This has been the case with the diffusion of gas in a vacuum, where k equals three. The formula for the spread of gas is

$$ds/dt = c/s^{k-1} = c/s^2 \quad . \tag{1.1}$$

Here c equals the mass of the gas and temperature and pressure are assumed to be constant.

If k is allowed to become variable, some interesting relations are revealed. Where k equals one, the unidimensional case of point to point information diffusion, the distance among the nodes becomes irrelevant.

$$ds/dt = c/s^{1-1} = c/s^0 = c/1 = c \quad . \qquad 1.2$$

This is the case when information is transferred within a network or through the mass media. The velocity of the spread of a message is no longer a function of physical distance among the interactants but instead is determined by other variables. These may include the number of nodes which must process the message or the capacity of the channel over which the message is sent.

Where k equals 2, the formula for the diffusion of a message becomes

$$ds/dt = c/s^{2-1} = c/s \quad . \qquad 1.3$$

This is the case for the spread of information on the surface of the earth, which is essentially a curved plane.

Where k equals three, the formula is

$$ds/dt = c/s^{3-1} = c/s^2 \quad . \qquad 1.4$$

This is diffusion in a cube and it is applicable to physical phenomenon, as the spread of a gas in a vacuum. More generally, the constant c may be replaced by a variable or series of variables which take into account the capacity of the channel over which the message is sent and the routing of that information.

Thus, a transform to make these models equivalent is needed. The function necessary to perform the operation between the two models will be power transforms, where the exponent of the distance relations among the nodes of the network are altered according to a function that will allow for precise prediction of diffusion.

The probabilities of an object being adopted or a message being received at a given point is also a function of the dimensionality of the space and the distance from the initial node. In a three dimensional space it is

$$p = c/s^2 \quad . \qquad 2.1$$

The probability decreases to zero as an iverse square, becoming assomtotic with zero as $S \to \infty$. This is shown in figure 2.1.

In a two dimensional space,

$$p = c/s \quad . \tag{2.2}$$

The probability decreases to zero linearly as s becomes large. This is shown in figure 2.2.

In the unidimensional case,

$$p = c/s^0 = c/1 = c \quad . \tag{2.3}$$

Distance per se, is irrelevant and the probability is constant with respect to distance. This is shown in figure 2.3. However, other factors such as the structure of the network determine the probability of receiving the information. That is, the presence or non-presence of a link and the routing parameters will determine the probability and rate of diffusion.

This paper began with a sociomatrix of distances, S. Information should diffuse among the individual nodes in S according to the equations presented above, i.e., the probability of receiving a message is inversely related to distance in a space of two or more dimensions. Thus, S should be raised to the inverse of the k-1 power, where k equals the underlying rank or dimensionality of S. This dimensionality, k, may be viewed somewhat differently in the matrix case. Rather than simply physical space, the dimensionality in this case represents the true rank of the matrix in which the nodes are arrayed. Thus, in the general case,

$$p = CS^{-1(k-1)} = CS^{-1(\ell)} \quad . \tag{3.0}$$

where,

k = dimensionality or rank of the matrix in which the network nodes are arrayed,

C = a matrix of scaling constants

S = the distance matrix

$\ell = k - 1$

In a special two dimensional case as communication among a group or cities or countries,

$$p = CS^{-1} \quad . \qquad 3.1$$

And in the case of point-to-point (model to model) communication,

$$p = CS^{0} = C \quad . \qquad 3.2$$

It should be readily apparent that the probability of receiving a message is again determined by the structure of the network rather than by physical distance in the unidimensional case.

This analysis has made the assumption that networks exist only in a unidimensional configuration. This is probably not the case. Due to switching in the network, distances do appear to exist. Generally, these distances resemble to a small degree the actual physical distances among the nodes (Schwartz, 1977). Thus, the dimensionality will be determined through the empirical analysis suggested later in the paper. This suggests that the equations derived above may represent the idealized case only and that l is somewhat greater than zero.

In order to test the above models, three sets of data are necessary. 1. A spatial model, such as the matrix of intercity distances D presented below (See Table One). 2. A matrix of the frequency of the use of one or more mediated network(s) among these nodes, S. This may be a matrix of the frequency of long distance telephone calls per unit time among the cities in the matrix or such things as the frequencies of air traffic, mails or telegraphs. 3. The final set would be over-time data on the adoption of some innovation by each node.[3] Ideally, the information about this innovation should have been transferred among the nodes solely through the mediated networks. This could be the sales data for any product which is not advertised in the broadcast media or the national print media, such as illicit drugs.

Other factors may determine the actual relationship between these matrices (D and S). These may include the cost of communicating in the network (telephone rates among the cities), the frequency of direct interpersonal contact, and the cost to adopt the innovation.

It is worth noting that the proposed study uses aggregate level data (Rogers, 1978). The unit of analysis is not the individual but the entire social system which makes up each node. In this case, that would be the total population of each city. This allows for more accurate prediction of the adoption rates among the node's of the network because random individual variance and the effects of other communication channels are randomized. Additionally, it makes possible the analysis suggested here.

In order to demonstrate the utility of these models to diffusion, the network matrix S must be compared to the physical distance matrix D. That is, the physical distance factor must be controlled out of the network. This may be done in two ways. One, the matrices of distances (S) between the nodes for the physical space and the network are transformed to spatial coordinate matrices (S* & D*). This has been done in table two for the cities in table one (See Table Two). Next, these matrices (S* & D*) are rotated to a least-square congruence. Then, the differences between S* and D* are determined by simple subtraction. These procedures are generally known as metric multidimensional scaling (Woelfel and Fink, forthcoming). A computer program, GALILEO TM IV (Woelfel, et al., 1976) performs the necessary operations for this analysis. They are summarized by equation 4.1.

$$D^* - S^* = a \qquad 4.1$$

where,

D^{*} = spatial coordinates of D

S^{*} = spatial coordinates of S

a = the difference between D^{*} and S^{*}.

a is a vector which indicates the difference in location for each node between D and S. Therefore, it is an indication of the degree of independence between D and S, such that the greater the a, the more independent the two matrices.

Alternatively, the power transform between S and D can be determined, in this case ℓ.

$$D^{\ell} = S \qquad 4.2$$

The procedures for performing these transformations are discussed by Dinkelacker (personal correspondence). From the earlier discussion it was suggested that the measured ℓ should be between zero and one ($0<\ell<1$). This suggests that some degree of physical distance will be taken into account in the network matrix or what Schwartz (1977) calls the topological characteristics of the network. In this case, the measured ℓ will describe the degree of structure which is taken into account by physical distance. Thus, as $\ell \rightarrow 0$, the network will be independent of distance and as $\ell \rightarrow 1$, the actual physical distance will have a greater influence on the network structure. Over a series of different networks, it would be expected that there would be a high inverse correspondence between a (4.1) and ℓ (4.2).

These models were developed to accurately describe the diffusion process. With over-time data on the relative frequency of adoption for each node it can be determined where the innovation enters and how it spreads through the social system. Without network information the best estimate of diffusion would be that the node closest to the entry port would have the highest rate of adoption and therefore the next highest relative frequency for that measured point in time. By examining the overall lag in relative frequencies for the entire system we can describe the diffusion process and determine how well the distance matrix D describes this process.

The adoption frequency curve for each node should resemble the classic S-shaped curve the form of which has been precisely described by Barnett (1975, 1978). It is the lag in time of this curve that is of interest. The probability of adoption should resemble a reverse U curve; initially low, increasing to a peak, and then, growing small. The lag in this will correspond closely to the lag in the actual adoption curve. It may be substituted for the frequency curve in nodes with small populations.

Since the nodes in S are networked, they should provide better predictions of diffusion than D. Nodes which are closer to the entry port in S than in D should experience an increase in the frequency of adoption at an earlier lag than would be predicted by D. Additionally, the degree of change in lag time would be accurately preducted by a. By examining the direction of change from the entry port and the node's value of a, the lag in the adoption curve for each node can be determined.

A Partial Empirical Example:

In order to perform an example of the analysis suggested above, the author obtained the frequencies of airline traffic (number of passengers for the twelve months ending June 30, 1978) among the sixteen cities in table one (Matrix D). They are presented in table three and represent matrix S as specified on page 5.

Table three about here

These data were then multiplied by a scaling constant $1/s_{ij} \times 2.307 \times 10^7$ so that they would be inversely scaled and adjusted to a standard or common metric with the physical distances in D which were measured in kilometers. The value 2.307×10^7 was obtained as the ratio of the means of the non-zero cells in D and S. By multiplying each cell in S by this value the two means became equivalent. The multiplication of the values in S by a scaling constant in no way distorts the data and in fact simplifies the analysis by removing the effects of the differential scaling metrics from the data. This matrix, S' is presented in table four.

Table four about here

S' was the next transformed to a set of spatial coordinates, S*, which is presented in table 5. An examination of the characteristic roots of S' (the eigenvalues presented with S*) reveals an interesting finding. 40.2% of the total variance in S' is accounted for by negative (imaginary) roots, indicating that the spatial representation is highly non-Euclidean. A warp factor of 3.04 was also obtained, further verifying this point. Warp equals the ratio of the sum of the positive roots to the trace (sum of all roots) and thus provides a convenient measure of the degree to which the space is non-Euclidean. A indepth discussion of warp is provided later in the paper and by Woelfel, et al. (1978). This finding is quite significant and will be discussed shortly.

Table five about here

A graphic representation of the first two real (positive) dimensions is presented in figure one. Although there is considerable variance on the imaginary (negative) dimensions of the spatial configuration, these two dimensions account for 49.3% of the total variance. A visual examination of figure one readily suggests that the frequency of airline travel may be described as a star-type network with tendencies toward a tree-type configuration, although among the nodes at the center (hub) a mesh-type network may be the best descriptive label (Schwartz, 1977; Katz & Kahn, 1977). In order to travel by airplane from New Orleans to Phoenix or Seattle, the nodes at the periphery or the points of the star, one must go through one of the central switching nodes such as Chicago. Also, the plot suggests that Atlanta serves as a tree node or an intermediate switching facility, taking passangers from New Orleans and Miami and rerouting their travel prior to their reaching the more central nodes. Among the central nodes (Chicago, Cleveland, Dallas, Denver, New York, Los Angeles, San Francisco and Washington), there appears to be a mesh-type network with each node having direct link to each other.

[Figure One About Here]

Matrices S* and D* were next compared by translating D* to a common origin with S* and then rotating them to a least-squares congruence. The differences between the two spatial configurations are presented in table six. It is equivalent to vector a in equation 4.1. Table six shows the degree of departure from the physical distances that exist in the network of airline traffic among the sixteen cities. The negative sign indicates that there is greater difference between S* and D* on the negative, imaginary dimensions rather than on the positive, real ones. This is because there is considerably more variance on the imaginary dimensions in the network data (40.2%) than in the physical data (14.0%).[4] Differences on these dimensions must also be minimized, accounting for the discrepancies between the matrices and the negative signs in table six.

The root mean square (RMS) difference between the two spaces was 1846.00 units. This value is 91.5% of the mean value in D and S', indicating a considerable degree of difference and suggesting that only a small proportion of the airline network can be attributable to physical distance. When the direction of change is examined with the plot and the matrices S' and D along with the magnitudes of difference in table six, it becomes clear that the changes are departures from the spatial configuration and toward the network configuration. Thus, in a case where information or innovations would spread primarily by the airlines, physical location would provide little predictive power in determining a node's probability of learning a bit of information or its rate of adoption.

[Table Six About Here]

The transformation proposed earlier was not attempted because it does not appear to be correct. The power transformation (equation 4.2) does not have an imaginary term necessary to add the loadings on the negative dimensions. The model further suggested that the same transformation could be applied to each element in D. This does not appear to be the case. An examination of table six indicates the differences among nodes seem to be independent. For example, San Francisco changed -628.14 units while Los Angeles changed 1803.63 units and both moved toward the hub of the network. This transformation grew out of the assumption implied by equation 1.2 that the network was completely and equally interconnected or a fully mesh-type network. This is not the case with airline traffic. The model, however, may provide a more adequate solution in a fully mesh-type network as among the cities at the hub of this one.

The model further assumes that the nodes are of equal mass, i.e., each node has an equal potential or capacity for linking and using the network. This is not the case. The cities at the hub of the network seem to be those with greater population. Population may be an excellent indicator of a node's mass. In this case, it does seem to predict the potential for use of the airline traffic network by a node. Thus, it may be suggested that a weighting factor, population, should be incorporated into this model.

This discussion need not imply the total inadequacy of the proposed model. The absolute magnitude of the transform appears to be correct. ℓ (equation 4.2) would be quite small, approachig zero. This appears to be the case because the RMS difference between S* and D* is large relative to the mean value of S' and D, indicating that considerable foreshortening of the vector lengths in D is necessary to make it equivalent to S. In summary, the suggested transformation was not attempted because a large proportion of the variance in the sociomatrix was attributable to imaginary roots. This indicates that a complex function would be more appropriate. It was also not attempted because of some erroneous assumptions which make the model useful only in the case of a mesh-type network.

Network data by definition is non-Euclidean, i.e., at least one of the characteristic roots of S' must be imaginary. The reason for this is that any three nodes which vary in centrality must violate the rule of triangular inequalities. The exception is the completely and equally interconnected network. Any three points (nodes) can be said to form a Euclidean triangle if and only if the sum of any two of the distances among them (the link) does not exceed the third. For any set of k points (S*), those points represent a Euclidean configuration of and only if the triangular inequalities rule is not violated for any three of the points. When the triangle inequalities rule is violated for any triple of points, the result is a Riemann manifold, and which is represented by a coordinate system in which some of the dimensions are imaginary. The locations of the non-Euclidean relations among the points may be determined by formula 5.1.

$$d^2_{jk} = d^2_{ij} + d^2_{ik} - 2d_{ij}\ d_{ik}\ \cos\theta \qquad 5.1$$

In cases, where, cos $\theta \leq /1.0/$, the relations may be considered Euclidean. In cases where, cos $\theta > /1.0/$, the relations among the three concepts may be considered non-Euclidean. It is from this later case that the complex eigenroots result (Woelfel, et al. ,1978).

For a network, take the three nodes a, b, and c. C is a central switching facility or a liaison in a social org.nization. A and b both send information to c, where it is processed and then sent along to the other node, b or a. In terms of communication distance, both a and b are close to c, but very far apart from each other. In other words, the abc triangle has two short legs (ac,bc) and one very long one (ab), such that the sum of the angles exceeds 180° and the cos θ exceeds 1.0. Thus, this triad cannot exist in a two-dimensional space without a complex dimension to foreshorten the ab leg.[5]

This conclusion is of some significance because in his discussion of the use of multidimensional scaling (MDS) for sociometric research, Lankford (1974) concluded that MDS was totally inadequate for clique identification. In the reported analysis he used a metric form similar to the method employed here. However, he did not indicate that he took into account the imaginary dimensions (those with negative toots). Additionally, did he not present an eigenvector nor specify nodes' loadings on the obtained coordinate system. In the example reported here, over 40% of the variance occurred on those dimensions whose roots were negative. Any analysis which did not take this variance into account would clearly be inadequate and lead to erroneous conclusions. Thus, any future sociometric research using MDS should use a computer program capable of dealing with complex eigenroots such as GALILEO IV(Woelfel, et al., 1976).

As empirically demonstrated, no one model may adequately describe all network data and provide transformation to the physical configuration. The data presented here while providing a number of interesting insights about network analysis may not be generalizable to other networks. The proposed transformational model was inadequate for this data set because it did not have a complex term or weighting factor for the node's mass. Future research will attempt to take these factors into account. It should be noted that the model originally proposed may prove more adequate in cases of fully mesh-type networks--say long distance telephone calls among the sixteen cities in this example. This research is also planned.

In summary, this paper proposed a mathematical model to describe the relationship between social networks and spatial models. Airline traffic data was used to demonstrate an application of this model. The model proved inadequate because it did not take into account the complex nature of the sociomatrix and the topological characteristics of the network itself. The analysis demonstrated that both theoretically and empirically, network data is non-Euclidean. This finding has important implications for all sociometric research. Future research is planned to take into account the necessary modification in the transformational model.

FOOTNOTES

1. This assumes that the relationship among the nodes is relatively stable over time. In the case where the relations among the nodes are dynamic this process of change can be precisely described by gathering matrix S at a number of points in time and calculating the changes between S_{t_i} and S_{t_j}.
2. Generally, the problems with these procedures are often the manners in which the data is entered into the analysis as well as the mathematical properties of the procedures themselves. For example, the matrix manipulation of Festinger meets the property of the real number system but only information on whether or not a link is present entered into the analysis. Thus, there is insufficient variance to make full use of the technique.
3. The selection of the time frame for the frequencies of the use of the network and the adoption of some innovation may prove to be a problem. These measures should take place quite frequently, i.e., they should be aggregated over a short interval of time. This interval should be less than or equal to 1/2 of the shortest cycle in order to completely index the minimum cycle through discrete measurements in time (Arundale, 1977). Arundale's advise, however, is of little utility for the study of a given innovation. The reason is that the frequency of the cycle is unknown before the process begins. However, for certain variables such as travel amongst certain nodes, seasonal periodicities are known and may prove to be a useful guide in the selection of the invertal of time over which the data can be aggregated.

4. The negative eigenroots in the physical data may be attributable to measurement error such as measuring the distance from city limit to city limit rather than point-to-point (airport to airport or city hall to city hall).

5. One must assume either that there is some communication between a and b or that a maximum value exists for the attribute (discrepancy) of the ab link. Without such a value, $ab = \infty$ and calculations cannot be performed. In the airline traffic network, this value would have been 2.307×10^7.

TABLE ONE: Airline Distances Among 16 Selected U.S. Cities (one unit = one kilometer)

	CITY	Atlanta	Boston	Chicago	Cleveland	Dallas	Denver	Detroit	Los Angeles	Miami	New Orleans	New York	Phoenix	Pittsburgh	San Francisco	Seattle	Washington
1.	Atlanta	0															
2.	Boston	1508	0														
3.	Chicago	944	1369	0													
4.	Cleveland	891	886	496	0												
5.	Dallas	1160	2496	1292	1649	0											
6.	Denver	1950	2846	1480	1974	1067	0										
7.	Detroit	869	986	383	145	1607	1860	0									
8.	Los Angeles	2310	4177	2807	3297	2005	1337	3191	0								
9.	Miami	972	2019	1911	1749	1788	2777	1354	3764	0							
10.	New Orleans	682	2186	1340	1487	713	1741	1511	2692	1076	0						
11.	New York	1204	302	1147	652	2211	2624	1258	3944	1757	1884	0					
12.	Phoenix	2562	3701	2338	2814	1427	943	2719	574	3189	2117	3451	0				
13.	Pittsburgh	838	777	660	185	1721	2124	330	3437	1625	1478	510	2941	0			
14.	San Francisco	3442	4343	2990	3485	2386	1524	3364	558	4174	3099	4137	1051	3643	0		
15.	Seattle	3511	3979	2795	3260	2705	1643	3118	1543	4399	3381	3874	1792	3440	1091	0	
16.	Washington	874	632	961	492	1907	2404	637	3701	1485	1554	330	3191	309	3929	3849	0

TABLE TWO: Spatial Coordinates for 16 Selected U.S. Cities

	1	2	3	4	5	6	7	8
1. Atlanta	-645.26	-672.22	800.83	-100.36	-100.56	-85.76	23.37	-100.35
2. Boston	-1723.14	635.29	177.62	-243.67	98.63	-163.42	75.98	-25.45
3. Chicago	-389.42	349.90	6.46	-97.94	29.77	-66.98	1.91	-7.54
4. Cleveland	-890.51	355.89	49.71	-170.01	43.58	-117.32	13.75	-33.92
5. Dallas	285.95	-594.45	-295.84	182.05	-3.83	131.22	-218.84	-40.18
6. Denver	1058.39	219.01	-241.49	75.06	-49.94	47.78	511.23	347.09
7. Detroit	-774.21	254.17	-83.72	-69.71	15.93	-48.35	414.95	-79.14
8. Los Angeles	2314.28	-409.31	644.48	13.06	-405.64	-43.56	-210.73	-85.20
9. Miami	-1360.94	-1273.00	-333.95	-224.15	-270.11	-197.49	169.33	-33.06
10. New Orleans	-351.95	-913.83	-294.36	48.94	-104.99	21.69	-241.34	-125.81
11. New York	-1534.93	361.91	144.51	-235.85	66.13	-162.01	-224.32	-199.23
12. Phoenix	1799.56	-400.04	-374.29	277.10	-183.34	176.53	-154.81	32.21
13. Pittsburgh	-1055.21	260.44	41.55	-188.93	28.80	-132.94	2.83	-53.26
14. San Francisco	2627.84	290.86	-289.99	368.51	-68.52	257.68	-63.66	156.33
15. Seattle	2010.26	1476.12	7.66	727.66	855.51	638.33	-115.91	383.12
16. Washington	-1370.69	59.24	40.82	-361.75	48.57	-255.40	16.26	-135.59
EIGENVALUES (ROOTS)OF EIGENVECTOR MATRIX--	32997093.42	6819619.60	1690325.50	1168683.48	1051710.81	729715.24	711340.63	397642.79
PERCENTAGE OF VARIANCE ACCOUNTED FOR BY INDIVIDUAL FACTORS-	83.69	17.29	4.28	2.96	2.66	1.85	1.80	1.00
PERCENTAGE OF VARIANCE ACCOUNTED FOR BY INDIVIDUAL FACTORS IN THEIR OWN SPACES-	72.37	14.95	3.70	2.56	2.30	1.60	1.56	0.87

SUM OF ROOTS 39425214.64

WARP FACTOR 1.15

	9	10	11	12	13	14	15	16
1. Atlanta	-23.96	-23.22	105.69	111.88	132.60	116.21	128.01	931.79
2. Boston	1.61	-2.17	4.58	57.70	47.97	40.64	-22.96	-97.41
3. Chicago	0.29	-0.00	-2.47	22.16	14.45	23.29	0.38	-29.23
4. Cleveland	-5.43	-5.62	22.43	55.35	47.59	46.13	8.55	-47.16
5. Dallas	-18.40	-16.59	86.16	-13.04	-24.17	-79.58	-50.86	21.55
6. Denver	102.82	96.80	-464.85	-256.03	-232.58	1.26	26.12	89.33
7. Detroit	-20.27	-19.46	90.20	79.22	145.79	20.28	5.39	-104.71
8. Los Angeles	-16.77	-16.78	71.93	118.48	134.25	279.89	391.06	-980.07
9. Miami	4.36	2.49	-26.25	117.45	151.53	291.62	330.24	9.89
10. New Orleans	-37.06	-34.92	167.44	94.18	92.40	51.84	87.55	-0.72
11. New York	-53.37	-50.87	238.69	184.07	163.84	60.00	6.33	-105.25
12. Phoenix	4.26	4.59	-16.85	-58.45	-52.36	3.92	95.42	14.34
13. Pittsburgh	-10.21	-10.24	43.63	75.83	69.53	65.10	28.62	-52.11
14. San Francisco	35.55	34.68	-155.79	-186.01	-190.01	-118.28	-43.80	52.14
15. Seattle	69.12	70.16	-292.11	-574.16	-667.84	-931.65	-1051.27	284.91
16. Washington	-29.31	-28.81	127.55	171.32	167.01	129.28	61.19	12.73
EIGENVALUES (ROOTS) OF EIGENVECTOR MATRIX --	23477.58	21778.55	-462286.09	-564772.84	-692257.71	-1097198.59	-1410748.47	-1958909.27
PERCENTAGE OF VARIANCE ACCOUNTED FOR BY INDIVIDUAL FACTORS-	0.06	0.05	1.17	1.43	1.75	2.78	3.57	4.96
PERCENTAGE OF VARIANCE ACCOUNTED FOR BY INDIVIDUAL FACTORS IN THEIR OWN SPACES-	0.05	0.35	7.49	9.16	11.23	17.79	22.88	31.77

TABLE THREE: Airline Traffic Among 16 Selected U.S. Cities: Number of Passages for 12 Months Ending June 30, 1978

CITY	Atlanta	Boston	Chicago	Cleveland	Dallas	Denver	Detroit	Los Angeles	Miami	New Orleans	New York	Phoenix	Pittsburgh	San Francisco	Seattle	Washington
1. Atlanta	0															
2. Boston	16926	0														
3. Chicago	32507	41370	0													
4. Cleveland	10050	13505	37203	0												
5. Dallas	20462	10556	36885	7705	0											
6. Denver	8258	10378	41643	5603	24691	0										
7. Detroit	15815	18855	59546	20818	11571	9090	0									
8. Los Angeles	19562	33840	105346	18258	39720.	43335	29826	0								
9. Miami	31497	25575	41020	12752	11903	10082	20515	26797	0							
10. New Orleans	15774	6050	19411	3501	23170	7536	5866	14326	15180	0						
11. New York	76949	184968	183012	56010	49338	39547	74322	164384	180912	31974	0					
12. Phoenix	3342	6120	31480	5878	10292	18933	9712	38889	3020	1958	23120	0				
13. Pittsburgh	10318	15270	29874	4284	6178	5656	12863	13991	14644	3804	64958	4099	0			
14. San Francisco	11040	26551	47388	8841	17523	25145	16251	110692	14092	8301	111092	17653	7603	0		
15. Seattle	4715	5269	14711	2219	6021	13060	3562	49581	4189	2627	16164	7030	1753	39601	0	
16. Washington	31945	63766	54025	15283	19618	17185	24062	41488	29852	10412	187202	7775	15802	32113	9259	0

TABLE 4: AIRLINE NETWORK DATA TRANSFORMED BY $1/s \times 2.307 \times 10^7$

Atlanta	0.0															
Boston	1363.44	0.0														
Chicago	710.56	558.29	0.0													
Cleveland	2295.46	1707.18	620.58	0.0												
Dallas	1128.12	2184.73	625.20	2994.48	0.0											
Denver	2793.78	2223.95	553.68	4117.99	934.33	0.0										
Detroit	1458.02	1222.71	387.58	1107.36	1993.25	2537.70	0.0									
Los Angeles	1178.88	221.47	219.16	1264.24	581.36	532.92	772.84	0.0								
Miami	731.32	902.04	1808.69	562.91	1937.88	2288.54	1123.51	860.51	0.0							
New Orleans	1462.64	3813.47	562.91	1808.69	996.62	3061.39	3933.43	1610.29	1520.31	0.0						
New York	299.91	124.58	1188.10	6588.79	468.32	583.67	29.99	140.73	126.88	722.09	0.0					
Phoenix	6902.54	3769.64	126.88	412.95	2242.40	1218.10	2376.21	592.90	7638.48	11781.84	998.93	0.0				
Pittsburgh	2235.48	1518.01	733.63	3924.21	3735.03	4078.77	7692.54	1649.50	1575.68	6065.10	355.28	5629.07	0.0			
San Francisco	2090.14	869.74	772.84	5384.54	1317.30	918.19	1418.80	207.63	1637.97	2779.93	207.63	1305.76	3033.70	0.0		
Seattle	4893.14	4378.68	1568.76	2609.22	3831.93	1767.16	6475.74	466.01	5506.80	8782.74	1428.03	3280.55	13161.43	583.67	0.0	
Washington	722.09	362.20	426.79	1508.78	1176.57	1342.67	959.71	555.99	772.84	2214.72	122.27	2966.80	1460.33	717.48	2491.56	0.0

TABLE FIVE: SPATIAL COORDINATES OF AIRLINE TRAFFIC NETWORK FOR 16 SELECTED U.S. CITIES

		1	2	3	4	5	6	7	8
1.	Atlanta	-1304.26	-1165.89	65.30	-640.10	802.74	-211.48	397.78	146.52
2.	Boston	-277.28	352.07	222.34	-1081.64	-223.58	-309.64	-26.18	132.05
3.	Chicago	231.80	80.37	395.39	587.34	112.22	-488.57	-272.38	-95.34
4.	Cleveland	347.57	-176.44	3835.84	113.50	-57.39	15.77	-21.82	-17.70
5.	Dallas	105.95	-216.39	-268.50	1211.20	466.77	190.56	122.71	83.72
6.	Denver	806.36	-108.36	-890.05	894.58	-758.33	14.85	52.14	-9.67
7.	Detroit	-756.78	1156.76	452.75	-208.13	399.01	663.13	-245.57	-105.70
8.	Los Angeles	375.09	7.44	231.36	208.24	-236.42	152.37	28.97	-27.27
9.	Miami	-1755.92	-1303.36	285.39	-890.81	-515.33	371.34	36.82	15.53
10.	New Orleans	-3947.55	-4188.93	-38.86	746.83	-55.85	-35.08	-103.38	-34.41
11.	New York	115.88	79.39	-2764.38	-257.07	290.09	22.44	136.53	4.47
12.	Pheonix	4170.55	4531.93	193.65	238.45	33.84	7.53	-15.96	-17.15
13.	Pittsburgh	-4967.13	3972.37	-158.91	4.41	-127.99	-149.79	22.29	-14.78
14.	San Francisco	583.59	-4.81	-1722.78	-270.16	-78.38	-50.41	-298.56	-30.19
15.	Seattle	6233.51	-2951.81	-39.60	-316.43	21.77	-14.81	-16.68	-25.67
16.	Washington	38.62	-60.27	201.06	-250.21	-73.16	-178.22	203.28	-4.38

EIGENVALUES (ROOTS) OF EIGENVECTOR MATRIX--

103271780.40	67205541.90	26843898.01	6169930.69	2100690.27	1075711.93	474694.09	70880.50

PERCENTAGE OF VARIANCE ACCOUNTED FOR BY INDIVIDUAL FACTORS-

151.44	98.55	39.36	9.04	3.08	1.57	0.69	0.10

PERCENTAGE OF VARIANCE ACCOUNTED FOR BY INDIVIDUAL FACTORS IN THEIR OWN SPACES-

49.83	32.43	12.95	2.97	1.01	0.51	0.22	0.03

SUM OF ROOTS 68191540.59

WARP FACTOR 3.03

TABLE FIVE CONTINUED

	9	10	11	12	13	14	15	16
1. Atlanta	-34.16	-350.37	109.19	-758.80	190.45	92.38	-1129.16	320.27
2. Boston	-33.11	-4.65	-450.04	405.70	566.69	653.57	-272.16	659.36
3. Chicago	22.52	189.69	-7.76	53.23	-623.38	1169.13	110.07	1633.59
4. Cleveland	4.30	19.44	11.20	-5.28	-2.05	-2024.95	1163.70	2250.90
5. Dallas	-19.82	-159.50	-256.32	127.74	-280.23	711.92	747.86	986.61
6. Denver	1.95	63.19	-156.03	-494.24	506.94	125.60	436.74	1250.85
7. Detroit	25.42	148.85	-30.54	85.65	469.33	952.71	719.36	85.66
8. Los Angeles	6.93	-11.24	343.16	55.48	126.15	1087.54	74.28	1580.00
9. Miami	-3.93	4.09	-99.66	156.20	-882.93	422.47	-1569.42	-43.76
10. New Orleans	8.28	48.38	66.33	233.06	261.45	-694.23	-2593.73	-3289.60
11. New York	-1.78	86.84	48.08	392.42	-187.15	-1800.17	631.47	2270.87
12. Pheonix	4.14	21.30	23.49	1.35	-43.20	-528.33	-4078.62	-2284.34
13. Pittsburgh	3.61	13.13	67.21	-104.62	-146.79	-245.30	2918.27	-4054.86
14. San Francisco	7.88	-39.23	71.82	-84.33	-14.04	-895.57	392.80	1707.90
15. Seattle	6.17	35.87	27.98	-13.00	-87.26	96.88	2618.53	-4391.84
16. Washington	1.59	-65.80	231.90	261.84	146.02	876.34	-170.26	1275.38
EIGENVALUES (ROOTS) OF EIGENVECTOR MATRIX--	4101.51	-228534.64	-504818.30	1333731.89	-2253222.98	-14335751.04	-45769645.07	-74591781.66
PERCENTAGE OF VARIANCE ACCOUNTED FOR BY INDIVIDUAL FACTORS	0.00	0.33	0.74	1.95	3.30	21.02	67.12	109.39
PERCENTAGE OF VARIANCE ACCOUNTED FOR BY INDIVIDUAL FACTORS IN THEIR OWN SPACES	0.00	0.16	0.36	0.95	1.62	10.31	32.92	53.65

TABLE SIX

DIFFERENCES BETWEEN PHYSICAL SPACE AND AIRLINE NETWORK

1.	Atlanta	757.15
2.	Boston	196.70
3.	Chicago	-1715.96
4.	Cleveland	1831.48
5.	Dallas	-1344.39
6.	Denver	-1151.51
7.	Detroit	401.03
8.	Los Angeles	1803.63
9.	Miami	-222.98
10.	New Orleans	2663.18
11.	New York	2542.61
12.	Phoenix	3473.80
13.	Pittsburg	2571.32
14.	San Francisco	-628.14
15.	Seattle	2887.94
16.	Washington	-229.89

THE ROOT MEAN SQUARE DIFFERENCE BETWEEN ALL NODES IN SPACE 1 AND THEIR COUNTERPARTS IN SPACE 2
1846.00

FIGURE ONE: Two Dimensional Representation of Airline Traffic Network

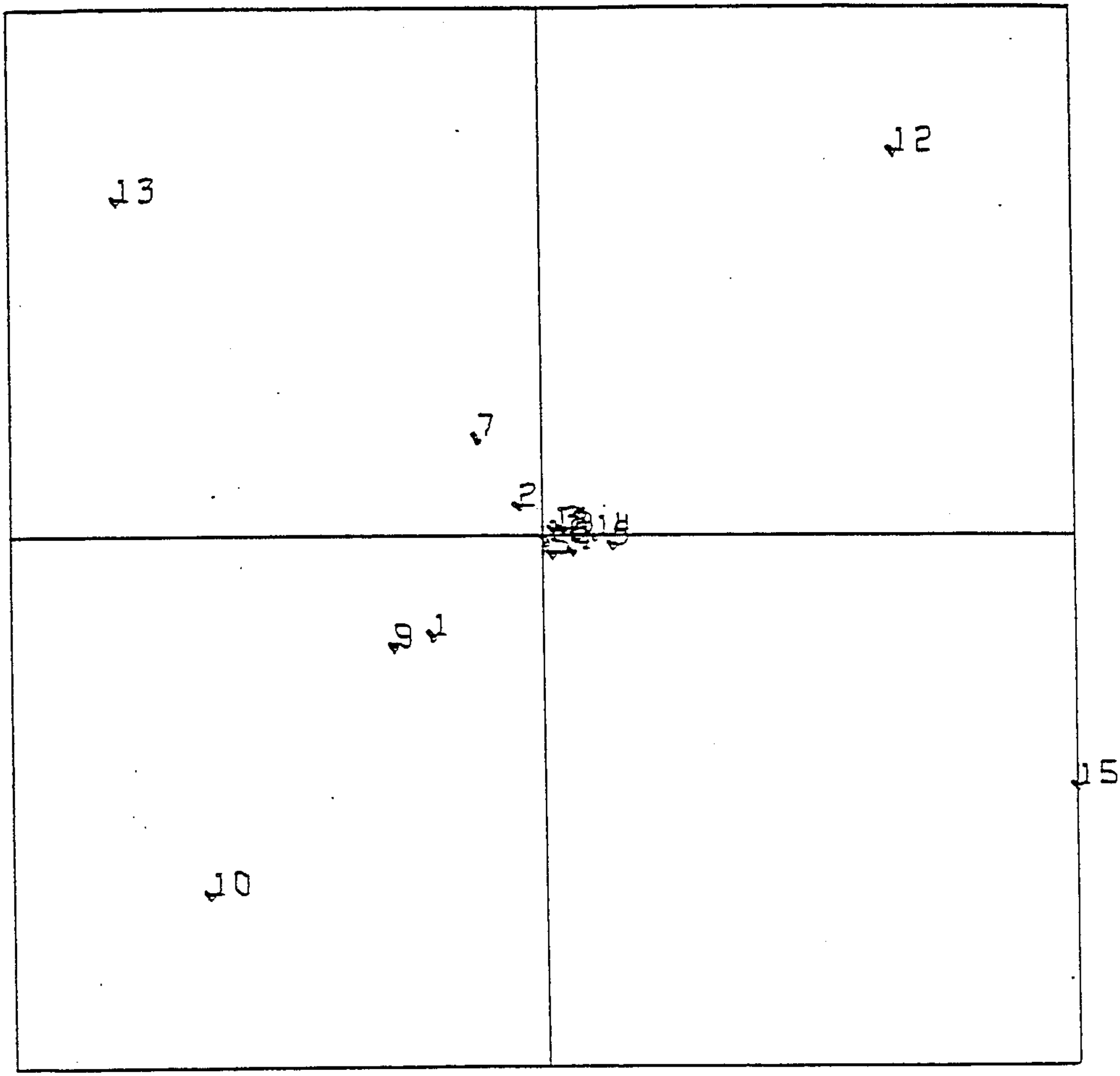


FIGURE 2.1

The Probability of Receiving a
Message in a 3-Dimensional Space

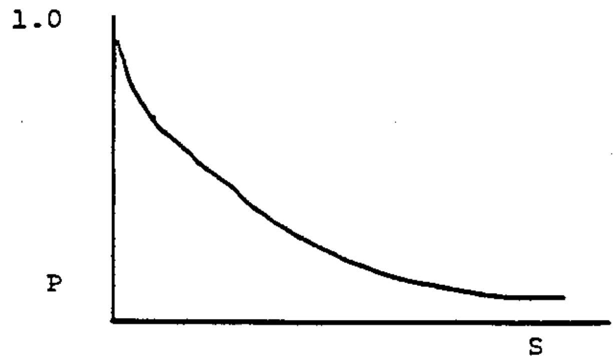


FIGURE 2.2

The Probability of Receiving a
Message in a 2-Dimensional Space

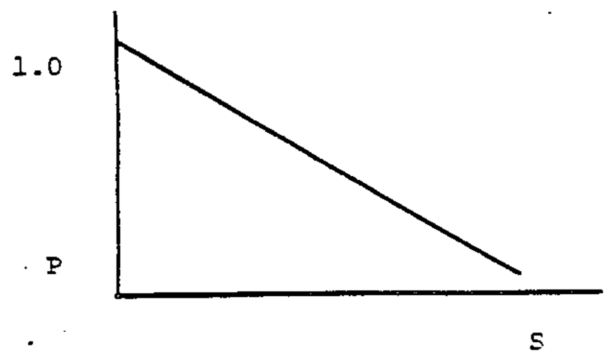
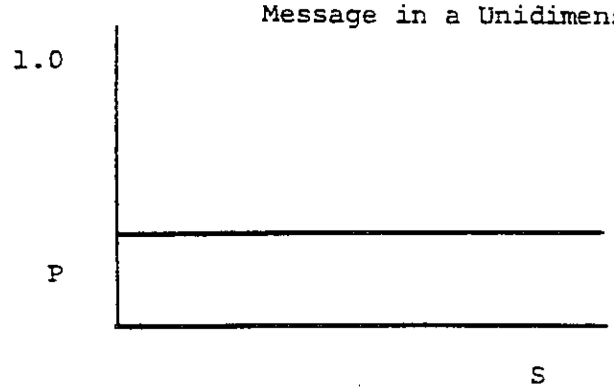


FIGURE 2.3

The Probability of Receiving a
Message in a Unidimensional Space



*These equations may be rewritten with the distance that a message will difuse as a function of time for a given dimensionality as presented in pages 2-3.

REFERENCES

Arundale, R.B., "Sampling across time for communication research: a simulation." In Hirsch, P.M., P.V. Miller and F.G. Kline (Eds.) Strategies for Communication Research. Beverly Hills, Ca.: Sage, 1977.

Barnett, G.A., "The diffusion of complex innovations." Unpublished research report, Department of Communication, Michigan State University; East Lansing, 1975.

Barnett, G.A., "An associational model for the diffusion of complex innovations." Paper presented at the annual meeting of the International Communication Association, Chicago, 1978.

Danowski, J.A., "Communication network analysis and social change: group structure and family planning in two Korean villages." In Chu, G., S. Rahim & D.L. Kinkaid (eds..) Communication for Group Transformation in Development. Honolulu: East-West Communication Institute, 1976.

Dinkelacker, J.W. "Pair comparison data re-expressions to achieve symmetry." personal correspondence.

Dodd, S.C., "Diffusion is predictable: testing probability models for laws of interaction." American Sociological Review 20:392-401, 1955.

Festinger, L., "The analysis of sociograms using matrix algebra." Human Relations 2:153-158, 1949.

Forsyth, E. & Katz, L. "A matrix approach to the analysis of sociometric data." Sociometry, 9: 340-347, 1946.

Gillham, J.R. & J. Woelfel, "The Galileo System of measurement: preliminary evidence for precision, stability, and equivalence to traditional measures." Human Communication Research 3(3): 222-234, 1977.

Goldstein, K.M., S. Blackman, & D.J. Collins, "Relationship between sociometric and multidimensional scaling measures." Perceptual and Motor Skills, 23:639-643, 1966.

Hägerstrand, T. Innovation Diffusion as a Spatial Process. Chicago: University of Chicago, 1953.

Katz, D. & R.L. Kahn, The Social Psychology of Organizations. New York: Wiley, 1978.

Korzenny, F. "A theory of electronic propinquity: mediated communication in organizations." Paper presented at the annual meetings of the International Communication Association, Berlin, 1977.

Lankford, P.M. "Comparative analysis of clique identification methods." Sociometry 37(2): 287-305, 1974.

Moreno, J.L., Who Shall Survive? Washington, D.C.: Nervous and Mental Disease Publishing Company, 1934.

Pitts, F.R. "Recent trends in social network analysis: a bibliography." Paper presented at the seminar on Communication Network Analysis, East-West Communication Institute, Honolulu, 1979.

Richards, W.D., "Network analysis in large complex systems." Paper presented to International Communication Association, New Orleans, 1974.

Roberts, K.H. & C.A. O'Reilly, "Organizations as communication structures: an empirical approach." Human Communication Research 4(4) 283-293, 1978.

Rogers, E.M., "Network analysis of the diffusion of innovations: family planning in Korean villages." Stanford, Ca.: Stanford University, Department of Communication, 1976.

Rogers, E.M., "The rise and fall of the dominant paradigm." Journal of Communication, 28(1) 64-69, 1978.

Rogers, E.M. with F.F. Shoemaker, Communication of Innovations: a Cross-Cultural Approach. New York: Free Press, 1971.

Schwartz, M., Computer-communication Network Design and Analysis. Englewood Cliffs, N.J.: Prentice-Hall, 1977.

Woelfel, J., G.A. Barnett, & J.W. Dinkelacker, "Metric multidimensional scaling in Reimann space." Paper presented to Psychometric Society, Hamilton, Ontario, 1978.

Woelfel, J. & E.L. Fink, The Galileo System: A Theory of Social Measurement and its Application. New York: Academic, forthcoming.

Woelfel, J., E.L. Fink, K.B. Serota, G.A. Barnett, R. Holmes, M. Cody, J. Saltiel, J. Marlier, & J.R. Gillham, Galileo IV--A Program for Metric Multidimensional Scaling. Honolulu: East-West Communication Institute, 1976.