A GRAPH THEORY INTERPRETATON
OF NODAL REGIONS

By John D. Nystuen and Michael F. Dacey

The authors are, respectively, Assistant Professor of Geography in the University of Michigan and Assistant Professor of Regional Science in the Uuiversity of Pennsylvania. This study was formulated and the computations were obtained while both authors were at the University of Washington, Seattle. Professor Nystuen has been the primary contributor to the present statement. The responsibility for this statement is, of course, shared jointly.

THE PURPOSE OF THIS PAPER is to describe a procedure for ordering and grouping cities by the magnitude and direction of the flows of goods, people, and communications between them. Current theories of nodal regions and central place hierarchies provide the bases for the recognition of region-wide organization of cities into networks. These two theories were developed by students who recognized that the direction and magnitude of flows associated with social processes are indicators of spatial order in the regional structure of urban society. Whether the flow is local and to the city's hinterland, or regional and to the rank ordering of cities, the notion of central or nodal point is dependent upon the levels of strongest associations within the total flow.[1]

The present problem is to develop a method capable of quantifying the degree of association between city pairs in a manner that allows identification of the networks of strongest association. These associations may be in terms of interactions that occur directly between two cities, or indirectly through one or more intermediary cities. The magnitude of the combined direct and indirect associations is measured by an index that is related to certain concepts of graph theory. This index is used to identify the degree of contact between city pairs and it provides a quantitative basis for grouping cities. The resulting subgroups of cities are analogous to nodal regions. When each city in a study region is assigned to a subgroup, it is possible to specify the rank ordering of cities and to evaluate the functional relations of the nodal hierarchy.

In this paper, pertinent geograrphic and graph theoretic concepts are discussed and are then used as a basis for deriving the method of isolating nodal regions. While this method is illustrated by the use of intercity telephone calls in Washington state, the techniques are quite general and may be adapted to

---

[1] Berry, B.J.L. and W.L. Garrison, "A Note on Central Place Theory and the Range of a Good," *Economic Geography*, Vol. 34 (1958), pp. 304–311. Ullman, E.L., "A Theory for Location of Cities," *American Journal of Socialogy*, Vol. 46 (1941), pp. 853–864. Whittlesey D., "The Regional Concept and the Regional Method," *American Geography: Inventory and Prospect*, (P.C. James and C.F. Jones, eds.) Syracuse University, Syracuse, 1954.

many types of phenomena. A particular phenomenon is suitable for this type of analysis when it may be viewed as a relationship or flow that links objects that are properly mapped as points. In the present illustration, cities are conceptualized as punctiform elements in a telephone network. Other suitable areas of application include the flow of information or material products between business firms in a metropolitan area, the flow of mail or freight between cities in a region, the interpersonal relations between the inhabitants of a city or the political structure that connects federal, state and local governments.

Relationship to Existing Theory

Cities may be viewed as nuclei of specialized activities which are spatially concentrated and functionally associated. Each activity has its own set of associations outside the city. To account for the many different external connections of each specialization, general statements concerning urban associations must be multi-dimensional. Accordingly, urban hinterlands are normally defined by establishing a boundary from a composite of the spatial range of several central place functions, such as the trade area of the local newspaper, the extent of wholesale drug distribution, bus passenger volumes, governmental jurisdiction, and similar indices of central place functions.

Long-distance telephone communications may be considered a single index of this multi-dimensional association among cities. A grouping of cities on the basis of telephone data defines only a network of telephone traffic centers. The validity of interpreting these telephone traffic centers as an accurate indicator of multifunctional associations depends upon a correspondence of the hinterlands which are developed with those obtained from studies which evaluate many types of contacts. The authors are willing to accept that telephone flows are one of the best single indices of all functional contacts. It has an advantage over the use of a series of indices because it obviates weighing the individual contributions of the several indices.[2]

The Nodal Region or Hinterland

Nodal regions are defined by evaluating the external contacts of small areal units. Each of these areal units is assigned to that place with which it has the dominant association. Usually, this will be a nearby city, and this city is defined as the central place or nodal point for the unit areas oriented to it. The aggregation of these unit areas, in turn, is called the nodal region.

This does not deny the existence of other flows or associations to and from each areal unit. Such flows do exist so that each areal unit is connected to many other cities. Newspaper circulation, for example, may be dominated by the local daily while the nearest metropolitan paper may also be well represented, and *The New York Times* may find its way into a few homes in the area. Also, many sporadic contacts with former hometown papers may be present.

---

[2] Hammer, C. and F.C. Iklé, "Intercity Telephone and Airline Traffic Related to Distance and the 'Propensity to Interact'," *Sociometry*, Vol. 20 (1957), pp. 306–316. Harris, C.D., *Salt Lake City: A Regional Capital*, University of Chicago, Chicago, 1940. Ullman, E.L., *Mobile: Industrial Seaport and Trade Center*, University of Chicago, Chicago, 1940.

Nevertheless, the "dominant association" remains the critical concept in defining a nodal structure. The remaining non-dominant associations are not used, even though the magnitudes of some of these associations may be relatively large.

The Hierarchy of Cities

The nodal rgion describes the relationship between the hinterland, which is areal, and the central or nodal city, which is punctiform. Clearly, there is no loss of generality by considering only paired contacts between points. In the hinterland concept, the areal units may be abstracted to the level of points so that the association is in terms of many points being linked to a single central point.[3]

The hinterland of a major metropolitan center, such as Chicago, may encompass a large region and incorporate many of the region's functions. The strongest of the flows between Chicago and its hinterland are point to point associations of the cities within the region. At this scale, the relationship between nodal regions and the hierarchy of central places becomes clear. The major hinterland of Chicago is defined by its dominant association with many smaller metropolises. Each of these centers, in turn, is the focus of association from other, smaller centers within its immediate vicinity. These associations incorporate lower-order functions than those establishing direct associations to Chicago. In this fashion city regions are nested together, intimately dependent upon the range of the functions which define the associations at each level.

A hierarchy of cities of this type may be reduced to an abstract network of points and lines. The points represent the cities while the lines represent the functional associations. Though a myriad of lines exists in the network, there is present a basic structure of strongest associations which creates the nested nodal regions and the hierarchy of cities. Both the direct and indirect associations are important in these intercity structures. In terms of the direct associations, for example, a wholesale establishment may receive orders directly and ship directly to some points within the system. Alternately, the associations are indirect when the orders are accumulated at various levels of the hierarchy and proceed upward to the regional headquarters. In the same manner, the outbound shipments from the central city proceed down the ranks to intermediary levels through middlemen, rather than directly to every point in the region.

Many associations are of this indirect type. For instance, political control moves up and down the ranks, rather than through direct communication between the national party leaders and the ward leaders. Most commodities are assembled and distributed through a hierarchical structure within the organization. This results, in part, from the economies of moving large lots over long distances and, in part, from the better control it affords over the operation. In evaluating the entire fabric of urban society, it is evident that subtle, indirect influences and associations are frequently exerted by one location on another. A system of analysis which accounts for both the direct and the indirect as-

---

[3] For example see: Isard, W. and D.J. Ostroff., "General Interregional Equilibrium," *Journal of Regional Science*, Vol. 2 (1960) pp. 67-74.

sociations between cities is appropriate.

In summary, the nodal region is defined on the basis of the single strongest flow emanating from or moving to each of the unit areas in the vicinity of a central place. The region is delimited by the aggregation of these individual elements. The hierarchy of central places is determined by the aggregation of the smallest central places which are dependent upon a single, larger center for the functions they lack. This nesting of cities defines the organization of networks of cities and the position of each city within the network. Such nesting depends upon the available bundle of functions and the relative dominance of bundles.

In this study we start with the cities and towns of a large area. Then, the structure of association among the cities is specified by assigning each city to one of several subgroups. By considering the system as a set of points and lines, where the lines represent the association between points, certain theorems of linear graphs become available for the analysis of the functional association of cities within an area.

A GEOGRAPHICAL APPLICATION OF SOME GRAPH THEORY CONCEPTS

Graph theory is a mathematics of relations. By specifying certain properties of the relations between cities and accepting the point-line abstraction of graph theory, certain theorems become available for analyzing intercity flows.[4] Consider the cities in a region as a set of points. Consider, also, a line joining a pair of points whenever there exists a certain flow between the cities they represent. The finite collection of points and lines, where each line contains exactly two points, is a linear graph of the relations established by the flows.

Some Characteristics of Linear Graphs

A point is called *adjacent* to another point if it is connected to it by a line. The network of lines is the only information contained within the graph. Scalar distance and direction, the most striking aspects of geographical maps, are not defined for a graph. If the relationship is of equal value for every connected pair, the graph is a *binary graph*. Most graph theory relates to this type of construction which simply indicates whether a line (a relation) exists or does not exist between any pair of points. The connections, however, may be considered to have intensity. *Intensity* is displayed on the graph by assigning a value to the lines.

Orientation of a relation between two points is displayed on the graph by an arrowhead $a \rightarrow b$, and read "*a* is related to *b*." A graph which specified

---

4 Some general statements of graph theory are: König, D., *Theorie der Endlichen und Unenlichen Graphen.* Leipzig, 1936 (reprinted by Chelsea Publishing Co., New York, 1950); Berge, C., *Theorie des Graphes et Ses Applications.* Dunod, Paris, 1958; Harary, F., "Unsolved Problems in the Enumeration of Graphs," *Publications of the Mathematical Institute of the Hungarian Academy of Sciences,* Vol. 5 series A (1960) pp. 63-95;—, "Some Historical and Intuitive Aspects of Graph Theory," *Siam Review* Vol. 2 (April 1960) pp. 123-131. The utility of graph theory for geographic analysis has been demonstrated by Garrison, W.L., "Connectivity of the Interstate Highway System," *Papers and Proceeding of The Regional Science Association,* Vol. 6 (1960) pp. 121-137.

orientation is called a *directed graph* or *digraph*. The relationship between two points on a directed graph need not be symmetrical and, when intensity of the connection is defined, the intensity may be different for each direction.

A path from the points a to e is a collection of points and lines of the form, $a, a \to b, b, b \to c, \cdots, d, d \to e, e$, where the points $a, b, \cdots, e$, are distinct. A *sequence* is a collection of points and lines from a to e in which the intermediate points need not be distinct. A graph is *weakly connected* if there exists a path between each pair of points, disregarding orientation. The points in a *component* of a graph are weakly connected and are not connected to any other points in the graph. The *degree* of a point is the number of points to which it is adjacent. In a directed graph a point has an out-degree and an in-degree depending on the orientation of the lines incident to it.

Matrix Notation

For every linear graph there is an *adjacency matrix* which completely describes the graph, and *vice versa*. The matrix notation is convenient for arithmetic manipulation. Every point in a graph is represented by a row and a column of the matrix. The element, x_{ij}, of the adjacency matrix takes the value of the line; if it exists, between the points i and j; if the line does not exist, the value of the x_{ij} is O.

The diagonal elements, x_{ii}, of the adjacency matrix represent the relation of each point to itself. This relationship may or may not be defined. When it is not defined, all elements of the main diagonal are, by convention, put equal to zero.

Properties of the Dominant Relations Between Cities

The geographic theory reviewed above suggests that within the myraid relations existing between cities, the network of largest flows will be the ones outlining the skeleton of the urban organization in the entire region. The term "largest" implies an oriented relation because a flow between a pair of cities may be the largest in terms of one city but not necessarily in terms of the other city. The relation "largest flow" may have various definitions, such as the largest out-flow, in-flow, or total flow. The present example uses the number of out-going intercity telephone messages from each city to every other city in the study area. It is possible to construct a directed graph of these relations. Using the principle of dominant association, a single out-directed line is assumed to be associated with each point. When number of telephone messages are used to measure intensity of intercity associations, this assumption is easily accepted because for any city the largest volume to any one city is typically several times greater than the next largest message flow. An assumption of this type is tenable only for intercity relations which may be ranked or have a unique, largest interaction. In other situations, nodal region is most likely an inappropriate concept.

The collection of largest flow lines between city pairs defines a network of orientation among the points. Where each point has a largest flow, that largest flow may be found by simple inspection of a matrix of flows between

all pairs, and it is the maximum element in each row when the matrix displays number of messages from the row city to the column city. The present intention is to use this notion of largest flow to aggregate cities associated with a central place. The resulting aggregation is said to be composed of the "subordinates" of the central city. The problem is the recognition of a "central city." In order to establish a "dominate center" three additional properties of the "largest flow" relation are now identified.

One property states that a city is "independent" if its "largest flow" is to a "smaller city." A small city remote from large metropolitan centers may display this type of independence because its largest flow is to an even smaller, nearby city. Conversely, in the same region a large satellite city closely associated with a metropolitan center does not have this independence because its largest flow is to the metropolis. So, to identify independent cities a measure of size is required. Size may be externally assigned, e.g., by population of each city; or it may be internally assigned, e.g., by the total volume of messages to or from all cities in the region. In the example below size is assigned in accordance with the total in-message flow from all cities in the study region. This value is the column total of the matrix of flows between all pairs of cities. In these terms, an "independent" or "central city" is defined as one whose largest flow is to a smaller city: A subordinate city is a city whose largest flow is to a larger city. This assumes no ambiguities arise to obscure the dominate (largest) city of a pair. This occurs when largest flows are reflexive, that is, two cities whose largest out-connections are to each other.

A second property is transitivity. This property implies that if a city a is subordinate to city b and b is subordinate to c, then a is subordinate to c.

A third property stipulates that a city is not a subordinate of any of its subordinates. A graph showing this relation is called *acyclic*. It is easily seen that an acyclic graph contains a *hierarchy*.

Two Theorems

The largest flow from every subordinate city is called the *nodal flow*. These flows form the *nodal structure* of the region and (for the particular relation under study) this skeleton displays the functional association of the cities in a region. This structure is analogous the nodal region and contains a hierarchy of centers. It is important to recognize that in this nodal structure the out-contact from at least one point is zero. This particular case is called a *terminal point* and in terms of an urban structure, this type of point is interpreted as a central city.

The following statements are useful deductions concerning the graph of a nodal structure. Figure 1 illustrates the resulting concepts.

(1) The components of a nodal structure partition the set of cities.

Proof. Each city is represented by a point in a graph of a nodal structure. Each point is weakly connected to every point in its component and to no point not in its component. A trivial component is an isolated point. Each point is, therefore, assigned to one and only one component. Such an assignment of a set is a partitioning.

(2) Each component of a nodal structure has a unique central city (terminal

MATRIX OF NUMBER OF MESSAGES BETWEEN CITY-PAIRS

TO CITY

		a	b	c	d	e	f	g	h	i	j	k	l
	a	0	⑦⑤	15	20	28	2	3	2	1	20	1	0
	b*	69	0	45	50	58	12	20	3	6	35	4	2
	c	5	㊝	0	12	40	0	6	1	3	15	0	1
	d	19	㊅⑦	14	0	30	7	6	2	11	18	5	1
FROM	e*	7	40	48	26	0	7	10	2	37	39	12	6
	f	1	6	1	1	10	0	㉗	1	3	4	2	0
CITY	g*	2	16	3	3	13	31	0	3	18	8	3	1
	h	0	4	0	1	3	3	6	0	12	㊳	4	0
	i	2	28	3	6	43	4	16	12	0	㊈⑧	13	1
	j*	7	40	10	8	40	5	17	34	98	0	35	12
	k	1	8	2	1	18	0	6	5	12	㉚	0	15
	l	0	2	0	0	7	0	1	0	1	6	⑰	0
Column Total		113	337	141	128	290	71	118	65	202	311	91	39

Largest flow circled.
Largest flow
determined by the
number of out-going
messages.

*Largest flow from these cities is to a "smaller" city where
"size" is determined by the column totals.

GRAPH OF THE NODAL STRUCTURE BETWEEN CITIES

Graph of a,b,..., l cities in
Region G.
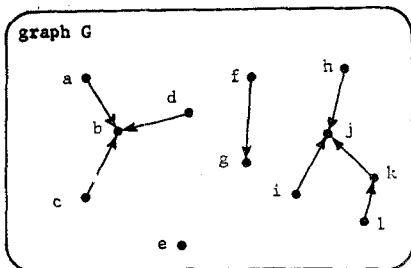
Adjacency Matrix of Graph G

	a	b	c	d	e	f	g	h	i	j	k	l	
a		1											
b													*
c		1											
d		1											
e													**
f							1						
g													*
h										1			
i										1			
j													*
k										1			
l												1	

Blank spaces represent zero elements.

*Terminal point. **Trivial terminal
point.

FIGURE 1. Graph of a Nodal Structure in a Region. (Hypothetical)

point).

Proof. Every path has at least one subordinate point and also a point to
which all points on the path are subordinate by the transitive property of the
relation. If this point is subordinate it must be adjacent to a point not on the
path because the relation is acyclic. Upon extending the path, an end point
with zero out-degree will be found in the component because at least one point
is in the component and is not subordinate.[5] This end point is the only terminal

---

[5] Here, only finite graphs are considered.

point to which all points on the related paths are subordinate because no branching occurs on any path (every subordinate point has an out-degree of one).

Now assume distinct paths in a component and extend the paths to their terminal points. Any other point connected to a point in one extended path has no other connection because each point has an out-degree no greater than one. If their terminal points were distinct, the subgraphs associated with distinct paths by the method described would not be connected in any way. This contradicts the fact that all elements in a component are weakly connected. Therefore, each extended path must have the same terminal point.

Interpretation of the Nodal Structure

The nodal structure may be used to distinguish groups of cities that have maximum direct linkages and the rank order of these cities may be calculated. The hinterland of the central city may also be determined by mapping the cities in a nodal structure and then drawing a line just beyond the cities which are most distant from the central city. In accordance with existing theory, the hinterland or nodal region contains the area in which the maximum association or flow is toward the central or nodal city. In addition, this plotting shows the hierarchy over which the central city is dominant.

An Extention of the Theory to Indirect Associations

The operations and structure that has been described evaluates only direct contacts between city-pairs. It does not incorporate the indirect associations, and these, conceivably, could be very influential in determining functional associations. Admittedly, direct contacts should receive the greater weight, but some evaluation of the indirect channels between city-pairs would seem appropriate because of the indirect associations which occur within a hierarchy. Indirect associations may be evaluated by using matrix manipulations to adjust the nodal structure. It is postulated that the increment of indirect association or influence decreases with increases in the length of the channel.

Power Series of the Adjacency Matrix

The first step in accounting for indirect influence is to adjust the raw data matrix so that the direct association between each city-pair is some proportion of the total association of the largest center in the area. This is accomplished by obtaining the maximum column total of the adjacency matrix ($\max_j \sum x_{ij}$) and dividing every x_{ij} element by this summation. Put: $y_{ij} = x_{ij}/\max_j \sum x_{ij}$. The following inequalities result for a graph of n points:

(1) $\qquad\qquad 0 \leqq y_{ij} < 1 \qquad\qquad\qquad (i, j = 1, 2, \cdots, n)$

(2) $\qquad\qquad 0 < \sum_j y_{ij} \leqq 1 \qquad\qquad\qquad (j = 1, 2, \cdots, n).$

The maximum column total equals *1*.

The linear graph corresponding to this adjacency matrix has the appropriate, positive, decimal loading. Let the adjacency matrix be called Y. In terms of

linear graphs, the power expansions of Y have interesting interpretations. The matrix Y^2, which is obtained by $Y \cdot Y$ under usual matrix multiplication, describes a graph when all sequences have a length of 2. The length of a sequence is the number of lines it contains. Further, the loading of the lines of each sequence of length 2 are obtained by multiplication. Since the initial loadings are decimal values, an attenuated value is associated with a contact that proceeds from point i to j through a sequence of length 2. The sum of all such two-step sequences from i to j is the value of all possible indirect contacts of length 2.

This assertion may be demonstrated as true by considering the meaning of the summation:

$$(3) \qquad a_{ij} = \sum_k y_{ik} y_{kj} \qquad\qquad (k = 1, 2, \cdots, n)$$

and where a_{ij} is an element in Y^2.

The y_{ik} is the loading on the line from point i to point k in the graph and the y_{kj} has the same meaning for the link from k to j. The only terms which enter the summation are those where a sequence of length 2 exists. When a link from or to the kth point does not exist, the whole term is zero. The a_{ij} is the total value of all sequences of length 2. In a similar manner it may be shown that the elements of Y^3 specify the attenuated value of all sequences of length 3, and so on. The meaning of the following summation is clear:

$$(4) \qquad B = Y + Y^2 + Y^3 + \cdots + Y^n + \cdots$$

The element, b_{ij}, of B represents the total direct and indirect influence from i to j.

Some examples may be useful. Given the cities $a, b, \cdots, n$, a typical sequence from a to e might be $a \to b \to c \to d \to e$. Imagine an activity in city a as having influence on a respondent in b, this b in turn contacts a respondent in c, and continuing until a small response in e is affected. The probability of such a chain of occurrence depends, in part, on the magnitude of the flows in every link of the sequence. In general, the longer the sequence, the more remote is the probability of a response and when a response occurs it is less intense.

Alternatively, the flow of influence may be re-channeled through the same city more than once. For example, a sequence may have the form $a \to b \to a \to e$. All such summations are included in the matrix B.

The summation of the power expansion of Y is not demonstrated to be the correct form of the attenuation of flows in a sequence. It is extremely doubtful that the matrix B is the most appropriate measurement of the total direct and indirect influences. It is essentially a measure of chance indirect contact. The distribution of actual indirect association is very likely not at all random but rather concentrated in certain flow channels, in which case the matrix B would be an underestimate of indirect influence. It does, however, have a greater appeal than the matrix Y which incorporates only the direct influences. The choices of the particular power expansion is dictated by the ease of its compu-

tation. Several other methods may also be appropriate.[6]

Computation of the Power Series of the Adjacency Matrix

A convenient method of computing the matrix B is to use the following identity:

$$(5) \qquad (1 - Y)^{-1} = 1 + Y + Y^2 + \cdots + Y^n + \cdots$$

and then:

$$(6) \qquad B = (1 - Y)^{-1} - 1 ,$$

where the 1 is the identity matrix. The inverse, $(1 - Y)^{-1}$, is known to exist if the inequalities (1) and (2) hold.

The Nodal Structure of Matrix B

The nodal structure of matrix B is established by isolating the network of largest flows in the same manner as was described for the direct associations. Because the associations enumerated in matrix B are adjusted for both direct and indirect flows, it is expected that a more reasonable structure is obtained.

AN EXAMPLE

Washington State was chosen as the study area. The utility of the nodal structure concept is evaluated by choosing a set of cities in this area and then determining the nodal structure that prevails. The nodal structure which emerges should resemble the known hinterland and ranking of the major cities in the area. Certain cities outside of the State were included in the study in order to examine the role they play in the network of city associations. Portland, Oregon and Vancouver, British Columbia were especially important additions.

The associations were defined by the number of long distance telephone messages between city-pairs during one week in June, 1958.[7] Certain cities were omitted from the study due to characteristics of the data and in order to limit the size of the study.

Many pairs of neighboring cities have direct dialing service and in these instances the intercity calls were not recorded in long distance data. Dormitory towns for Seattle and several "twin cities" such as Aberdeen-Hoquiam, Chehalis-Centralia and Pasco-Kennewick had direct-service exchange. This is not a serious deficiency in the data because such cities very likely function as a single point in the state-wide network, and one of the "twin cities" in each pair could be used in the study. Certain fairly large cities north of Seattle and along the Puget Sound were omitted for lack of data. These cities were serviced by a different telephone company. Because each year the telephone companies simul-

---

[6] For examples see, Luce, R.D. and D. Perry, "A Method of Matrix Analysis of Group Structure," *Psychometrika*, Vol. 14 (1949) pp. 95–116; and Katz, L., "A New Status Index Derived from Sociometric Analysis," *Psychometrika*, Vol. 18 (1953) pp. 39–44.

[7] We are indebted to the Pacific Bell Telephone and Telegraph Company and especially to Mr. Homer Moyer, a Seattle officer of that company, for this information.
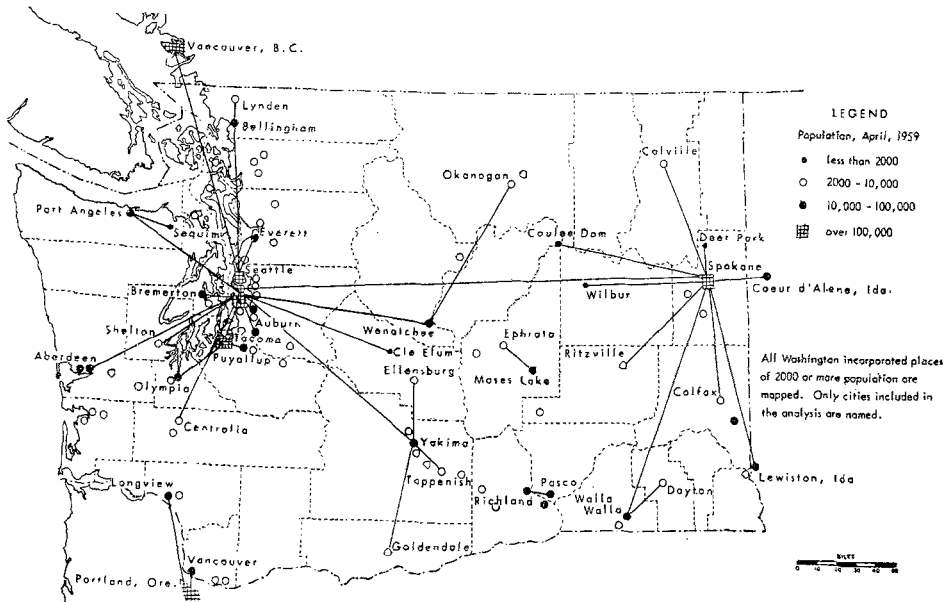
FIGURE 2. Nodal Structure Based on Telephone Data, State of Washington, 1958.

taneously take a one-week sample of intercity telephone calls, comparable data exist but there was no attempt to obtain them.   Finally, all cities above a certain population size were not included in order to restrict the size of the study.   Some small towns were chosen, however, in an effort to obtain samples of hierarchies with directed paths of length 2 or more.   The map in Figure 2 identifies the cities in the study.   With the advantage of hind-sight, it might have been preferable to have included more small towns in the study.

TABLE I.   A Portion of the 40 × 40 Table of Number of Messages Between City-Pairs, for One Week of June 1958.

From City	To City									
	Code	01	02	03	04	08	13	26	27	40
Aberdeen	01	—	24	50	0	246	3671	54	4	1005
Auburn	02	26	—	35	0	8	7654	42	0	163
Bellingham	03	55	27	—	782	24	2494	101	3	356
Lynden	04	4	0	2250	—	4	357	9	0	110
Longview	08	329	15	32	0	—	1911	87	4	4773
Seattle	13	3427	4579	3843	308	1268	—	6168	269	16781
Spokane	26	61	32	119	6	85	9991	—	3842	3838
Couer d'Alene	27	0	4	4	0	6	254	5104	—	141
Portland, Ore.	40	802	210	304	22	4190	22179	3310	98	—

Largest column total—Seattle 154,192.

Certain channels of communication are omitted because they are not recorded in the long distance data. Direct-line calls are an example. Probably the only large volume, direct-line in the state links Seattle with Olympia, the state capital. If these data were included, the maximum association of Olympia might shift from Tacoma to Seattle. Though beyond the scope of this study, interesting results could be obtained if all channels of communication were included, such as radio, telegraph, mail and messenger service.

Table I is an example of the raw data tabulations. Forty cities were used in the study.[8] The entire table is the adjacency matrix of the almost completely connected graph of associations—there are a few zero entries. The row totals are the total out-contacts while the column totals are the total in-contacts. The direction of the message flow is read from the "row" city to the "column" city. The main diagonal entries are zero by convention.

TABLE II. A Portion of the Matrix B (Direct and Indirect Associations).

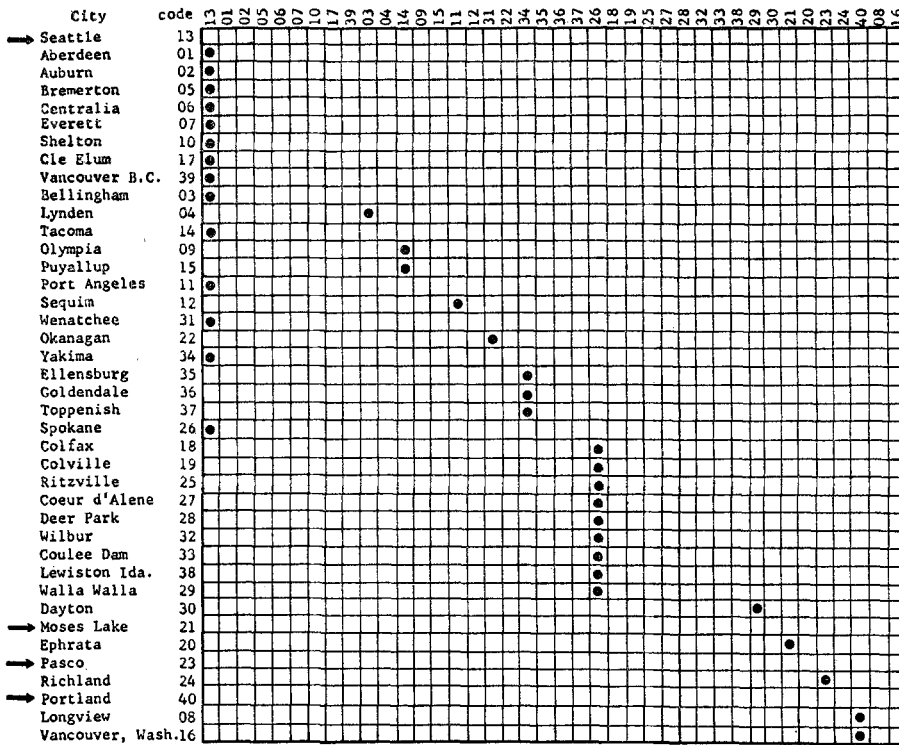| From City		To City								
	Code	01	02	03	04	08	13	26	27	40
Aberdeen	01	—	.248(4)	.395(4)	.551(6)	.166(3)	.245(2)*	.479(4)	.325(5)	.726(3)
Auburn	02	.303(4)	—	.364(4)	.109(5)	.108(4)	.508(2)*	.500(4)	.107(5)	.171(3)
Bellingham	03	.402(4)	.232(4)	—	.513(3)	.180(4)	.165(2)*	.739(4)	.247(5)	.254(3)
Lynden	04	.328(5)	.790(6)	.148(2)*	—	.307(5)	.239(3)	.716(5)	.689(7)	.753(4)
Longview	08	.221(3)	.150(4)	.252(4)	.341(6)	—	.131(2)	.699(4)	.325(5)	.316(2)*
Seattle**	13	.227(2)	.303(2)	.253(2)	.204(3)	.870(3)	—	.409(2)	.188(3)	.111(1)
Spokane	26	.568(4)	.421(4)	.953(4)	.536(5)	.688(4)	.649(2)*	—	.252(2)	.260(2)
Couer d'Alene	27	.650(6)	.332(5)	.340(5)	.560(7)	.459(5)	.191(3)	.335(2)*	—	.103(3)
Portland, Ore.**	40	.563(3)	.185(3)	.237(3)	.176(4)	.278(2)	.140(1)	.224(2)	.725(4)	—
Column Total		.548(2)	.588(2)	.613(2)	.866(3)	.585(2)	.102(0)	.229(1)	.311(2)	.563(1)

**Terminal point. *Nodal flow.

Remark: Figures are rounded to three significant digits. Data were processed to 8 significant figures. The value in parentheses represents the number of zeros before the first significant digit.

Table II is the adjacency matrix B which evaluates both the direct and indirect associations between the cities.[9] The nodal structure contained in this matrix was determined by (1) identifying the nodal flow, (2) ranking the cities by their total incoming associations (column totals), (3) assigning an orientation from cities with smaller total associations to one with a larger total association and (4) identifying the non-oriented cities as the center of its hierarchy. Figure 2 shows the results. Figure 3 is the adjacency matrix of the nodal structure derived from the direct and indirect associations.

City	code
→ Seattle	13
Aberdeen	01
Auburn	02
Bremerton	05
Centralia	06
Everett	07
Shelton	10
Cle Elum	17
Vancouver B.C.	39
Bellingham	03
Lynden	04
Tacoma	14
Olympia	09
Puyallup	15
Port Angeles	11
Sequim	12
Wenatchee	31
Okanagan	22
Yakima	34
Ellensburg	35
Goldendale	36
Toppenish	37
Spokane	26
Colfax	18
Colville	19
Ritzville	25
Coeur d'Alene	27
Deer Park	28
Wilbur	32
Coulee Dam	33
Lewiston Ida.	38
Walla Walla	29
Dayton	30
→ Moses Lake	21
Ephrata	20
→ Pasco	23
Richland	24
→ Portland	40
Longview	08
Vancouver, Wash.	16

Arrows indicate cities which are terminal points.

FIGURE 3. Matrix of the Nodal Structure.

CONCLUSIONS

The techniques defined in this paper will divide a set of cities into sub-groups which specify a central place and its subordinate hierarchy. The association between cities is not the only system which may be defined as a network of points and lines. Nations or states may be thought of as points with migrations or commodity flows as lines. The important step in the employment of abstract linear graph analysis is the assignment of plausible meaning to the points and lines, preferably in terms of some real world phenomena. The usefulness of the attributes and the interpretation of the resulting hierarchy depends on the correspondence between an empirical example using graph theory analysis and other knowledge of the phenomena. The procedure described in this paper may be employed in a variety of ways, but the application is valid only when significant theoretical conclusions are produced and verified empirically.

Implications of the Nodal Structure in Washington State

The nodal regions that are suggested by the nodal structure agree, in general, with expectation. Seattle is the dominant center with nested hierarchies

defined around Spokane and Yakima. Portland forms a system of its own by capturing nearby Washington State cities. The two small but independent hierarchies defined on Pasco and Moses Lake are most interesting.

Ephrata and Moses Lake are located on the boundary between two large hinterlands where it is postulated that self-reliance or independence is most likely to appear.[10] In addition, these two cities constitute an anomaly because, while Ephrata is an old city, Moses Lake was recently created by government fiat.

The small hierarchy with Pasco as the central point was anticipated by Ullman when he evaluated the growth centers of the western United States:[11]

> "One hypothesis that occurs to me for the future is that Pasco-Kenneick-Richland... might develop as the subregional shopping center, supplanting the dominance of older (and more attractive) Yakima and Walla Walla."

The effect of the national border is clear. Vancouver is subordinate to Seattle and it does not dominate any city in the study, even though it is a large city and is much nearer to Lynden and Bellingham than is Seattle. It is probable that Vancouver would have been a terminal point if other Canadian cities had been included within the study. This is not a defect in the method. The results are only an evaluation of the associations between the cities in the study.

The nodal region of Tacoma, a large city south of Seattle, is also anticipated by theory. Tacoma is dominant in a nearby region. This region is off-center, in the direction away from the larger city of Seattle. The dominance of Seattle re-asserts itself at even greater distances so that Aberdeen and Centralia are directly associated with Seattle, rather than by a two-link path through the closer and larger city of Tacoma[12] This and the other agreements with existing theory and accepted empirical evidence demonstrate the utility of the nodal structure for analyzing city associations.

Further Graph Theory Applications

Given a set of cities in an area and a measure of association between them, a set of hierarchies has been obtained. Even more information is desirable. Spokane is obviously the second most important central place in Washington state, yet it is subordinate to Seattle in a hierarchy while the much smaller places of Moses Lake and Pasco dominate their respective systems. Intuitively, a second in command position in a large organization is more important than the primary position in a tiny organization. Some measure of this difference in status is desirable. A further application of graph theory to this problem is suggested in a paper by Harrary.[13] His ideas are adapted to this problem by the present authors in a further study of city associations.

---

[10] Hoover, E.M., *The Location of Economic Activity*, McGraw-Hill, New York, 1948. Isard, W., *Location and Space Economy*, John Wiley and Sons, New York, 1956. Lösch, A., *The Economics of Location*, Yale University, New Haven, 1954.

[11] Ullman, E.L., *Growth Centers of the West*, University of Washington, Seattle, 1955, p. 48.

[12] Hoover, E.M., *op. cit.* Isard, W., *op. cit.* Lösch, A., *op. cit.*

[13] Harary, F., "Status and Contrastatus," *Sociometry*, Vol. 22 (1959) pp. 23-43.