AN INTRODUCTION TO GRAPH THEORETICAL METHODS IN GEOGRAPHY

K.J.Tinkler

ISSN 0306 - 6142

ISBN 0 90224668 2

CATMOG(Concepts and Techniques in Modern Geography)

1 An introduction to Markov chain analysis - L. Collins

2 Distance decay in spatial interactions - P.J. Taylor

3 Understanding canonical correlation analysis - D. Clark

4 Some theoretical and applied aspects of spatial interaction shopping models - S. Openshaw

5 An introduction to trend surface analysis - D. Unwin

6 Classification in geography - R.J. Johnston

7 An introduction to factor analytical techniques - J.B. Goddard & A. Kirby

8 Principal components analysis - S. Daultrey

9 Causal inferences from dichotomous variables - N. Davidson

10 Introduction to the use of logit models in geography - N. Wrigley

11 Linear programming: elementary geographical applications of the transportation problem - A. Hay

12 An introduction to quadrat analysis - R.W. Thomas

i3 An introduction to time-geography - N.J. Thrift

14 An introduction to graph theoretical methods in geography - K.J. Tinkler

15 Linear regression in geography - R. Ferguson

16 Probability surface mapping. An introduction with examples and Fortran programs - N. Wrigley

17 Sampling methods for geographical research - C. Dixon & B. Leach

18 Questionnaires and interviews in geographical research - C. Dixon & B. Leach

19 Analysis of frequency distributions - V. Gardiner & G. Gardiner

20 Analysis of covariance and comparison of regression lines - J. Silk

21 An introduction to the use of simultaneous-equation regression analysis in geography - D. Todd

22 Transfer function modelling: relationship between time series variables - Pong-wai Lai *(in preparation)*

23 Stochastic processes in one-dimensional series: an introduction - K.S. Richards *(in preparation)*

24 Linear programming: the Simplex method with geographical applications - James E. Killen *(in preparation)*

Price each: 75p or S1,50 *except Catmogs 16 s 19 which are £1.25 (S2.50)*

The series is produced by the Study *Group* in Quantitative *Methods, of the Institute* of *British Geographers. For details of membership* of *the Study Group,* write to *the Institute of* British *Geographers, 1 Kensington Gore, London SW7. The series is published by GEO ABSTRACTS, University of East Anglia, Norwich NR4 7JT, to whom* ail *other enquiries should be addressed.*

CONCEPTS AND TECHNIQUES IN MODERN GEOGRAPHY

AN INTRODUCTION TO GRAPH THEORETICAL METHODS IN GEOGRAPHY

by

Keith J. Tinkler
(Brock University, Ontario, Canada).

CONTENTS

Acknowledgements

I am indebted to Neil Wilson, formerly Teaching Assistant in Geography at Brock University, and Thomas A. Jenkyns of the Mathematics Department for their careful review of an early version of the manuscript. An anonymous referee made valuable comments on a later draft.

Phyllis Riesberry has withstood a siege of drafts and corrections and I am grateful to her for her patience. I thank David Bell for drawing the diagrams.

Acknowledgements to Publishers and Authors

We thank - R.J. Wilson for permission to reproduce Table 1 from Table 1 in An Introduction to Graph Theory, 1972, Longman Group Ltd. L.J. WOOD and the editor of Geografiska Annaler, Series B, for permission to reproduce Figures 3 and 4 and Tables 1 and 2 from "Functional Structure of a rural market system", 1975, Geografiska Annaler, p. 112 et seq.. R. SYMANSKI and M.J. WEBBER and the Executive Director of the Association of American Geographers, for permission to reproduce Figures 1 and 2, p. 205 from "Complex Periodic Market Cycles" in Annals, Association of American Geographers, 1974.

I INTRODUCTION

(i) What is a graph?

A graph consists of a set of points and a set of relationships between pairs of points. A pair of such points may be termed a line, link, connection or relation from the first point to the second one. In geographical applications the points may be places, towns, road junctions, telephone exchanges, regions, etc. and the relationships between them can be defined as is appropriate to the problem. Relationships fall in three main categories: (1) equal and reciprocal : town i is joined to town j (and vice-versa) by a road, (2) unequal and reciprocal : telephone exchange i calls exchange j with 20 calls a day whereas exchange j calls i with only 10 calls, (3) non-reciprocal : in one-way systems cars can go from junction i to junction j but not reciprocally (except perhaps by a circuitous route). Category (1) gives rise to graphs, and categories (2) and (3) give rise to directed graphs or digra hs. Because geographical problems frequently deal with a finite number of p aces, and with the inter-relationships between them, it follows that graph theory has much to offer in elucidating the form and structure of geographic space.

Graph theory is primarily a theoretical framework for looking at a problem and in this respect differs from most volumes in this series which are primarily concerned with a single well-defined technique. The literature using graph theoretical methods naturally reflects this characteristic and consequently there is no single article that exemplifies all the techniques discussed here, and there is no reason why there should be. Most of the basic ideas of graph theory can be exemplified by small diagrams, but most of the useful applications involve applying these ideas to very large and complex systems for which a visual display and solution are impossible for all practical purposes. The reader should not assume that because simple examples are used here for obvious reasons of space that therefore graph theory has nothing to offer. For most of these examples we have a totally unfair advantage : the ability, visually, to comprehend the totality of graph. However, given a graph in its abstract form : as a set of elements together with relationships between the elements we do not have this advantage; and this is the form in which a large empirical example is usually found. For example, given the graph: Fig 1 we can immediately comprehend that it is in one piece so to speak but given the following list of lines:

2 - 3
4 - 5
1 - 3
4 - 3
3 - 5

it takes much longer to discover this fact. Given a long list it is often difficult to draw the graph corresponding to Figure 1 and even more troublesome to analyse it by eye. Consequently graphs are usually represented as matrices using the data in the list and the algebra of matrices is used to define and discover the various properties of graphs. Throughout this

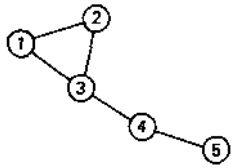


Fig. 1

treatment it will be understood that a graph, G, and its matrix form A (the adjacency matrix A of the graph G, defined as A = A(G)) are interchangeable. It is simplest to regard both the visual graph and the numerical matrix as different representations of an abstract entity, an abstract graph, comprising a set of points and the set of relationships between the points.

(ii) Prerequisites and Literature

The only prerequisites for this volume are a willingness to work through all the examples, to read a representative selection of the literature for each section and to get to grips with the matrix algebra required to operate graph theory. Additional familiarity with matrix algebra will be useful but is not strictly necessary.

Of the basically mathematical introductions to graph theory Wilson (1972) is to be recommended but Harary (1969), Ore (1963), Berge (1962), Flament (1963), Busacker and Saaty (1965) are all alternatives. An extensive treatment of digraphs is given to Harary, Norman and Cartwright (1965). Within Geography, Kansky (1963) is frequently cited, although only Chapter II deals with strictly structural notions and there is no treatment of matrix methods.

A similar introduction but including the powers of an adjacency matrix is found in Garrison and Marble (1965) which has only recently become widely available through reprinting in Eliot Hurst (1974). More recently, Campbell (1971) has provided a brief introduction to digraph analysis as a basis for examining inter-industry relationships. Other than these, most articles have taken from graph theory just those notions needed for a particular problem. In this respect Garrison (1960), Nystuen and Dacey (1961) and Pitts (1965) are early classics in the field. A comprehensive review of the uses of graph theory in geography can be found in Tinkler (1979).

(iii) Some definitions : Types of graphs

In Wilson's terminology a graph with n points and q links (pairs of connected points) is termed a simple graph whenever every distinct pair of points is joined either exactly once, or not at all, and where no point is linked directly to itself, (Wilson 1972).

A graph is defined when either or both of the above conditions do not hold with respect to single connections and self-connections. Graphs turn up frequently in geographical applications because it is often necessary to represent two places as being joined not just by one link but by several. For example, two towns may be connected by road, rail and air links and so a proper representation would show three links between the places; or two towns may make several telephone calls to one another. In later sections, in considering the graph of a transportation structure as a framework for possible movements as well as a mere spatial form, a place may be defined as being connected to itself. Consequently a simple graph is just a special, and useful, kind of graph. The term graph will be used to refer to either type, while the adjective simple will be used when simple graph is specifically needed.

Mathematicians conceive of graphs as relationships (lines, edges, arcs, links) between pairs of points (nodes or vertices). In this form a graph is abstract. In many geographical applications, however, the points are places and the links are transportation or communication links between them. When we 'draw' the graph on paper we are making a 'map' of some reality. A mathematician calls this an embedding of a graph in another space, in this case a Euclidean plane surface, a small region on the earth's surface approximately, or a sheet of paper). This particular embedding is very common and gives rise to many fundamental problems in graph theory exemplified by the next class of graphs.

A plane graph is a graph which is embedded in the plane in such a way that no lines in the graph intersect. The ends of lines may only join at points of the graph, (Fig. 2).

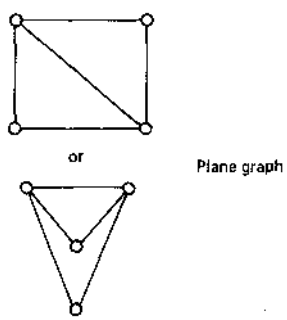


Fig. 2

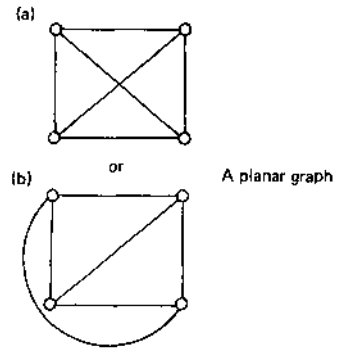


Fig. 3

A planar graph is a graph whose lines may be represented in the plane without intersections, i.e. as a plane graph. Hence Fig. 3(a) is planar but not plane. Any planar graph may be drawn as a plane graph and any plane graph is planar. There is no need to labour the point but it is important to realise that Fig. 3(a) is not non-planar. Some authors (e.g. Golledge 1968) fail to realise that some simple graphs drawn with an intersection could in fact be drawn in the plane without such an intersection, e.g. Fig 3(a) becomes Fig. 3(b), and thus despite first appearances are planar.

A non-planar graph is any graph which cannot be drawn as a plane graph without crossings, which is why Fig. 3 is planar but Fig. 4 is non-planar.

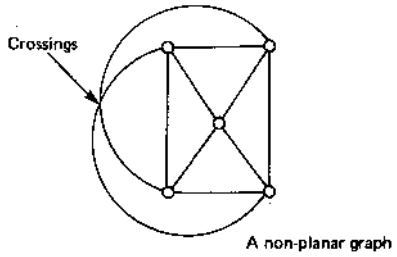
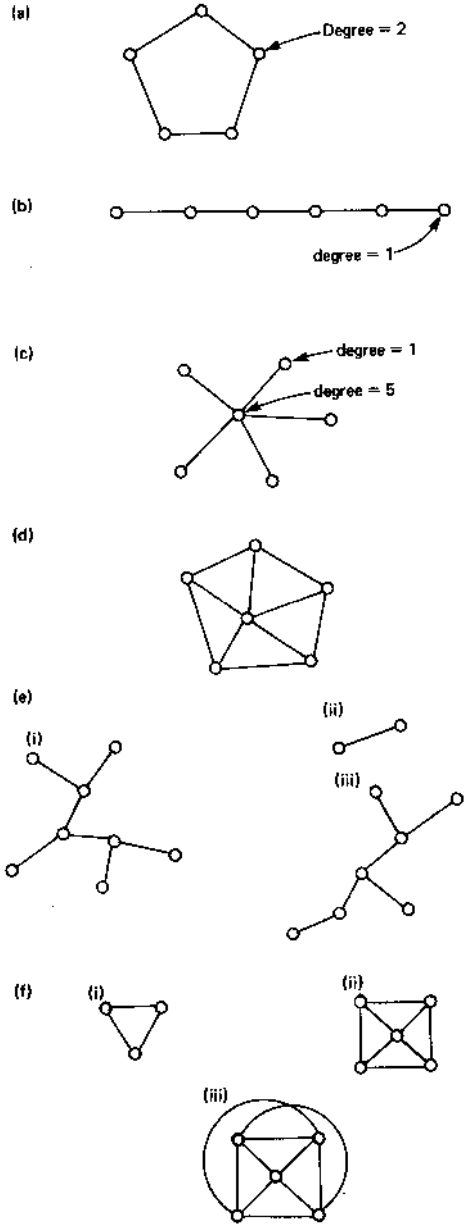


Fig. 4

A simple graph with five points and ten links can never be drawn in the plane without at least one crossing. If a G is planar and simple it is useful to know that it has a plane representation where all the edges are represented by straight line segments. In practice for all matrix operations concerning graphs the distinction between planar and non-planar graphs is immaterial, and for a given large graph it is not always an easy problem to determine whether or not it is planar.

Many simple structures are found so frequently that they have been recognised and named. While they are not always very interesting on their own they often occur as parts of larger structures or graphs and it is convenient to have a name to reference them. In general it will be understood that the structure defined can be drawn for as many points as we like. The degree of a point is the number of links that connect to it, as exemplified in Fig. 5(a), (b) and (c).

The diagrams show examples of rings, stars, wheels, trees, complete graphs, nullgraphs and regular graphs. The last three categories f, g and h may require some explanation. A complete graph (f) is a graph in which every pair of points is connected by a link, and is of course defined for any number of points. A complete graph which is a sub-graph (see below) is termed a clique. A null graph (g) on n points has n isolated nodes and no links. While trivial in itself, the idea is needed in describing the structures which may arise when certain operations are performed on graphs, and in ensuring that two graphs are conformable for addition and multiplication. A regular graph (h) is one in which every point has the same number of links, say δ, incident at it. The graph is then said to be regular of degree δ. In practice few graphs encountered in actual problems are like these ideal examples but certain parts of them may be. James, Cliff, Haggett and Ord (1970) have devised an index that can discriminate amongst groups similar in type to these. Tinkler (1972(b)) has characterised a broad class of graphs with a systematic radial structure.



(a) a ring (polygon or simple circuit) on n points and n links

for n >= 3

(b) line (simple path) on n points and (n-1) links,

for n >= 2

(c) a star on n points, a hub and n-1 satellites and n-1 links,

for n >= 3

(d) a wheel on n points - a union of a ring and a star,

for n >= 5,
for n = 5, see f(ii).

trees on
(i) 8 points
(ii) 2 points
(iii) 7 points

Note a star is a particular kind of tree, and so is a line.

Complete graphs (cliques - see text) on

(i) 3 points
(ii) 4 points
(iii) 5 points, etc.

Fig. 5 Types of Graph

A sub-graph of a graph, G, is defined as any collection of points and lines taken from the graph G. Fig. 6(a) is a graph, G, and Fig. 6(b) is a sub-graph of G based on the points (1,2,5,6,8,9,10) together with the links which inter-connect those points. This particular sub-graph is a wheel centred on point 8. Fig. 6(c) is another sub-graph based on points (9,13,10,4,5,6,12) and their interconnecting links. This sub-graph consists of two disconnected pieces, and both parts are both lines and trees (see Fig. 5(b) and (e)). In a real-world example G might be the graph of a

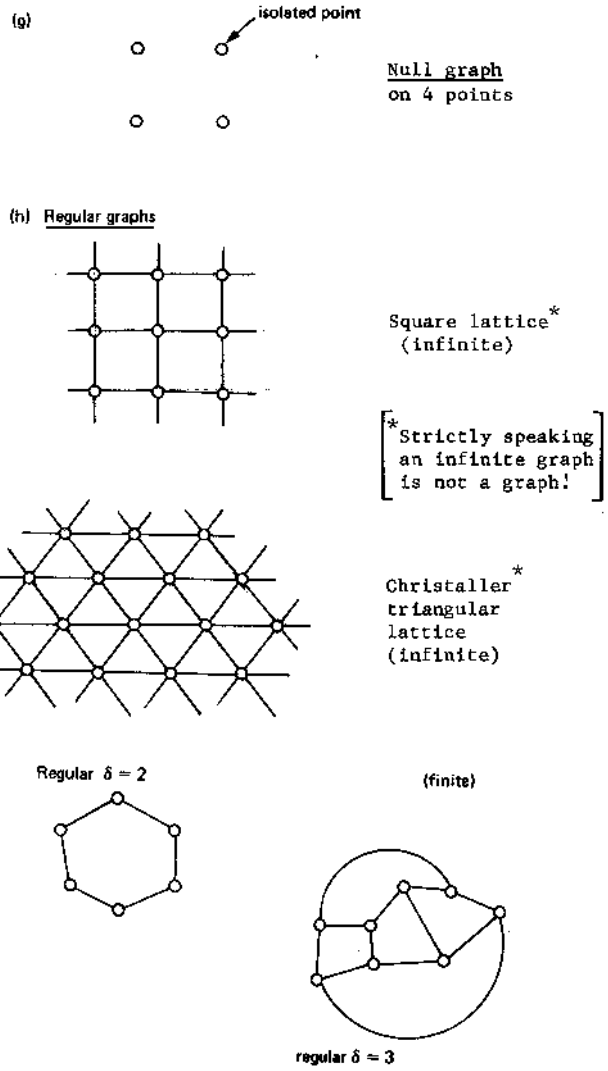


Fig. 5 (continued)

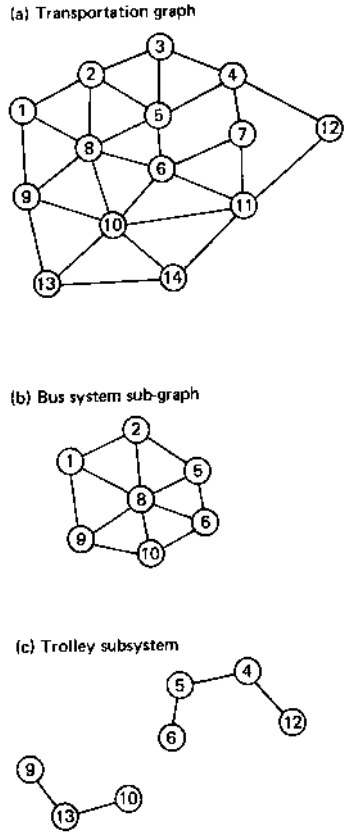


Fig. 6 A Transportation system and its subgraphs

transportation system and sub-graph Fig. 6 (b) might be the bus system whereas the sub-graph 6 (c) might represent the graph of the street trolleys The remaining links, and the points at each end of them may represent links on which there is only private transportation. A sub-graph may be derived just as well by taking a collection of links from the graph together with the points to which they are connected.

For example the links on which there is only private transportation define the following sub-graph of G, Fig. 7.

A super graph is the reverse concept of a sub-graph. Fig. 7 may be called a graph G one of whose super-graphs is Fig. 6 (a). A graph of the

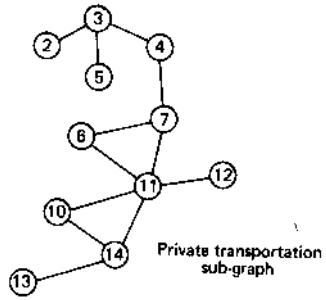


Fig. 7 Private transportation sub-graph of Figure 6

railway system of any country has as a super-graph the graph of all transport links in the country. The general idea of a super-graph is that a graph may be in fact a sub-graph of some larger graph.

So far all links in graph have been considered as two-directional but in many cases we want to define one-way relationships, for example stream networks and one-way road systems. In the former case a stream such as Fig. 8(a) is drawn as a digraph in Fig. 8(b).

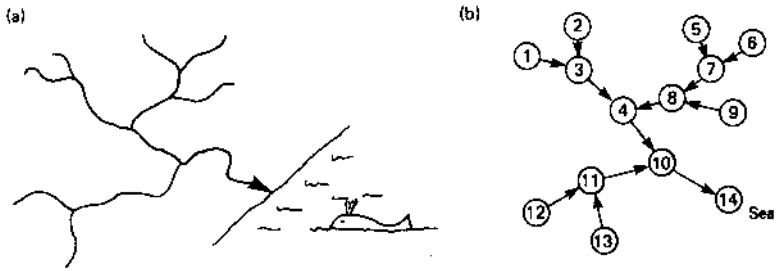


Fig. 8 A stream network and its digraph

In a stream obviously all links are uni-directionally downstream but in a road system this need not be so; for example Fig. 9 contains one-way and two-way links and the whole includes circuits so that following certain links will bring a traveller back to his starting point.

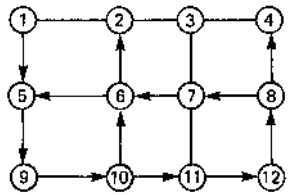


Fig. 9 One-way road system as a graph

It is salutary to reflect briefly on the number of graphs which are structurally distinct from each other for each n. Table 1, simplified, is taken from Wilson, p.162.

Table 1

n	1	2	3	4	5	6	7	8
graphs	1	2	4	11	34	156	1044	12346
trees	1	1	1	2	3	6	11	23
connected* planar graphs	1	1	2	6	20	105	?	?
digraphs	1	3	16	218	9608	2×10^{6}	9×10^{8}	2×10^{12}

(Wilson, 1972) *For definition see Chapter IV of this monograph.

It is clear that one soon encounters an astronomical number of structurally distinct graphs. The numbers remain large even when, for a particular n, the numbers of graphs with a fixed number of lines is considered. For example, Fig. 24 (a) which has n=6 points and q=7 lines has 24 possible, different, structures and a graph with n=9 and q=14 would have 15615 different representations: Obviously in these circumstances it is natural to ask what the 'average' structure of an n=9, q=14, graph would look like. Section III (iv) addresses itself to this question in terms of the frequency distribution of the number of lines incident at any one point; a limited introduction to a difficult topic.

(iv) Deriving graphs in actual examples

There is really no end to the possible applications of graphs to geographical problems - the main requirement is a willingness to try to see a problem in graph theory terms and to apply a little ingenuity. The reader is urged to see applications of graph theory in the literature where apart from the 'obvious' examples of transportation systems they have also been applied to trying to simplify the definition of atmospheric pressure surfaces, (Warntz and Waters, 1975), to industrial structure (Schmidt, 1975, Campbell, 1974), and to patterns of through valleys in heavily glaciated terrain, (Haynes, 1975). Rural periodic market systems, Jackson (1972), Tinkler (1973), Symanski and Webber (1974), and Wood (1975) have yielded a particularly rich variety of examples. By representing the routes of travelling vendors between rural markets Symanski and Webber (1974) define digraphs (Fig. 17), which may then be analysed by the methods of this booklet to reveal patterns of possible interactions between markets. Similarly Good (1975) has examined the efficiency of vendor routes in Buganda using a modification of a travelling salesman algorithm. This springs from the ideas of traversability mentioned in Section V. Wood (1975) in a paper summarised briefly in Section VI has examined the articulation of a market system by domestic market visits using connectivity analysis, (Section IV). In a theoretical analysis Tinkler (1973) conceives of each market as lying at the centre of its own market area, and being surrounded by a variable number of other market areas with which it has a boundary in common. The assignment of days to the markets (and market areas) so that no two mutually adjacent markets operate on the same day is the same problem as colouring a planar graph - a topic outlined in Section V. In a final example Jackson (1972) derives valued digraphs to represent the expected flow of crops between rural markets with different price levels. Indeed, the power of graph theory lies not so much in simply drawing a representation on a piece of paper, as from the application of the powerful algebraic methods discussed in later sections.

(v) Matrix representation of a graph

Drawing a graph on a sheet of paper to represent a situation does little more than clarify and simplify the pattern. In addition, as mentioned in the introduction, a visual analysis is only possible for small examples and we need reliable methods for obtaining the same results with very large examples. To achieve this a graph is represented as a matrix. The points of the graph are labelled arbitrarily and these labels are used to index the rows and columns of a matrix. A number is placed at the intersection of the i^{th} row and the j^{th} column if the i^{th} and j^{th} points are joined by a link, and a 0 otherwise. The number represents the number of links between the two places. A simple graph has 1 or 0 whereas a graph may have any non-negative integers. For simple graphs and graphs if i is joined to j then j is joined to i and the same entry appears in two places. Fig. 10 (a) shows a simple graph and 10 (b) shows its matrix, A, usually termed an adjacency matrix or sometimes a connection matrix. Notice that for a graph (simple or general) the entry $A_{ij} = A_{ji}$ in all cases and the matrix is said to be symmetric about the main diagonal; for example entry $A_{23} = 1 = A_{32}$, and entry $A_{16} = 0 = A_{61}$, and so on. Conventionally an adjacency matrix has 0's on the main diagonal but the adjacency matrix is symmetric, or is not, irrespective of the values along the main diagonal. Diagonal entries are used to represent self-connections. Their use and meaning

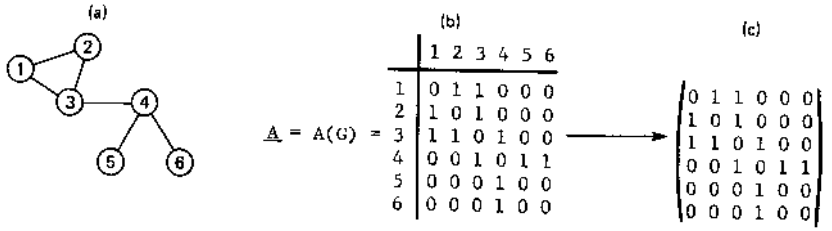


Fig. 10 Deriving the matrix of a simple graph

are discussed in Section II.

Any graph may be represented as a matrix and conversely any symmetric matrix whose entries are non-negative integers can be interpreted as a graph. However, for simple graphs the matrices will always have entries of either 0 or 1, so called (0,1) matrices. A graph, however, will lead to matrices with (0, integer) entries because two places may be linked by more than one link. Fig. 11 (a) and (b) illustrates this.

The adjacency matrix of a simple digraph is constructed in the same fashion but the matrix is not symmetric. For example for a stream system: Figure 12 (a) to (d) and the matrix is not symmetric. If there are one-way and two-way links then parts of the matrix will be symmetric and parts will not. Therefore, the matrix is asymmetric, for example Fig. 13 (a), (b). As shown in Figs. 10(c), 12(d) and 13(b) the usual representation of the matrix is with curved brackets, the labelled rows and columns are implicitly understood, and of course the matrix is square.

In some cases we may wish to place some numerical weight on a link in a structure. So far these weights have been 0,1 or integer, but the principle is easily extended to real-valued weights. A weight with a fractional part expressed to m decimal places may be converted to an integer by multiplying by 10^m. Assuming all weights in the matrix are expressed to the same accuracy then all we need do is multiply all weights by 10^m and the problem reverts to that of a graph or digraph. Each link in the graph or digraph can then be thought of as representing a smaller unit than the original one. In practice most weights will be integer.

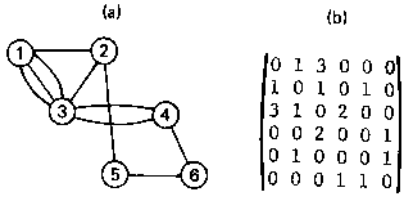


Fig. 11 A graph and its matrix

It makes no difference in matrix calculations how the labels are assigned to points of the graph but for a large empirical example it helps if some order is imposed on the labelling process. Often labels are set up by working systematically through the system, as a diffusion might spread throughout it. Another method that may be of help in structural studies is to label points in descending order according to the number links connected to them. For example Figs. 14 (a), (b). It is easy to see that for this example the places with many connections are themselves linked together (States 1,2,3) while the places with few connections, nodes 5 to 10, are not connected amongst themselves.

II MATRIX OPERATIONS AS PROCESSES IN GRAPHS

Once a graph has been represented by a matrix then operations on the matrix can be used to reveal properties of the graph, some of them obvious and others much less obvious. The important issue is that every matrix operation performed on the graph must be interpreted properly in terms of what it 'does' to the graph. It is convenient to learn the operations of matrix algebra in terms of what they mean for graphs.

(i) Matrix addition and subtraction

Two matrices may be added by simply adding their corresponding elements. To be conformable for addition two matrices must have an identical number of rows and columns. Suppose we have two different graphs with the same number of nodes, and we are able to identify the nodes in each graph as identical, as for example in Fig. 15, then by taking the adjacency matrix for each graph and adding we obtain an adjacency matrix for the addition of the two graphs, G_3. (The addition and union of two graphs may mean different things in some graph theory texts, Wilson (1972), Harary (1969)).

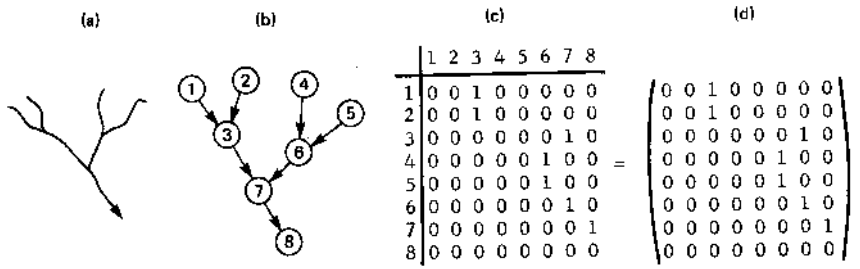


Fig. 12 The asymmetric matrix of the digraph of a stream network

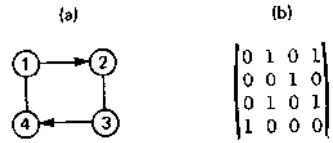


Fig. 13 An asymmetric matrix

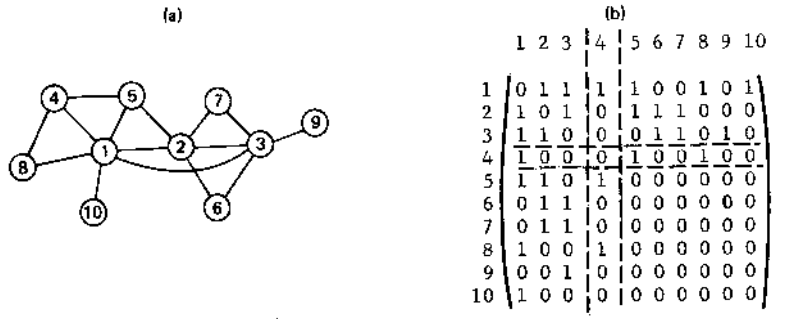
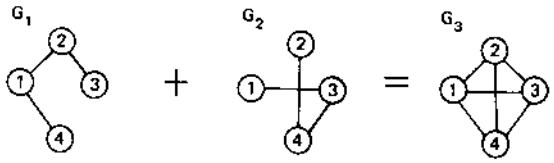


Fig. 14 Canonical ordering of a matrix of a graph



$$\underline{A}(1) + \underline{A}(2) = \underline{A}(3)$$

$$\begin{pmatrix}0&1&0&1\\1&0&1&0\\0&1&0&0\\1&0&0&0\end{pmatrix} + \begin{pmatrix}0&0&1&0\\0&0&0&1\\1&0&0&1\\0&1&1&0\end{pmatrix} = \begin{pmatrix}0&1&1&1\\1&0&1&1\\1&1&0&1\\1&1&1&0\end{pmatrix}$$

(Add corresponding elements in the matrices.)

Fig. 15 Adding graphs and matrices

We might regard G_3 of $\underline{A}_{(3)} = A(G_3)$ as the graph or matrix respectively of two different transport systems G_1 and G_2 defined over the same set of cities (the nodes). The systems may be graphs or digraphs, and there is no need for the result to be just (0,1) in form and if in some transport mode a city has no links then it appears as an isolated node; in the adjacency matrix it has a row and column of zeros. For example Figure 16.

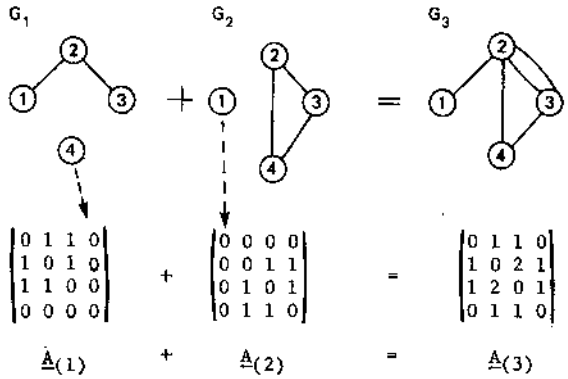


Fig. 16 Adding graphs and matrices

Matrix subtraction is just the reverse process so that taking the examples in Fig. 16, $G_2 = G_3 - G_1$ in graphs and $\underline{A}_{(2)} = \underline{A}_{(3)} - \underline{A}_{(1)}$ in matrices gives:

$$\begin{pmatrix}0&0&0&0\\0&0&1&1\\0&1&0&1\\0&1&1&0\end{pmatrix} = \begin{pmatrix}0&1&1&0\\1&0&2&1\\1&2&0&1\\0&1&1&0\end{pmatrix} - \begin{pmatrix}0&1&1&0\\1&0&1&0\\1&1&0&0\\0&0&0&0\end{pmatrix}$$

As with normal subtraction if two matrices are chosen at random on the right then the result may have negative elements.

The obvious principle is that we can 'add' graphs (and their matrices) whenever the set of nodes involved is the same - which makes obvious sense - but with the proviso that, as in G_1 of Fig. 16, we can augment the node set of the graph by adding in nodes with no links so that both graphs to be added have the same node set. In G_1 of Fig. 16 node 4 augments the set of nodes 1, 2 and 3 to ensure conformability to G_2. In G_2 node 1 augments the set 2, 3 and 4.

Symanski and Webber, (1974, p 205-6) provide an example of graph addition through matrix addition representing the different patterns of movement of trading vendors in periodic market systems. The node sets defined are the markets which may be visited and the directed links in the k^{th} matrix represent the path of the k^{th} vendor, (Fig. 17).

(ii) Multiplying or dividing a matrix by a scalar

A scalar is merely an ordinary common or garden number but is so called to distinguish it from matrices which are collections of numbers. Multiplying a matrix by a scalar is the process of multiplying each element of the matrix by that scalar. If the scalar is an integer n then multiplying an adjacency matrix A by n is equivalent to adding A to itself n times. If the scalar is a real number then it is equivalent to scaling every element of the matrix by the same weight. If the scalar is a fraction smaller that 1 the same argument applies - it is again a weight on each element of the matrix, obviously if the weight is zero then the whole matrix reduces to zero entries - a system of n nodes with no links. At this stage these applications would appear to be trivial or nonsensical but they do in fact arise in more sophisticated studies arising from matrix operations dealt with in later sections (see Section III). Some articles relevant to this are Garrison (1960), Nystuen and Dacey (1961), Harvey (1972), Stutz (1973) and Tinkler (1974).

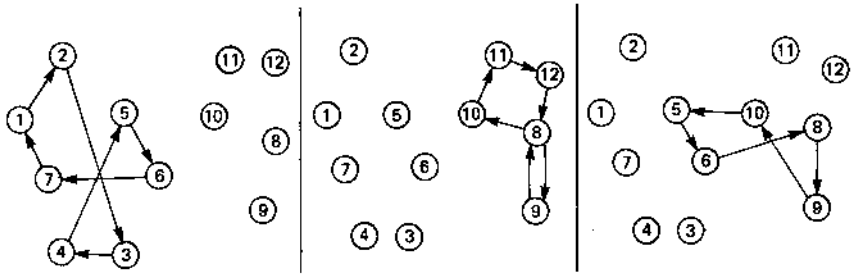
(iii) The Null matrix, 0, and the Identity matrix, I

These matrices fulfill a similar role in matrix algebra to the numbers 0 and 1 in ordinary algebra. They are square matrices defined on whatever n is appropriate to the problem in hand. As graphs they are trivial, but they are essential to matrix operations and interpretations. The null matrix is an n by n matrix full of zeros. It is equivalent to a graph of n nodes with no links. The identity matrix is an n by n matrix with 1's down the principal diagonal and zeros elsewhere. It is equivalent to a graph of isolated nodes but with a self-connection defined for each node. There are no links between the nodes. When written with other matrices the null and the identity matrices are assumed to take on appropriate dimensions to ensure conformability.

(iv) Matrix multiplication

Matrix multiplication is the basis for many methods for manipulating graphs. Most of these methods involve multiplying a matrix by itself but a few involve multiplying a matrix by a vector or by another matrix. In order to multiply two matrices the matrices must be conformable but the rules for conformability are different to those for addition. To make this clear let a matrix $\underline{A}$ with n rows and m columns by an $\underline{A}_{n,m}$ matrix and similarly for a matrix $\underline{B}_{m,n}$. Matrices are conformable for multiplication if, when written in the order intended for multiplication:

$$\underline{A}_{n,m} \cdot \underline{B}_{m,n} = \underline{C}_{n,n}$$



$\underline{A}$ =

	1	78	12
1	0100000		
	0010000		
	0001000		
	0000100	O	
	0000010		
	0000001		
7	1000000		
		O	O
12			

$\underline{B}$ =

	1	78	12
1			
		O	O
7			
8			01100
		O	10000
			00010
			00001
12			10000

$\underline{C}$ =

	1	78	12
1			
	O	O	O
			010000
			000100
7	O		000000
8			000010
			000001
			100000
12	O	O	O

$\underline{M} = \underline{A} + \underline{B} + \underline{C}$ =

	1	78	12
1	0100000		
	0010000		
	0001000	O	
	0000100		
	0000020		
	0000001	10000	
7	1000000	00000	
8		00002100	
		00010100	
	O	10000010	
		00000001	
12		00010000	

Fig. 17 Routes of travelling vendors in rural periodic markets (after Symanski and Webber, 1974, with permission)

the two inner indices (m,m) are identical in size. In the example they are and the reverse in this example is also true, (n,n), viz:

$$\underline{B}_{m,n}\ \underline{A}_{n,m} = \underline{D}_{m,m}$$

However, as indicated the result is different in size. In the first example the product is $\underline{C}_{n,n}$ whereas in the second the product is $\underline{D}_{m,m}$. The rule is that if the inner indices are the same and permit multiplication then the outer indices indicate the size of the product. Matrices for graphs are always square and hence we only need to check that the matrices being multiplied are of equal size. By these rules, an $\underline{A}_{n,n}$ matrix multiplied by a $\underline{B}_{n,n}$ matrix yields a product $\underline{C}_{n,n}$. As mentioned most operations on graphs using matrices involve self-multiplication so that automatically the matrices are square and conformable. A vector is simply a matrix with one row, a row vector say $\underline{r}_{1,n}$ or a matrix with one column, a column vector say $\underline{c}_{n,1}$. To be conformable with a matrix $\underline{A}_{n,n}$ we see that

$\underline{r}_{1,n}\ \underline{A}_{n,n} = \underline{r}^{*}_{1,n}$ i.e. another row vector following the rules given above whereas:

$\underline{A}_{n,n}\ \underline{r}_{1,n}$ is not conformable for multiplication. On the other hand: $\underline{A}_{n,n}\ \underline{c}_{n,1} = \underline{c}^{*}_{n,1}$ another column vector and again:

$\underline{c}_{n,1}\ \underline{A}_{n,n}$ is not conformable for multiplication.

For convenience vectors are usually written in lower case boldface, (underlined in hand-written work), to make them easy to distinguish in reading equations, from matrices, usually written in capitals boldface, (underlined when hand written).

We now come to matrix multiplication in detail; it is described first and interpreted later. The formal definition using summation, Σ, notation is that:

$$\underline{C}_{i,j} = (\underline{A}.\underline{B})_{i,j} = \sum_{k=1}^{k=n} \underline{A}_{i,k}\ \underline{B}_{k,j}.$$

To find the i,j^{th} element of the product $\underline{C}$ where i is row index and j is the column index we take the i^{th} row of the first matrix and the j^{th} column of the second matrix. We now multiply together their corresponding elements and add the results. The result is the i,j^{th} element of the product.

Specifically, and in this case letting A, the adjacency matrix of a graph, be both the first and the second matrix we write:

$\underline{A}.\underline{A}. = \underline{A}^2$ - by analogy to ordinary algebra the resulting matrix is $\underline{A}^2$.

Let $\underline{A}$ be $\underline{A}_3$ from Fig. 16:

$$\begin{pmatrix} 0&1&1&0 \\ 1&0&2&1 \\ 1&2&0&1 \\ 0&1&1&0 \end{pmatrix} \times \begin{pmatrix} 0&1&1&0 \\ 1&0&2&1 \\ 1&2&0&1 \\ 0&1&1&0 \end{pmatrix} = \begin{pmatrix} ? \end{pmatrix}$$

To calculate the 1,1 element of A2 take the first row of the first matrix A and the first column of the second matrix, again A; ignore their nature as rows or columns now and place side by side viz:

1st row	1st column
0 x 0	= 0
1 x 1	=1
1 x 1	= 1
0 x 0	= 0
	Add 2

multiply the corresponding elements and add. We could also represent this process as:

$$(0 \times 0) + (1 \times 1) + (1 \times 1) + (0 \times 0) = 2$$

Now simply follow these rules to obtain all 16 elements of $\underline{A}^2$. For example $\underline{A}^2_{1,2}$ is:

$$\underline{A}^2_{1,3} = \begin{matrix} (0 \times 1) + (1 \times 0) + (1 \times 2) + (0 \times 1) = 2 \\ (0 \times 1) + (1 \times 2) + (1 \times 0) + (0 \times 1) = 2 \end{matrix}$$

and $\underline{A}^2_{1,4} = (0 \times 0) + (1 \times 1) + (1 \times 1) + (0 \times 0) = 2.$

The first row of $\underline{A}^2$ is therefore:

$$\begin{pmatrix} 2 & 2 & 2 & 2 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{pmatrix}$$

The remaining elements are:

$$\underline{A}^2_{2,1} = (1 \times 0) + (0 \times 1) + (2 \times 1) + (1 \times 0) = 2$$

$$\underline{A}^2_{2,2} = (1 \times 1) + (0 \times 0) + (2 \times 2) + (1 \times 1) = 6$$

$$\underline{A}^2_{2,3} = (1 \times 1) + (0 \times 2) + (2 \times 0) + (1 \times 1) = 2$$

$$\underline{A}^2_{2,4} = (1 \times 0) + (0 \times 1) + (2 \times 1) + (1 \times 0) = 2$$

$$\underline{A}^2_{3,1} = (1 \times 0) + (2 \times 1) + (0 \times 1) + (1 \times 0) = 2$$

$$\underline{A}^2_{3,2} = (1 \times 1) + (2 \times 0) + (0 \times 2) + (1 \times 1) = 2$$

$$\underline{A}^2_{3,3} = (1 \times 1) + (2 \times 2) + (0 \times 0) + (1 \times 1) = 6$$

$$\underline{A}^2_{3,4} = (1 \times 0) + (2 \times 1) + (0 \times 1) + (1 \times 0) = 2$$

$$\underline{A}^2_{4,1} = (0 \times 0) + (1 \times 1) + (1 \times 1) + (0 \times 0) = 2$$

$$\underline{A}^2_{4,2} = (0 \times 1) + (1 \times 0) + (1 \times 2) + (0 \times 1) = 2$$

$$\underline{A}^2_{4,3} = (0 \times 1) + (1 \times 2) + (1 \times 0) + (0 \times 1) = 2$$

$$\underline{A}^2_{4,4} = (0 \times 0) + (1 \times 1) + (1 \times 1) + (0 \times 0) = 2$$

and therefore:

$$\underline{A}^2 = \begin{pmatrix} 2 & 2 & 2 & 2 \\ 2 & 6 & 2 & 2 \\ 2 & 2 & 6 & 2 \\ 2 & 2 & 2 & 2 \end{pmatrix}$$

We now give the interpretation which can be placed on these elements. The i,j^{th} element in $\underline{A}^2$ indicates the number of 2-step routes between i and j (nodes in the graph) that are possible in the graph. In general $\underline{A}^k$ indicates the number of k-step routes in the graph for each pair of nodes. Consider for example the element $\underline{A}^2_{2,3}$ which is formed from the row (1,0,2,1) and the column (1,2,0,1). Writing these side by side vertically:

row 2	column 3	row 2 interpretation	col. 3 interpretation	through possible	routes
1	1	from 2 to 1	from 1 to 3	yes	1
0	2	from 2 to 2	from 2 to 3	no	0
2	0	from 2 to 3	from 3 to 3	no	0
1	1	from 2 to 4	from 4 to 3	yes	1
				total	2

and interpreting each element then since the row is 2 then the column positions in this row, as it lies in the matrix, indicate the number of routes available from node 2 to all nodes in the system. For column 3 the row positions indicate routes from all nodes in the system to node 3. Placed side by side we see that they can be used to show the availability of routes between nodes 2 and 3 (the values of i and j being considered) via all other nodes in the system. The row vector element 1 shows that one route exists from node 2 to node 1 and the column vector element 1 shows that one route is possible from node 1 to 3, hence there is a total of one route, (1 x 1), available. 1i the case of node 2 to node 2 there is no route available, 0, hence no onward route to node 3 is possible. For routes from node 2 to 3 via node 3 it is again seen that none are possible because there are none from 3 to 3. The remaining possibility is from 2 to 3 via node 4 and in this case a through route is possible. Adding all these terms gives the total number of possible routes which is 2. When the two terms being multiplied are larger than 0 or 1 the interpretation is that if there are say 3 routes from i to k and then 2 routes from k to j then the 3 routes on the first part combine with the 2 in the second part giving a total of 2 x 3 = 6 possible routes.

The discussion has been long and tedious, but understanding it is critical to an appreciation of graph theoretical methods.

The interpretations extend easily to higher powers of $\underline{A}$ for the following reason. Consider the multiplication of $\underline{A}^2$ and $\underline{A}$ to give $\underline{A}^3$. The entries of $\underline{A}^2$ give the number of two step routes between each i and j. By substituting into the arguments given above for say the $A^2_{2,3}$ element then from:

$$\underline{A}^2 \cdot \underline{A} = \underline{A}^3$$

$$\begin{pmatrix} 2 & 2 & 2 & 2 \\ 2 & 6 & 2 & 2 \\ 2 & 2 & 6 & 2 \\ 2 & 2 & 2 & 2 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 2 & 1 \\ 1 & 2 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix} = \begin{pmatrix} & & ? & \\ & & & \\ & & & \\ & & & \end{pmatrix}$$

2nd row of $_A{}^2$	3rd. col. of A		row	col.	possible routes
2	1	Total 2 step routes 2- 1 1- 3		yes	2
6	2	Total 2 step routes 2- 2 2- 3		yes	12
2	0	Total 2 step routes 2- 3 3- 3		no	0
2	1	Total 2 step routes 2- 4 4- 3		yes	2
					Tr

giving $\underline{A}^3_{2,3} = 16$.

The reader should check that $_A3$ is:

$$\underline{A}^3 = \begin{pmatrix} 4 & 8 & 8 & 4 \\ 8 & 8 & 16 & 8 \\ 8 & 16 & 8 & 8 \\ 4 & 8 & 8 & 4 \end{pmatrix}$$

The general rule is therefore that $\underline{A}^k$ contains the number of k-step routes between all pairs of nodes i and j. Of course the original $\underline{A}$ can be written in power form as $\underline{A}^1$, agreeing with this definition.

While this information can be obtained visually in simple graphs, the number of nodes and the number of steps required do not have to be very large for the process to become impossibly tedious and very prone to error. In most practical examples use is made of computer routines to do the matrix multiplication.

(v) Multiplying a matrix by a vector

In fact we have already done this because to compute a row in the product $\underline{A}^2$ we took the first row of $\underline{A}$ and combined it in turn with every column of the second matrix $\underline{A}$ to produce each element in the row vector which is the first row of $\underline{A}^2$. Symbolically -

$$\begin{pmatrix} x & x & x \\ & & \end{pmatrix} \cdot \begin{pmatrix} x & x & x \\ x & x & x \\ x & x & x \end{pmatrix} = \begin{pmatrix} x & x & x \\ & & \end{pmatrix}$$

The procedure for pre-multiplying a matrix A by a row vector $\underline{v}$, $v.A = c$, is simply this procedure. The procedure for post-multiplying A by a column vector v' is analogous: combine each row in A in turn with the column vector v'to produce each element in the resultant column vector r

$$\underline{A}\,\underline{v}' = \underline{r}.$$

$$\begin{pmatrix} x & x & x \\ x & x & x \\ x & x & x \end{pmatrix} \begin{pmatrix} x \\ x \\ x \end{pmatrix} = \begin{pmatrix} x \\ x \\ x \end{pmatrix}$$

The procedure has already been followed to compute the first column of $\underline{A}^2$ from $\underline{A}$ and $\underline{A}$. If v and v' are vectors with ones in every location the two operations above yield the column sums and row sums respectively of $\underline{A}$. This method can be used with row and column vectors filled with 1's to obtain the degrees of the nodes of a graph.

For example using the matrix $\underline{A}_{(3)}$, (Fig. 16), and a column vector of 1's, $\underline{1}$:

$$r = \underline{A}\,\underline{1} = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 2 & 1 \\ 1 & 2 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \\ 4 \\ 2 \end{pmatrix}$$ is the vector, $\underline{r}$, of out-degrees

and when $\underline{1}'$ is a row vector of 1's:

$$c = \underline{1}'\underline{A} = (1, 1, 1)\begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 2 & 1 \\ 1 & 2 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix} = (2, 4, 4, 2)$$

is the vector, c, of in-degrees. The vectors r and c may be different when the matrix is the adjacency matrix of a digraph, and will be identical for the adjacency matrix of a graph.

(vi) 0's or 1's in the main diagonal of A

In structural studies it makes little sense to connect a node to itself with a loop but in the previous section it is clear that the existence of a route implies movement along it, that is to say movement in the graph. In these circumstances connecting nodes to themselves, placing 1's in the main diagonal, allows a 'movement' to remain at a node as a move equivalent to moving one step - a sort of graph-theoretical stopover: In this case the 'routes' are counted with the understanding that a route may remain at a node for one or several moves. One might wish to examine the results of allowing certain nodes to be retentive in this fashion, while others are not. A notation to distinguish between 0's and 1's in the main diagonal is $_0\underline{A}$ and $_1\underline{A}$. 0's are assumed if the pre-subscript is omitted. Alternatively $_1\underline{A}$ may be written as $(\underline{I} + \underline{A})$ where $\underline{A}$ is understood to be $_0\underline{A}$. With 1's in the main diagonal the numbers in the powers of $_1\underline{A}$ will necessarily change so that $_1\underline{A}^2$ is now, (from Fig. 16),

$$ {}_1\underline{A}\cdot{}_1\underline{A} = \begin{pmatrix}1&1&1&0\\1&1&2&1\\1&2&1&1\\0&1&1&1\end{pmatrix}\cdot\begin{pmatrix}1&1&1&0\\1&1&2&1\\1&2&1&1\\0&1&1&1\end{pmatrix} = \begin{pmatrix}3&4&4&2\\4&4&6&4\\4&6&7&4\\2&4&4&3\end{pmatrix} = {}_1\underline{A}^2 $$

The reader should check for himself the values of ${}_1\underline{A}^3$.

A problem that can occur with 0's in the main diagonal can be exemplified by the following example: (from Fig. 12(a)) -

$$ \underline{A} = \begin{pmatrix}0&1&0&1\\1&0&1&0\\0&1&0&1\\1&0&1&0\end{pmatrix}, \underline{A}^2 = \begin{pmatrix}2&0&2&0\\0&2&0&2\\2&0&2&0\\0&2&0&2\end{pmatrix}, \underline{A}^3 = \begin{pmatrix}0&4&0&4\\4&0&4&0\\0&4&0&4\\4&0&4&0\end{pmatrix}, \underline{A}^4 = \begin{pmatrix}8&0&8&0\\0&8&0&8\\8&0&8&0\\0&8&0&8\end{pmatrix} $$

Inspection of these examples shows that while node 3 can be reached in two steps from node 1 it cannot be reached in 3 steps, although it is again reachable in 4 steps. Hence if 0's are placed in the main diagonal $\underline{A}^n$ records routes of exactly n steps. If, however, 1's are placed in the main diagonal then $\underline{A}^n$ records that a given i and j are connected, or not, by at most an n-step route, i.e. it includes records of shorter steps too, by virtue of the stop-over property. For example if:

$$ {}_1\underline{A} = \begin{pmatrix}1&1&0&1\\1&1&1&0\\0&1&1&1\\1&0&1&1\end{pmatrix}, \text{ then } {}_1\underline{A}^3 = \begin{pmatrix}7&7&6&7\\7&7&7&6\\6&7&7&7\\7&6&7&7\end{pmatrix} $$

and in this case node 3 is reachable from node 1 in 6 different ways. This is achieved by combining holding moves at nodes 1 and 3 with the two ways to get to 3 from 1, plus holding moves on either route to 3 at nodes 2 and 4 respectively; one such sequence is 1-)2-42-43 where the 2-42 is the 'holding' move.

(vii) Matrix multiplication involving digraphs - asymmetric matrices

All the results obtained above hold just as well for digraphs as for graphs except that the matrices involved, and their products, are no longer symmetric matrices. To illustrate briefly consider the digraph, D, Fig. 18, with adjacency matrix A = A(D),

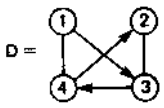


Fig. 18 A digraph

$$ \underline{A} = \begin{pmatrix}0&0&1&1\\0&0&1&0\\0&1&0&1\\1&1&0&0\end{pmatrix} $$

Then to find the number of 4-step routes from say node 1 to node 2 we need $\underline{A}^4=\underline{A}^2\underline{A}^2$

$$ \underline{A}\,\underline{A} = \begin{pmatrix}0&0&1&1\\0&0&1&0\\0&1&0&1\\1&1&0&0\end{pmatrix}\cdot\begin{pmatrix}0&0&1&1\\0&0&1&0\\0&1&0&1\\1&1&0&0\end{pmatrix} = \begin{pmatrix}1&2&0&1\\0&1&0&1\\1&1&1&0\\0&0&2&1\end{pmatrix} = \underline{A}^2 $$

$$ \underline{A}^2\underline{A}^2 = \begin{pmatrix}1&2&0&1\\0&1&0&1\\1&1&1&0\\0&0&2&1\end{pmatrix}\cdot\begin{pmatrix}1&2&0&1\\0&1&0&1\\1&1&1&0\\0&0&2&1\end{pmatrix} = \begin{pmatrix}1&4&2&4\\0&1&2&2\\3&4&1&2\\2&2&4&1\end{pmatrix} = \underline{A}^4 $$

and the reader is invited to check by hand in the original graph that there are indeed 3 routes of length 4 from node 1 to node 2, and to find, for example, the 4 distinct routes from node 1 to node 4. Note that there is no need to obtain the intermediate power $\underline{A}^3$ and that $\underline{A}^8$ is obtained just as easily via $\underline{A}^4\underline{A}^4=\underline{A}^8$. It is a fortunate fact that in matrix algebra a matrix A and its powers $\underline{A}^n$ are said to commute with each other i.e. $\underline{A}^6=\underline{A}^4\underline{A}^2=\underline{A}^2\underline{A}^4$ and the order of the multiplication is irrelevant. This makes intuitive sense, the number of six-step routes between i and j in a graph must always be the same however they are computed!

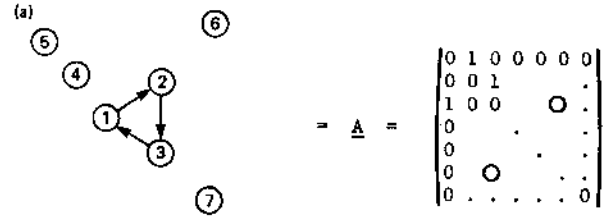
(viii) Matrix multiplication of two different matrices or graphs

In powering a single matrix, A, we examined possible routings within the A system. In applying the same basic technique of multiplication to two or more matrices, A, B, C etc., we can examine routings available in a multi-mode system where routes are followed first in one system, A, and then in another, B, and so on.

To be conformable two matrices of graphs must be defined over the same node set (p.17). This is no problem for powers of an adjacency matrix since the nodes remain the same. Conformability is ensured by augmenting the node sets where necessary as in matrix addition, see p.14. Let the first matrix, A, be the adjacency matrix of one transport mode over a set of nodes, perhaps an inner city one-way rapid transit bus system. Let the second, B, be the adjacency matrix of a second transport mode; perhaps a suburban railway system. Note that to ensure conformability A is augmented with the isolated nodes 4, 5, 6, 7, whereas B involves all the nodes. If now we write :

$$ AB=C $$

then we can still resort to the previous interpretations.



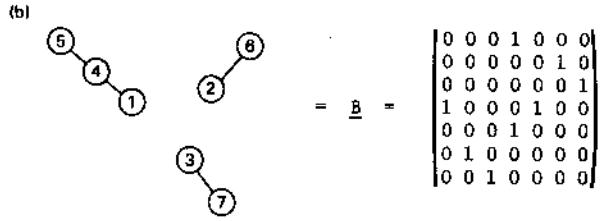


Fig. 19 A two mode transportation system

In full we have :

$$\underline{A}\,\underline{B} = \begin{pmatrix} 0&1&0&0&0&0&0\\ 0&0&1&0&0&0&0\\ 1&0&0&0&0&0&0\\ 0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0 \end{pmatrix} \cdot \begin{pmatrix} 0&0&0&1&0&0&0\\ 0&0&0&0&0&1&0\\ 0&0&0&0&0&0&1\\ 1&0&0&0&1&0&0\\ 0&0&0&1&0&0&0\\ 0&1&0&0&0&0&0\\ 0&0&1&0&0&0&0 \end{pmatrix} := \underline{C} = \begin{pmatrix} 0&0&0&0&0&1&0\\ 0&0&0&0&0&0&1\\ 0&0&0&1&0&0&0\\ 0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0 \end{pmatrix}$$

and for example taking $\underline{C}_{1,6}$

1st row A		6th col. B		
0	x	0	to 1 to 1 in A, from 1 to 1 in B no	0
1	x	1	from 1 to 2 in A, from 2 to 6 in B yes	1
0	x	0	no	0
0	x	0	no	0
0	x	0	no	0
0	x	0	etc. no	0
0	x	0	no	0
				-
			Total	1

but the interpretation is that the first move is made in graph A and the second in graph B. The only two-step route from A to B starting in node 1 is to node 6 as the reader may verify. Note that the 2-step route from 1 to 3 is not recorded since this requires 2 steps in the A network. In general, therefore, this notion can be extended so that if we investigate $(\underline{A}^2\underline{B})$ we should be looking at the result of 2 successive moves in the A network followed by one move in the B network; $(\underline{A}^2\underline{B}^2\underline{A})$ would record the results of 2 steps in A, 2 steps in B followed by another step in A, and so on. Of course additional networks, defined on the same node set could be included and as a final refinement calculating say $\underline{A}\,(\underline{A}+\underline{B})^2\,\underline{B}$ would give the results of one move in A followed by two moves in the graph consisting of A and B together followed by a move in B alone. A full discussion may be found in Tinkler (1975).

(ix) Powers of A as a non-conserved diffusion system

The availability of routes, and their numerical computation in applications of these ideas would seem to imply their utilisation. A technical argument (Tinker, 1973, 1974) shows that in fact any diffusion interpretation placed on powers of A is isomorphic to a non-conserved diffusion where total inflow to a node does not necessarily balance with total outflow. This behaviour characterizes the diffusion of information as opposed to physical entities and the powers of A are analagous to a generalised Hagerstrand model in which the telling rate varies from node to node instead of being constant. A classical conserved diffusion requires that A be converted to a transition matrix before the powering commences; (see Section III (ii) p. 32).

(x) The Solution Matrix, ${}_0\underline{A}^m$

If we have an adjacency matrix ${}_0\underline{A}$ then the solution matrix of ${}_0\underline{A}$ is given by ${}_0\underline{A}^m > 0$ where m is the smallest integer for which the inequality is true. (A matrix equality or inequality must be true for all the elements of the matrix.) The problem is to determine whether a matrix has a solution matrix and if so to determine the smallest m. Alao (1970) gives a full discussion for ${}_0\underline{A}$ and shows in the most general case, when ${}_0\underline{A}$ is the adjacency matrix of a digraph, that m is difficult to determine. When ${}_0\underline{A}$ is the adjacency matrix of a graph (simple or otherwise) then ${}_0\underline{A}$ has a solution matrix if (1) the graph is connected, (a term fully explained in Section VI but intuitively corresponding to the idea that the graph is in 'one piece'), and (2) the graph contains an odd cycle. For details

and definitions of cycles see Section VI, but the principle is that if it is possible, for some node, to return to the same node in an odd number of steps then there is an odd cycle. Most empirical networks contain odd cycles, so if they are graphs, not digraphs, (their matrices are symmetric), then $_0\underline{A}^m$ can be found by continuing the powering process until $_0\underline{A}^m > 0$ is satisified. As Alao points out there is no easy way to find m, and the odd cycle criterion is not sufficient to guarantee the existence of $_0\underline{A}^m$ if $_0\underline{A}$ is from a digraph.

However, the smallest possible odd cycle is one of length one and therefore the main diagonal of $\underline{A}$ will carry a 1 in at least one position, ($_1\underline{A}$ has one in every position). In these circumstances $_1\underline{A}^m$ does always exist and the upper limit on m is n-1 where n is the number of nodes in the system. When m is smaller than n-1 then m is the graph diameter, Section VI, p.41. Alao dismisses the usefulness of this situation on the grounds that entries of $_1\underline{A}^k$ no longer have the simplest interpretation possible when the main diagonal is zero. This is true but if the 'stop-over' interpretation (above) is remembered then one can insert 1's and proceed happily. It can be shown in fact that while the numbers in $_1\underline{A}^k$ are different to those in $_0\underline{A}^k$ they tend to the same proportions provided that $_0\underline{A}$ has a solution matrix, and the difference in terms of proportions becomes vanishingly small the larger k is made. If $_0\underline{A}$ fails to have a solution matrix then it will be necessary to average several successive $_0\underline{A}^k$. The proportions obtained from this averaging will tend to the proportions in $_1\underline{A}^k$ for the same k values. A fuller discussion is found in Tinkler (1972).

(xi) Sums of powers of adjacency matrices

One way to take into account all possible routings of all possible lengths in a graph is to add all the matrices obtained e.g. A, A2, A3 etc. so that:

$$\underline{S} = \underline{A} + \underline{A}^2 + \underline{A}^3 + \underline{A}^4 + \underline{A}^5 + \ldots\ldots\ldots$$

($_0\underline{A}$ is assumed), and $\underline{S}$ is termed the accessibility matrix.

The problem with this approach is two-fold. On the one hand it is not obvious at what power level, k, the summation should stop, and on the other because the higher powers contain much larger numbers than the smaller powers they consequently dominate the total sum. (In fact the entries in successive powers increase in an approximately geometric fashion - and for matrices of graphs the common ratio is larger than 1 and smaller than the largest marginal sum of A, the node with the largest number of links into or out of it. For a full discussion see Tinkler (1972 (a) and Tinkler (1974)). The excessive weight of the higher power sums in S is contrary to our geographical expectations, where we expect a decreasing influence with increasing distance, not an increasing one. Because of this problem direct sums of this sort are not usually made and accessibility, if it is studied in these terms, is analysed with respect to a single matrix Ak. The problem still remains, which m or k to choose? There are three rules of thumb. The first is to stop at the solution number (see the last section). The second one is to choose k n - 1 where n is the number of nodes in the graph. This ensures that, if the graph is strongly connected (see Section IV), the two most separated points are connected by at least one route. In fact this is a fail-safe and so

thirdly we can usually substitute a smaller, more exact number, d, called the diameter of the graph, (see Section IV for definition). The latter two choices are usually made on the assumption that $\underline{A}^{n-1}$ or $\underline{A}^d$ will contain no zero elements, which takes us back to the last section. For safety, therefore, it may be advisable to use $_1\underline{A}$ rather than $_0\underline{A}$. Because the results of either raised to a large power and viewed as proportions tend to the same values eventually it is perhaps better to choose $_1\underline{A}$ since this converges to the final equilibrium values more rapidly.

The earliest study of accessibility in the geographical literature is that of Garrison (1960) who analysed a portion of the Interstate Highway using the methods outlined in this section. For each node he computed an accessibility index by summing the numbers in each row of the final summation matrix, $\underline{S}$. A similar study was performed by Pitts (1965) for thirteenth century Russian trade routes. He raised the $_0\underline{A}$ matrix to $_0\underline{A}^{ll}$ where ll was the diameter of the network and then summed the rows of the matrix to obtain the accessibility index for each node. The results were then expressed as proportions. The higher the row sum is, the more outgoing routes are available from that node, and the higher the 'centrality' or 'accessibility' is judged to be for that node. Another historical study, modelled on Pitts, is that of Carter (1969) for the Serbian Oecumene.

(xii) Weighted sums of powers of adjacency matrices

To counteract the influence of the excessive weight of the higher powers in a direct sum such as:

$$\underline{S} = \underline{A} + \underline{A}^2 + \underline{A}^3 + \ldots$$

a modified procedure is to use a weight on each matrix. The most commonly used is of the form:

$$\underline{S} = r\underline{A} + r^2\underline{A}^2 + r^3\underline{A}^3 + \ldots + r^n\underline{A}^n + \ldots \qquad \ldots.$$

where r is suitably chosen. The suitable choice of r requires some care. Obviously to provide attenuation r must be initially smaller than 1, but it must be much smaller than this as the following argument shows. The last section pointed out that the entries in $\underline{A}^n$ are approximately λ times larger than those in $\underline{A}^{n-1}$ where λ is a fairly constant value and for large n does tend to a steady value (called the principal eigenvalue of A). As a consequence any choice of r, reducing the entries of A, must at least counteract the growing tendency due to λ. An upper limit on λ is provided by the largest marginal sum of A. For example if A is:

$$\begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 2 & 1 \\ 1 & 2 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix} \begin{matrix} 2 \\ 4 \\ 4 \\ 2 \end{matrix}$$

2 4 4 2

then 4 is the largest marginal sum. Successive powers of A will not grow by a factor larger than 4 and so can be counteracted BY a factor of 1/4. Genuine attenuation can then be achieved by choosing the attenuation factor, r, to be smaller than 1/4, and larger than 0. If we call δ_{max} the

maximum sum of the matrix A then the proper range for r is:

$$0 < r \leq \delta_{max}^{-1}$$

(A full discussion can be found in Tinkler (1974) which discusses the literature on this issue.)

Having chosen r then it is clear that eventually some r^n is virtually zero and as r^n is acting as a scalar on $\underline{A}^n$ then eventually the contribution of $r^n\underline{A}^n$ tends asymptotically to a null matrix, $\underline{0}$, full of zeros. Consequently the matrix series is convergent and $\underline{S}$ is a finite sum. The sum may be computed directly, continuing until some $r^n\underline{A}^n$ is sufficiently close to zero, or it may be summed by computing a matrix inverse, (Tinkler, 1974). Incidentally one should note that if $\underline{Y} = r\underline{A}$ then the following forms are equivalent :

$$\underline{S} = r\underline{A} + r^2\underline{A}2 + r^3\underline{A}3 + \ldots = \sum_{n=1}^{n\to\infty} r^n\underline{A}^n$$

$$\underline{S} = \underline{Y} + \underline{Y}^2 + \underline{Y}^3 + \ldots = \sum_{n=1}^{n\to\infty} \underline{Y}^n, \text{ where } \underline{Y} = r\underline{A}$$

Of course if in some study step lengths larger than some index p are impossible or inappropriate then the series sum may be arbitrarily terminated at index p :

$$\underline{S} = \underline{Y} + \underline{Y}^2 + \underline{Y}^3 + \ldots + \underline{Y}^p.$$

An example of the correct use of the infinite series sum $\underline{S} = \underline{Y} + \underline{Y}^2 + \underline{Y}^3 + \ldots$ is to be found in Nystuen and Dacey (1961).

III 1ST ORDER PROPERTIES OF GRAPHS

(i) Nodality

The nodality (degree, local degree, valency are alternative terms) of a point is simply the number of links in the graph incident at that node. Given the adjacency matrix A of a graph the ith node has a nodality given by the sum of the ith row of A. If 1 is a column vector of ones then (A 1) r is the degree vector ofthe graph, it contains the nodality of every point or node. When A is the adjacency matrix of a digraph then A 1 r defines the out-degrees of the nodes, outgoing links, and 1' A = c where 1' is a row vector of 1's defines the in-degrees of the nodes, incoming links. In itself nodality looks too simple to be interesting despite a hallowed geographical tradition embodied in the map-reading ritual at school where the canonical invocation is to a 'town important because it is a focus of routeways'. However, for a recent appeal to simple nodality, see Burghart, (1969, p. 436).

In fact because nodality can be linked precisely to random processes in a graph its use can be justified in operations where it might otherwise seem trivial.

(ii) The degree vector of a graph and the fixed vector of the transition matrix of the same graph.

An adjacency matrix $\underline{A}$ can be reduced to a transition matrix $\underline{T} = T(\underline{A})$ by the simple expedient of dividing the elements in the i^{th} row of $\underline{A}$ by the i^{th} row sum of $\underline{A}$:

$$\underline{A} = \begin{pmatrix} 0 & 2 & 1 \\ 2 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \begin{matrix} 3 \\ 3 \\ 2 \end{matrix} \quad \underline{T} = T(\underline{A}) = \begin{pmatrix} 0 & .66 & .33 \\ .66 & 0 & .33 \\ .50 & .50 & 0 \end{pmatrix}$$

$\underline{T}$ has a fixed vector = (3/8,3/8,1/4)

The structure of the graph can then be studied using random processes and the full and powerful methods involved in Markov Chain Analysis for which the reader is referred to CATMOG No. 1 (L. Collins). Fortunately it is possible to state a theorem which relates, under certain circumstances, the degree vector of a graph, the row sums of A, to the unique fixed point probability vector of T = T(A).

Theorem (Tinkler (1973))

Let r be a vector of row sums of $\underline{A}$, i.e. $\underline{A}\,1 = r$ and c is a vector of column sums i.e. $1'\underline{A} = c$, then if $r = c \neq 0$ (where the vector inequality holds for every pair of elements) then $r = c$ is a fixed vector of $\underline{T}$, i.e. $r\,\underline{T} = r = c\,\underline{T} = c$ and may be converted to a probability vector by suitable normalization.

This theorem is used to justify many procedures in which the degree of a node is used as a measure of its importance in the graph and is a nice confirmation of the geographer's ancient commitment to nodality. The interpretation is that, at the i^{th} node, the expected amount of a process circulating at random in the structure, is exactly equal to the i^{th} degree divided by the sum of the degrees for the whole graph. Any graph, by virtue of its symmetry, satisfies the theorem and its degree vector, normalised to sum to 1, may be taken as the fixed vector of its transition matrix. A digraph may, or may not, have an adjacency matrix with equal in- and out- degree vectors.

(iii) Mean nodality, $\bar{\Delta}$, of a graph, the β index in disguise.

Obviously a weak measure of connectivity on a graph with n nodes is the mean nodality defined as:

$$\bar{\Delta} = \frac{\sum_{i=1}^{i=n} r_i}{n}$$

, $\underline{r}$ as defined in the last section. The numerator in this expression is exactly twice the number of links, q, in the graph since for nodality the two ends of each link are counted. Hence the mean nodality is $\frac{2q}{n}$, which is just twice the B index described by Kansky (1963). As more links are added to the system for fixed n the indices rise but neither Δ or β is very discriminating. Applied to rail, road and airline networks it does little more than point up the obvious and the same β (or $\bar{\Delta}$) can characterise radically different network structures, a criticism made, with examples, by James, Cliff, Haggett and Ord (1970).

TABLE 2 Uganda road network 1921 after Gould, 1967.
See Figure 20.

18 points, 17 edges, mean nodality 1.8889.

Degree	Probability	Obs. Freq.	Exp. Freq.	Cum. Obs. Prob.	Cum. Exp. Prob.	Difference	
0	0.119706	0.000000	2.154713	0.000000	0.119706	0.119706	
1	0.288293	8.000000	5.189268	0.444444	0.407999	0.036445	
2	0.304971	7.000000	5.489474	0.833333	0.712970	0.120364	*
3	0.187482	2.000000	3.374677	0.944444	0.900452	0.043993	
4	0.074688	0.000000	1.344383	0.944444	0.975140	0.030695	
5	0.020358	0.000000	0.366453	0.944444	0.995498 **	0.051054	
6	0.003909	1.000000**	0.070359	1.000000	0.999407	0.000593	
7	0.000536	0.000000	0.009652	1.000000	0.999943	0.000057	
8	0.000053	0.000000	0.000950	1.000000	0.999996	0.000004	
9	0.000004	0.000000	0L000067	1.000000	1.000000	0.000000	
10	0.000000	0.000000	0.000003	1.000000	1.000000	0.000000	
11	0.000000	0.000000	0.000000	1.000000	1.000000	0.000000	
12	ditto	ditto	ditto	ditto	ditto	ditto	
13							
14							
15							
16							
17							

* maximum difference

** According to the cumulative expected probability a node of degree 5 is expected with probability 0.0055, i.e. less than 1% of the time. The polarisation of the net is judged significant.

Reference

Gould, P.R., 1967, "Geographical Interpretation of Eigenvalues", Trans. Inst. Brit. Geogr., No. 42, p. 53-86.

Werner and Smart (1973).

IV HIGHER ORDER PROPERTIES OF GRAPHS

The section on matrix operations, particularly matrix multiplication, has introduced the idea of higher than 1st order properties in graphs, i.e. those properties which only become apparent when information on n-step adjacencies is available. In this section extensive use is made of n-step adjacencies and the reader is assumed to be familiar with them and their calculation, but, first some definitions.

(i) Routes, paths, chains and cycles

A path in a graph is a sequence of edges or links, defined by successive pairs of nodes, which are all distinct and any successive pair in the sequence have exactly one node in common. For example in Fig. 21 a path from 1 to 3 is the sequence 1→2→3, but another path is 1→5→4→3.

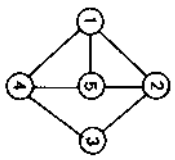


Fig. 21 A graph to illustrate routes, etc.

The routes discussed earlier are less restrictive. They are edge or link sequences in which links may be repeated so that a route from 1 to 3 might be 1→2→5→2→3, and in the case where the matrix has ones in the main diagonal so that the graph is strictly drawn as :

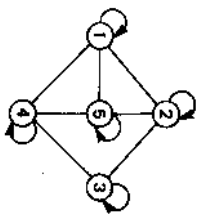


Fig. 22 A graph with self-loops to illustrate routes, etc.

a route can be 1→2→2→5→2→2→3. The length of a path or a route is the number of links in it. For the examples given the lengths are respectively, 2, 3, 4 and 6. A path is closed if the first and the last nodes are the same. A chain is a path in which no link or node repeats itself and a closed chain is termed a circuit or cycle. (These are definitions from Wilson, but different authors have widely different terminology.)

(ii) The Cyclomatic number, μ, and the α index

An independent cycle is one such that it contains no smaller cycle. The number of independent cycles is called the cyclomatic number, μ, or the first betti number, and is defined for a graph G as :

$$\mu = q - n + p$$

where q and n are number of links and nodes in the graph and p = 1 if the graph is connected, or equals the number of components in the graph if it is not. (Connectivity and components are discussed below.)

If we consider a tree, Fig. 23, with n nodes and n - 1 links it is clear that the addition of any new link between two nodes not currently

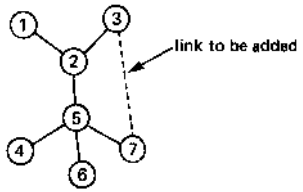


Fig. 23 The effect on μ of adding a link to a graph

linked directly, will produce a cycle, for example between nodes 3 and 7. For the tree,

$\mu = 6 - 7 + 1 = 0$, ... no circuits,

with the additional link,

$\mu = 7 - 7 + 1 = 1$ .... as observed.

The addition of a new link either splits the existing cycle into two if it has more than three nodes in it, e.g. linking 2 and 7, or creates a new cycle outside it. In either case a new cycle is created. Consequently μ is a crude index of connectivity in a graph. While of importance to graph theory as a mathematical study it is not especially useful in geographic work since it is unrelated to the size of the graph. It has been used or mentioned by Garrison & Marble, (1965), Warntz & Waters, (1975),

Symanski & Webber, (1974), Haynes, (1975). Modifications of the index are given by Kansky (1963) in which it is related to the maximum number of cycles in planar and non-planar simple graphs respectively.

simple graphs

(planar) $\alpha = \dfrac{\mu}{2n - 5}$

(non-planar) $\alpha = \dfrac{\mu}{N - n - 1}$

q, n, N as defined previously.

In this form the index is a little more discriminating although it has not attracted a wide practical use. Examples of its use can be found in the pavers just mentioned.

The gamma index γ is similar to the α index, indeed Garrison and Marble (1965) found a correlation coefficient of +0.998 between them for a sample of 22 empirical graphs. This would indicate one can serve for the other. For completeness we define them :

Simple graphs

(planar) $\gamma = \dfrac{q}{3n - 6}$

(non-planar) $\gamma = \dfrac{q}{N}$

where N = (1/2) (n2 - n), and q and n are as defined above. The gamma index is seen to be a simple ratio of the observed number of links to the total possible in either a planar or a non-planar graph.

(iii) Connectivity and distance

The connectivity of a graph is easy to determine at a glance for a small graph drawn on paper. It may be defined in the following way. A graph is strongly connected or connected if there exists a path between every pair of nodes in the graph. For a graph the symmetry ensures that if node i is connected to node j then j is connected to i. For digraphs this need not be the case. When it is true the digraph is strongly connected. The question of whether any particular pair of nodes is connected or not can be settled by calculating the powers of the adjacency matrix of the graph. In a simple example connectivity can be determined at a glance. For a complex empirical graph, perhaps a graph of eighty rural markets in which connections are determined by whether two markets are visited by a householder in the sample period, such an approach is impossible, (L.J. Wood, 1975), see below VI, Fig. 29.

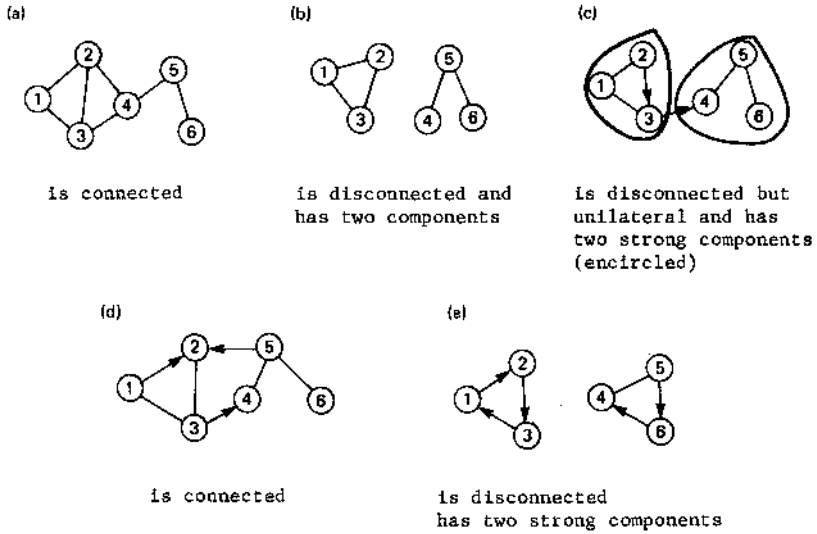


Fig. 24 Types of connectivity in graphs and digraphs

If a graph or digraph has n nodes then the maximum power of its adjacency matrix we need to inspect is the $n - 1^{th}$. This is because by reference to a line graph with n nodes we see that the two end nodes cannot be more than n - 1 lines apart. In practice we will not usually have to power as far as that. At the same time as the powering process we can incidentally determine the distance between every pair of nodes. The distance (geodesic, shortest path) between any pair of nodes is the length of the shortest path between them. Suppose two nodes i and j are separated by a distance k then by the interpretation of matrix powers the ij^{th} entry of $\underline{A}^{k-1}$ must be zero and similarly for smaller powers. The entry ij becomes positive for the first time in $\underline{A}^k$ if i and j are k links apart. By keeping track of when each ij position first becomes positive in some power of A a distance matrix D can be compiled. The computations will be illustrated for digraph Fig. 24 (c).

$$\underline{A} = \begin{pmatrix} 0&1&1&0&0&0\\ 1&0&1&0&0&0\\ 1&0&0&1&0&0\\ 0&0&0&0&1&0\\ 0&0&0&1&0&1\\ 0&0&0&0&1&0 \end{pmatrix}, \underline{A}^2 = \begin{pmatrix} 2&0&1&1&0&0\\ 1&1&1&1&0&0\\ 0&1&1&0&1&0\\ 0&0&0&1&0&1\\ 0&0&0&0&2&0\\ 0&0&0&1&0&1 \end{pmatrix}, \underline{A}^3 = \begin{pmatrix} 1&2&2&1&1&0\\ 2&1&2&1&1&0\\ 2&0&1&2&0&1\\ 0&0&0&0&2&0\\ 0&0&0&2&0&2\\ 0&0&0&0&2&0 \end{pmatrix}, \underline{A}^4 = \begin{pmatrix} 2&1&3&3&1&1\\ 3&2&3&3&1&1\\ 1&2&2&1&3&0\\ 0&0&0&2&0&2\\ 0&0&0&0&4&0\\ 0&0&0&2&0&2 \end{pmatrix}$$

$$\underline{D}_{(1)} = \begin{pmatrix} 0&1&1&x&x&x\\ 1&0&1&x&x&x\\ 1&x&0&1&x&x\\ x&x&x&0&1&x\\ x&x&x&1&0&1\\ x&x&x&x&1&0 \end{pmatrix}, \underline{D}_{(2)} = \begin{pmatrix} 0&1&1&2&x&x\\ 1&0&1&2&x&x\\ 1&2&0&1&2&x\\ x&x&x&0&1&2\\ x&x&x&1&0&1\\ x&x&x&2&1&0 \end{pmatrix}, \underline{D}_{(3)} = \begin{pmatrix} 0&1&1&2&3&0\\ 1&0&1&2&3&0\\ 1&2&0&1&2&3\\ x&x&x&0&1&2\\ x&x&x&1&0&1\\ x&x&x&2&1&0 \end{pmatrix}, \underline{D}_{(4)} = \begin{pmatrix} 0&1&1&2&3&4\\ 1&0&1&2&3&4\\ 1&2&0&1&2&3\\ x&x&x&0&1&2\\ x&x&x&1&0&1\\ x&x&x&2&1&0 \end{pmatrix}$$

In $\underline{D}$, 0 = diagonal positions which remain zero
x = undetermined distances
$\underline{D}$ is updated after each power.

From the diagram we can see that nodes 1 and 6 are four links apart and they are the two nodes farthest apart, hence the powering can stop at $\underline{A}^4$. However, we will not know this fact in a large empirical graph so the previous rule enjoins us to continue to n - 1 = 6 - 1 = 5 so we calculate $\underline{A}^5$:

$$\underline{A}^5 = \begin{pmatrix} 4&4&5&4&4&1\\ 5&3&5&4&4&1\\ 4&1&3&5&1&3\\ 0&0&0&0&4&0\\ 0&0&0&4&0&4\\ 0&0&0&0&4&0 \end{pmatrix} \quad \text{with } \underline{D}_{(5)} = \begin{pmatrix} 0&1&1&2&3&4\\ 1&0&1&2&3&4\\ 1&2&0&1&2&3\\ x&x&x&0&1&2\\ x&x&x&1&0&1\\ x&x&x&2&1&0 \end{pmatrix}$$

While the entries in $\underline{A}^5$ are larger than those in $\underline{A}^4$ no new positions have become positive and this is confirmed by $\underline{D}_{(5)}$ which is identical to $\underline{D}_{(4)}$. When some $D_{(n)} = D_{(n+1)}$ the powering can stop. The x locations indicate that no distances are defined between these node paths, i.e. they are not connected. In the case of a graph the distance matrix is symmetric so that for Fig. 24(a) the distance or short path matrix is:

$$\underline{D} = \begin{pmatrix} 0&1&1&2&3&4\\ 1&0&1&1&2&3\\ 1&1&0&1&2&3\\ 2&1&1&0&1&2\\ 3&2&2&1&0&1\\ 4&3&3&2&1&0 \end{pmatrix} \quad \text{but for Fig. 24(b) is} \quad \begin{pmatrix} 0&1&1&x&x&x\\ 1&0&1&x&x&x\\ 1&1&0&x&x&x\\ x&x&x&0&1&2\\ x&x&x&1&0&1\\ x&x&x&2&1&0 \end{pmatrix}$$

An equivalent way to obtain the distance matrix D which is especially useful when A has cost or distance values on the links is to use 'modified arithmetic' in the matrix power computations. The modified operations are:

for a x b	substitute	a+ b
for a + b	substitute	min (a,b),

and the latter operation may apply to a term with several elements:

$$a + b + c + \quad + h = \min (a, b, c, \ldots\ldots\ldots , h).$$

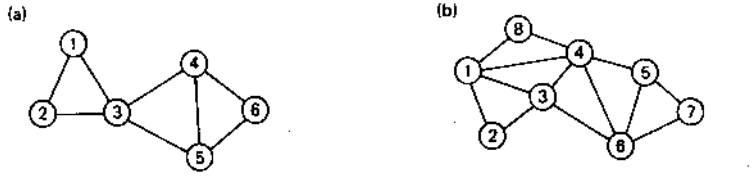
In these operations of course a, b, c, ... etc. are the elements of the matrix in hand. In addition all zero elements in the initial matrix are set equal to infinity, ∞.

A number of indices can be derived from the distance matrix. Taking each ith row the maximum number in each row, indicating the distance to the most distant node from i, is the Konig number or Associated Number. Taking the set of all Konig numbers for a graph the largest is the diameter of the graph, agreeing with the previous definition and the smallest Konig number is the radius of the graph. The row sums of D, specifically D.1 have been used as measures of access since the smallest sum indicates the place closest to all other places in simple topological distance terms. Dividing each row sum by (n-1) yields the average distance from that node to all others. Remembering that (D.1) is a column vector, then summing all these elements, or in matrix notation 1' (D.1) where 1' is a row vector of ones, gives a measure that has been termed the dispersion of the graph. Computing these various indices for the distance matrix of Fig. 24(a) we get the following values:

node	Konig	nodal row	access sums	mean distance
1	4	11		2.2
2	3	8		1.6
3	3	8		1.8
4	2	7		1.4
5	3	9		1.8
6	4	13		2.6
		56		

diameter = 4 = maximum Konig number dispersion = 56
radius = 2 = minimum Konig number

The following authors illustrate the use of the short path matrix and a variety of the derived indices; Garrison (1960), Pitts (1965) and Carter (1969). Reed (1970) investigates changes in the average node distances when certain nodes and links are deleted from the network. Kissling (1969) uses the measure of nodal access applied to a least cost-distance matrix. James, Cliff, Haggett and Ord (1970) have used the frequency distribution of distances in the short path matrix to construct a very sensitive index (S-I) which is able to discriminate between broad structural classes of graphs similar to those identified in Section II above. The same approach may be applied to the frequency distribution of the short path distances for each node to discriminate the structural positions of nodes within a network.



(a) {3} is a cutpoint, (b) {4, 3} is a cutpoint set but so are the sets: {1,3} {5,6} {1, 4} {4, 6} however {3, 6} is not although {3, 4, 6} is.

Fig. 25 Graphs to illustrate cutpoints and bridges

(iv) Components of a graph, articulation points and bridges

When the distance matrix indicates that certain nodes are not connected then graph has two or more strong components. A strong component is a sub-graph which is strongly connected. Fig. 24 (b) (c) and (e) all have two strong components. Fig. 24 (a) and (d) have one strong component and the graphs are said to be strongly connected. Werner (1971) compared the number of components in a developing transport system to the numbers expected in a comparable series of random graphs. Connectivity analysis may be combined with two additional notions in the analysis of graph structure. An articulation node or cutpoint is a point whose removal together with the incident links will result in a graph becoming disconnected. A cutpoint set is a collection of points that give the same result. Fig. 25 illustrates both concepts. A bridge is a single link whose removal disconnects the graph, (the nodes at either end are not removed), and a bridge set is a collection of links with the same property. There are no bridges in the examples above but in (a) bridge sets are {1-2, 2-3} and {3-4, 3-5}, and in (b) {4-5, 4-6, 3-6} is a bridge set. Removing a single point is equivalent to removing the bridge set consisting of all links incident to that point.

In the empirical case one might want to either identify cutpoint or bridge sets or test whether specific elements, or sets of elements have these properties. In either case the selected set is deleted and then the distance matrix is computed to identify the components of the graph. In this way critical elements in the system can be identified, or the criticality of elements can be tested. Reed (1970) used this approach by removing critical nodes in the Indian National Airline system and examining the effect on the structure. A systematic method for investigating an empirical

would be to remove nodes in descending order of nodality, or sets of nodes based on the same principle. Wood (1975) provides an example of just such a procedure for investigating the structure of a market visiting network for both nodes and links, (see Section VI, Fig. 29).

Bridge sets arise in an apparently different connection in dealing with flows in a network. If values attached to the directed links of a digraph are used to describe the maximum capacity of links with respect to flows going from a node i to a node j then the Ford-Fulkerson max-flow min-cut theorem (1956) states that the maximal possible flow from i to j through the network is equal to the total capacity in that bridge set, out of all possible bridge sets separating i and j, whose total capacity is the smallest. It is a characteristic of graph theory that an apparently simple idea leads on to very powerful results. A lucid treatment of this topic may be found in Wilson (1972). The Ford-Fulkerson algorithm has been applied by Gauthier (1968), Kissling (1969) and Muraco (1972) following on from simpler accessibility analyses.

V Extensions

Given the basic elements of graph theory reviewed so far it is possible to outline a number of directions which yield applications in broadly geographical fields. The first extension involves nodal analysis in which a particular type of accessibility matrix from Section II is used as the basis for an investigation of dominant telephone flows amongst the towns of an urban hierarchy. The relationship that this methodology bears to Mean First Passage Times in Markov Chain Analysis (see CATMOG No. 1) is discussed in Tinkler (1972(a), 1976). A second direction concerns the existence of certain routes and optimal routes within networks - the issue of traversability, and the third concerns coverings and colourings of graphs in graphs - notions fundamental to optimising location and allocation problems for graphs. For the latter two directions Scott (1971) provides a detailed review.

(i) Nodal Analysis

Some mention should be made of Nodal Analysis since it has become widely referenced as a graph theoretical technique, (Nystuen and Dacey, 1961; Clayton, 1974; Soja, 1968, p. 46, 47; Taafe and Gauthier, 1973, p. 149). The method may be applied to any interaction matrix to identify nodal flows. A flow is nodal if a given centre's largest flow is to a centre higher in the city hierarchy than itself. Usually a city's hierarchical position is measured by the total flows incoming to the city - the appropriate column sum in the interaction matrix. A city is a terminal node in the nodal flow system if its largest flow is to a centre lower in the hierarchy than itself. The system is constructed from the interaction matrix by identifying the largest value in each row. If the column sum for that value is larger than the column sum indexed by the row number then a 1 is entered in the nodal graph matrix, and a zero otherwise. Using a matrix of telephone flows in Kenya, 1967, between exchange regions, the analysis is performed below. (Nystuen and Dacey (1961) also give a worked example).

TELEPHONE TRUNK CENSUS JULY 1967 (5 Days Total 30 hrs.)

KENYA

	Eldoret	Kericho	Kisumu	Kitale	Mombasa	Nairobi	Naivasha	Nakuru	Nanyuki	Nyeri	Nodal ?	If yes, to where?
Eldoret	0	14	79	189	1	172	10	166	-	6	No	-
Kericho	33	0	222	8	8	396	14	270	1	-	Yes	Nairobi
Kisumu	121	148	0	32	35	757	3	188	2	5	Yes	Nairobi
Kitale	265	3	56	0	7	233	-	88	-	5	Yes	Eldoret
Mombasa	-	3	-	2	0	1315	1	33	-	2	Yes	Nairobi
Nairobi	284	230	633	110	1747	0	386	1153	334	577	No	-
Naivasha	6	7	1	2	2	420	0	272	-	1	Yes	Nairobi
Nakuru	109	190	191	69	30	1800	209	-	21	16	Yes	Nairobi
Nanyuki	4	-	2	1	2	410	1	76	0	221	Yes	Nairobi
Nyeri	4	1	2	1	10	805	1	37	183	0	Yes	Nairobi
Column Totals	826	596	1186	414	1842	6308	625	2278	541	833		

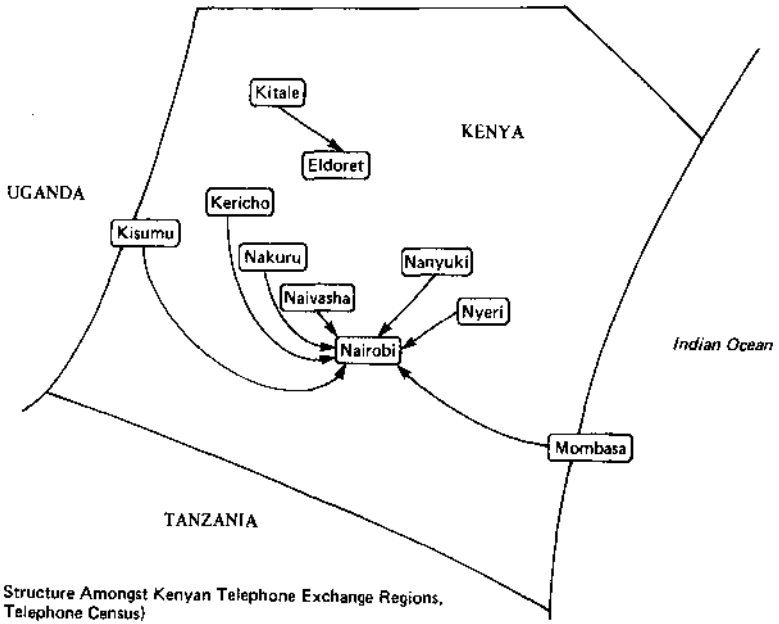
(Basis for the hierarchy of cities.)

The highest entry in each row, the potential nodal flow, is underlined. Note that Nairobi to Mombosa is not a nodal flow, nor is Eldoret to Kitale.

We get a nodal flow matrix :

$$\begin{pmatrix} 0&0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&1&0&0&0&0 \\ 0&0&0&0&0&1&0&0&0&0 \\ 1&0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&1&0&0&0&0 \\ & & & & & 0 & & & & \\ & & & & & 1 & & & & \\ & & & 0 & & 1 & 0 & & & \\ & & & & & 1 & & & & \\ & & & & & 1 & & & & \end{pmatrix}$$

which is the adjacency matrix of the digraph, Fig. 26.



Nodal Structure Amongst Kenyan Telephone Exchange Regions, (1967 Telephone Census)

Fig. 26

The method was originally applied by Nystuen and Dacey to the sum of a power series expansion of a weighted adjacency matrix, as discussed on p. 31. It may, however, be applied to a raw interaction matrix, as in the example just given. Clearly, in the example, the polarising effect of Nairobi is evident, but Eldoret at the northern end of the former White Highlands is able to maintain a local nodal position.

(ii) Traversability

An Eulerian path is a closed path (i.e. starting and ending at the same node) which includes every link in the graph just once. Euler's eponymous solution to this problem is credited with creating graph theory. The necessary and sufficient condition is that the degree or nodality of every node in the graph be even.

An apparently similar problem, but one that offers many unsolved problems in graph theory, is the Hamiltonian path problem in which it is

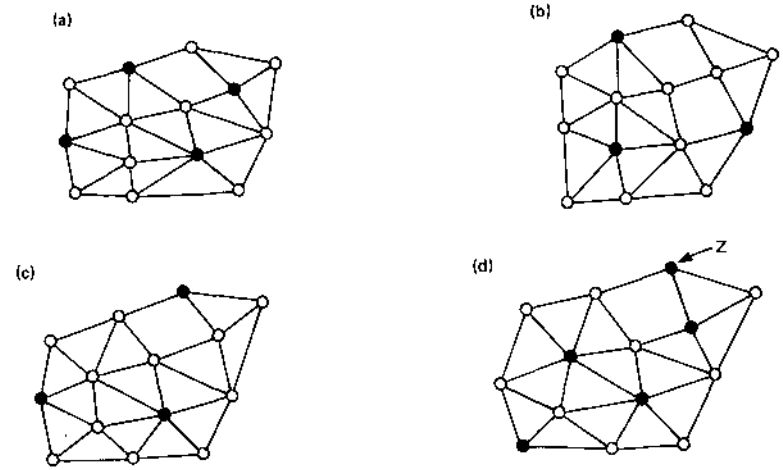


Fig. 27 Graph to illustrate types of cover

required to traverse the graph with a closed path which visits each node exactly once. There is no general solution to this problem, the simplest result being that if the graph has n nodes and every node has degree equal or larger than (1/2)n then a Hamiltonian path exists. A related problem that is commonly encountered in scheduling and allocation problems is to select, out of a large number of Hamiltonian paths, where each link in the system has a fixed cost value attached, that which has the lowest total cost associated with it. This is the travelling salesman problem and it is usually posed for complete, valued-graphs in which there are (n-1): possible Hamiltonian paths. An extension of the problem with an application to periodic market rings is given by Harvey, Hocking and Brown, (1974). Good (1975) has attempted to delineate optimal paths for travelling traders in Buganda using ideas based on travelling salesman algorithms.

(iii) Covering a graph

A point cover of a graph G is a set of nodes such that all other nodes in the graph are adjacent to at least one of the nodes in the set. Fig. 27 (a), (b), (c) illustrates three different independent point covers. The covers are termed independent since no pair of nodes in the cover are themselves adjacent in the graph. Figure 27 (d) shows a point cover which is not independent. Often but not always, dropping the independence requirement reduces the number of points in the cover; in this example it is not

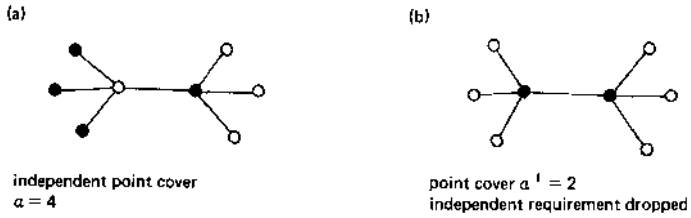


Fig. 28

so but the next one shows the type of structure in which this can occur, Fig. 28 (a), (b).

In practice it is often required to find a minimum point cover - in allocation problems the need is to use as few facilities as possible and to locate them as close as possible to all remaining nodes. We need therefore a minimum point cover. A set X is minimal with property P if and only if X has property P and no proper subset of X has property P. Hence Fig. 27 (a), (b) and (c) are all minimal whereas Figure 27 (d) is not minimal since node z can be deleted without affecting the property P, i.e. the covering of the remaining nodes. A minimum set with property P is a set with the smallest possible number of elements in it; i.e. the minimum or minima of all the minimal sets. Hence Figure 27 (b) and (c) are both different minimum sets whereas Figure 27 (a) is simply minimal.

We have spoken of point covers in terms of the basic adjacency matrix A. However, if we calculate $_1A^n$ the matrix contains information on adjacencies up to and including the n^{th} step. We can then take this to be the matrix of the graph we are interested in and construct point covers in which points of the graph are guaranteed to be within n-steps of a node in the point cover. (The probable multiplicity of many of the n-step adjacencies is, of course, irrelevant.)

(iv) An algorithm for minimal independent point covers

Most algorithms to obtain minimum independent point covers are extremely laborious. It is easy to find a minimal independent point cover but hard to find a minimum one, i.e. that which is smallest of all the minimal covers. The procedure is to use a reasonably 'good' method to find a minimal cover and to hope that it is close to being a minimum cover.

Step 1 - Compute A for the graph, this is the current A matrix.
Step 2 - Sum the rows of the current A matrix, enter points with row sums of zero into the cover, provided they have not already been entered. Deleted rows do not count as zero row sums.

Step 3 - Choose the point with the largest row sum, (in case of ties make an arbitrary choice). Enter into the cover.
Step 4 - Delete from A the point chosen, i.e. set that row and column to crosses, and alsodelete from A the rows and columns of all points to which it was connected, similarly setting these to crosses.
Step 5 - If A is now a null matrix enter any currently isolated points (zero row sums), and the point cover is complete. Otherwise go to Step 2 and use the current A matrix, i.e. with the various rows and columns deleted.

As just mentioned, the procedure works just as well on A^n and the procedures are, of course, quite independent of planarity. This particular algorithm is not foolproof and pathological graphs can be devised that have point covers that the algorithm does not find, much smaller than those it does find.

(v) Colouring a graph

A proper colouring of a graph is an assignment of distinct labels - colours - to the points of a graph with the property that no two points adjacent in the graph have the same label (colour). The minimum number of colours that can be used to achieve this requirement is termed the chromatic number, $\chi(G)$, of the graph G. A straightforward algorithm to do this colouring with a usually good approximation to χ (G) is based on the previous algorithm. Use the previous algorithm to obtain a minimal point cover. These points are labelled colour 1. Return now to the original A for the graph and delete simply those rows and columns corresponding to points in the cover obtained. Using this reduced matrix, A', re-use the algorithm to obtain a point cover amongst the remaining nodes. Label these nodes colour 2. Reduce the A' matrix to A" by removing the rows and columns for points coloured with 2 and continue in like fashion until no nodes remain to be coloured. The number of colours used will be a good estimate of the chromatic number χ (G), and will usually only differ at most by +1 from the actual value, (Saaty, 1972). An explicit discussion of colouring in the geographical literature is that by Tinkler (1973), where rural periodic markets and market areas surrounding them are conceived of as a planar graph. Two rural markets are adjacent when they share a common boundary. The four-day market week typical of West Africa is interpreted as a natural response to the four-colour conjecture for planar graphs. The conjecture states that any planar graph can be coloured with only four colours. It was long ago proved that any planar graph is five-colourable, and has been proved for planar graphs with up to 51 points (Stromquist, 1975). A computer proof for an arbitrary number of points has recently been given by Appel and Haken (1976). For graphs with certain structural properties χ is known. For example, the null graph on n-nodes is trivially 1-colourable, trees and graphs with no cycles of odd length are two-colourable. A graph with an odd cycle necessitates three colours and one containing as a subgraph a wheel on n = even points necessitates four colours. (See Section VIII, p.54, for an application of the algorithm to Figure 29). The chromatic number, χ, is defined with reference to 1-step adjacencies but it might also be defined for n-step adjacencies within the same graph by using A^n matrices; for example, to locate facilities which have n-step avoidance restraints amongst themselves.

VI AN APPLICATION : KISII DISTRICT, KENYA.

Wood (1975) has used graph theoretical concepts to examine the functional structure and integration of rural markets in Kisii District, Kenya. In total 83 rural markets were identified, but two isolated ones, and one with only one connection, were excluded from the analysis. For a period of 14 days, market visits by the mothers in 103 households were recorded. Two markets were deemed to be connected if they were visited in the sample period by a member of the household. In this way the pattern of connections between markets was established and gives rise to Fig. 29, which simply records connections. In fact many links represent multiple connections. In either case the system is obviously highly polarised. An analysis may be based on the binomial method in Section V, p. 34, where n=80 and q=517. Mean nodality, Δ, is therefore 12.925 and the variance equation gives :

$$\sigma^2 = \frac{2q\,(n-2)}{n^2} = \frac{1034(78)}{6400} = 12.60$$

giving a standard deviation of 3.55. Table 3 shows the observed frequency distribution of nodality and obviously it contains several nodes lying well outside a three standard deviations limit. Kisii, for example, has a nodality in terms of visits of 78 and is directly connected to 38 of the 79 other places. The behaviour patterns displayed are adjudged very significantly polarised.

Table III

Frequency Distribution of Nodality in Kisii District Market Visiting Network.

	Nodality Class								
	0-4	5-9	10-14	15-19	20-24	25-29	30-34	35-39	40+
Frequency	15	23	13	15	4	2	3	2	3

Maximum = 78
Minimum = 3
Mean = 12.93

A short path analysis forms the basis for a connectivity analysis. Table 4 tabulates the Konig number frequency distribution and for the network of both domestic marketers and traders* the diameter of the system

*Wood was able to determine from the questionnaires that some mothers were trading items part-time for profit - traders - as distinct from purely domestic buying and selling - marketers. Full time itinerant traders were not included in the analysis, see Wood (1975), footnote 11.

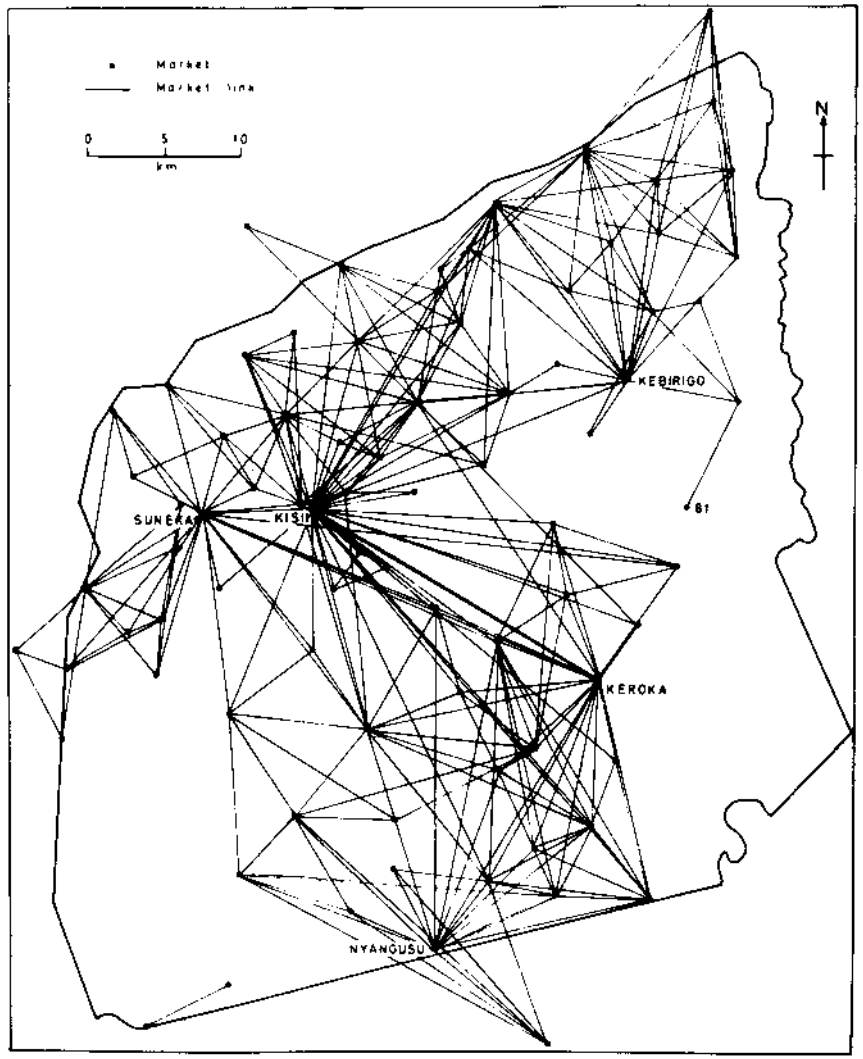


Fig. 29 Inter-market connections in Kisii District, Kenya (from Wood, 1975, with permission)

is only 6, (despite the 80 markets), which indicates a high degree of integration. As expected, Kisii is closest to all points; only three links at most from any place. When the traders are removed from the network a slight shift occurs : the diameter moves to 7 and Kisii has a Konig number of 4. This indicates that the traders do knit the network together more closely, but apart from 3 markets which become isolated when the trader links are removed, their removal from the system does not split it into major regional components.

Table 4 - The distribution of Konig numbers (from Wood, 1975, with permission).

Konig number	marketer and trader network	marketer network alone
3	1	0
4	41	13
5	32	33
6	6	25
7	0	6
Total	80	77

An analysis of the subsystems comprising the network was made using the procedures described in Section VI, p. 37. In the absence of useful official data to establish a hierarchy of markets in terms of market use, recourse was had to the hierarchy established by the total numbers of visits received as recorded by the survey. Kisii received 78 visits, 8 markets received between 28 and 56 visits and 71 markets had less than 25 visits. Wood tested the structural integrity of the system as successively more nodes were removed. First Kisii was removed, then Kisii and the next largest, and so on for the first nine nodes. The significant results are recorded in Table 5 and in Figure 30.

Table 5 - The decomposition of the Kisii market system

Extraction number	Nodes extracted (cumulative)	Disintegration produced by extraction within:[a] marketer and trader network	marketer network
1	21	None	None
2	+57	None	Isolates small central component (Fig. 30B)
3	+35	None	Market 51 isolated
4	+38	Produces two major components (Fig. 30A); market 50 isolated	Produces two major components (Fig. 30C) market 50 isolated
5	+23	None	Splits northern component into two
6	+15	None	None
7	+10	None	Market 2 isolated
8	+64	None	Market 80 isolated (Fig.30D)
9	+11	None	None

a The disintegration of the networks is cumulative, i.e. the isolation of market 51 in extraction 3 for the marketer network means that this market is isolated in all subsequent extractions. (From Wood 1975, with permission)

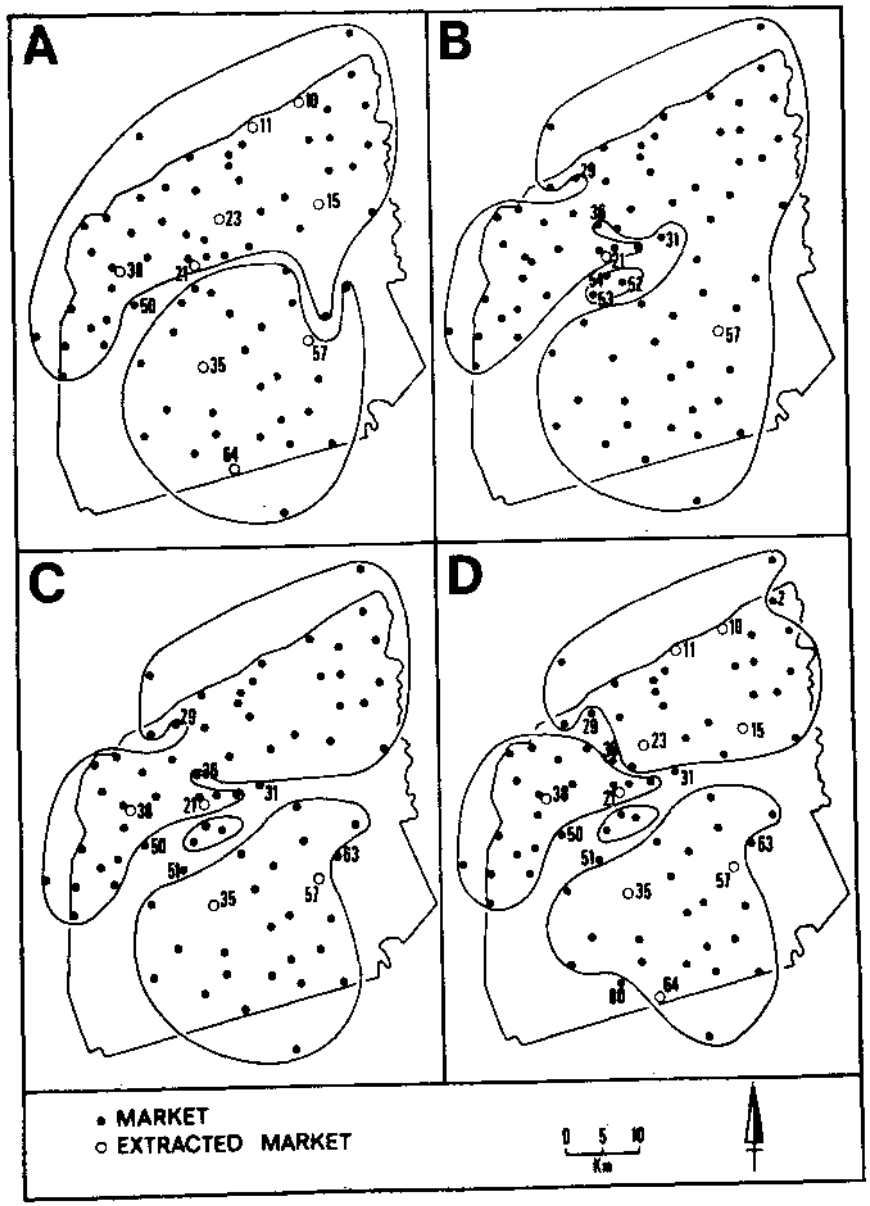


Fig. 30 Disintegration of the market visiting structure with the removal of various market nodes (from Wood, 1975, with permission)

The analysis demonstrated the highly connected nature of the marketer and trader network and its independence from an integration caused just by part-time traders or one or two critically placed, frequently visited, markets. Interestingly, while emphasising the tightly woven structure caused by individual trading, the one major regional decomposition that was revealed, see Fig. 30, can be interpreted as a split between the subsistence agriculture practiced in the south and the more intensive cash cropping in the north. Four major markets lying along zonal boundaries appear to have an integrating effect. A similar analysis for the marketer network alone, (i.e. with trader links removed), had similar results, Table 5, column 4, although in this case the northern (cash crop) component splits in two when the first five nodes were extracted. The integrating role of the trader is therefore seen to be relatively limited. An application of the graph colouring algorithm, Section V, yields a chromatic number of 8. (This analysis is not reported in Wood (1975), but was made with the market matrix kindly made available by Wood). Therefore, despite its complexity, 8 colours (days) would be sufficient to ensure that no householder, on visiting one market in his market set would miss a market on the same day in the remaining markets in his set. Since the 8th. colour is used to resolve only one conflict, and the 7th. to resolve only four conflicts it may be concluded that the system can work with almost total efficiency on a six day week. (Kisii is a stronghold of Seventh Day Adventists and Saturday is hardly ever used as a market day, (Wood, 1975)).

Wood's analysis reveals the utility of a straightforward graph theoretical approach and yields useful answers to problems of spatial patterning and interactions which would not be achieved easily by other methods.

VII PROBLEMS AND LIMITATIONS

The possible limitations of graph theory were pointed out by some of the earliest users - Garrison (1960), Nystuen and Dacey (1961) and Garrison and Marble (1965). The latter concluded that "the present investigations, while illuminating certain points, have failed to provide an entirely satisfactory set of results. Part of the problem appears to lie in the line of attack which was chosen; the graph theoretical model of network structure has proved to be strong in some areas, but disappointingly weak in others ....". Burton (1961) was more hopeful when he stated: "It is conceivable that a body of useful theory could be built up around the application of graph theory to geographical problems". If the field has not moved as fast as Burton anticipated this may be because a rather limited interpretation was put upon graph theory and what it had to offer Geography. This restricted view within Geography may have stemmed from an over-reliance on the summaries provided in Garrison (1960) and Kansky (1963); both angled at a transportation audience. The results obtained by Garrison and Marble (1965) undoubtedly derive from the elementary form of analysis that they undertook. No such limitation is evident in the work of Gauthier (1968) who applied the Ford-Fulkerson theorem to network flows in Argentina, yet this same theorem may be proved directly from the basic idea of a bridge set outlined in Section V. As the dates of many of the references indicate, the more fruitful and imaginative applications have come in graph theory's second decade in geography.

Of course there are limitations. Graph theory is a theoretical system for finite state problems and it is not well adapted to dealing with problems in the continuous plane. A problem that arises in dealing with combinatorial structures in general is the complexity of the algorithms needed to find the

optimal solutions to most of the meaningful problems. A tradeoff must be made between the speed of the algorithm and the quality of the solution and a good recent review of algorithms arising from graph theoretical problems is Christofides (1975). Nevertheless, it is evident that graph theory and related combinatorial theory in transversals, matchings and matriods has an enormous potential for finite state problems, in geography as in other subjects. As the application discussed in Section VI illustrates a good deal of structural clarity is achieved by fairly simple methods.

BIBLIOGRAPHY

A. Graph Theory Texts

Berge, C., 1962, *The theory of graphs and its applications.* John Wiley.

Busacker, R.G. and Saaty, T., 1965, *Finite graphs and networks: An introduction with applications.* McGraw-Hill.

Christofides, N., 1975, *Graph theory: an algorithmic approach.* Academic Press.

Flament, C., 1963, *Applications of graph theory to group structure.* Prentice-Hall.

Harary, F., 1969, *Graph theory.* Addison-Wesley.

Harary, F., Norman, R.Z., and Cartwright, D., 1965, *Structural models: an introduction to the theory of directed graphs.* John Wiley.

Ore, Oystein, 1963, *Graphs and their uses.* Random House.

Wilson, R.J., 1972, *Introduction to graph theory.* Oliver and Boyd.

B. Texts or Papers in Geography with considerable instructional component in graph theory.

Alao, N., 1973, Some aspects of network theory. *Geographica Polonica.* 25, pp 107-137.

Campbell, J., 1971, Growth pole theory, digraph analysis and inter-industry relationships. *Tidjschrift V. Econ. en Soc. Geografie,* 63, pp 79-87.

Haggett, P., and Chorley, R.J., 1969, *Network analysis in geography.* Edward Arnold.

Garrison, W., 1960, Connectivity of the Interstate highway system. *Papers, Reg. Sci. Ass.,* 6, pp 121-37.

Garrison, W., and Marble, D.F., 1965, A prolegomenon to the forecasting of transportation development. Technical Report 65-35. Prepared for the U.S. Army. (Reprinted as 'Graph theoretic concepts', in: *Transportation geography: comments and readings,* ed M. Eliot Hurst, (McGraw-Hill) 1974, p 58).

Hay, A., 1973, *Transport for the space economy.* Macmillan, see Chapter 4.

Kansky, K.J., 1963, The structure of transport networks. university *of Chicago, Dept. of Geography, Res. Paper,* 84.

Nystuen, J.D., and Dacey, M.F., 1961, A graph theory interpretation of nodal regions. *Papers Reg. Sci. Ass.,* 7, pp 29-42.

Taafe, E.J., and Gauthier, H.L., 1973, *Geography of transportation.* Prentice-Hall, especially chapters 4, 5 and 6.

C. Examples of 'rash theoretic applications in geography and further theoretics developments.

Alao, N., 1970, A note on the solution matrix of a network. *Geographical Analysis,* 2, pp 83-88.

Appel, K. and Haken, W., 1976, Every planar map is four colorable. *Bulletin of the American Mathematical Society,* 82(5), pp 711-2.

Burghart, A., 1969, The origin and development of the road network of the Niagara Peninsula, Ontario, 1770-1851. *Annals, Assoc. Amer. Geogr.,* 59(3), pp 417-440.

Burton, Ian, 1963, The quantitative revolution and theoretical geography. *Canadian Geographer,* 7, pp 151-162.

Carter, F.W., 1969, An analysis of the medieval Serbian Oecumene: a theoretical approach. *Geografiska Annaler,* 51B, 1, p 39.

Clayton, C., 1974, Communication and spatial structure. *Tijdschrift voor Econ en Soc. Geografie,* 65(3), pp 271-7.

Collins, L., 1974, An introduction to Markov Chain Analysis. Catmog No. 1, *Geo Abstracts,* Norwich.

Cummings, L.P., Manly, B.J., Weinand, H.C., 1974, Measuring association in link-node problems. **Geoforum,** 13, pp 43-51.

Cummings, L.P., Manly, B.J., Weinand, H.C., 1974, Measuring association in link-node problems: a reply. *Geoforum,* 17, p 92.

Ford, L.R. Jr. & Fulkerson, D.R., 1956, Maximal flow through a network. *Canadian Journal of Mathematics,* 8, pp 395-404.

Gauthier, H.L., 1968, Least-cost flows in a capacitated network: a Brazilian example. in: *Geographic Studies of Urban Transportation and Network Analysis,* ed F. Horton, Northwestern University Studies in Geography, No. 16.

Golledge, R.G., 1968, Methods in Australian trade and transportation studies. in: *Studies in Australian Geography,* ed Dury, G.H., and Logan, M.I., (Heinemann) pp 303-327.

Good, C.M., 1975, Periodic markets and travelling traders in Uganda. *Geographical Review,* 65(1), pp 45-72.

Hay, A.M., 1974, Measuring association in link-node problems: a comment. *Geoforum,* 17, pp 90-92.

Haynes, V., 1975, in: Cirque morphometry. *Area,* 7(1), p 69.

Harvey, M.E., 1972, The identification of development regions in developing countries. *Economic Geography,* 48, pp 229-43.

Harvey, M.E., Hocking, R.T., and Brown, J.R., 1974, The chromatic travelling-salesman problem and its application to planning and structuring geographic space. *Geographical Analysis,* 6, pp 33-52.

Jackson, R.T., 1972, Economic Rationality amongst African rural traders. *Geografiska Annaler,* 54B, pp 85-94.

James, G.A., Cliff, A.D., Haggett, P., and Ord, J.K., 1970, Some discrete distributions with applications to regional transport networks. *Geografiska Annaler,* 52B, pp 14-21.

Kissling, C.C., 1969, Linkage importance in a regional highway network. *Canadian Geographer,* 13(2), pp 113-129.

Muraco, W.A., 1972, Intraurban accessibility. *Economic Geography,* 48(4), pp 388-405.

Pitts, F.R., 1965, A graph theoretic approach to historical geography, *Professional Geographer,* 17(5), pp 15-20.

Reed, W.E., 1970, Indirect connectivity and hierarchies of urban dominance. *Annals, Association of American Geographers, 60,* pp 770-85.

Saaty, T., 1972, Thirteen colorful variations on Guthries conjecture. *Amer. Math. Monthly,* 79(1), pp 2-43.

Scheidegger, A.E., 1967, Random graph patterns of drainage basins. *International Association for Scientific Hydrology,* Publication No. 76, pp 415-25.

Schmidt, C.G., 1975, Firm linkage structure and structural change - a graph theoretical analysis. *Economic Geography,* 51(1), pp 27-36.

Scott, A.J., 1971, *Combinatorial programming, spatial analysis and planning.* Methuen.

Smart, J.S., 1969, Topological properties of channel networks. *Bull. Geol. Soc. Am.,* 80, pp 1757-74.

Soja, E.W., 1968, *The Geography of Modernization in Kenya,* Syracuse Geographical Series, No.2

Stutz, F.P., 1973, Accessibility and the effect of scalar variations on the powered transportation connection matrix. *Geographical Analysis,* 5, pp 61-66.

Symanski, R., and Webber, M.J., 1974, Complex periodic market cycles. *Annals, Association of Amer. Geogr.,* 64(2), pp 208-213.

Stromquist, W., 1975, The four-color theorem for small maps. *J. of Combinatorial Theory,* Series B, 19, pp 256-268.

Tapiero, C.S. and Boots, B.N., 1974, Structural inference in transportation networks, *Environment and Planning* A, 6, pp 411-18.

Tinkler, K.J., 1972(a), Physical interpretation of eigen functions of dichotomous matrices. *Trans. Inst. Brit. Geog.,* 55, A-46.

Tinkler, K.J., 1972(b), Bounded planar networks: a theory of radial structures. *Geographical Analysis,* 4(1), pp 5-33.

Tinkler, K.J., 1973, The topology of rural periodic market systems. *Geografiska Annaler,* 55, Ser. B., pp 121-133.

Tinkler, K.J., 1974, On summing power series expansions of accessibility matrices by the inverse method. *Geographical Analysis,* 6, pp 175-178.

Tinkler, K.J., 1975, Access in several networks used successively. **Area,** 7(4), pp 296-299.

Tinkler, K.J., 1976, On functional regions and indirect flows. Geographical *Review*, 65(4), pp 476-492.

Tinkler, K.J., 1979, Graph Theory in *Progress in Human Geography*, 3, pp 85-116.

Warntz, W., and Waters, N., 1975, Network representation of critical elements of pressure surfaces. *Geographical Review*, 65(4), pp 476-492.

Werner, C., 1971, Formal study of the evolution of the railway network in Schleswig-Holstein. in: *Studies in Geography*, ed F.E. Ian Hamilton, (Northwestern University), 18.

Werner, C., and Smart, J.S., 1973, Some new methods of topologic classification of channel networks, *Geographical Analysis*, 5, pp 271-295.

Wood, L.J., 1975, The functional structure of a rural market system. *Geografiska Annaler*, 57B, pp 109-18.