Environment and Planning A, 1977, volume 9, pages 727-738

An evaluation of lattice solutions to the problem of corridor location

M F Goodchild
Department of Geography, The University of Western Ontario, London, Ontario, Canada
Received 3 September 1976, in revised form 3 May 1977

Abstract. Many route-location problems can be regarded as the minimisation of some accumulated cost, impact, or similar friction per unit length. Analogies can be found with paths of light and other least-time paths, and with geodesics. The problem is often solved by finding a path through a lattice of sample costs, with the use of modified shortest-path algorithms. Lattice paths do not converge to continuous-space paths. The differences are shown to depend on the set of permitted moves in the lattice, with the use of three cases. The continuous-space problem is solved for surfaces described by simple functions and for choropleth surfaces, and compared with lattice solutions. Three heuristic approaches for large problems are reviewed, with emphasis placed on regular spatial aggregation.

Introduction

The selection of rights of way for highways, transmission lines, and other utilities inflames public opinion more than almost any other locational issue. The need for a corridor is largely established at and beyond the points which it must connect, so that land along the route must be sacrificed to a distant rather than a common good. Furthermore the location of each segment of the route is not independent of other segments, which makes it difficult to accommodate local objection on a pragmatic basis.

The problem has been formalised (Hopkins, 1973; Turner and Miles, 1971; OECD, 1973) as one of finding that path between a given origin and a destination, that optimises an objective function based on construction, environmental, social and other costs, and possibly benefits. This does not solve the decisionmaking problem, of course, since it may be just as difficult to reach consensus on the criteria for route selection as on the route itself. But the ability to make this conceptual connection objectively is clearly useful to any party in the planning process.

It is convenient to regard the cost per unit length of construction as a surface over the study area. The optimum path will then tend to follow the valleys and passes of the surface in such a way that the integrated cost over the length of the path is minimum, although it is clear that the path of least total cost is not necessarily the one with the lowest maximum elevation. Algebraically, the path is defined by

$$\min \int_{\text{path}} z \, \mathrm{d}s \, ,$$

where z is the local cost per unit length of path, and $\mathrm{d}s$ is an increment along the path. Clearly the optimum path on a flat surface is a straight line.

Several problems are analogous to this one. Angel and Hyman (1976) review the problem of minimum-time trip paths in cities,

$$\min \int_{\text{path}} \frac{\mathrm{d}s}{v} \, ,$$

where v is the local velocity; whereas the path of light in a medium with local refractive index μ is defined by

$$\min \int_{\text{path}} \mu \frac{ds}{c} ,$$

where c is the velocity of light in a vacuum.

Analytic solutions of the problem are limited to a few simple cases, such as the radially symmetric z functions of Angel and Hyman (1976), or the binary problems treated in optics. More general problems must be solved by numerical integration. In cases where z or its equivalent is everywhere known, the construction due to Huygens (see, for example, Warntz, 1965) may be used, although it is difficult to adapt the method to a digital computer. But for a number of practical reasons the minimum-cost problem has been approximated by representing the surface as a rectangular lattice of sample values, and by finding the path as a series of moves between nodes in the lattice (Hopkins, 1973; Turner and Miles, 1971; OECD, 1973). First, there need be no restrictions on the forms of surfaces handled in this way. Second, the use of sample values is probably more consistent with methods of data collection than is a continuous surface. And third, sampling introduces a level of spatial resolution which can be adjusted to the needs of the study and the width of the proposed corridor.

Adopting this form of numerical integration allows the problem to be handled as a case of finding the shortest path through a network, by discrete dynamic programming. The links between nodes define the moves permitted in the lattice, which may be vertical, horizontal, or diagonal. Figure 1 shows a typical square array of costs, in reality measures of impact on agricultural land from a proposed highway, and the minimum-impact path between two points (Owens, 1975). In this case moves were permitted to eight nearest neighbours from each node.

This paper examines several aspects of the lattice technique. There are regularities, not present in the general shortest-path problem, which permit modifications to conventional algorithms, with consequent economies in storage and computing costs. The paper also examines the relationships between lattice solutions and paths on continuous surfaces, and the extent to which one approach provides an approximation to the other. The final section describes heuristic methods suitable for problems with large numbers of nodes.

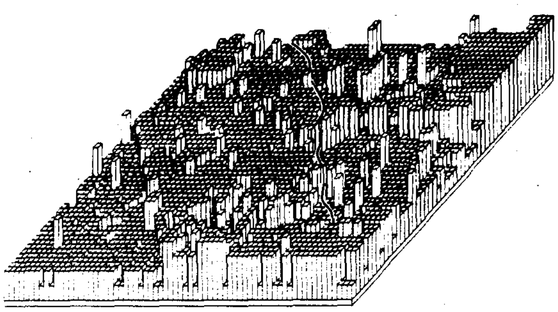

Figure 1. Example cost surface and optimum path.

Shortest-path algorithms

Dreyfus (1969) has reviewed methods for finding the shortest path between two given nodes and identifies the algorithm due to Dijkstra (1959) and Whiting and Hillier (1960) as the most efficient. The basic steps are as follows:

Step 1. Label the origin node 'reached' and set its minimum path cost to zero. Label all other nodes 'unreached'.

Step 2. Examine all links which directly connect 'reached' nodes to 'unreached' nodes. In each case add the link cost to the minimum-path cost at the 'reached' node to obtain a temporary minimum-path cost at the 'unreached' node.
Step 3. Find the minimum such temporary path cost in the lattice, label the corresponding node 'reached', and make the path cost permanent.
Step 4. If the new 'reached' node is not the destination, return to *step 2.* The 'reached' nodes and their optimum paths form a tree which builds outwards from the origin in obvious analogy to the Huygens construction or the spreading of light.

Two potential economies occur in the case of a lattice. First, since the links occur systematically, they need not be identified explicitly. Costs can be determined for nodes, and ascribed to links as the average of the respective node costs, or in some other systematic fashion. Second, it is likely that ties will occur between temporary path costs. In such cases *step 3* can be executed repeatedly without returning to *step 2.* With these modifications, the algorithm is as follows:
Step 1. Label the origin node 'reached' with path cost zero. Label all other nodes 'unreached'.
Step 2. Build an 'open node' table of all 'unreached' nodes directly connected to 'reached' nodes, and compute temporary path costs to them.
Step 3. Examine the 'open node' table for all nodes which tie for minimum temporary path cost.
Step 4. Label the tied nodes 'reached' and delete them from the 'open node' table.
Step 5. Add to the 'open node' table any 'unreached' nodes directly connected to the newly deleted nodes, with temporary path costs.
Step 6. If the destination is 'unreached' return to *step 3.*

Two independent values must be stored for each node. The first is the node cost; once the node has been 'reached' this can be replaced by the minimum path cost to the node. The second is the 'reached'/'unreached' label: the 'reached' state must be qualified by the direction of the last move, to allow the optimum path to the destination to be recovered at the end.

Table 1 shows some typical execution times for a CDC Cyber 73 system, with the use of a square array of nodes, and origin and destination at diagonally opposite corners. Times are directly proportional to the number of nodes, at 625 μs per node, for the rook's case of vertical and horizontal moves, and proportional to $n^{3/2}$ for the queen's case, which includes diagonal moves.

Table 1. Typical execution times (in seconds) for a CDC Cyber 73 system, for a square array of nodes and the origin and destination of diagonally opposite corners.

Nodes	Rook's case	Queen's case
400	0·25	1·05
900	0·59	3·51
2500	1·70	16·34
4900	3·26	49·91
10000	6·52	125·89

Characteristics of lattice paths
Each lattice node can be conveniently represented as a two-element vector $\boldsymbol{x}$ whose elements take integer values. The ith node on a path through the lattice is written as $\boldsymbol{x}_i$, and the next incremental move as $\boldsymbol{m}_i$, such that

$$\boldsymbol{x}_{i+1} = \boldsymbol{x}_i + \boldsymbol{m}_i \ .$$

In the rook's case $\boldsymbol{m}$ can take values of $(\pm 1, 0)$ or $(0, \pm 1)$. The queen's case adds the moves $(\pm 1, \pm 1)$, and the knight's case $(\pm 2, \pm 1)$ and $(\pm 1, \pm 2)$.

Let the cost at node $\boldsymbol{x}$ be $Z(\boldsymbol{x})$. Then the minimum-cost path problem can be written:

Given an origin node $\boldsymbol{x}_1$, a destination $\boldsymbol{x}_k$, and a set of costs $Z(\boldsymbol{x})$, find k and a set of nodes $\boldsymbol{x}_2 \ldots \boldsymbol{x}_{k-1}$ to minimise

$$\sum_{i=1}^{k-1} (\boldsymbol{m}_i \cdot \boldsymbol{m}_i)^{1/2} \tfrac{1}{2}[Z(\boldsymbol{x}_i)+Z(\boldsymbol{x}_{i+1})] ,$$

subject to $\boldsymbol{x}_{i+1} = \boldsymbol{x}_i + \boldsymbol{m}_i$, $i = 1, k-1$, and each $\boldsymbol{m}_i$ is one of the permitted set of moves. $(\boldsymbol{m}_i \cdot \boldsymbol{m}_i)^{1/2}$ is the length of the ith link in the path. It is operationally more convenient if the path cost is taken as a sum of node costs rather than link costs, or

$$\sum_{i=1}^{k} Z(\boldsymbol{x}_i)(\boldsymbol{m}_i \cdot \boldsymbol{m}_i)^{1/2} .$$

The relationship between a lattice path and the equivalent path on a continuous surface depends on the move set from which each $\boldsymbol{m}_i$ is selected. Consider a section of path of unit length in a direction ϕ on a flat, continuous surface of constant z, and suppose that the s permitted move directions in the lattice form an ordered set of directions $\theta_1 \ldots \theta_s$. In order to connect the same origin and destination a lattice path will make moves in at most two directions, θ_j and θ_{j+1}, such that $\theta_j \leqslant \phi < \theta_{j+1}$. By way of proof, consider a segment in some direction θ_k, and suppose that some θ_i exists such that $\phi \leqslant \theta_i < \theta_k$. Then a segment can be drawn at θ_i beginning at the same node as the segment at θ_k, and this new segment must rejoin the path before the destination. It follows that the segment at θ_k can be replaced by a shorter segment at θ_i and that the first path was not optimal.

It is easy to show that the distance moved in each lattice direction is given by

$$a_j = \frac{\sin(\theta_{j+1}-\phi)}{\sin(\theta_{j+1}-\theta_j)} ; \qquad a_{j+1} = \frac{\sin(\phi-\theta_j)}{\sin(\theta_{j+1}-\theta_j)} .$$

The length of the lattice path computed by summing lattice moves is thus longer by a factor

$$\epsilon(\phi) = \frac{\sin(\theta_{j+1}-\phi)+\sin(\phi-\theta_j)}{\sin(\theta_{j+1}-\theta_j)} .$$

This lattice-path elongation is independent of the fineness of the lattice, depending only on the relationship between ϕ and the move directions.

Furthermore, any permutation of moves in a lattice results in a path of the same length. The execution of the requisite number of moves in two permitted directions θ_j and θ_{j+1} can thus result in a large number of different paths of equal length, as illustrated in figure 2. The greatest deviation from the continuous-space path will occur when all moves in one direction are executed first, giving a maximum deviation of

$$\delta(\phi) = \frac{\sin(\theta_{j+1}-\phi)\sin(\phi-\theta_j)}{\sin(\theta_{j+1}-\theta_j)} ,$$

measured perpendicular to the continuous-space path. Both error measures, ϵ and δ, are maximised when ϕ bisects the angle between θ_{j+1} and θ_j.

Table 2 shows the maximum values of ϵ and δ for three sample cases. Four move directions are permitted in the rook's case, 0, $\frac{1}{2}\pi$, π, and $\frac{3}{2}\pi$. The queen's case adds odd multiples of $\frac{1}{4}\pi$, and the knight's case, solutions of $\tan^{-1}(\pm\frac{1}{2})$ and $\tan^{-1}(\pm 2)$. Figure 3 shows the variation in each measure over one quadrant. For example, it is possible for a path defined by means of queen's and knight's moves to give a 2·6% overestimate of path length and a deviation of 11·8%, however fine the lattice.

Consider now a general continuous-cost surface, with an exact solution to the minimum-path problem defined by

$$\min \int_{\text{path}} z \, ds .$$

Writing $\tan\phi = dy/dx$ for the instantaneous path direction, we have

$$\min \int_{\text{path}} z(1+\tan^2\phi)^{1/2} dx .$$

But the problem solved by the lattice approach reduces to

$$\min \sum z(1+\tan^2\phi)^{1/2}\epsilon(\phi)\Delta x .$$

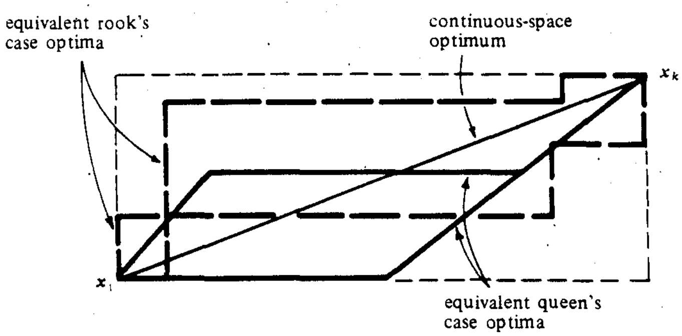

Figure 2. Relationship between discrete-space and continuous-space optima.

Table 2. Maximum elongation and deviation for simple move sets.

	Elongation ϵ	Deviation δ
Rook's case	1·4142	0·5000
Queen's case	1·0824	0·2071
Queen's + knight's case	1·0261	0·1178

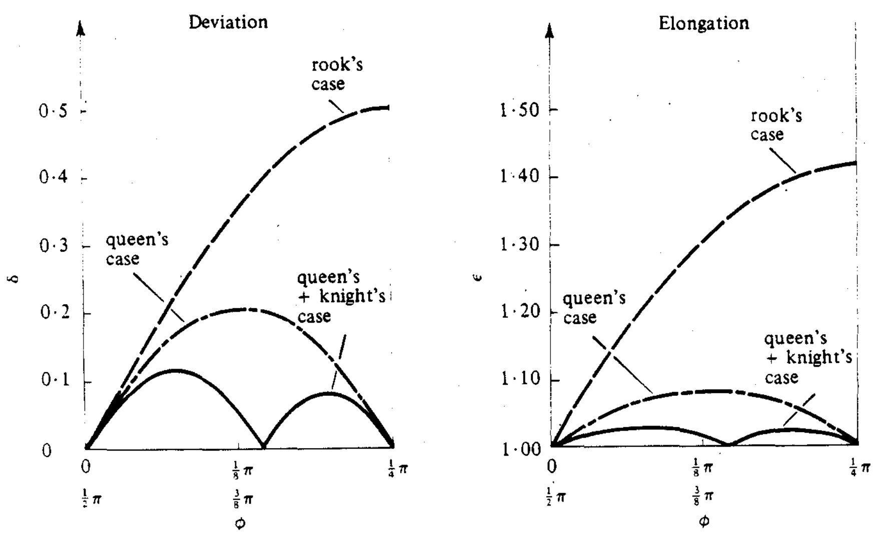

Figure 3. Variation in deviation and elongation with ϕ.

where Δx is the horizontal projection of each increment. It is clear that lattice paths do not converge to the corresponding continuous-space paths as Δx tends to zero.

Figure 4 illustrates lack of convergence for three lattice cases. It is possible to solve the continuous-space problem analytically in the case when z is a linear function of x alone, and the result is shown as the 'exact' solution. The development of the solution is as follows:

Consider a path in a direction ϕ in an area with cost per unit length $z-\Delta z$, and let this path cross into an area of cost z, in which its direction is $\phi+\Delta\phi$. Assume that the boundary between the areas is parallel to the y axis, so that z varies with x alone. Then by analogy to Snell's law of optical refraction (see for example Jenkins and White, 1957, page 5; Wardrop, 1969; Angel and Hyman, 1976, page 12) we have

$$\frac{\sin(\phi+\Delta\phi)}{\sin\phi} = \frac{z-\Delta z}{z} ;$$

and in the limit as Δz tends to zero,

$$\frac{d\phi}{dz} = \frac{-\tan\phi}{z} .$$

Writing $\tan\phi = dy/dx$ and $z = a+mx$ we have the solution

$$\frac{dy}{dz} = \frac{k}{m(z^2-k^2)^{1/2}} ,$$

where k is a constant of integration (compare for example Angel and Hyman, 1976, page 22).

Integration gives the solution

$$y = \frac{k}{m}\ln[z+(z^2-k^2)^{1/2}]+c , \qquad \text{or} \quad y = \frac{k}{m}\ln\{a+mx+[(a+mx)^2-k^2]^{1/2}\}+c .$$

Both k and c can be determined by substituting the known origin and destination points. In particular an origin at (0, 0) and destination at (x_0, y_0) give

$$y_0 = \frac{k}{m}\ln\frac{a+mx_0+[(a+mx_0)^2-k^2]^{1/2}}{a+(a^2-k^2)^{1/2}} ,$$

which can be solved numerically for k.

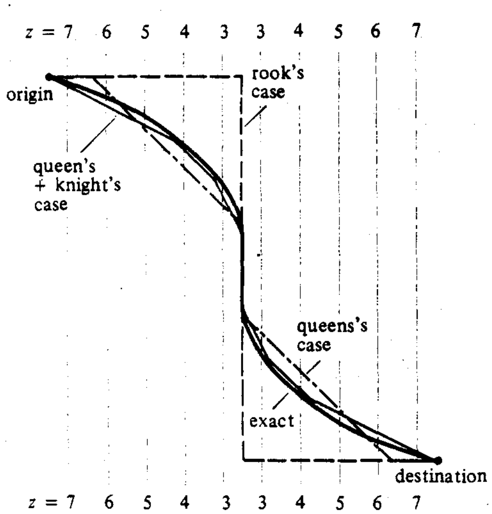

Figure 4. Optima on a simple nonuniform cost surface.

Two kinds of error, elongation and deviation, were identified for lattice paths on flat cost surfaces. On general surfaces the presence of varying degrees of elongation leads to a lack of convergence between lattice- and continuous-surface solutions. But the ties between the costs of different paths, which led to possible deviation error on flat surfaces, tend to be resolved on general surfaces, so that this problem is only encountered when the cost surface shows substantial plateaux. For example, the lattice solutions in figure 4 are unique optima, in contrast to those in figure 2.

Refraction at discontinuities

Many geographic variables are best represented choroplethically, by attaching values to areas outlined by boundaries. The resulting surface appears as a series of plateaux, with sharp discontinuities between them. As before, we identify the height of the surface as a generalised cost per unit length of construction, and consider the location of paths of minimum cost.

Let z_i be the cost in the ith homogeneous area, and ψ_i the angle between the optimum path and a normal to the boundary at the point where the path crosses to area $i+1$. Then by analogy to Snell's law,

$$\frac{\sin\psi_{i+1}}{\sin\psi_i} = \frac{z_i}{z_{i+1}} .$$

Now let continuous space be replaced by a lattice with permitted move directions $\theta_1 \ldots \theta_s$ measured from a line drawn perpendicular to the discontinuity. Let the length of path in area A in direction θ_j be a_j, in area B in direction ϕ_k be b_k. Then we have the constraints

$$\sum_j a_j \cos\theta_j = D_1 , \qquad \sum_k b_k \cos\phi_k = D_2 , \qquad \sum_j a_j \sin\theta_j + \sum_k b_k \sin\phi_k = D_3 ,$$

where D_1 and D_2 are the perpendicular distances of origin and destination from the discontinuity, and D_3 is the distance between origin and destination parallel to it. We wish to minimise

$$Z = z_A \sum_j a_j + z_B \sum_k b_k .$$

Applying Lagrange multipliers λ_A, λ_B, and μ to the three constraints, respectively, and differentiating with respect to the a_j and b_k gives the conditions

$$z_A + \lambda_A \cos\theta_j + \mu \sin\theta_j = 0 , \qquad j = 1, 2, \ldots ,$$

$$z_B + \lambda_B \cos\phi_1 + \mu \sin\phi_k = 0 , \qquad k = 1, 2, \ldots .$$

Since there are only three unknowns, there can be no more than three nonzero *a*s or *b*s. So in one area the path follows a permitted direction, and in the other it is a combination of two.

Figure 5 shows an example of the refraction of a route according to a continuous-space solution, obeying Snell's law, and three lattice paths with different permitted move directions. In each of the lattice cases the path follows one permitted direction in the high-cost medium. No two solutions cross the discontinuity at the same point.

Extension of Snell's law to more than two regions is straightforward, since the law can be applied wherever a path crosses a boundary (Werner, 1968). Several further points are of interest, however. The first arises when a path must reach a destination behind a high-cost region. There is no refracted path to the destinations in figure 6(a); rather, the optimum path runs around the end of the high-cost region and requires a bend, in the absence of a corresponding discontinuity, by analogy to the process of optical diffraction. Diffraction also occurs at sharp changes in boundary direction [figure 6(b)]. It is interesting that particle- and wave-related phenomena are both involved in the analogy between optics and path location.

From Snell's law the angle of incidence in the first region, θ_1, may not exceed $\sin^{-1}(z_2/z_1)$. Optically, such paths will be reflected at the discontinuity. In the location problem such paths cannot be optimum routes to any point in the second region.

A process analogous to optical scattering occurs at any highly convoluted boundary, such as a meandering river bank. The rapid changes in boundary direction create even more rapid changes in path direction beyond the boundary.

Choroplethic maps are one case where it is possible to find exact solutions in continuous space, without resorting to some form of numerical integration. By searching all directions from a given origin and solving Snell's law at each boundary, it is possible to find a path passing through a given destination. However, in a complex system there may be more than one path between origin and destination, including both refractive and diffractive processes. Points of diffraction create major

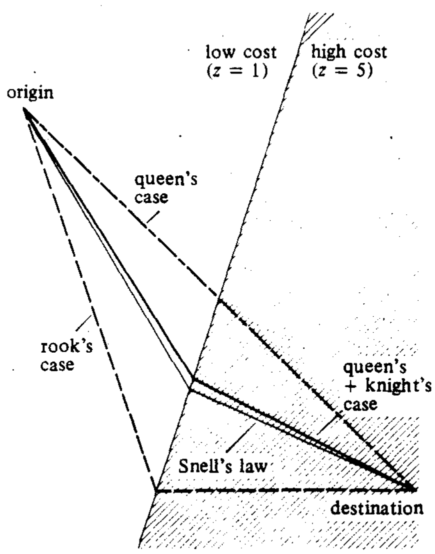

Figure 5. Refraction at a cost discontinuity.

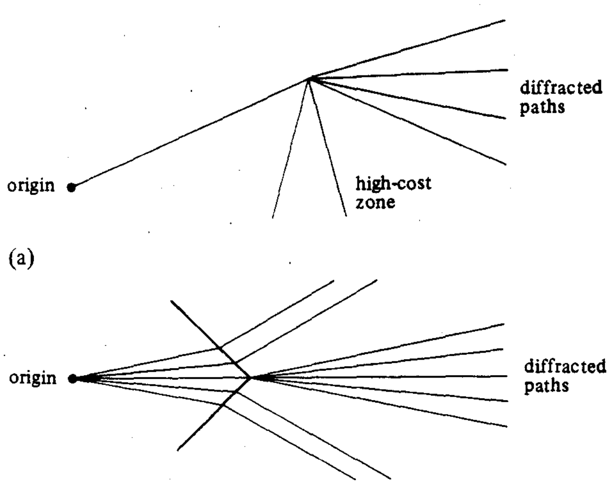

Figure 6. Route diffraction: (a) at an edge; (b) at a vertex.

problems since each one expands the number of possible paths. And although each path which intersects the destination is locally optimum, and valid optically, the set of such paths must be searched to find the globally optimum least-cost path.

Approaches for large problems

The capacity of most computing systems is reached by problems of the order of 10^5 nodes, which is a severe limitation in the application of the technique to realistic problems. Larger problems can be handled, but incur a disproportionate increase in computing costs. In this section I shall consider methods for reducing the number of nodes to within computing-system limits. The result is an approximate, or heuristic, solution.

The first and simplest method is by truncation of the study area. In any application there must be some definition of the area over which costs are to be evaluated. Although in general it is possible for the optimum path to extend outside the study area, it is implicitly assumed that this will not happen, and the study area is defined as that region within which the optimum path probably lies. By defining the area as narrowly as possible, one can reduce the number of nodes in the problem or achieve greater spatial precision for a given number of nodes. But there would be no indication even after solution of the problem if the true optimum path lay outside the study area for part of its length. No heuristic solution can ever indicate its own suboptimality.

The second and third approaches involve aggregation of nodes, or equivalently a reduction in sampling intensity. The density of the original lattice is likely to be influenced by a number of factors. First, it may be related to the width of the planned corridor, since narrow corridors clearly require a greater spatial precision. Second, the spatial variability of cost is important. If cost is highly regionalised (Matheron, 1965), showing little local variation, the sampling intensity can be correspondingly light. Aggregation produces the most acceptable results when these two influences conflict, when the spatial precision required for corridor location is higher than that needed for adequate sampling. Cells can be aggregated in a constant spatial pattern, or by taking account of local cost variation. In the former case each level of aggregation is itself a lattice, but in the latter, aggregate cells are made much larger in areas of constant cost.

Newkirk (1976) has explored the latter approach. All neighbouring nodes of constant cost are replaced by a single node, and the optimum path is then found for a smaller, irregular network. Two disadvantages should be noted. First, the change to an irregular network incurs a penalty in computing costs. Second, the spatial precision of the solution is that of the aggregated network, and is much lower in areas of constant cost. Newkirk has developed methods for adjusting and smoothing the final solution to reduce this problem.

A spatially constant aggregation by a factor A is defined as the replacement of each square of A by A nodes in the original maximum-precision lattice by a single node with a value equal to the sum of the nodes in the square; p rows and q columns are thus reduced to p/A and q/A respectively. If the cost surface is sufficiently regionalised, showing little systematic variation within each aggregate cell, then optimum paths in the two lattices will be very similar.

In the first step of the procedure, nodes are aggregated by a factor A_1, and a path is found between a specified origin and a specified destination. The nodes on the path, together with those within a certain number of nodes d of the path, are then partially disaggregated to a new level A_2 and the problem is solved again. The procedure is repeated through n steps, with a final complete disaggregation $A_n = 1$. The final solution is thus at the maximum spatial precision. It will in general diverge from the result of an exact one-step solution, depending on the scale of regionalisation of the cost surface.

The first step is in pq/A_1^2 nodes, of which roughly $(p+q)/A_1$ occur in the solution. In the second step, in which the inclusion of d extra nodes on either side of the first optimum is allowed for, the number of nodes after partial disaggregation is $2d[(p+q)/A_1](A_1/A_2)^2$. In general, at step $k > 1$ there are $2d(p+q)(A_{k-1}/A_k^2)$ nodes in the problem and $(p+q)/A_k$ in the solution.

From the computational point of view, the aggregation factors should be arranged to minimise the number of steps, by making the number of nodes equal to the system capacity N at each step:

$$\frac{pq}{A_1^2} = \frac{2d(p+q)A_{k-1}}{A_k^2} = N .$$

Thus

$$A_1 = \left(\frac{pq}{N}\right)^{1/2} ; \qquad A_k = \left[\frac{2d(p+q)}{N}\right]^{1-(1/2)^{k-1}} \left(\frac{pq}{N}\right)^{(1/2)^k} ;$$

$$A_n = 1 ; \qquad n = \frac{1}{\ln 2} \ln \frac{2 - \ln(pq/N)}{\ln[2d(p+q)/N]} .$$

These relationships only provide rough guidelines for the choice of the A_k, however, since in reality each p/A_k and q/A_k must be an integer.

The method was developed for a study of a proposed power transmission corridor between Nanticoke and London, Ontario (Potts, 1975; Newkirk and Troughton, 1974).

Summary and conclusions

The problem of optimal horizontal alignment for a transportation or communication corridor has frequently been studied as one of finding the least-cost route through a weighted lattice. This approach has several advantages. It places no restrictions on the distribution of cost over the study area, introduces an explicit level of spatial resolution to the study, and makes the problem finite and tractable. However, the lattice approach implies two forms of approximation. The precision of the solution depends on the fineness of the lattice, and on the set of move directions permitted. The ideal solution is only found when both tend to infinity. Solutions found with a finite move set do not tend asymptotically to the ideal solution as the fineness of the lattice increases. Furthermore, it is impossible to examine the effects of increasing the move set beyond a few simple cases without at the same time considering longer and longer moves which reduce the spatial precision of the lattice.

Two indices, deviation and elongation, were introduced in order to measure the degree of divergence between lattice and exact solutions. Divergence occurs in two ways. First, lattice solutions tend to be nonunique, especially when areas of constant cost are present in the study area. This effect can be lessened by making a random, rather than a systematic, resolution of ties in the solution algorithm. Second, the objective function minimised in the lattice solution contains an implicit set of weights related to the move set. The divergence which results can only be reduced by enlarging the set of permitted moves.

A useful analogy exists between the selection of optimum paths and the properties of light, especially for cost surfaces which consist of homogeneous areas separated by sharp discontinuities. However, the refractive analogy proposed by Werner (1968) as a solution method is only viable for special cases. In general, diffractive as well as refractive processes are present. Diffraction creates difficulties, as it requires an expansion of the set of paths to be searched, and is overly sensitive to the spatial precision of the cost surface. Furthermore, solutions to the optical problem are not necessarily unique.

The success of any lattice method is difficult to assess, since the continuous-space ideal is usually hypothetical. The indices evaluated in this study for constant-cost surfaces place crude upper limits on lattice-path error, depending on the set of permitted moves. The difference between lattice solutions to the same problem for different move sets also gives some indication of likely error. Consider, for example, the index

$$\gamma = \frac{Z_{\text{rook's}} - Z_{\text{queen's + knight's}}}{Z_{\text{queen's}}} .$$

The value of γ will be zero if and only if no improvement in cost occurs between the smallest and the largest simple-move sets, a strong indication that the lattice solutions are very similar to a continuous-space solution.

In general, the divergence of the rook's case is such that it should never be accepted as a solution. Since solution costs usually increase by an order of magnitude from one move set to the next, a useful practical strategy is to proceed in a stepwise fashion, terminating when a further expansion of the move set produces little improvement in path cost, or when the marginal benefit from better approximation to the ideal solution is outweighed by the marginal cost of solving the problem with a larger move set. Fineness of the lattice, on the other hand, is usually controlled more by the spatial resolution of available data, or by the proposed corridor width.

The lattice approach to the corridor location problem raises many of the questions which are inherent in any heuristic method. It is difficult to obtain an accurate measure of the degree of suboptimality of the solution without solving the infinite-move-set/infinitely-fine-lattice problem, particularly since in this case the fineness of the lattice is probably limited by the resolution of the data. Greater precision can only be achieved at greater cost, either in data collection or in computer time, so that a researcher can only hope to produce the best possible solution, given externally imposed constraints.

References

Angel S, Hyman G M, 1976 *Urban Fields* (Pion, London)

Dijkstra E W, 1959 "A note on two problems in connexion with graphs" *Numerische Mathematik* **1** 269-271

Dreyfus S E, 1969 "An appraisal of some shortest-path algorithms" *Operations Research* **17** 395-412

Hopkins L D, 1973 "Design method evaluation—an experiment with corridor selection" *Socio-economic Planning Sciences* **7** 423-436

Jenkins F A, White H E, 1957 *Fundamentals of Optics, 3rd Edition* (McGraw-Hill, New York)

Matheron G, 1965 *Les Variables Régionalisées et leur Estimation* (Masson, Paris)

Newkirk R T, 1976 *A Planning System for Environmental Resource Allocations* unpublished PhD thesis, Faculty of Engineering Science, University of Western Ontario, London, Ontario, Canada

Newkirk R T, Troughton M J, 1974 "Computer based environmental assessment applied to electricity transmission corridors" *Proceedings GIRMS* Department of Geography, Wilfrid Laurier University, Waterloo, Ontario, Canada **5** 31-45

OECD, 1973 *Optimization of Road Alignment by Computer* (OECD Road Research Group, Paris)

Owens G B, 1975 *Evaluating a Highway Location Model: the London 402 Controversy* unpublished MA thesis, Department of Geography, University of Western Ontario, London, Ontario, Canada

Potts J M, 1975 *Concept and Location Algorithm for a Multi-Purpose Corridor* unpublished MA thesis, Department of Geography, University of Western Ontario, London, Ontario, Canada

Turner A K, Miles R D, 1971 "The GCARS System: a computer-assisted method of regional route location" *Highway Research Record* **348** 1-15

Wardrop J G, 1969 "Minimum-cost paths in urban areas" *Strassenbau und Strassenverkehrstechnik* **86** 184-190 (reprinted in Angel and Hyman, op cit, pages 155-161)

Warntz W, 1965 "A note on surfaces and paths and applications to geographical problems" DP-6, Michigan Inter-University Community of Mathematical Geographers (reprinted in Angel S, Hyman G M, 1976, pp 121-134)

Werner C, 1968 "The law of refraction in transportation geography: its multivariate extension" *Canadian Geographer* **12** (1) 28–40
Whiting P D, Hillier J A, 1960 "A method for finding the shortest route through a road network" *Operational Research Quarterly* **11** 37–40