Fuzzy Sets and Systems 53 (1993) 241–252
North-Holland

A fuzzy decision model for the design of rural natural gas networks

J. Davidson
Department of Civil Engineering, University of Manitoba, Winnipeg, Manitoba, Canada R3T 2N2

W. Pedrycz
Department of Electrical and Computer Engineering, University of Manitoba, Winnipeg, Manitoba, Canada R3T 2N2

I. Goulter
Department of Civil Engineering and Building, University College of Central Queensland, Rockhampton, Queensland, Australia

Received March 1991
Revised September 1991

Abstract: A fuzzy decision model to assist in computer synthesis of network geometry for design of rural natural gas distribution systems is proposed. The membership functions for individual objectives involved in the problem, such as minimization of cost and generation of *generally rectilinear* geometry, are derived. The method combines Saaty's priority method, a regression analysis of the membership function and a piecewise linear approximation of the function obtained through regression. The procedure is demonstrated by application to an example network.

Keywords: Fuzzy decision-making; conflicting objectives; priority method; cost; distribution systems; gas; layouts; networks; rectilinearity.

1. Introduction and problem formulation

The focus of this work is computer synthesis of network geometry in the context of rural natural gas networks and as such it extends previous studies [1, 2, 3]. The problem belongs to the field of graph theory where it is known as the Minimal Rectilinear Steiner Graph problem. It is generally considered to be a member of the set of NP-complete problems, for which there are no known algorithmic solutions [4].

The technique used in previous work to generate layouts for gas distribution systems was a heuristic procedure coupled with a rule base to improve, according to some experience based criteria, the layouts generated by the procedure. The procedure and the rule base together produce layouts that are completely rectilinear. The techniques employed in this work transform these rectilinear networks to networks that are *generally rectilinear* or almost completely rectilinear. In this context the term *generally rectilinear* implies that some diagonal lines are permitted in the network.

The specification that the networks should be *generally rectilinear* has evolved from common engineering practice. Rural pipe networks are different from urban systems in which geometry is almost always constrained by existing right-of-ways. The relative absence of 'right-of-way' constraints on the design of rural systems means that an infinite number of geometric layouts are possible.

These systems are, however, often required to be rectilinear with respect to the existing survey grid because of the need to locate the system at some later date for crossings or modifications. Since complete rectilinearity is a costly restriction, and to some degree unnecessary, the restriction is not strictly observed in practice. The designer must therefore decide on the degree of rectilinearity to be incorporated in a particular layout. Thus from a formal standpoint, the problem can be considered a one-stage decision making process in which two objectives of highly contradictory or conflicting character are involved. In addition the rectilinearity objective is not defined precisely and, after appropriate calibration, may be conveniently represented as a fuzzy set. Alternative layouts (ranging from completely rectilinear to

Correspondence to: Prof. W. Pedrycz, Department of Electrical and Computer Engineering, University of Manitoba, Winnipeg, Manitoba, Canada R3T 2N2.

completely diagonal) have varying degrees of membership in this fuzzy set. The problem of designing *generally rectilinear* layouts can therefore be modelled as a fuzzy decision process with the two following conflicting objectives:

1. Minimization of the cost of the pipe network.

2. Production of a layout with *generally rectilinear* geometry. It should be noted that as it currently operates in practice, the design process, is quite unsystematic. This unsystematic approach represents one way of dealing with the exponential number of alternative solutions which might be considered during the process. This work attempts to impose a systematic approach to the design problem while avoiding the danger of having to completely enumerate all possible solutions.

2. Development of the fuzzy decision model

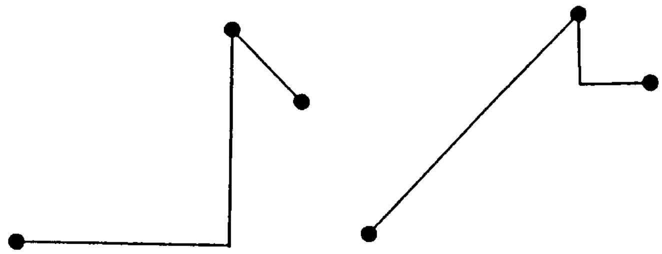
2.1. Preliminary assumptions

Determination of the membership function: In the first instance it is necessary to formulate a quantitative measure of rectilinearity to serve as the universe of discourse for the membership function. A range of rectilinearity from 0 to 1 was selected where 0 is a completely diagonal layout and 1 is a completely rectilinear layout. Intermediate values in this range can be interpreted as the proportion of the network that is rectilinear.

The first step in deriving this quantitative measure of rectilinearity was to assume that the set of completely diagonal lines and the set of completely rectilinear lines are crisp sets. The assumption that the sets are crisp allows fewer fuzzy sets to be used in the model. This reduction in the number of fuzzy sets reduces the number of membership functions that would have to be derived.

The assumption is not strictly correct since there often exists lines which are not strictly rectilinear but are so close to being rectilinear that for practical purposes they could be regarded as such. The cost associated wih the extra length required to make these lines rectilinear is so small in comparison to the total system cost, that in terms of the objective of minimizing cost it is not significant whether these lines are regarded as rectilinear. Thus the assumption of a crisp set for completely rectilinear lines can be justified on the basis that there is no significant impact on the costs of the networks produced by the assumption.

A simple method to determine the degree of rectilinearity of a design outcome is to count the number of segments in a layout that are rectilinear and divide this number by the total number of segments. However, this approach is not sensitive to lines of different length. The lack of sensitivity to lines of different lengths can be explained with reference to the simple networks shown in Figure 1. Both networks in Figure 1 are composed of three line segments. In Figure 1(a) the longer line is diagonal and in Figure 1(b) the pair of longer lines corresponding to or replacing the diagonal in Figure 1(a) is rectilinear. Since both systems are composed of two rectilinear lines and one diagonal, the degree of rectilinearity defined by the above approach is the same, at 0.67 (2/3). Yet it is clear from inspection that Figure 1(b) has a greater length of rectilinear pipe and is the more rectilinear of the two layouts. The solution to this inability to differentiate between lengths of lines that are



(a) (b)

Fig. 1. Two systems with different degrees of rectilinearity.

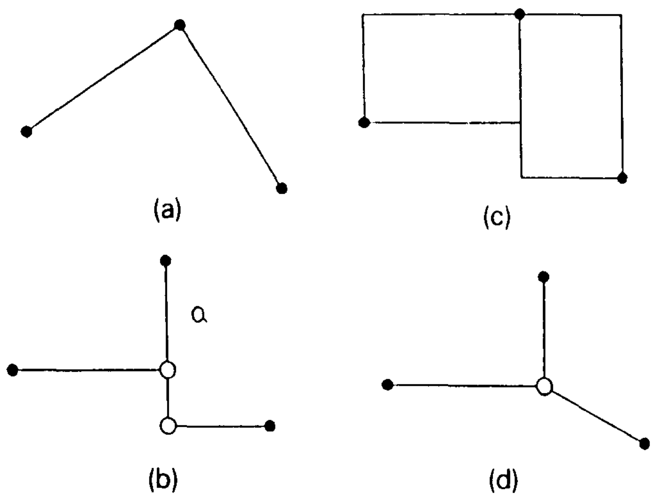


Fig. 2. Maximizing common segments.

rectilinear is to weight each line by the ratio of its length to the total length of the network. Thus an overall degree of rectilinearity can be described accordingly:

$$R = \frac{\sum_{i=1}^{n} b_i L_i}{L_t} \tag{1}$$

where

R stands for the degree of rectilinearity;

b_i is a Boolean (two-valued) predicate such that:

$$b_i = \begin{cases} 1 & \text{if the segment is rectilinear,} \\ 0 & \text{if the segment is diagonal;} \end{cases}$$

L_i is the length of segment i; and

L_t is the total length of the system.

The total length L_t is calculated using the 'rectilinear' length of the network. For those sections of the network which are already rectilinear the length used in calculating the total length of the network is the actual length of the segments. For those portions of the network containing diagonal lines a different approach is used to obtain the contribution to the total length of the network. The total length of the rectilinear distances corresponding to the diagonal line, i.e., the pair of lines for which the diagonal is the hypotenuse, are used rather than the actual length of diagonal lines.

This procedure has two advantages. Comparisons between various alternative layouts are more meaningful since the denominator, the L_t term, is the same for each alternative. It also becomes unnecessary to repeatedly recalculate L_t when numerous alternatives with diagonal lines are compared.

Preserving common segments: As stated previously, the overall objective of the decision model is minimization of the total cost of the system while keeping the geometry *generally rectilinear*. Minimizing the total cost of the system is not necessarily the same as minimizing the total length of the system. Pipes of different diameters have different associated costs. A system with a lower total length may cost more than one with greater total length if the system with less length requires more pipe of large diameter because of inefficient hydraulics. This additional requirement for adequate hydraulic performance influences the total system cost.

The heuristic procedure that is used to produce the initial rectilinear layout attempts to produce systems that are efficient in terms of both total system length and hydraulics. One of the techniques that is used in the procedure to minimize total system length is to maximize the length of common segments in the system. Figure 2 illustrates the principle of common segments. In Figure 2(a) two diagonal lines connect 3 nodes. Figure 2(b) shows the alternative rectilinear links between the three nodes. The minimal rectilinear solution, shown in Figure 2(c) uses the common segment labelled '*a*' to achieve minimal total length. The segment is a common segment because it is part of (common to) two different rectangles describing the alternative rectilinear routes in Figure 2(b).

In order to make the problem more tractable in the formulation the following simplifying assumption is made with respect to common segments. It is assumed that common segments are essential to the efficiency of a system and that no diagonal line which eliminates a common segment in a system will be allowed in that system. The assumption greatly simplifies the process of incorporating diagonal lines by eliminating many potential candidate layouts.

The elimination of these potential diagonal candidates 'because' of this simplification occurs for the following reasons. Define a dummy node as a node which does not contribute any load to the system. Since the line segments in a system are, by convention, all considered to be straight lines, dummy nodes are required to create the elbows and branches in the rectilinear network. Once the above simplification is applied the only rectilinear lines which are candidates for substitution by diagonal lines are those pairs of rectilinear lines which form a 90 degree elbow at a dummy node that has only two incident lines. In Figure 2 the sinks (nodes contributing to the load) are represented as solid circles and the dummy nodes are represented as circles that are not filled in. Figure 2(d) shows the only potential location of a diagonal line that will not interfere with the common segment '*a*' in the minimal rectilinear layout of Figure 2(c). Note that once a desirable diagonal line is identified, it replaces both the rectilinear lines for which it is the hypotenuse and the dummy node at the elbow of these two lines.

Removing redundant nodes: The substitution of diagonals into the system can be further simplified if it is recognised that a dummy node with two incident lines where the two incident lines form a 180 degree angle is redundant. Nodes of this type are created by the procedure that generates the initial rectilinear system. The formal procedure used to generate these dummy nodes is described in [2]. Note that the dummy nodes in Figure 2(c) were generated from Figure 2(b) by this initial layout procedure.

The identification of redundant dummy nodes can be demonstrated by examination of Figure 3(a). In Figure 3(a) the nodes 15, 16, 19 and 20 are redundant nodes that were created by the same heuristic initial layout procedure. (Node 20 is located at the tee near the source. The label has been omitted from the figure for clarity.) These nodes serve no purpose at this stage of the network design process (they are used only to generate the rectilinear layout) and should be eliminated for the following two reasons:

1. Redundant nodes may result in sub-optimal modifications. It may be possible to substitute a larger diagonal if the redundant node is removed.

2. Redundant nodes can produce too many possible layouts. Recall that the proposed design procedures attempts to limit the number of solutions to a manageable level and to provide enough alternatives to ensure a solution that satisfies the objectives and constraints. Exclusion of redundant nodes reduces the possibility of a combinatorial explosion in the number of alternative layouts to be considered.

2.2. Initial decision model

In this section the decision model which takes into account the goals and constraints of the design process is formulated.

Calculating incremental values: As a first step consider a crude decision model consisting of two membership functions on a common universe of discourse. The two membership functions correspond to the two objectives, i.e., the produce a *generally rectilinear* solution and to minimize cost. A more precise description of the second objective as it was implemented in the model is maximization of savings.

The degree to which each of these objectives is satisfied corresponds to the degree of membership in the respective fuzzy set. Since both objectives are defined in the same space, the intersection of the two membership functions defines the membership function of a fuzzy set which satisfies both objectives. The maximum of this membership function represents the best compromise between the two conflicting objectives.

The process as it applies to the network

Table 1. Incremental values for example in Figure 3 (Total length of system = 48300)

Point	Length	Cost	iR[a]	iC[b]
13	7182.9	1132.7	0.149	0.143
18	6661.8	1818.5	0.138	0.228
12	6618.2	1812.4	0.137	0.228
14	6407.3	1874.5	0.133	0.236
11	4504.5	1303.6	0.093	0.164

[a] Incremental change in degree of rectilinearity.
[b] Incremental cost.

problem is best explained using an example. Refer again to the rectilinear layout in Figure 3(a). As mentioned earlier the rectilinear lines which can be replaced by diagonal lines occur as pairs of lines which form a 90 degree elbow about a dummy node. The potential candidate, i.e., candidate links for replacement by a diagonal, will be referred to by the number of the dummy node at the elbow. In Figure 3(a) the candidates are 13, 18, 12, 14 and 11. (The source is represented as a circle with the letter 'S' inscribed inside.) The first step in the procedure is to identify all the potential candidates and compute the length of the pair of rectilinear lines that form the elbow for each of the candidates. Table 1 shows the candidates identified by corner dummy node number in the column labelled 'point' and the corresponding length in the column labelled 'length'.

It was stated previously that minimizing the cost of the system is not necessarily the same as minimizing the length of the system. However at this point another simplifying assumption is made. The assumption is that the reduction in cost associated with substituting a diagonal is proportional to the resulting decrease in the length of the system. (This assumption is a heuristic to guide the procedure rather than an absolute control on the solution strategy.) The length of each diagonal is determined and subtracted from the rectilinear length in column 2 of Table 1 to produce the values in the column labelled 'cost'.

The total length of the system then is calculated. Each of the values in the length column are divided by the total length of the network to produce the incremental change in the degree of rectilinearity of the system associated with substituting a diagonal for that candidate pair of diagonal lines. These values are located in column 3 of Table 1 under the label 'iR'. These incremental rectilinearity values are consistent with the degree of rectilinearity as determined in equation (1) and are used as follows. The system shown in Figure 3(a) is completely rectilinear and has a degree of rectilinearity equal to 1. The candidate lines associated with node 13 have an iR value of 0.149. If a diagonal line is substituted at node 13 the degree of rectilinearity of the system is 0.851, i.e., 1–0.149. Similarly, if the diagonals associated with 13 and 18 were substituted with diagonals the degree of rectilinearity would be 0.713, i.e. 1–0.149–0.138.

In a similar manner the values in the 'cost' column are summed to arrive at the maximum saving that is feasible by substituting diagonals. To create membership functions the following normalization procedure is applied. To derive values for the membership function for maximum savings, the calculated values for savings (based on decreased system length) must be normalised such that the maximum value (the total) is not greater than 1, since a membership function have a value greater than 1. The individual values in the cost column are then divided by the total possible savings to produce a normalized incremental saving in cost. These values are placed in the column labelled 'iC'. The value of the cost membership function for any decision is the sum of the incremental savings values (iC) corresponding to the diagonals that have been substituted into the system to that point. Therefore if all the diagonals are substituted into the system the value of the cost membership function is the total of column iC which is 1 corresponding to the maximum possible savings.

The ranking criterion: Let a candidate node be defined as a node at which two rectilinear lines which are candidates for replacement by a diagonal meet. The candidate nodes therefore represent the potential decisions to replace two line segments by a diagonal. Recall that replacement of the two lines represented by a candidate node also removes the node as the corner it constitutes is eliminated. The number of possible ways, i.e., the order, in which the candidate nodes (pairs of rectilinear lines) can be replaced is equal to $n!$ where n is the number of candidate nodes. A combinatorial explosion therefore results for any n greater than a very small value. To avoid these problems with combinatorics a single ordering scheme based on the following ranking criterion is imposed.

Candidate nodes with smaller values of iR are preferred over those with larger values. Application of this ranking procedure produces a series of layouts generated by sequential substitution of the diagonals into the system in order of the ranking criterion. Note that the most desired ranking order is one that produces the most alternatives in the range where the two membership functions overlap. In the early stages of the development of this model, this range was assumed to lie roughly between 0.75 and 1.0. By selecting smaller values if iR first, more alternative layouts with higher degrees of rectilinearity are produced. The large number of alternatives with a high degree of rectilinearity generated by the procedure ensures that a large number of the generated alternatives lie in the desired range of 0.75–1.0.

Note that the effect of decisions is cumulative. The nature of this cumulative process is shown in Table 2 for the example in Figure 3. The smallest value of iR is always zero, corresponding to no candidate nodes being selected, i.e., no diagonal decision being made. The first line in Table 2 represents this stage. In this first stage the value of R, representing the rectilinearity of the system, is 1.0 as the system is completely rectilinear with no diagonal replacement having occurred. The savings that can be achieved by selecting a particular diagonal replacement (as represented by the choice of a candidate node) for the example in Figure 3 are shown in column C of Table 2. In this first row in which no candidate nodes have been chosen, the C value is 0 because no savings are realized when no diagonals are substituted.

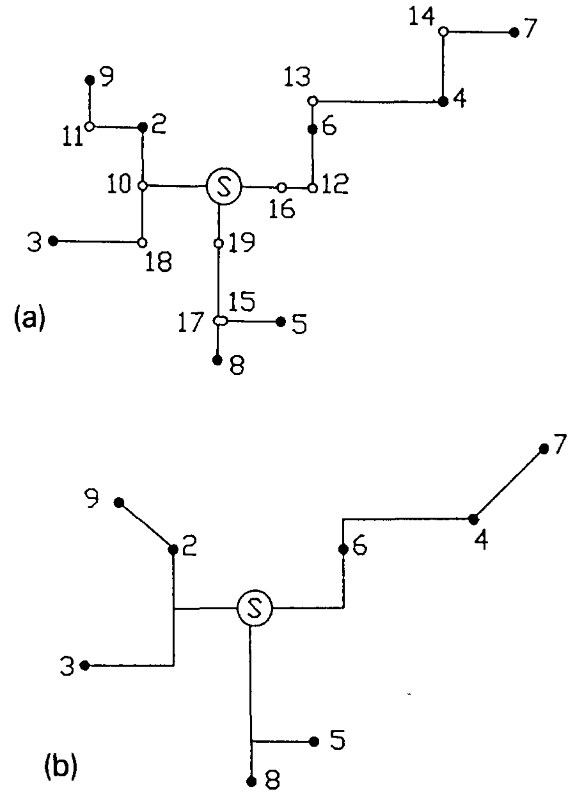


Fig. 3. Example problem.

Table 2. Cumulative values for example in Figure 3

Decision[a]	R^{b}	C^{c}	G^{d}	D^{e}
Null	1	0	0	0
11	0.907	0.164	0.936	0.164
14, 11	0.774	0.400	0.348	0.348
12, 14, 11	0.637	0.628	0	0
18, 12, 14, 11	0.499	0.856	0	0
13, 18, 12, 14, 11	0.350	0.999	0	0

[a] Selection of candidate nodes.
[b] Degree of rectilinearity.
[c] Saving in system cost.
[d] From equation (2).
[e] $\min\{C, G\}$.

Since candidate node 11 has the smallest iR value (see Table 1) that node is selected first according to the ranking criterion described previously. The degree of rectilinearity, or R value, arising from the decision is calculated by *subtracting* the iR value for the chosen node from the R value arising from the previous decision. Similarly the saving, or C value, arising from a decision is calculated by *adding* the iC value for the chosen node to the C value arising from the previous decision. For the first choice in the example, namely node 11, the previous decision was no decision and the corresponding previous rectilinearity and cost values were 1.0 and 0 respectively. The cumulative impact of these successive decisions (choice of candidate nodes on the R and C values is shown in Table 2.

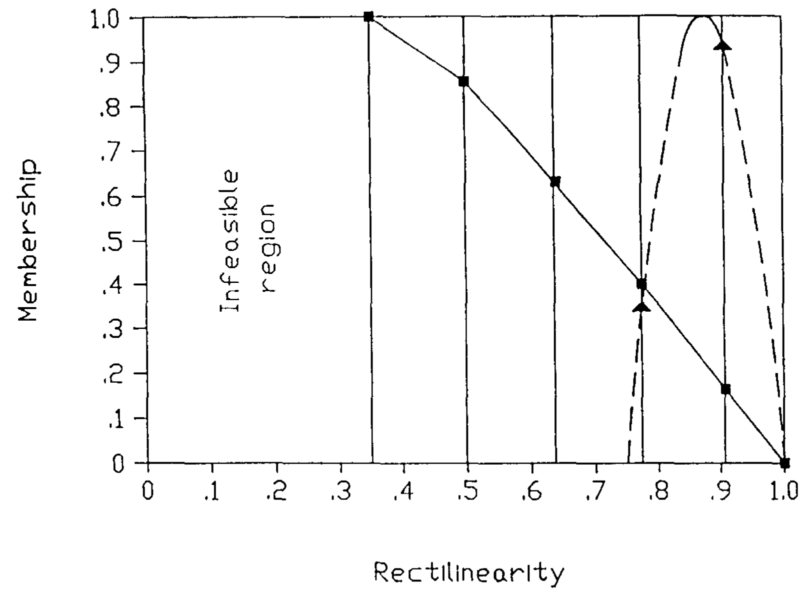


Fig. 4. Membership function for example in Figure 3 (■ saving, ▲ generally rectilinear).

The plotting of the membership function for this model is also best described by demonstration using the result obtained above for the example in Figure 3. Figure 4 shows the sets of feasible values of R and C plotted as small squares on a graph in which R is the horizontal axis and C is the vertical axis. These points are connected to produce the membership function of the fuzzy set representing maximum savings. Examination of this graph shows that the maximum saving is fully realized at the point where the degree of rectilinearity is equal to 0.35. All values to the left of this point are infeasible because no more candidate nodes exist, i.e., no more diagonal replacement is possible once the value of 0.35 has been reached.

A potential rectilinear membership function for a 'generally rectilinear' layout for the example problem is shown in Figure 4. The actual nature of the membership function is not known at this time. The membership function that is proposed for the example assumes that a 'generally rectilinear' layout has a degree of rectilinearity greater than 0.75 and the most desirable layout occurs between 0.75 and 1.0. This function is purely theoretical, which is acceptable at this time since the example problem is presented only to illustrate, in general terms, the mechanism of the proposed decision model. In the actual model the rectilinearity membership function is not required. Rather a method to derive the intersection of the rectilinearity and savings membership functions, which is the important part of the process, without having to specify the membership functions explicitly, is used. This procedure is described in the next section.

The parabolic membership function in Figure 4 should not be confused with the rectilinearity measure that has been developed to this point. The membership function models "generally rectilinear" solutions, which in this instance are assumed to exist only within a narrow region of the universe of discourse, as shown in Figure 4.

At this point a parabolic function shown below is used as a membership function for *generally rectilinear* geometry:

$$G = \max\left(1 - \left(\frac{R - 0.875}{1.125}\right)^2, 0\right). \tag{2}$$

The two membership functions in Figure 4 are used as follows. Using equation (2), G values (shown in Table 2) are calculated from the values in the R column. These G values are plotted as triangles in the graph in Figure 4. The D value, which is calculated by taking the minimum of the C and G values, forms the basis of the decision. The most desirable set of candidate nodes, i.e., the most desirable replacement decision is the set with the highest D value. In the example the preferred or most desirable solution, has a D value of 0.348 corresponding to both nodes 11 and 14 being substituted by diagonals. This solution corresponds to the layout shown in Figure 3(b).

2.3. Calibrating the model

The Saaty method: The method described by Saaty [5] was employed to determine the character of the membership function for the set formed by the intersection of the membership functions of the two objectives of rectilinearity and cost. Define this set by the set of values of D. [In the example problem this set was also referred to by the values of D. However the function $\min\{G, C\}$ used to derive the values of D in the example was selected for demonstration purposes only.] Saaty's method assigns a degree of preference to different alternatives based on a pairwise comparison of the alternatives. The method also provides a measure of consistency of the pairwise comparisons.

To determine the decision function, four different configurations of nodes were used. In each case the heuristic procedure of Davidson and Goulter [2] was used to generate an entirely rectilinear layout with an R value of 1.0. All the potential candidate links for substitution by diagonals were identified. These candidates were then ranked in order of preference and a series of layouts based on substituting the diagonals in the ranking order generated. A separate layout was generated at each step as each diagonal was substituted and the degree of rectilinearity, as defined by equation (1), was calculated for each layout.

Since the initial layout is totally rectilinear the ranking criterion causes nodes to be removed in order of their increasing reduction in rectilinearity. The layouts are produced successively in an order of highest to lowest degree of rectilinearity. Figure 5 shows a series of layouts and corresponding R values for a typical network. The layouts are shown in the order in which they are generated by the procedure.

After all layouts have been generated they are arranged in order of preference. The order of preference in this case is nominally the degree to

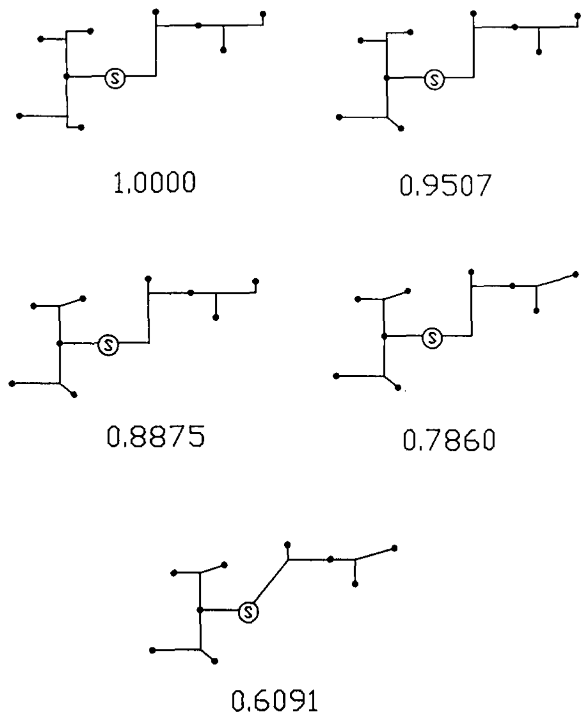


Fig. 5. Candidate layouts generated by diagonal replacement for a typical network.

which the layout satisfies both objectives – minimizing cost and maintaining *generally rectilinear* geometry. In application the layout selected first is that which has a mix of diagonal and rectilinear lines most closely resembling the ratios found in layouts produced in practice. This preference scheme is obviously subjective but it is based upon field experience. In this study networks were selected preferentially by the first author who has such experience.

The remaining layouts from the complete set generated by diagonal replacement are selected successively in decreasing order of preference and ranked accordingly. In the example test case the sequence of R values corresponding to networks selected in a decreasing order of preference was 0.7860, 0.8875, 0.9507, 0.6091 and 1.0. It can be seen that the layout with an R value of 0.7860 corresponding to a relatively high degree of rectilinearity was the most preferred while the completely rectilinear layout with an R value of 1.0 was least preferred.

The next step in the procedure is to obtain the values for elements in the Saaty matrix. The Saaty matrix is a square matrix with dimensions equal to the number of layouts that are to be

Table 3. Elements of the Saaty matrix

	0.7860	0.8875	0.9507	0.6091	1.0000
0.7860	1.0000	1.3333	2.0000	4.0000	10.0000
0.8875	0.7500	1.0000	1.3333	2.0000	4.0000
0.9507	0.5000	0.7500	1.0000	1.2500	2.5000
0.6091	0.2500	0.5000	0.8000	1.0000	2.0000
1.0000	0.1000	0.2500	0.4000	0.5000	1.0000
Largest eigen value	5.0561				
Eigen vector	1.0000	0.5880	0.4014	0.2799	0.1350

compared. The rows and columns of the matrix represent layouts corresponding to the sequence of R values derived in the previous step, and shown in Table 3. The value derived for a particular matrix element is the degree of preference for the layout with an R value corresponding to the row label in comparison with the layout with an R value corresponding to the column label. A high value for a matrix element indicates that the layout corresponding to the row label is preferable to the layout corresponding to the column label. The elements on the diagonal represent a layout compared with itself. The value for these elements is therefore always equal to 1.

The labels for the rows and columns of the matrix are arranged in order of decreasing preference. The layouts corresponding to row labels will all be less preferable than the layouts corresponding to column labels for elements below the diagonal. Therefore all values below the diagonal should all be less than one.

A simple method was derived to determine the values for elements below the diagonal in a single column. The layout corresponding to the R value assigned to that column is assumed to have a degree of preference equal to 1. The values for elements in the lower portion of that column, i.e., below the diagonal, are assessed relative to this assumed value.

The method is best explained using the example in Figure 5. Consider column 2 in the matrix in Table 3. This column corresponds to the layout with an R value equal to 0.8875. Assume this layout to have a degree of preference equal to 1. The elements below the diagonal in this column, i.e., the lower three elements, will have their values derived by pairwise comparison with this layout. The rows corresponding to these three elements correspond to layouts with R values of 0.9507, 0.6091 and 1.0. The entry in row 2 and column 3 is 0.75 whichs means that the layout corresponding to $R = 0.9507$ (the row) is 'only 75% as good' as the layout correspond to $R = 0.8875$ (the column). Once all the values for the elements below the diagonal have been assessed in this way, the values for the elements above the diagonal can be calculated based on the relationship $a(i, j) = 1/a(j, i)$.

Like the preference ranking of alternatives, the specification of the values of the elements of Saaty's matrix is subjective. However, as with preference ranking this specification can be based upon field experience and in this study the values were again specified using the relevant field experience of the first author.

Saaty's method assumes that some inconsistency, or error, usually occurs in the pairwise comparisons. The eigen vector of the matrix corresponding to the largest eigen value has elements that are proportioned to minimize inconsistency. In this study the elements of this eigen vector correspond to the preference for each of the different feasible layouts. The order of the elements in the eigen vector corresponds to the order used in the columns and rows of the matrix. For the example the eigen values or preferences of 1.0, 0.5880, 0.4014, 0.2799 and 0.1350 show in Table 3 correspond to the layouts with R values of 0.7860, 0.8875, 0.9507, 0.6091 and 1.0 respectively.

The largest of the eigen values represents a measure of the consistency of the pairwise comparisons. For a perfectly consistent set of pairwise comparisons this eigenvalue is equal to n, the dimension of the matrix. A small difference between the eigen value and n, indicates a highly consistent matrix. In the example (see Table 3) the dimension of the matrix is 5 by 5 and the largest eigen value is 5.0561. The small, almost negligible, difference between these two values indicates a high degree of consistency.

As mentioned Saaty's method provides the preference values in the form of an eigen vector. This eigen vector provides the proportions for these values and does not provide the actual values. To assign values to the elements of the

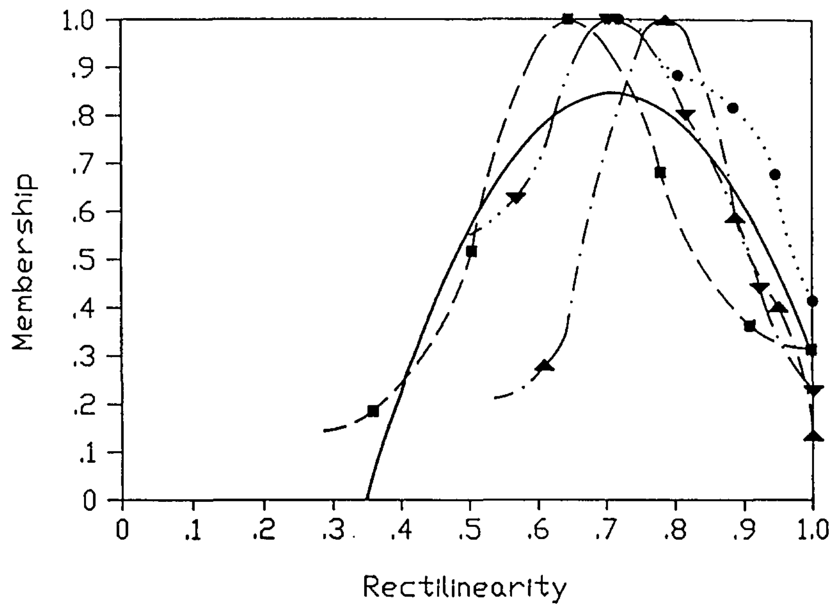


Fig. 6. Derived membership functions for example problem (···●··· Test 1, --■-- Test 2, ·–▲–· Test 3, ···▲··· Test 4).

eigen vectors the element with the largest value, the first element in each case, is assigned a value of 1 and the others scaled proportionally.

As mentioned at the beginning of this section four different configurations of nodes were used as test cases. The results from each of these test cases is shown in Figure 6. A line is drawn through the data points to produce a separate membership function for each test case. A high degree of variation between the membership functions for each of the test cases is evident from Figure 6. The variation may be due to one of the three following reasons:

1. The membership function may depend on the geometric configuration of the nodes. Each pattern of nodes may require a membership function with a slightly different shape because the savings function described in the preliminary model is not necessarily the same for all patterns of nodes.

2. One or more of the assumptions involved in determining the degree of rectilinearity of a layout is incorrect.

3. The subjective judgement that is used to obtain the test data is not consistent.

Approximation of the membership function: It should be noted that the objective in approximating the membership function based on all the test data is to establish if a single membership function can be found which can be used for all problems, i.e., all configurations of nodes. For practical purposes, the membership function is only required to select the most desirable solution from a set of alternatives. In actual practice only one design solution is implemented for a problem. The goal is, therefore, to find an approximate membership function that can select the best alternative regardless of the slight differences between this approximate function and the actual membership functions for individual problems. In this way universality is achieved for the small loss in precision and repetitive calculations for each problem can be eliminated.

The test data suggest a membership function for decision that is a generally parabolic in shape and concave downward. Quadratic regression was therefore performed to fit the R value and its square with the membership values provided using Saaty's method. The resulting function is given below and shown graphically in Figure 6.

$$D = -6.483R^2 + 9.202R - 2.420. \quad (3)$$

It can be seen in Figure 6 that the function does not fit the data well. Two indices of significance, the F statistic and the coefficient of multiple correlation (r^2), were calculated. An F

value of 15.99 and an r^2 value of 0.6399 were obtained. These values support the conclusion that the function does not fit the data particularly well. Although the values obtained using Saaty's method are not reproduced accurately by the function, the function still gives the highest membership value to the layout in each test case that was assigned the highest value by Saaty's method. Therefore in spite of its inability to reproduce the values obtained by Saaty's method accurately, the function still selects the layout that was most preferred in each of the test cases.

Linear approximation of the membership function: A simple membership function is derived based on the function obtained through quadratic regression. Initially the function is comprised of two straight lines. The first line rises from the point where the regression line crosses the horizontal axis ($R = 0.3485$, $D = 0$) and reaches a maximum of $D = 1$ at the same point on the horizontal axis as the maximum of the regression line ($R = 0.7097$). In this case the second line is a straight line connecting the maximum to a point where the regression line would cross the horizontal axis if the universe of discourse were extended beyond the point $R = 1$. The line terminates at the point $R = 1$, $D = 0.1963$. These two lines produce a crude linear approximation of the regression line.

As defined to this point the function does not return values in the range $R \geq 0$ $(0 \leq R < 0.3485)$. It is assumed that layouts in this region are less preferable than those in the remainder of the universe of discourse and become increasingly less desirable toward the origin. This assumption is modelled by a line that passes from the origin to the point on the left hand side of the function where the D value

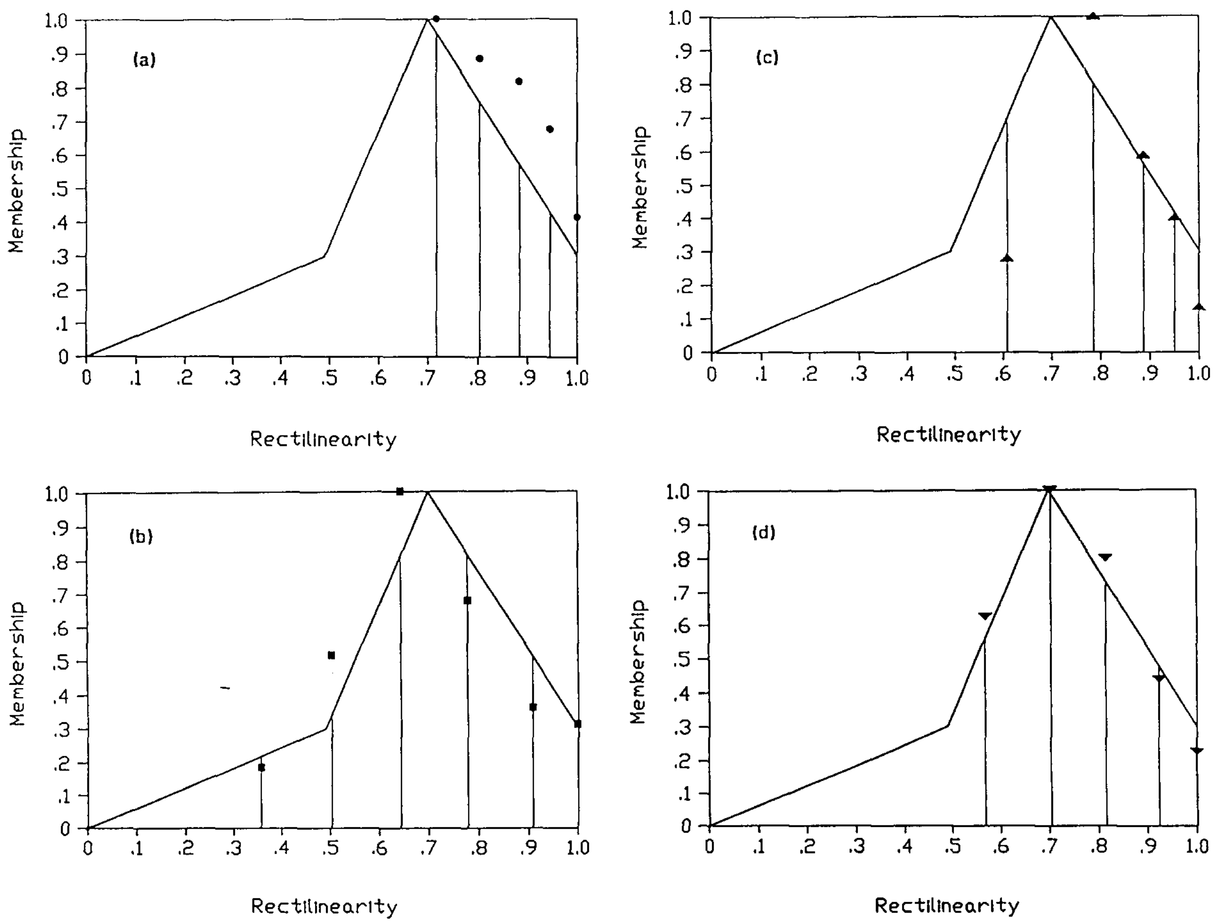


Fig. 7. Comparison of assumed linear membership function with actual membership values (Test 1–4 in (a)–(d)).

equals the value of the minimum on the right hand side ($R = 0.4194$, $D = 0.1963$). The point established at the leftmost point supported by the alpha-cut at $D = 0.1963$, the lowest value of D on the right hand side of the maximum (i.e., at $R = 1$). Recall that since the extreme left is the most undesirable, and perhaps even infeasible, region of the universe of discourse, all the D values on the extreme left should be lower that the lowest D value to the right of the maximum. This piecewise linear membership function can be stated mathematically as

$$D = \begin{cases} 0.4680R & \text{if } 0 \leqslant R \leqslant 0.4194, \\ 2.7685R - 0.9648 & \text{if } 0.4194 < R \leqslant 0.7097, \\ -2.7685R + 2.9648 & \text{if } 0.7097 < R \leqslant 1. \end{cases} \tag{4}$$

Figure 7 shows the proposed membership function plotted against the test data for each of the test cases. The order of preference for layouts is reproduced by the function in two of the four test cases. In both of the two test cases in which the sequence is not reproduced, only one layout was found to be out of sequence. In all the test cases the layout that was selected as most desirable on the basis of field experience corresponds to the layout given the highest value by the linear approximation of the membership function.

3. Conclusion

A method to generate network geometry for rural gas distribution systems in accordance with the conflicting objectives of minimizing cost and producing *generally rectilinear* geometry is proposed. The method models the selection of the most preferred layout as a fuzzy decision. The membership function for the fuzzy set corresponding to the intersection of the two objectives was obtained using a combination of Saaty's Method, linear regression and linear approximation. The membership function was found to select the preferred layout in a small sample of test cases. However, due to conflicting test data the function could not reproduce the sequence of preference for layouts in some of the test cases.

During the development of the model assumptions were made regarding the method to determine the degree of rectilinearity of a network, a value which serves as the universe of discourse. Further investigation toward alternative methods to determine the degree of rectilinearity may produce more consistent data to derive and calibrate the membership function. The method of determining alternatives and ranking alternatives involved additional assumptions which also suggest areas for further investigation.

Acknowledgements

The authors wish to thank the Canadian Natural Sciences and Engineering Research Council (NSERC) for the financial support given to this project.

References

[1] J. Davidson and I. Goulter, Microcomputer workstation for design of rural natural gas distribution systems, *Microcomputers in Civil Engineering* **4**(1) (1989) 61–73.

[2] J. Davidson and I. Goulter, Rule-based design of layout of rural natural gas systems, *J. Comput. in Civil Engineering, ASCE* **5**(3) (1991) 300–314.

[3] J. Davidson and I. Goulter, Heuristic procedure for layout design of rural natural gas distribution systems, *J. Comput. in Civil Engineering, ASCE* **5**(3) (1991) 315–332.

[4] M.R. Garey and D.S. Johnson, The rectilinear Steiner tree Problem is NP-complete, *SIAM J. Appl. Math.* **32**(4) (1977) 826–834.

[5] T.L. Saaty, A scaling method for priorities in hierarchical structures, *J. Math. Psychology* **15** (1977) 234–281.