Engineering Optimization
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/geno20

CALCULATING MAXIMUM ENTROPY FLOWS IN MULTI-SOURCE, MULTI-DEMAND NETWORKS

A. YASSIN-KASSAB [a], A. B. TEMPLEMAN [a] & T. T. TANYIMBOH [a]
[a] Department of Civil Engineering, University of Liverpool, Liverpool, L69 3BX, UK

Version of record first published: 27 Apr 2007.

Eng. Opt., 1999, Vol. 31, pp. 695–729
Reprints available directly from the publisher
Photocopying permitted by license only

CALCULATING MAXIMUM ENTROPY FLOWS IN MULTI-SOURCE, MULTI-DEMAND NETWORKS

A. YASSIN-KASSAB, A. B. TEMPLEMAN* and T. T. TANYIMBOH

Department of Civil Engineering, University of Liverpool, Liverpool L69 3BX, UK

(Received 21 October 1997; In final form 7 September 1998)

Previous work has shown how maximum entropy flows in a water distribution network can be calculated by maximizing a nodal entropy function. This requires the use of numerical nonlinear optimization. In the case of single-source networks a much simpler path entropy formulation exists which permits solutions to be obtained quickly by manual calculations not requiring numerical optimization. This paper extends the path entropy formulation to general multi-source, multi-demand networks and develops a simple manual calculation method for maximum entropy flows in these general networks. The method is quick, non-iterative and does not directly involve the use of numerical optimization. Examples are presented and discussed.

Keywords: Networks; water supply; entropy; reliability; uncertainty

1. INTRODUCTION

This paper presents a new quick method for calculating maximum entropy flows in a looped water distribution network which has multiple source and multiple demand nodes. A key feature of the problem is that very limited data is assumed to be available: only source flow rates, demand flow rates and the topology of the network with arc flow directions are assumed to be known. Data on arc properties such as lengths, diameters of pipes and roughness properties are assumed not to be available. Because of this very limited data there is insufficient

*Corresponding author.

information to permit an accurate physical analysis of the pipe network to be performed which would generate accurate pipe flow rates and a corresponding pressure regime. Under these circumstances, if an analysis of the flows in the pipe network is required it must be based upon only limited data and then has the characteristics of a most-likely or least-biased estimate of the performance.

In a sequence of papers, Tanyimboh and Templeman [8, 9, 10] have examined how such least-biased estimates of pipe flows can be obtained. The key to this problem is the use of the maximum entropy formalism [5] and Shannon's informational entropy function [7] to ensure that pipe flows estimates made in the presence of uncertainty are rigorously founded. In Ref. [8], the least-biased flow estimates of water-distribution networks were found by expressing them in the form of a constrained entropy maximization problem in which the conditional entropy formula of Khinchin [6], which is a more general form of Shannon's entropy function [7], was related to the **nodes** of the network. This results in an optimization problem needing a computer solution and suggests the need for a much simpler method. Subsequently, Tanyimboh and Templeman [9] produced a very quick and rigorous algorithm for calculating maximum entropy flows in a single-source network. This algorithm is based upon the **path** entropy concept. According to Laplace's principle of insufficient reason, which can itself be interpreted as a consequence of the maximum entropy formalism [5], all paths from the source to a demand node must supply that node equally. Unfortunately, that algorithm for single-source networks is unable to handle general network problems, leading to the need for the present paper which extends the path entropy idea to multi-source, multi-demand general networks.

A possible means by which the simple single source algorithm might be extended to multiple sources by means of a super-source concept was explored in part by Tanyimboh and Templeman [9]. In a discussion on that paper, Walters [11] has pointed out that the super-source idea as propounded in Ref. [9] is actually incorrect. He shows [11] how it should correctly be used, although the method is unwidely. This present paper shows how multiple source networks may be rigorously handled by a relatively simple algorithm, and is based on the work of Ref. [12].

In this paper, first, a brief resumé of the algorithms used by Tanyimboh and Templeman [9] for single-source networks is presented with an illustrative example. This is followed by a full description of the proposed method of calculating maximum entropy flows for multi-source, multi-demand general networks. Then, an algorithm of the proposed method is presented, followed by examples solved and checked by maximizing the nodal entropy function of the network [8]. Finally, the results are discussed and some conclusions are drawn.

2. SINGLE-SOURCE, MULTI-DEMAND NETWORKS

The following is a description of the method proposed by Tanyimboh and Templeman [9] for calculating maximum entropy flows in single-source networks. It is illustrated by means of an example. The method consists of three algorithms: a node numbering algorithm, a node weighting algorithm and a flow distribution algorithm.

The single-source network example shown in Figure 1, which is taken from Ref. [9], is used to demonstrate the above algorithms. First, all nodes of the network are numbered according to the node numbering algorithm. The source node is given the number 1, then the rest of the nodes are numbered in an ascending sequence starting with any node for which all upstream nodes have already been numbered. The numbering of nodes 4 and 5 is arbitrary and may be interchanged.

The next step is to calculate the number of paths from the source to each node using the node weighting algorithm, and then to enclose that number, as a weight of the node, in a box next to it. This is done by assigning a weight of 1 to the source node, then, in ascending node numbering sequence, the weight of each node is equal to the sum of the weights assigned to all nodes immediately upstream of it. Consequently, the weight of node 2 is equal to the weight of node 1, and the weight of node 3 is the sum of the weights of nodes 1 and 2, which is 2. Similarly, the weight of node 4 equals the weight of node 1 plus the weight of node 3, and the weight of node 5 is the sum of the weights of nodes 2 and 3, which, in both cases, equals 3. It may be noted that the node numbering algorithm ensures that all nodes immediately upstream of the node being considered have been weighted.

Finally, the flow distribution algorithm is used to determine maximum entropy link flows. The total outflow at a node is shared

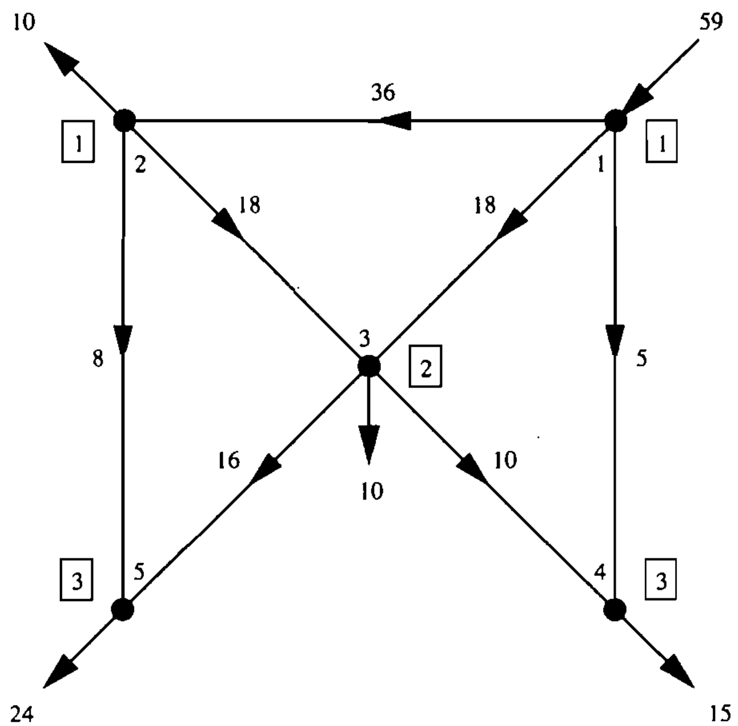


FIGURE 1 Maximum entropy flows for a single-source network with equal path flows to each demand node.

among the inflowing links at that node in proportion to the upstream nodal weights. The flow distribution algorithm operates in descending node number order. Therefore, starting with node 5, the flow in link 2–5 is obtained by multiplying 24, this being the total outflow at node 5, by the ratio 1/3 which is the ratio between the weights of nodes 2 and 5. The flow inlink 3–5 equals 24 multiplied, this time, by the ratio 2/3, this being the ratio between the weights of nodes 3 and 5. Similarly, considering node 4, the flow in links 1–4 and 3–4 can be obtained by multiplying 15 by the ratios 1/3 and 2/3 respectively. At this stage, the flow in the links ending at node 3 can be calculated. The total outflows at that node equal its demand plus the flows in links 3–4 and 3–5, resulting in 36 units. Consequently, the flow in links 1–3 and 2–3 will share the total outflows at node 3 equally due to the equality of the weights of the immediate upstream nodes of these two links. The only link left is link $1-2$ whose flow is equal to 36, this being the total outflows at node 2, multiplied by the ratio 1/1.

The above algorithms are rigorous for single-source networks. They give the same results as those given by maximizing the nodal entropy function [8], but they require no computing. They are much simpler and quicker than the nodal entropy method presented in Ref. [8].

3. MULTI-SOURCE, MULTI-DEMAND NETWORKS

3.1. Establishing the Principles

The method described above for single-source networks is inapplicable to multi-source networks. In a single-source network the demand at any node is numerically known and can be supplied only from the single source. It is therefore easy to allocate the source-demand flows, equally among all paths from the source to the demand node. Where there are several sources, although the total demand at a node is numerically known, the proportion received from each source is not known and cannot be allocated numerically among available paths. The key to solving this problem therefore lies in determining the proportions of flow received by a demand node from each of the sources.

The strategy adopted to investigate how these proportions can be quickly determined was to study a multi-source multi-demand network example for which maximum entropy flows had been determined by nodal entropy maximization. By comparing these link flows with algebraic path flows the principles for determining the proportions were identified.

Consider the 2-source network shown in Figure 2. The maximum entropy flows for that network were calculated by maximizing the nodal entropy function of the network [8]. The resulting flows are shown in Figure 3 with a maximum entropy value of 2.3885315. As in the example in Section 2, the multiple paths from each source to the same demand node must carry the same flows. Defining q_{ij} to be the path flow from source node i to demand node j, Figure 4 shows the unknown and equal path flows from each source to each demand node. For example, for demand node 5, there are three paths from

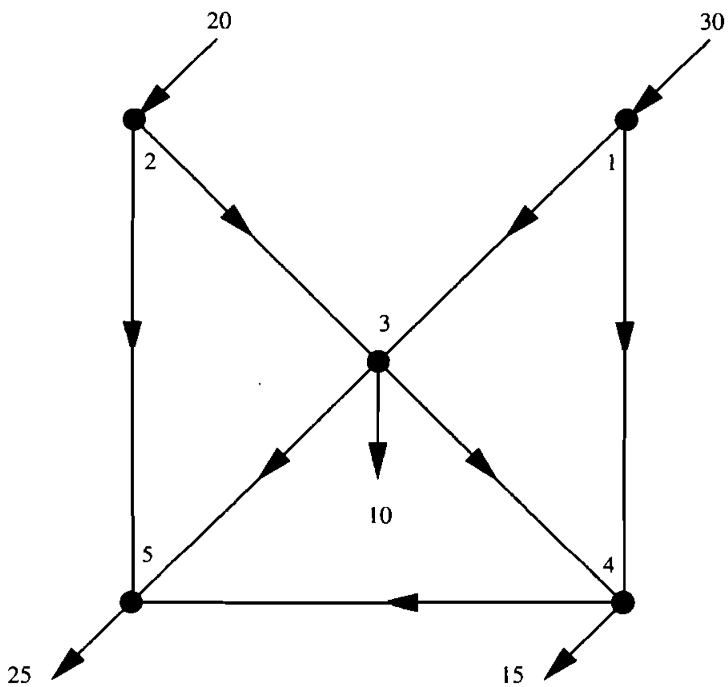


FIGURE 2 Two-source network.

each source supplying that node, those being 1−3−5, 1−3−4−5 and 1−4−5 from source node 1 (Fig. 4b) and 2−5, 2−3−5 and 2−3−4−5 from source node 2 (Fig. 4a). Demand node 4 receives two paths from source node 1 (Fig. 4d) and one path from source node 2 (Fig. 4c). Finally, only two paths supply node 3, one from each source as shown in Figures 4f and 4e. Now, equating the total flow in each link of Figure 3 to the sum of all path flows passing that link in Figure 4, the following equations can be written:

$$2q_{15} + q_{14} + q_{13} = 20.061912 \qquad \text{(for link } 1-3\text{)}$$
$$q_{15} + q_{14} = 9.938088 \qquad \text{(for link } 1-4\text{)}$$
$$2q_{25} + q_{24} + q_{23} = 16.268405 \qquad \text{(for link } 2-3\text{)}$$
$$q_{25} = 3.731595 \qquad \text{(for link } 2-5\text{)}$$
$$q_{25} + q_{15} + q_{24} + q_{14} = 17.996982 \qquad \text{(for link } 3-4\text{)}$$

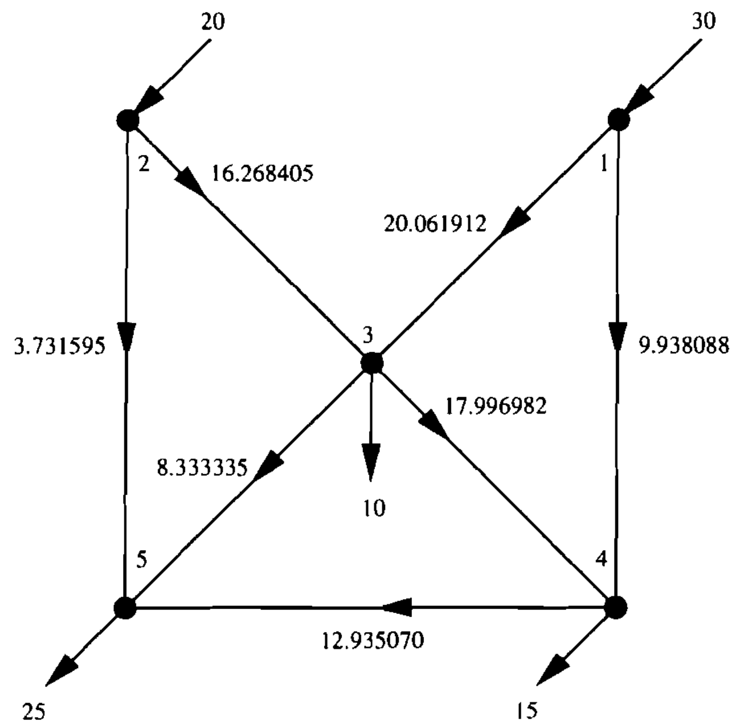


FIGURE 3 Maximum entropy flows for the network of Figure 2 obtained by maximizing the nodal entropy function of the network.

$$q_{25} + q_{15} = 8.333335 \qquad \text{(for link } 3-5\text{)}$$
$$q_{25} + 2q_{15} = 12.935070 \qquad \text{(for link } 4-5\text{)}$$

Solving the above equations gives: $(q_{13}, q_{14}, q_{15}, q_{23}, q_{24}, q_{25}) = (5.522086, 5.336350, 4.601738, 4.477916, 4.327299, 3.731595)$. At this stage, path probabilities p_{ij} can be obtained by normalising each path flow calculated above by its individual source flow as follows: $p_{ij} = q_{ij}/q_{0i}$, where p_{ij} is the probability that path flow q_{ij} supplied by source node i reaches demand node j, and q_{0i} is the external inflow at source node i. Applying that to the network being considered gives: $(p_{13}, p_{14}, p_{15}, p_{23}, p_{24}, p_{25}) = (0.1840695, 0.1778783, 0.1533913, 0.2238958, 0.2163650, 0.1865798)$. Those probabilities form two finite schemes, one for each source, and depend upon the flow probability of that source.

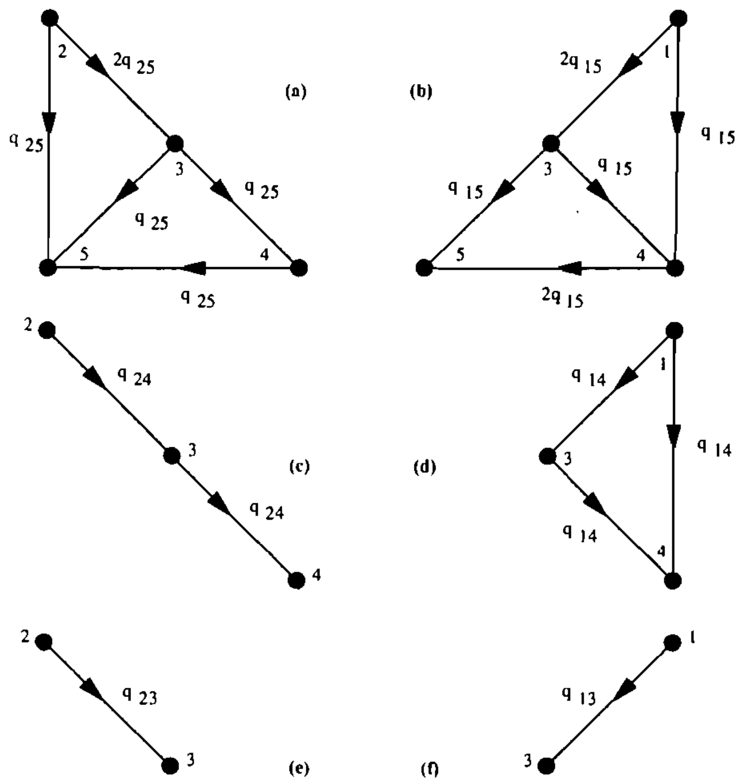


FIGURE 4 Equal path flows from each source to each demand node for the network of Figure 2.

Therefore, the conditional entropy function of Khinchin [6] is appropriate, which has the form:

$$S/K = -\sum_{i\in I} p_{0i} \ln p_{0i} + \sum_{i\in I} p_{0i} S_i \tag{1}$$

where S is the entropy (Shannon [7]), K an arbitrary positive constant, I the set of all source nodes, and p_{0i} the probability associated with the external inflow at source node i given by:

$$p_{0i} = \frac{q_{0i}}{\sum_{i\in I} q_{0i}} = \frac{q_{0i}}{T_0} \quad \forall i \in I \tag{2}$$

where q_{0i} is the external inflow at source node i and T_0 is the total supply or demand. In Eq. (1), S_i is the entropy of all paths supplied by source node i and is given by:

$$S_i = -\sum_{j \in Ji} \mathrm{NP}_{ij} p_{ij} \ln p_{ij} \quad \forall i \in I \tag{3}$$

where J_i is the set of all demand nodes supplied by source i; NP_{ij} is the number of paths from source node i to demand node j; and p_{ij} is the probability of the path flow from source node i to demand node j. Applying Eq. (1) to the network of Figure 2 and the path probabilities calculated above gives a maximum entropy value of 2.3885315 which compares exactly with the entropy value calculated by maximizing the nodal entropy function of the network [8]. This is further confirmation of the general correctness of the approach.

The path probabilities can now be investigated further. For each demand node in the network examine the ratio of the path probabilities to that node from each of the sources. For node 3:

$$p_{13}/p_{23} = 0.1840695/0.2238958 = 0.822121$$

For node 4:

$$p_{14}/p_{24} = 0.1778783/0.2163650 = 0.822121$$

For node 5:

$$p_{15}/p_{25} = 0.1533913/0.1865798 = 0.822121$$

These ratios are **identical** for all demand nodes in the system. Thus it appears that all demand nodes receive their flows from the two sources according to the same principle: the ratio of path probabilities from each source to a demand node is the same for all demand nodes.

The above result has been checked out on a large number of networks with different configurations, numbers of sources and number of demand nodes. It appears to be universally true. Indeed, with the benefit of hindsight, this result is quantitatively what Laplace's principle of insufficient reason might be expected to suggest: the ratios must be the same because there is no reason for them to be different. Note, however, that it is the ratio of path **probabilities** (not path flows) which is the same for all demand nodes, and also that the value of the ratio is not known *a priori*.

The above result forms the basis upon which a solution method for general multi-source, multi-demand networks can be developed. It is formalised in the following statement:

Principle A

The maximum entropy flows in multiple source networks are such that the ratio of the probabilities of path flows from any pair of sources to a demand node reachable from those sources is the same for every demand node supplied by those sources in the network.

No formal proof of Principle A is yet available, but no numerical counter-examples have been found. Principle A contains the concept of the reachability of a demand node from a pair of sources. This will be examined in detail later.

3.2. Development of a Solution Method

Principle A is employed in the proposed solution method. The precise value of the constant ratio of path probabilities is *a priori* unknown. It is therefore given the symbol α, whose value must be determined in the solution. The solution method is consequently named the alfa method. The network of Figure 2 and its path flow representation in Figure 4 is now used to demonstrate how solutions can be obtained.

First, the number of paths from each source to each demand node is calculated using the node numbering algorithm described earlier by considering each source node individually. Therefore, the following values can be obtained: $(NP_{13}, NP_{14}, NP_{15}, NP_{23}, NP_{24}, NP_{25}) = (1, 2, 3, 1, 1, 3)$. The equilibrium equations are then constructed at each demand node by realizing that the sum of all path flows supplying that node must be equal to the demand of that node. Thus:

$$3\,q_{15} + 3q_{25} = 25$$
$$2\,q_{14} + q_{24} = 15$$
$$q_{13} + q_{23} = 10$$

Normalising each path flow by its individual source flow and substituting them gives the above equations in path probability form:

$$90p_{15} + 60p_{25} = 25$$
$$60p_{14} + 20p_{24} = 15$$
$$30p_{13} + 20p_{23} = 10$$

According to Principle A, the following ratios can be identified:

$$\alpha = p_{15}/p_{25} = p_{14}/p_{24} = p_{13}/p_{23}$$

Substituting these ratios into the above equations gives:

$$p_{25} = 25/(60 + 90\alpha) \quad p_{15} = 25\alpha/(60 + 90\alpha)$$
$$p_{24} = 15/(20 + 60\alpha) \quad p_{14} = 15\alpha/(20 + 60\alpha)$$
$$p_{23} = 10/(20 + 30\alpha) \quad p_{13} = 10\alpha/(20 + 30\alpha)$$

At this stage, the normality condition at source node 1 may be written as:

$$\mathrm{NP}_{15}p_{15} + \mathrm{NP}_{14}p_{14} + \mathrm{NP}_{13}p_{13} = 1$$

Substituting the probabilities by their equations yields the following equation:

$$75\alpha/(60 + 90\alpha) + 30\alpha/(20 + 60\alpha) + 10\alpha/(20 + 30\alpha) = 1$$

which can be solved to give: $\alpha = 0.822121$. Note that the normality condition at source node 2 can be used for checking purposes. Finally, back-substituting α gives path probabilities and hence path flows which are exactly the same as those calculated earlier by maximizing the nodal entropy function of the network [8]. The final pipe flows can then be obtained by summing all the path flows passing each pipe in turn. The maximum path entropy value of the network can then be calculated using Eq. (1). It is exactly the same as that given by maximizing the nodal entropy function of the network [8].

The above method is capable of handling any general network with any number of sources. More examples are given later in this paper to demonstrate the general applicability of the method. The example of

Figure 2 had two sources and two sets of path probabilities. In a network with NS number of sources there will be NS sets of path probabilities, and therefore there will be (NS-1) path probability ratios, *i.e.* (NS-1) unknown αs. However, there are NS normality condition equations available in a network, each at a source. Any (NS-1) of these equations can be used to determine all the α values, and hence all the path probabilities, in the network. The remaining normality condition equation can then be used for checking purposes.

The above alfa method for calculating maximum entropy flows in general multi-source water networks has been formalised and cast into algorithms reflecting all the requirements needed in the method. An overview of the algorithms is presented next, followed by a full list of the proposed algorithms.

3.3. Overview of the Proposed Algorithms

It has been noted earlier that the proposed method requires the set of nodes reachable from a source to be identified for each source in the network. Also, the number of paths from each source to each demand node reachable from it has to be determined. The proposed algorithms reflect all these requirements and are described next by means of an example.

Figure 5 shows the two-source network of Figure 2 solved by the proposed algorithms. First, all the network nodes are numbered globally by ascending consecutive positive numbers starting with number 1 for any source in the network, then the rest of the source nodes followed by the rest of the nodes chosen in a random order as shown in Figure 5a. Numbering the two sources is therefore arbitrary and may be interchanged. Also, numbering of demand nodes 3, 4 and 5 may be interchanged.

The next task is to determine the set of nodes reachable from each source in the network. For a source i, all the immediate downstream nodes of that source are stored in a set J_i. The immediate downstream nodes of all the nodes in the set J_i are checked to determine if they already exist in the set before they are added to that set, to avoid double counting of the nodes. The process continues until all the last batch of nodes added to the set are terminal nodes which do not have any link outflows. The set J_i then contains all the node reachable from

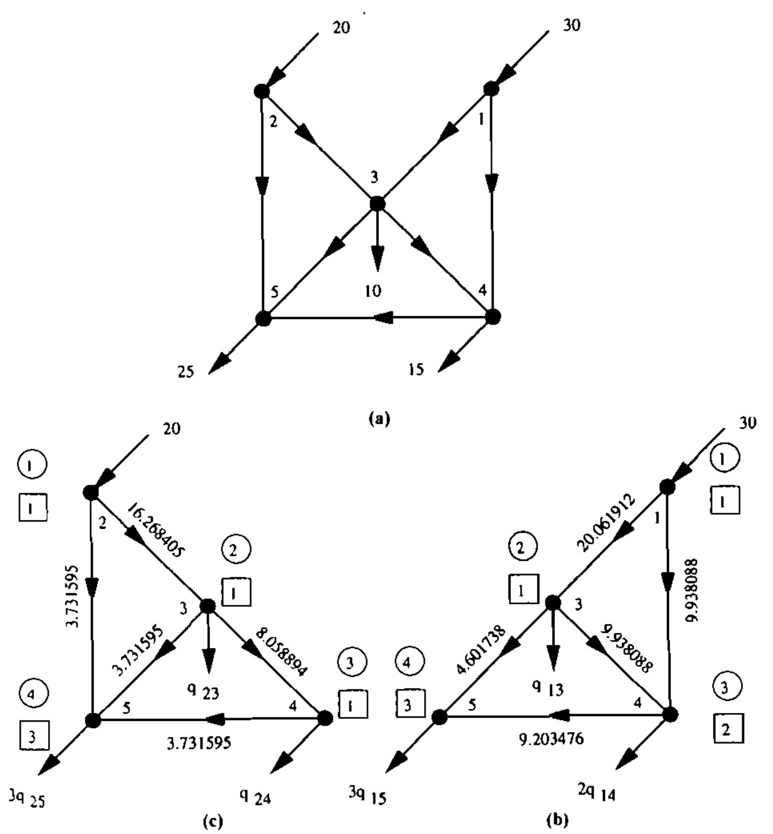


FIGURE 5 Maximum entropy flows for the network of Figure 2. (a) Global node numbering, (b) Sub-network SNK 1, (c) Sub-network SNK 2.

source node i. The above procedure is repeated for every source in the network until all the source nodes have been processed. For the example of Figure 5, nodes 3, 4 and 5 are all reachable from both sources. Neither source is reachable from the other.

Having obtained the set of nodes reachable from each source in the network, the set of source nodes I_j supplying demand node j in the network can easily be determined by arguing that for every demand node j in the network, if node j belongs to the set J_i, then source node i must belong to the set of sources I_j supplying demand node j. In the network of Figure 5, both sources supply all the three demand nodes.

At this stage, the proposed algorithms require NS sub-networks to be constructed, each corresponding to a source and its reachable nodes. The node numbering algorithm and the node weighting algorithm proposed by Tanyimboh and Templeman [9] and described earlier in this paper are then used for each sub-network to determine the number of paths from each source to each node reachable from it. Figures 5b and 5c show the two sub-networks of the network of Figure 5a, each corresponding to a source. For each sub-network, the numbers enclosed in circles next to the nodes are the local node numbers obtained by the single-source node numbering algorithm, and the numbers enclosed in boxes next to the nodes are the number of paths from the source of each sub-network to the corresponding node obtained by using the single-source node weighting algorithm. Note that the local node numbering scheme is used to determine the sequence of nodes to be weighted. However, the global node numbers matching the original numbers in Figure 5a are maintained for convenience.

The next step is to determine the path flow probabilities and their corresponding path flows from each source to each reachable demand node. For this purpose, path flow equilibrium equations at the demand nodes are constructed by equating the sum of all path flows supplying each demand node *i.e.*,

$$\sum_{i \in I_j} \mathrm{NP}_{ij}\, q_{ij} = q_{j0} \quad \forall j \in J \tag{4}$$

where I_j is the set of all source nodes supplying node j; J is the set of all demand nodes in the network; and q_{j0} is the demand at node j. Equation (4) can be expressed in terms of path flow probabilities by normalising each path flow by its individual source flow to give the following probability equilibrium equations:

$$\sum_{i \in I_j} \mathrm{NP}_{ij}\, q_{0i}\, p_{ij} = q_{j0} \quad \forall j \in J \tag{5}$$

in which q_{0i} is the external inflow at source node i; p_{ij} is the probability that path flow q_{ij} supplied by source node i reaches demand node j.

According to Principle A, the ratio of the probabilities of path flows from each pair of sources to a demand node reachable from this pair of sources is the same for every demand node supplied by both sources. In general, there are (NS-1) such unknown ratios in a network, where NS is the number of sources in the network. These ratios can be formalised by the following general equation:

$$\frac{p_{ij}}{p_{kj}} = \frac{\alpha_i}{\alpha_k} \quad \forall i,k \in I_j;\ \forall j \in J \tag{6}$$

where i and k are any pair of sources supplying demand node j in the network. The number of αs in Eq. (6) is obviously NS. Because the number of unknown αs required in a network is (NS-1), one of the αs in Eq. (6) has to be set to unity. In this paper, the unknown α corresponding to source node 1 is set to unity, *i.e.*, $\alpha_1 = 1$. Substituting Eq. (6) into Eq. (5) enables path flow probabilities to be expressed as functions of the (NS-1) unknown αs. The general formula for path flow probabilities expressed in terms of unknown αs can be obtained by substituting Eq. (6) into Eq. (5) and rearranging the resulting equation to give:

$$p_{ij} = \frac{q_{j0}\ \alpha_i}{\sum_{i \in I_j} \mathrm{NP}_{ij}\, q_{0i}\, \alpha_i} \quad \forall i \in I_j;\ \forall j \in J \tag{7}$$

in which $\alpha_i = 1$ for $i = 1$. Equation (7) contain therefore (NS-1) unknown αs. However, there are NS normality condition equations which can be constructed in a network.

They are:

$$\sum_{j \in J_i} \mathrm{NP}_{ij}\, p_{ij} = 1 \quad \forall i \in I \tag{8}$$

If path flow probabilities of Eq. (7) are substituted into (NS-1) of the normality condition equations of Eq. (8) there will be (NS-1) equations with (NS-1) unknown αs, which can be solved to determine all the values of the unknown αs. Back-substituting the resulting values of αs into Eq. (7) gives all the path flow probabilities in the network and hence their corresponding path flows which are given by:

$$q_{ij} = q_{0i}\, p_{ij} \quad \forall i \in I;\ \forall j \in J_i \tag{9}$$

Having determined all the path flows in a network, the final link flows which correspond to the maximum entropy value of the network can then be calculated by handling each sub-network in turn and calculating the flows in the links of the sub-network supplied by the corresponding source. All the flows carried by each link in all the sub-networks can finally be superposed to give the final maximum entropy flows in the network.

Consider the two sub-networks of Figures 5b and 5c which respectively correspond to source nodes 1 and 2 of the network of Figure 5a. The demands at the nodes are the flows supplied by the corresponding source, *i.e.*,

$$q_{j0,i} = \mathrm{NP}_{ij}\, q_{ij} \quad \forall i \in I;\ \forall j \in J_i \tag{10}$$

in which $q_{j0,i}$ is the demand at node j supplied by source node i. The total outflows at a node in any sub-network is distributed amongst all the immediate upstream links to that node in proportion with the weight of the upstream node to the corresponding link. The procedure operates from any terminal node in the sub-network in a descending local numbering order to ensure that the link outflows at the node have all been considered. Starting at the terminal node 5 in the sub-network of Figure 5b, the total outflow at node 5 equalling 13.805214 units is divided by 3, this being the number of paths from source node 1 to that node. The quotient is then multiplied by 1 and 2 respectively, these being the respective number of paths of nodes 3 and 4 which are the immediate upstream nodes to node 5. The products, respectively, are the flows in links 3–5 and 4–5 supplied by source 1. At this stage, the total outflow at node 4 which is equal to 9.203476 + 10.6727 = 19.876176 units is shared equally between links 1–4 and 3–4 due to the equality of the weights of the immediate upstream nodes to these two links. Finally, all the outflows at node 3 are carried by link 1–3 since there is only one link upstream to node 3.

The same procedures are applied for the sub-network of Figure 5c. The resulting link flows supplied by source 2 are shown in the figure. The final link flows can finally be obtained by summing the flows in Figures 5b and 5c for each link in turn. The resulting link flows have been shown to match the maximum entropy flows shows in Figure 3.

The following is a list of simple algorithms proposed to tackle the path-based problem of calculating maximum entropy flows in multi-source, multi-demand general networks. In these algorithms, NM is the number of nodes in a network and NS is the number of its sources. The rest of the symbols used in the algorithms are defined within the algorithms. Also, the global node numbering scheme is used throughout the algorithms unless stated otherwise.

Global Node Numbering Algorithm

1. Select any source in the network and number it with 1. Set n to 1.
2. Increase n by 1. Select any other source which has not been numbered and number it with n.
3. If n = NS go to step 4. Otherwise, go to step 2.
4. Increase n by 1. Select any node which has not been numbered and number it with n.
5. If n = NN, exit. Otherwise, go to step 4.

Source Reachability Algorithm

1. Set n to 0.
2. Increase n by 1. Select source node n.
3. Select all the nodes immediately downstream of node n and include their node numbers in an empty set J_n.
4. Select all the nodes immediately downstream of the nodes contained in the set J_n. The node number of each node selected, if any, is added to the set J_n if it does not already exist in the set.
5. If the number of nodes added to the set J_n in step 4 equals zero, go to step 6. Otherwise, go to step 4.
6. The set J_n contains all the nodes reachable from source node n. If n = NS, exit. Otherwise go to step 2.

Demand Node Reachability Algorithm

1. Set n to NS.
2. Increase n by 1. Select demand node n.
3. Set m to 1.
4. If the set J_m which is identified by the source reachability algorithm contains node number n, include m in set I_n.
5. Increase m by 1. If m = NS, go to step 6. Otherwise, go to step 4.
6. The set I_n contains all the source nodes supplying demand node n. If n = NN, exit. Otherwise, go to step 2.

Local Node Numbering Algorithm

1. Set n to 0.
2. Increase n by 1. Select source node n.
3. Construct a sub-network SNK_n containing source node n and all nodes contained in the set J_n.
4. Number the source node n locally with 1. Set m to 1.
5. Increase m by 1. Select any node in the sub-network SNK_n whose immediate upstream nodes have all been locally numbered and number it with m.
6. If m equals the number of nodes contained in the sub-network SNK_n, go to step 7. Otherwise, go to step 5.
7. If n = NS, exit. Otherwise, go to step 2.

Node Weighting Algorithm

1. Set n to 0.
2. Increase n by 1. Select the sub-network SNK_n corresponding to source node n.
3. Set m to the source local number, 1. Set NP_{nm} to 1.
4. Increase m by 1 and select the node whose local number in the sub-network SNK_n is m.
   Calculate NP_{nm}:

   $$NP_{nm} = \sum_{j \in NU_m \subset SNK_n} NP_{nj}$$

   in which $NU_m \sqsubset SNK_n$ represents the set of upstream nodes of inflow links at a node whose local number in the sub-network SNK_n is m, provided that these upstream nodes are in the sub-network SNK_n.
5. If m equals the number of nodes contained in the sub-network SNK_n, go to step 6. Otherwise, go to step 4.
6. If n = NS, exit. Otherwise, go to step 2.

Alfa Algorithm

1. Set n to NS.
2. Increase n by 1. Select demand node n.

3. Define the probabilities of path flows supplying demand node n from each source reachable to it as follows:

$$p_{in} = \frac{q_{n0}\alpha_i}{\sum_{i \in I_n} \mathrm{NP}_{in}\, q_{0i}\, \alpha_i} \quad \forall i \in I_n$$

4. If n = NN, go to step 5. Otherwise, go to step 2.
5. Set n to 0.
6. Increase n by 1. Select source node n.
7. Construct the normality condition equation at source node n as follows:

$$\sum_{j \in J_n} \mathrm{NP}_{nj}\, p_{nj} = 1$$

8. If n = NS-1, go to step 9. Otherwise, go to step 6.
9. Set $\alpha_1 = 1$.
10. Solve the equations constructed in steps 3, 7 and 9 to obtain all the unknown values of αs and all the path flow probabilities. Calculate S/K, if necessary.
11. Calculate the path flows as follows:

$$q_{ij} = q_{0i}\, p_{ij} \;\; \forall i \in I; \forall j \in J_i$$

12. Exit.

Flow Distribution Algorithm

1. Set n to 0.
2. Increase n by 1. Select source node n.
3. Set m to the number of nodes in the sub-network SNK_n including the source node n.
4. Calculate $T_{m,n}$:

$$T_{m,n} = \mathrm{NP}_{nm}\, p_{nm} + \sum_{k \in \mathrm{ND}_m \subset \mathrm{SNK}_n} Q_{mk,n}$$

in which m indicates the node whose local number in the sub-network SNK_n is m; $T_{m,n}$ is the total outflows at node m; $(\mathrm{NP}_{nm} p_{nm})$ is the local demand at node m supplied by source n; $\mathrm{ND}_m \sqsubset \mathrm{SNK}_n$ represents the set of immediate downstream nodes of inflow links at node m, provided that these downstream nodes are in the sub-network SNK_n; $Q_{mk,n}$ is the flow in link mk in the sub-network SNK_n supplied by source node n. Note that the total demand at the node whose local number in the sub-network SNK_n is m is not included in $T_{m,n}$

5. Calculate $Q_{jm,n}, \forall jm \in \mathrm{NU}_m \sqsubset \mathrm{SNK}_n$:

$$Q_{jm,n} = T_{m,n} \frac{\mathrm{NP}_{nj}}{\mathrm{NP}_{nm}}$$

in which $Q_{jm,n}$ is the flow in link jm in the sub-network SNK_n, $\forall jm\,\mathrm{NU}_m \sqsubset \mathrm{SNK}_n$, supplied by source node n.

6. Reduce m by 1. If $m = 1$, go to step 7. Otherwise, go to step 4.
7. If $n = \mathrm{NS}$, go to step 8. Otherwise, go to step 2.
8. Calculate $Q_{ij}, \forall ij \in IJ$:

$$Q_{ij} = \sum_{n=1}^{\mathrm{NS}} Q_{ij,n}$$

in which Q_{ij} is the final maximum entropy flow in link ij; $Q_{ij,n}$ is the flow in the corresponding link calculated in the sub-network SNK_n in which the link ij is a member.

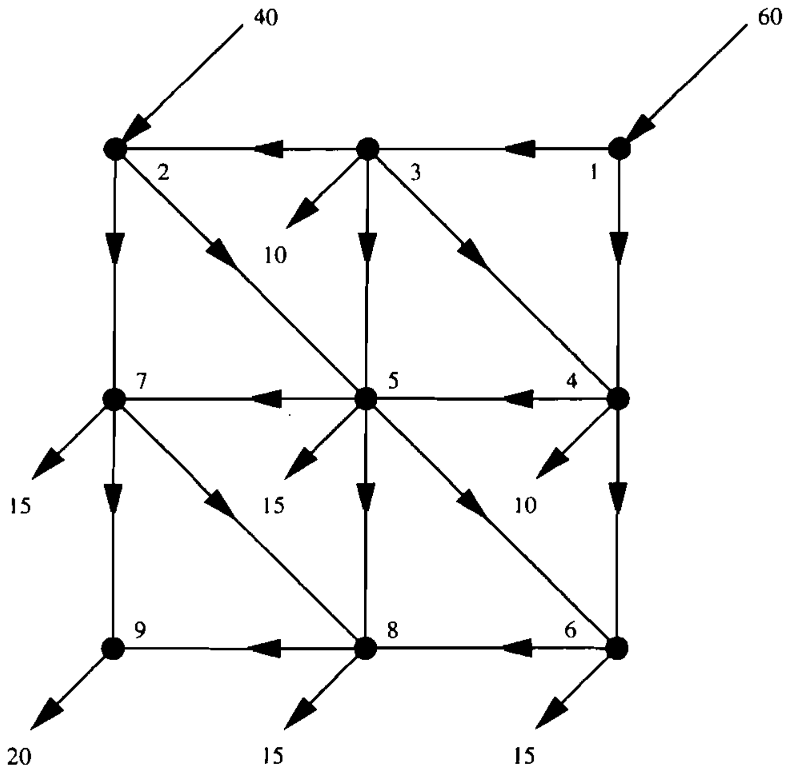
Exit.

4. NUMERICAL EXAMPLES

Figures 6 and 7 show two different general networks which are solved by the proposed path-based algorithms in order to obtain their maximum entropy flows. The two examples have been chosen carefully to demonstrate the general applicability of the proposed method to networks with different configurations and reachability conditions.

Example 1

The two-source network of Figure 6 used in this example has 9 nodes and 16 links. The network has the characteristic that some demand nodes are unreachable from one of the sources. First, the network nodes are numbered globally using the global node numbering algorithm. This is done by numbering the two sources first by 1 and 2, followed by the rest of the nodes in a random order as shown in Figure 6a.

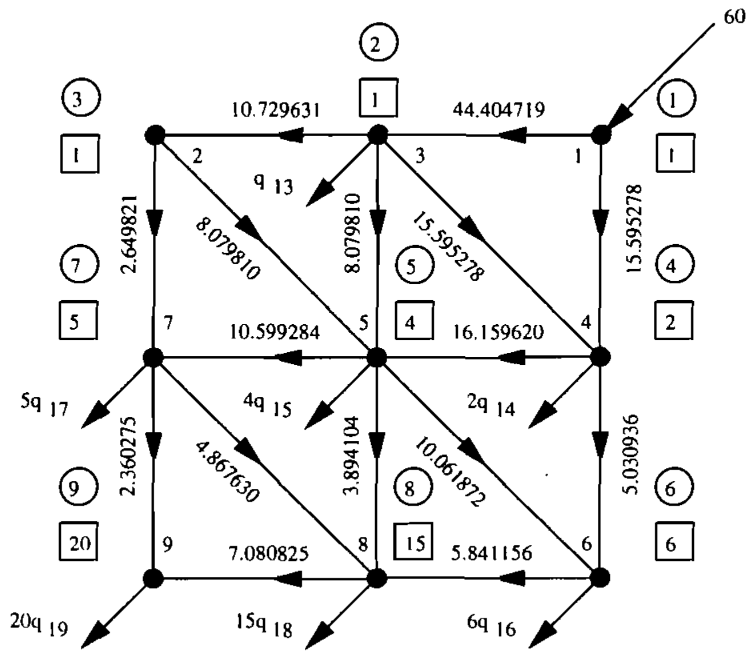


(a)

FIGURE 6 Maximum entropy flows for the network of Example 1. (a) Global node numbering, (b) Sub-network SNK 1, (c) Sub-network SNK 2, (d) Final link flows.

The source reachability algorithm and the demand reachability algorithm are then used to identify the set of nodes reachable from each source ($J_i, \forall i \in I$) and the set of source nodes supplying each demand node (I_j, $\forall j \in J$) respectively. Accordingly, the following sets can be obtained:

$$J_1 = \{3,4,2,5,6,7,8,9\}$$
$$J_2 = \{5,7,6,8,9\}$$
$$I_3 = I_4 = \{1\}$$
$$I_5 = I_6 = I_7 = I_8 = I_9 = \{1,2\}.$$

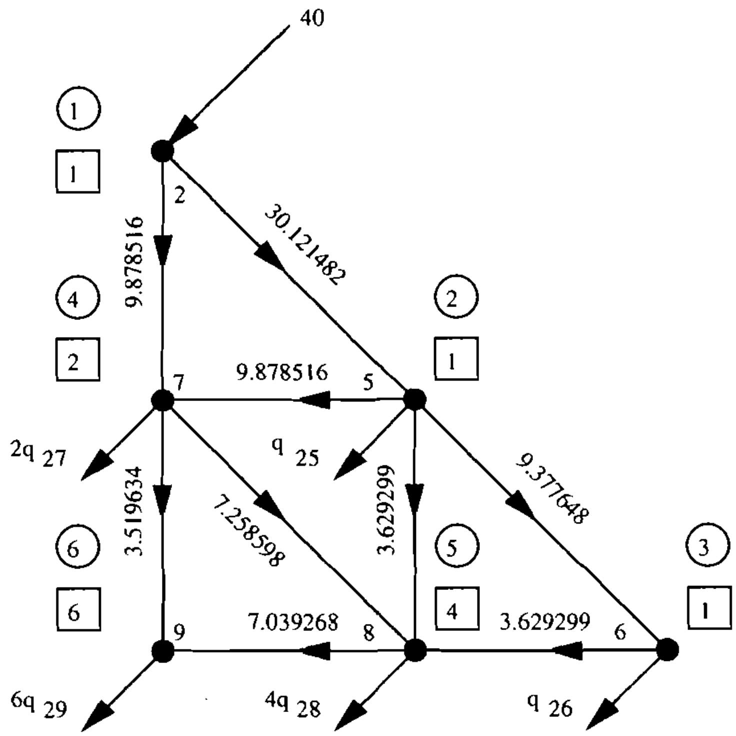


(b)

FIGURE 6 (Continued).

At this stage, the two sub-networks of Figures 6b and 6c, each corresponding to a single source and its reachable nodes, can now be constructed to determine the number of paths from each source to each demand node reachable from that source. For this purpose, the nodes for each sub-networks are first renumbered locally using the local node numbering algorithm, and then the number of paths from the corresponding source to each node in the sub-network is determined using the node weighting algorithm. Figures 6b and 6c show the local number for each node in each sub-network and the number of paths from the corresponding source enclosed in a circle and a box next to that node respectively.

Having defined the number of paths from each source to each reachable demand node, the path flow probabilities can now be

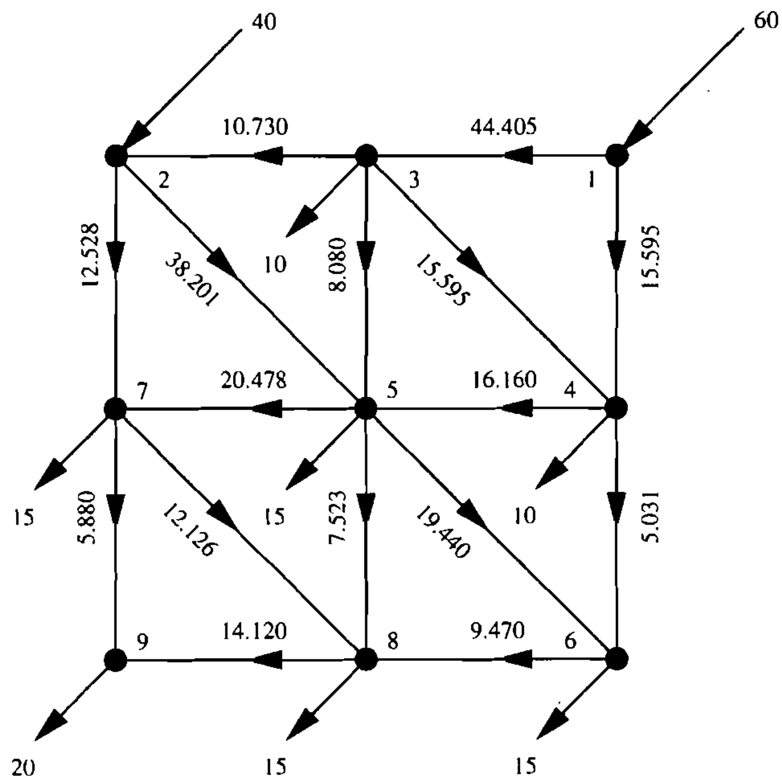


(c)

FIGURE 6 (Continued).

expressed in terms of αs using the alfa algorithm as follows:

$$p_{13} = (10\alpha_1/60\alpha_1) = 10/60 = 0.1666667$$
$$p_{14} = (10\alpha_1/12\alpha_1) = 10/120 = 0.0833333$$
$$p_{15} = 15\alpha_1/(240\alpha_1 + 40\alpha_2),\ p_{25} = 15\alpha_2/(240\alpha_1 + 40\alpha_2)$$
$$p_{16} = 15\alpha_1/(360\alpha_1 + 40\alpha_2),\ p_{26} = 15\alpha_2/(360\alpha_1 + 40\alpha_2)$$
$$p_{17} = 15\alpha_1/(300\alpha_1 + 80\alpha_2),\ p_{27} = 15\alpha_2/(300\alpha_1 + 80\alpha_2)$$
$$p_{18} = 15\alpha_1/(900\alpha_1 + 160\alpha_2),\ p_{28} = 15\alpha_2/(900\alpha_1 + 160\alpha_2)$$
$$p_{19} = 20\alpha_1/(1200\alpha_1 + 240\alpha_2),\ p_{29} = 20\alpha_2/(1200\alpha_1 + 240\alpha_2)$$



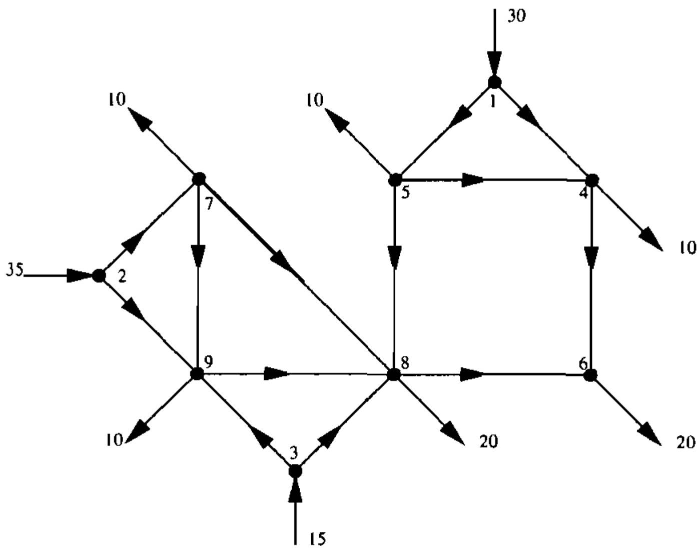
(d)

FIGURE 6 (Continued).

Also, the normality condition equation at source node 1 can be constructed as:

$$p_{13} + 2p_{14} + 4p_{15} + 6p_{16} + 5p_{17} + 15p_{18} + 20p_{19} = 1$$

which can be solved by substituting the above path flow equations into it and by setting α_1 to unity to give: $\alpha_2 = 5.5919906$. Back-substituting this value of α into the path flow probability equations gives all the path flow probabilities in the network, which can then be used to determine their corresponding path flows by multiplying each path flow probability by its individual source flow. All the resulting path flow probabilities and their corresponding path flows



(a)

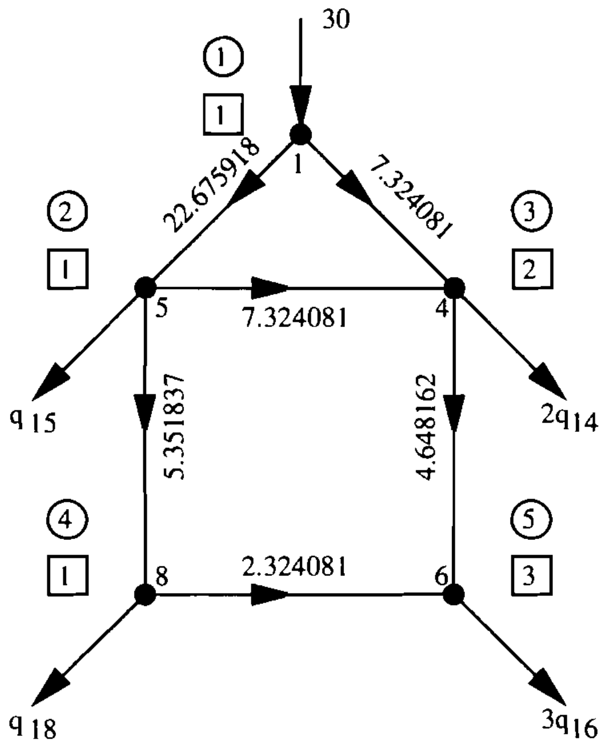
FIGURE 7 Maximum entropy flows for the network of Example 2. (a) Global node numbering, (b) Sub-network SNK 1, (c) Sub-network SNK 2, (d) Sub-network SNK 3, (e) Final link flows.

are given in Table I. The path entropy value of the network can now be calculated using Eq. (1) to give: $S/K = 3.748957$ which, corresponds exactly to the computed maximum nodal entropy value of the network flows.

The final part of the proposed method is to calculate the final link flows in the network using the flow distribution algorithm. Figures 6b and 6c show the resulting link flows supplied by source nodes 1 and 2 respectively. The final link flows in the network can be determined by superposing the link flows of the two sub-networks. The resulting final flows are shown in Figure 6d. They match the maximum entropy flows obtained by maximizing the nodal entropy function of the network [8].

Example 2

The network of this example shown in Figure 7 is a three-source network having 9 nodes and 13 links but with no source which reaches

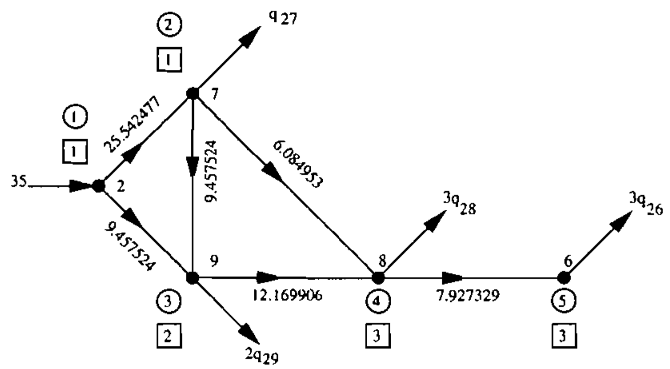


(b)

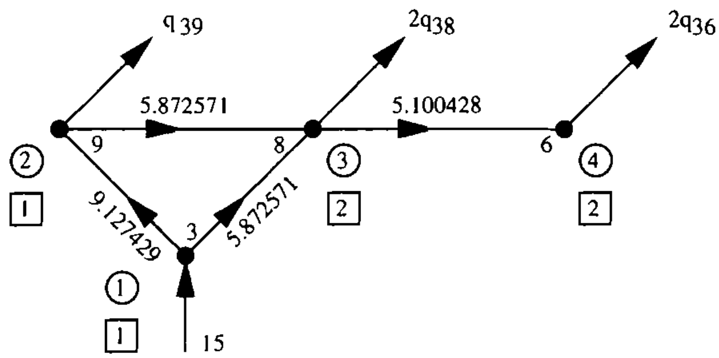
FIGURE 7 (Continued).

all demand nodes. The network nodes are globally numbered as shown in Figure 7a, and the following sets represent the sets of nodes reachable from each source and the sets of source nodes supplying each demand node respectively:

$$J_1 = \{4, 5, 6, 8\}$$
$$J_2 = \{7, 9, 8, 6\}$$
$$J_3 = \{9, 8, 6\}$$
$$I_4 = I_5 = \{1\}$$
$$I_6 = I_8 = \{1, 2, 3\}$$
$$I_7 = \{2\}$$
$$I_9 = \{2, 3\}.$$



(c)

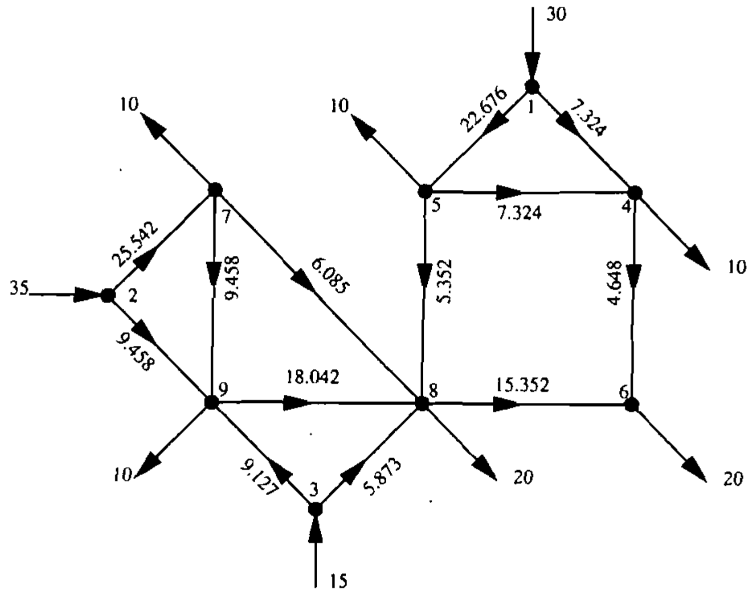


(d)

FIGURE 7 (Continued).

The three sub-networks, each corresponding to a source and its reachable nodes, can now be constructed as shown in Figures 7b, 7c and 7d with the local node numbers and the number of paths from the corresponding source to each node in each sub-network shown in circles and boxes next to the nodes. The next step is to set up the following path flow probability equations:

$$p_{14} = 10\alpha_1/60\alpha_1 = 10/60 = 0.1666667$$
$$p_{15} = 10\alpha_1/30\alpha_1 = 10/30 = 0.3333333$$
$$p_{16} = 20\alpha_1/(90\alpha_1 + 105\alpha_2 + 30\alpha_3)$$
$$p_{26} = 20\alpha_2/(90\alpha_1 + 105\alpha_2 + 30\alpha_3)$$



(e)

FIGURE 7 (Continued).

$$p_{36} = 20\alpha_3/(90\alpha_1 + 105\alpha_2 + 30\alpha_3)$$
$$p_{27} = 10\alpha_2/35\alpha_2 = 10/35 = 0.2857143$$
$$p_{18} = 20\alpha_1/(30\alpha_1 + 105\alpha_2 + 30\alpha_3)$$
$$p_{28} = 20\alpha_2/(30\alpha_1 + 105\alpha_2 + 30\alpha_3)$$
$$p_{38} = 20\alpha_3/(30\alpha_1 + 105\alpha_2 + 30\alpha_3)$$
$$p_{29} = 10\alpha_2/(70\alpha_2 + 15\alpha_3),\ p_{39} = 10\alpha_3/(70\alpha_2 + 15\alpha_3).$$

By setting α_1 to unity, there are two unknown αs in the above equations needing two normality condition equations to be constructed, which by considering source nodes 1 and 2, are:

$$2p_{14} + p_{15} + 3p_{16} + p_{18} = 1$$
$$3p_{26} + p_{27} + 3p_{28} + 2p_{29} = 1$$

Substituting the path flow probability equations into the above two normality condition equations and solving the resulting equations

TABLE I The path flow probabilities and the corresponding path flows for the two-source network of Figure 6 calculated by the alfa method

Source node i	q_{0i}	*Demand node j*	q_{j0}	*Path i–j*	NP_{ij}	p_{ij}	q_{ij}
		3	10	1–3	1	0.1666667	10.000000
		4	10	1–4	2	0.0833333	5.000000
		5	15	1–5	4	0.0323499	1.940995
1	60	6	15	1–6	6	0.0256990	1.541942
		7	15	1–7	5	0.0200707	1.204240
		8	15	1–8	15	0.0083579	0.501471
		9	20	1–9	20	0.0078676	0.472055
		5	15	2–5	1	0.1809005	7.236019
		6	15	2–6	1	0.1437087	5.748349
2	40	7	15	2–7	2	0.1122350	4.489400
		8	15	2–8	4	0.0467371	1.869482
		9	20	2–9	6	0.0439954	1.759817

gives: $\alpha_1 = 0.9745576$ and $\alpha_3 = 2.1945998$, which can be checked using the normality condition equation at source node 3. Back-substituting the above values of αs into the path flow probability equations gives all the path flow probabilities and hence their corresponding path flows in the network. The results are given in Table II with maximum path entropy value of 2.923974.

The final stage is to calculate the final link flows in the network using the flow distribution algorithm. Accordingly, the link flows for each sub-network supplied by the corresponding source are first determined as shown in Figures 7b, 7c and 7d respectively. The resulting

TABLE II The path flow probabilities and the corresponding path flows for the three-source network of Figure 7 calculated by the alfa method

Source node i	q_{0i}	*Demand node j*	q_{j0}	*Path i–j*	NP_{ij}	p_{ij}	q_{ij}
		4	10	1–4	1	0.1666667	5.000000
1	30	5	10	1–5	1	0.3333333	10.000000
		6	20	1–6	3	0.0774694	2.324081
		8	20	1–8	1	0.1009252	3.027756
		6	20	2–6	3	0.0754984	2.642443
2	35	7	10	2–7	1	0.2857143	10.000000
		8	20	2–8	3	0.9835740	3.442510
		9	10	2–9	2	0.0963592	3.372571
		6	20	3–6	2	0.1700143	2.550214
3	15	8	20	3–8	2	0.2214904	3.322357
		9	10	3–9	1	0.2169906	3.254858

sub-network link flows can then be superposed to give the maximum entropy link flows in the networks shown in Figure 7e.

5. DISCUSSION

Comments on the Values of α

The various values of α in the two examples are worthy of discussion and interpretation. In Example 1 setting $\alpha_1 = 1$ yields $\alpha_2 = 5.5919906$, *i.e.*, path probabilities from source 2 to a demand node are more than five times as large as path probabilities from source 1. Why is this so? Examination of Figures 6b and 6c shows that for any demand node there are considerably more paths to it from source 1 than from source 2. It appears that the high value of α_2 has the effect of magnifying the small number of paths from source 2 so that they are roughly equivalent to those from source 1. α_2 is of course aggregated over the entire network and over all nodes so that for any specific node exact uniformity is not achieved from both sources. Nevertheless, over the whole network the αs seem to be compensating for lack of uniformity in the network configuration and supply/demand regimes.

Example 2 seems to confirm this. Source 3 supplies very few nodes through very few paths so α_3 has a high value (2.1945988). Sources 1 and 2 supply more nodes through a greater number of paths. Consequently, α_1 and α_2 have similar values (1 and 0.9745576) of lower magnitude than α_3.

Comments on the Super-source Approach [9]

The super-source approach proposed by Tanyimboh and Templeman [9] as a possible extension of their single-source method to multiple sources is investigated next and compared with the proposed method. For this purpose, the two examples solved in Ref. [9] by the super-source approach are resolved in this paper by the alfa method. Figures 8 and 9 show these two examples with link flows for each example obtained by both the super-source approach and the alfa method.

In the first example, the super-source results shown in Figure 8a were obtained by replacing the two sources by one super-source, and then applying the single source method to the resulting super-source network. The resulting link flows have an entropy value of 2.367. The link flows shown in Figure 8b and obtained by the alfa method, however, have a higher entropy value of 2.45 suggesting that the results of Figure 8a are not optimal and the results of Figure 8b should be used instead.

The second example shown in Figure 9 uses the same network of Figure 8 but with the flow direction in link $1-2$ reversed. The link flows obtained by the super-source approach and shown in Figure 9a have an entropy value of 1.947 which is less than the entropy value of 2.154 corresponding to the link flows of Figure 9b which are obtained by the alfa method.

It is evident from the above two examples that the super-source approach of Tanyimboh and Templeman [9] does not produce optimal results, as Walters [11] has pointed out. The approach seems to ignore some path flows supplied by one source through the other source. This can be seen from the equal flows assigned to links $1-3$ and $2-3$ in Figure 8a. Although link $2-3$ carries six path flows, three of them are the same path flows carried by link $1-3$ and the other three path flows are those corresponding to source node 2. Also, the zero flow in link $2-1$ assigned by the super-source approach for the network of Figure 9a suggests that the path flows supplied by source node 2 through source node 1 have all been ignored. Consequently, the super-source approach of Tanyimboh and Templeman [9] for calculating maximum entropy flows in multi-source networks is invalid. The alfa method and its algorithms presented in this paper represent the correct way of solving these problem.

6. SUMMARY AND CONCLUSION

A rigorous and simple path-based method for calculating maximum entropy flows in multi-source, multi-demand general networks has been presented in this paper along with simple algorithms tackling all aspects encountered in the method. Two principles underlie the method: (1) the demand of any node in the network served by more than one path from any source should be divided equally amongst all paths supplying that demand node from that source; and (2) the maximum entropy flows in multi-source, multi-demand general networks are such that the ratio of the probabilities of path flows from

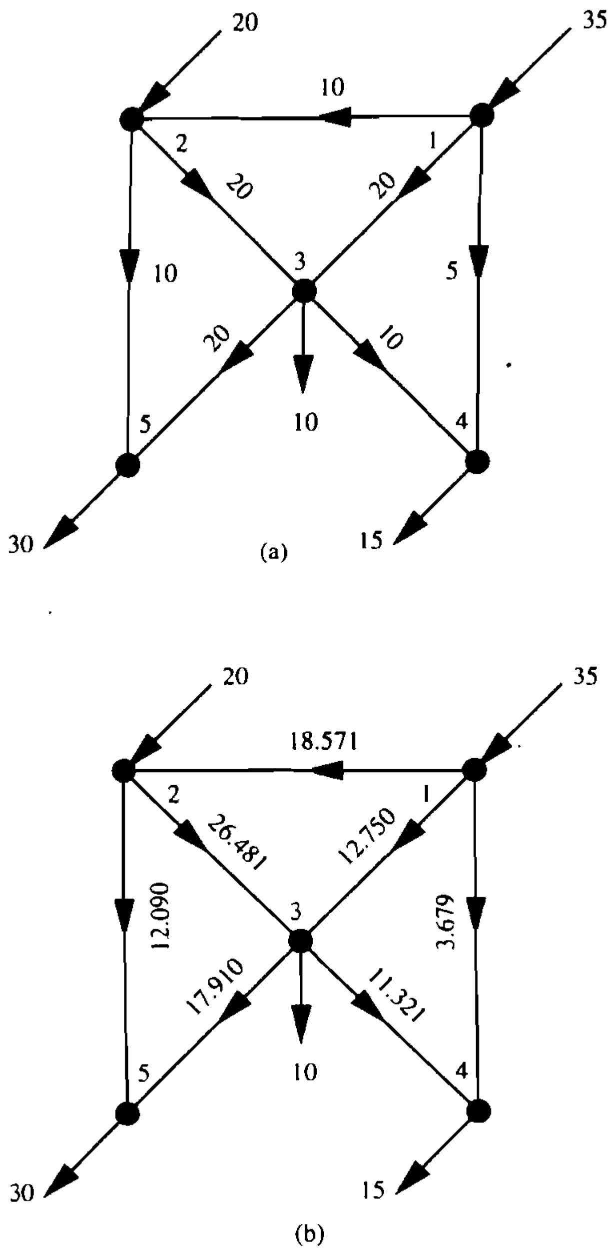


FIGURE 8 Maximum entropy flows for a two-source network adapted from Ref. [9]. (a) Ref. [9] results, (b) The alfa method results.

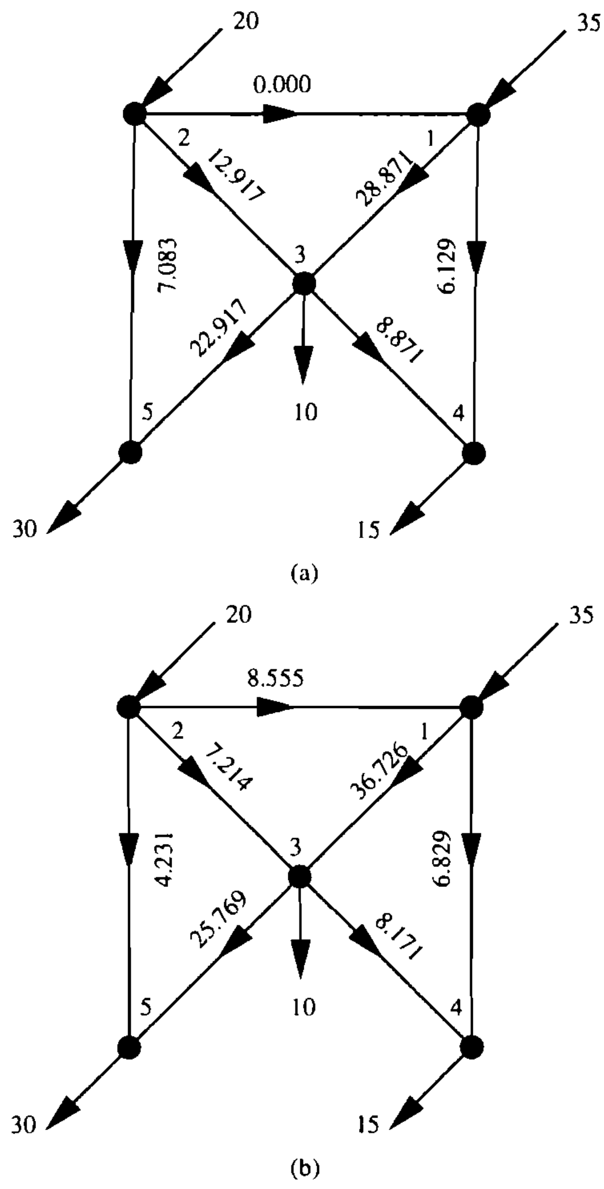


FIGURE 9 Maximum entropy flows for the two-source network of Figure 8 having the flow direction in link 1 − 2 reversed. (a) Ref. [9] results, (b) The alfa method results.

each pair of sources to a demand node reachable from this pair of sources is the same for every demand node supplied by this pair of sources in the network.

The proposed method is very quick and does not involve any mathematical programming. Also, it is not iterative and does not use the super-source approach in any way. Because the method is path-based, a network path entropy function has been developed to facilitate the proposed method using the conditional entropy formula for compound probability schemes.

Two examples exhibiting different characteristics which can be expected in general real networks have been solved by the proposed algorithms to test the general applicability of the method. It has been shown that the proposed algorithms are very efficient and easy to operate by hand calculations especially for small networks such as those used in the examples. Also, it has been demonstrated that the path flow probability ratios, α, defined by the method, are a way of making the flows supplied by each source to a reachable demand node as uniform as possible subject to available information. They depend on the supply and demand of the network nodes and are influenced by the layout of the network.

Also, in this paper, the super-source proposed by Tanyimboh and Templeman [9] as an extension of the single-source method to multiple sources for calculating maximum entropy flows in water networks has been found to be invalid and actually leads to less than optimal results. The proposed alfa method always gives optimal results and should be used instead.

Finally, the efficiency and simplicity of the proposed method, taking into account the applicability of flow entropy as a surrogate measure of reliability as established in Refs. [4] and [10], open up the possibility of a simple and quick method for designing reliable water distribution networks. If the maximum entropy flows calculated by the proposed path-based method are used in the linear programming phase of the method of Alperovits and Shamir [1], the gradient phase will not be needed and the problem of designing optimum and reliable water distribution networks becomes non-iterative and requires only linear programming.

References

[1] Alperovits, E. and Shamir, U. (1977). Design of optimal water distribution systems. *Water Resources Research*, **13**(6), 885–900.

[2] Awumah, K. and Goulter, I. (1992). Maximizing entropy-defined reliability of water distribution networks. *Engineering Optimization*, **20**(1), 57–80.

[3] Awumah, K., Goulter, I. and Bhatt, S. K. (1990). Assessment of reliability in water distribution networks using entropy-based measures. *Stochastic Hydrology and Hydraulics*, **4**(4), 325–336.

[4] Awumah, K., Goulter, I. and Bhatt, S. K. (1991). Entropy-based redundancy measures in water distribution network design. *ASCE J. Hydraulic Engineering*, **117**(5), 595–614.

[5] Jaynes, E. T. (1957). Information theory and statistical mechanics. *Phys. Rev.*, **106**, 620–630 and **108**, 171–190.

[6] Khinchin, A. I. (1953). The entropy concept in probability theory. *Uspekhi Mathematicheskikh Nauk*, **8**(3), 3–20. Translation in Khinchin, A. I. (1957). Mathematical foundations of information theory. Dover, New York, 1–28.

[7] Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical J.*, **27**(3), 379–423.

[8] Tanyimboh, T. T. and Templeman, A. B. (1993). Calculating maximum entropy flows in networks. *J. Operational Research Soc.*, **44**(4), 383–396.

[9] Tanyimboh, T. T. and Templeman, A. B. (1993). Maximum entropy flows for single-source networks. *Eng. Opt.*, **22**(1), 49–63.

[10] Tanyimboh, T. T. and Templeman, A. B. (1993). Optimum design of flexible water distribution networks. *Civil Engineering Systems*, **10**(3), 243–258.

[11] Walters, G. A. (1995) Discussion on "Maximum entropy flows in single-source networks". *Eng. Opt.*, **25**(2), 155–163.

[12] Yassin-Kassab, A. (1998). Entropy-based inference and calibration methods for civil engineering system models under uncertainty. *Ph.D. Thesis*, University of Liverpool.