THE VARIANCE LOCATION PROBLEM ON A NETWORK WITH CONTINUOUSLY DISTRIBUTED DEMAND (*)

by Mª Cruz López De Los Mozos ([1]) and Juan A. Mesa ([2])

Communicated by Brian Boffey

Abstract. – *Most location problems on networks consider discrete nodal demand. However, for many problems, demands are better represented by continuous functions along the edges, in addition to nodal demands. Several papers consider the optimal location problem of one or more facilities when demands are continuously distributed along the network, and the objective function dealt with is the median one. Nevertheless, in location of public services it is desirable to use an equity criterion. One of the latter is variance of distance distribution which has been studied only for discrete nodal demands. In this paper the variance problem has been generalized to the case where one allows the demand to arise discretely on the nodes as well as continuously along the edges. Properties and behaviour of the objective function are studied. Likewise we present an exact algorithm for solving this problem in a network, which reduces the complexity of the exhaustive procedure.*

1. INTRODUCTION

Most location problems on networks developed since the work of Hakimi [8, 9] assume discrete nodal demand, in which customer demands originate solely at the vertices of the network. However, as pointed out by several authors [3, 5, 15] in many real applications demands do not occur only at the vertices, but also along the edges. Restriction of demands to the vertices quite often is not a satisfactory approximation (see [5, 15]).

Some real world applications corresponding to this situation are the location of emergency or public services, or utility repair stations along a motorway. In these cases, demands are better represented by continuous functions along the edges in addition to nodal demands. As quoted in [3], the resulting problem of approximating internodal demands by a number of

artificial nodes can be intractable. More recent contributions have attempted to accommodate this feature (see [3-5, 15, 18]).

In all formulations concerned with continuous edge demand the objective is the minimisation of the weighted sum of distances from the facility to all points of the network, this being the median function. Such problems are called continuous median problems by Labbé [15].

Several results concerning the continuous median problem can be found in [5, 15, 18]. The concavity of the objective function along any edge contained in a cycle when the demand density is uniform is proved in [18]. In [5] the continuous median function is characterised in terms of the location of the service facility, and a linear algorithm is developed to solve the problem on tree networks. This algorithm is simplified in [15].

Other extensions of the problem involving continuous demands on edges are oriented to find the 2-median of a tree network [3], or solve the p-median problem on a chain graph [4].

The criterion used in the above formulations for the selection of optimal locations is one of various that may be employed to obtain efficiency. However, in many settings, most notably the public sector, this criterion is insufficient to generate acceptable decisions. More recently, increasing attention has been paid to equity aspects of location. This gives rise to several new location problems, in which an equity criterion is used based on the dispersion of the distance distribution from the facility to all users.

The introduction of equity measures into location theory was first discussed in [10], where two of them were described: the variance of distance travelled by all customers to the facility and the Lorenz curve. Maimon [17] proposed an $O(n)$ time algorithm to minimise the variance measure on a tree network whose demand occurs only at the n vertices. Kincaid and Maimon [13, 14] studied variance minimisation problems in triangular and in 3-cactus graphs, and Hansen and Zheng [11] have presented an $O(mn\log n)$ time algorithm for the variance problem in general networks.

The demand used in the model considered in the aforementioned problems is restricted to nodal locations. However, following the initial reasoning, in many applications, an equity criterion may be needed to be considered as the variance in a continuous demand context along the edges.

In this paper we attempt to combine the variance measure with an arbitrary spatial distribution of customers over the entire network. To this end, a density function will be associated with each edge which will represent the level of the demand at each point: and we will study the problem of finding a point in the network which minimises the variance function which will be called the

continuous variance problem. We present an exact algorithm for solving this problem (as well as the continuous median one) which is less complex than the exhaustive procedure. This algorithm is based on a decomposition of the problem within several subproblems where the objective is minimised, and the global optimum is chosen from the subproblem solutions. This procedure was used by Hooker [12] in a model where the demand is discrete, and where the objective function is convex in each subproblem. However, the presence of continuous edge demands introduces important differences with respect to the discrete case, since the objective function has no reason to be convex, as happens when dealing with median and variance problems.

The algorithm which we propose decomposes each edge of the network into closed segments (called "primary regions") in which the demand will be classified. In each primary region the continuous variance function will be expressed in terms of the contribution of each type of demand, in order to solve the subproblem restricted to such a region. We propose a method which calculates recursively the expression of the objective function along the edge, such that in each subproblem the data of the previous subproblem are used. In this way the effort to obtain the function to be minimised will be reduced.

The present paper is organised as follows. We begin Section 2 by introducing some definitions and notation, and by formulating the problem. In Section 3 we investigate the behaviour and properties of the continuous variance function. In Section 4 we determine the relationships between primary regions which will be needed in the following section, where we present the algorithm for solving both the continuous variance problem and the continuous median problem. Finally, in Section 6 some computational experience is provided.

2. PROBLEM FORMULATION

Let $N(V, E)$ be a general finite, connected and undirected network with vertex set $V = \{v_1, \ldots, v_n\}$ and edge set E, with $|E| = m$. Consider the edges to be rectifiable, and let l_{jk} be the positive length of each edge $c_{jk} = [v_j, v_k] \in E$. A point $x \in N$ may lie anywhere along an edge, and the distance $d(x, y)$ between two points $x, y \in N$ is determined by the length of the shortest path from x to y.

For any two points $x_1, x_2 \in c_{jk}$, let $[x_1, x_2]$ denote the subset of points of edge c_{jk} between x_1, x_2 inclusive, and let (x_1, x_2) denote the corresponding open set. Half-open sets are defined similarly.

Choosing an arbitrary vertex v_j of edge c_{jk} as the initial point, any variable point $x \in c_{jk}$ denotes both the point $x \in [v_j, v_k]$ as well as the length of the subedge $[v_j, x]$. Thus $x = 0$ means x coincides with v_j. In this way we assume a well-defined orientation in each edge.

Some positive weights w_i, p_{jk} are associated with each vertex v_i and each edge c_{jk}, respectively (p_e when the edge is denoted by c). For a subset of vertices $V' \subseteq V$ and a subset of edges $E' \subseteq E$ we define

$$w_V(V') = \sum_{v_i \in V'} w_i, \quad \text{and} \quad w_E(E') = \sum_{e \in E'} p_e.$$

We denote these by w_V and w_E when $V' = V$ and $E' = E$, respectively. Without loss of generality we may assume that the total weight of the network is

$$w(N) = w_V + w_E = 1.$$

To allow the demand to be distributed throughout the network we define, for each edge c_{jk}, a general density function for demand $f_{jk}(y)$ (f_e when the edge is denoted by c). This function $f_{jk}(y)$ is a continuous non-negative function in $y \in [0, l_{jk}]$ and has no impulses, such that the cumulative density function $F_{jk}(x) = \int_0^x f_{jk}(y)\, dy$ is a continuous function in $x \in [0, l_{jk}]$, with $F_{jk}(0) = 0$ and $F_{jk}(l_{jk}) = 1$.

Associated with each distribution of demand are the central moments

$$\mu_{jk} = \int_0^{l_{jk}} y f_{jk}(y)\, dy, \qquad \mu_{jk}^{(2)} = \int_0^{l_{jk}} y^2 f_{jk}(y)\, dy.$$

For each $x \in [0, l_{jk}]$, $\mu_{jk}(x)$ and $\mu_{jk}^{(2)}(x)$ represent the functions

$$\mu_{jk}(x) = \int_0^x y f_{jk}(y)\, dy, \qquad \mu_{jk}^{(2)}(x) = \int_0^x y^2 f_{jk}(y)\, dy.$$

The continuous median function for any $x \in N$ is given by

$$z_m(x) = \sum_{v_i \in V} w_i d(v_i, x) + \sum_{e_{jk} = [v_j, v_k] \in E} p_{jk} \int_0^{l_{jk}} d(x, y) f_{jk}(y)\, dy.$$

The variation equity criterion is measured by the variance of the distance travelled by all users of the facility with respect to the mean travel distance.

Consequently, we define the continuous variance function for any $x \in N$ by

$$z_\nu(x) = \sum_{v_i \in V} w_i [d(v_i, x) - z_m(x)]^2$$
$$+ \sum_{e_{jk} = [v_j, v_k] \in E} p_{jk} \int_0^{l_{jk}} [d(x, y) - z_m(x)]^2 f_{jk}(y) \, dy.$$

A point $x_\nu^* \in N$ is called a *continuous 1-variance* point if

$$z_\nu(x_\nu^*) \leq z_\nu(x), \quad \forall x \in N.$$

PROPOSITION 1: *For each $x \in N$, $z_\nu(x)$ satisfies*

$$z_\nu(x) = \sum_{v_i \in V} w_i d(v_i, x)^2 + \sum_{e_{jk} = [v_j, v_k] \in E} p_{jk} \int_0^{l_{jk}} d(x, y)^2 f_{jk}(y) \, dy - z_m(x)^2.$$

3. THE CONTINUOUS VARIANCE IN A PRIMARY REGION

In this section we provide a characterisation of $z_\nu(x)$ when x moves continuously within a closed subedge in which the distance function $d(v_i, x)$ is linear for all vertices of N. In the first place, we remark that, given $c_{jk} = [v_j, v_k]$, for each $z \in N$ the function $d(z, y)$ is continuous, piecewise linear and concave in $y \in [0, l_{jk}]$, therefore the functions $z_m(x)$ and $z_\nu(x)$ are continuous in $x \in c_{jk}$.

Each facility located at $x \in [v_j, v_k]$ induces a classification of nodal demand in which each vertex v_i is classified according to whether it is supplied by x *via* vertex v_j or *via* vertex v_k. However, such a classification changes in certain points of edge. These points, called edge bottleneck points were first introduced by Garfinkel *et al.* [7], and will play an important role here.

An interior point $\bar{x}_t$ of edge $[v_j, v_k]$ is said to be an edge bottleneck point with respect to vertex v_t if the travel distance $d(v_t, \bar{x}_t)$ *via* vertex v_j is the same as *via* vertex v_k (see Church and Garfinkel [6]). In such a case we have $d(v_t, v_j) + d(v_j, \bar{x}_t) = d(v_t, v_k) + d(v_k, \bar{x}_t)$. If $\bar{x}_t$ means $d(v_j, \bar{x}_t)$, we therefore have

$$\bar{x}_t = \frac{d(v_t, v_k) + l_{jk} - d(v_t, v_j)}{2}.$$

For simplicity's sake, we will say bottleneck point instead of edge bottleneck point. Let B_E be the set of bottleneck points of N. Since each vertex defines at most one bottleneck point on each edge, the cardinal of B_E is at most mn.

To study the properties of continuous variance function in each edge, we centre our interest on the portion of edge, or subedge, in which all distance functions are linear and monotone. This notion was introduced by Hooker [12] (who calls it "treelike segment"), and is the closed subedge delimited by two adjacent points of $V \cup B_E$. Henceforth such a segment will be called a primary region (see Berman [2] and Chiu [5]).

Within each primary region we will develop an algebraic expression for $z_V(x)$ by considering the contribution of each type of demand to the objective function. To this end we select an arbitrary edge $c_0 = [u, v]$, whose length is l_{uv}, and consider u as the initial point of the edge such that each $x \in c_0$ means $x = d(u, x)$ and therefore $l_{uv} - x = d(x, v)$.

The partition $\{V_u(x), V_v(x)\}$ of V induced by each $x \in [u, v]$ is given by

$$V_u(x) = \{v_i \in V \,/\, d(v_i, u) + x \leq d(v_i, v) + l_{uv} - x\}$$

and $V_v(x) = V \backslash V_u(x)$. Consequently, the partition $\{E_u(x), E_v(x), E_{[u,v]}(x), c_0\}$ of E is defined as follows

$$E_u(x) = \{c_{jk} = [v_j, v_k] \in E \backslash \{c_0\} \,/\, v_j, v_k \in V_u(x)\}$$

$$E_v(x) = \{c_{jk} = [v_j, v_k] \in E \backslash \{c_0\} \,/\, v_j, v_k \in V_v(x)\}$$

$$E_{[u,v]}(x) = \{c_{jk} = [v_j, v_k] \in E \backslash \{c_0\} \,/\, v_j \in V_u(x) \text{ and } v_k \in V_v(x)\}.$$

Since the facility x is located in $c_0 = [u, v]$, this partition classifies the edge demands of the set $E \backslash \{c_0\}$ into three types: an edge belongs to $E_u(x)$ (to $E_v(x)$) if its two vertices belong to $V_u(x)$ (to $V_v(x)$), and an edge belongs to $E_{[u,v]}(x)$ if one vertex belongs to $V_u(x)$ and the other vertex belongs to $V_v(x)$.

Let $\bar{x}_1 < \ldots < \bar{x}_k$ be the bottleneck points of edge $c_0 = [u, v]$, and let $\bar{x}_0 = 0$, $\bar{x}_{k+1} = l_{uv}$, (that is, $\bar{x}_0$ is equal to u and $\bar{x}_{k+1}$ is equal to v).

PROPOSITION 2: [11] *For $j = 1, \ldots, k + 1$, let $S_j = (\bar{x}_{j-1}, \bar{x}_j]$ be the half-open set associated with the primary region $[\bar{x}_{j-1}, \bar{x}_j]$, and let $S_0 = [\bar{x}_0, \bar{x}_0]$. Then, for each $j = 1, \ldots, k + 1$, the partition $\{V_u(x), V_v(x)\}$ of V is unchanged when x varies in S_j.*

According to this proposition, when $x \in S_j$ the sets $V_u(x), V_v(x)$ will be denoted by $V_u(S_j)$ and $V_v(S_j)$, respectively. Therefore, when $x \in S_j$ neither

$E_u(x)$, $E_v(x)$ nor $E_{[u,v]}(x)$ are changed, and similarly we may extend the above notation to these sets, which will be denoted by $E_u(S_j)$, $E_v(S_j)$, and $E_{[u,v]}(S_j)$, respectively.

DEFINITION 3: *Let S_j be a half-open set of edge $c_0 = [u,v]$, $(1 \le j \le k+1)$. For each edge $c_{ht} = [v_h, v_t] \in E$ we define the following functions in S_j:*

$$D_{e_{ht}}(x) = p_{ht} \int_0^{l_{ht}} d(x,y) f_{ht}(y)\, dy, \quad \forall x \in S_j$$

$$D_{e_{ht}}^{(2)}(x) = p_{ht} \int_0^{l_{ht}} d(x,y)^2 f_{ht}(y)\, dy, \quad \forall x \in S_j$$

$D_{e_{ht}}(x)$ and $D_{e_{ht}}^{(2)}(x)$ can be interpreted as the contribution of edge c_{ht} to the continuous median function $z_m(x)$ and to its square.

The expressions of these functions change according to the set of partition of E in which c_{ht} lies. When $c_{ht} = c_0$, the functions will be denoted by $D_{e_0}(x)$ and $D_{e_0}^{(2)}(x)$, respectively.

PROPOSITION 4: *For each $j = 1, \ldots, k+1$, when x varies in the primary region $[\bar{x}_{j-1}, \bar{x}_j]$, the continuous variance function $z_v(x)$ can be expressed as follows:*

$$z_v(x; S_j) = \sum_{v_i \in V_u(S_j)} w_i d(v_i, x)^2 + \sum_{v_i \in V_v(S_j)} w_i d(v_i, x)^2 + \sum_{e_{ht} \in E_u(S_j)} D_{e_{ht}}^{(2)}(x)$$
$$+ \sum_{e_{ht} \in E_v(S_j)} D_{e_{ht}}^{(2)}(x) + \sum_{e_{ht} \in E_{[u,v]}(S_j)} D_{e_{ht}}^{(2)}(x)$$
$$+ D_{e_0}^{(2)}(x) - z_m(x; S_j)^2, \quad \forall x \in [\bar{x}_{j-1}, \bar{x}_j]$$

where $\forall x \in [\bar{x}_{j-1}, \bar{x}_j]$

$$z_m(x; S_j) = \sum_{v_i \in V_u(S_j)} w_i d(v_i, x) + \sum_{v_i \in V_v(S_j)} w_i d(v_i, x) + \sum_{e_{ht} \in E_u(S_j)} D_{e_{ht}}(x)$$
$$+ \sum_{e_{ht} \in E_v(S_j)} D_{e_{ht}}(x) + \sum_{e_{ht} \in E_{[u,v]}(S_j)} D_{e_{ht}}(x) + D_{e_0}(x).$$

Proof: The Proposition holds for $x \in S_j$ from Proposition 2 and properties of both partitions of V and E. For $x \in [\bar{x}_{j-1}, \bar{x}_j]$ the result is obtained by continuity. □

In the following, we will develop each term of $z_\nu(x; S_j)$ in order to obtain an algebraic expression for the continuous variance function over the j-th primary region.

- <u>Nodal demands</u>

  This contribution is identified by the first two terms in $z_\nu(x; S_j)$, which define the function $z_\nu^{(2)}(x; S_j)$, and by the first two terms in $z_m(x; S_j)$, which define the function $z_\nu(x; S_j)$.

  Developing both functions, and applying $d(v_i, x) = d(v_i, u) + x$ if $v_i \in V_u(S_j)$, and $d(v_i, x) = d(v_i, v) + l_{uv} - x$ if $v_i \in V_v(S_j)$, we may write:

  $$z_\nu(x; S_j) = a_j + g_j\, x$$
  $$z_\nu^{(2)}(x; S_j) = b_j + 2c_j\, x + w_\nu\, x^2$$

  where the corresponding coefficients are given by

  $$g_j = w_\nu(V_u(S_j)) - w_\nu(V_v(S_j))$$
  $$a_j = z_\nu(u, V_u(S_j)) + z_\nu(v, V_v(S_j)) + w_\nu(V_v(S_j))l_{uv}$$
  $$b_j = \sum_{v_i \in V_u(S_j)} w_i d(v_i, u)^2 + \sum_{v_i \in V_v(S_j)} w_i[d(v_i, v) + l_{uv}]^2$$
  $$c_j = z_\nu(u, V_u(S_j)) - z_\nu(u, V_v(S_j)) - l_{uv}w_\nu(V_v(S_j))$$

  with $z_\nu(u, V_u(S_j)) = \displaystyle\sum_{v_i \in V_u(S_j)} w_i d(v_i, u)$ and $z_\nu(v, V_v(S_j))$

  $$= \sum_{v_i \in V_v(S_j)} w_i d(v_i, v).$$

- <u>Edge demands</u>

  For any given edge $e_{ht} = [v_h, v_t]$ of E, we obtain four expressions of $D_{e_{ht}}(x)$ and $D_{e_{ht}}^{(2)}(x)$, depending on whether the demand considered occurs on $e_{ht} \in E_u(S_j)$, on $e_{ht} \in E_v(S_j)$, on $e_{ht} \in E_{[u,v]}(S_j)$ or on $e_{ht} = e_0 = [u, v]$. In all cases we suppose that v_h is the initial point of edge e_{ht} such that

  $$\forall y \in e_{ht} \quad d(v_h, y) = y, \quad \text{and} \quad d(v_t, y) = l_{ht} - y. \tag{1}$$

  <u>$e_{ht} = [v_h, v_t] \in E_u(S_j)$</u>

  In this case, $\forall y \in e_{ht}$, $d(y, x) = d(y, u) + x$. Let $\xi_{ht} \in [v_h, v_t]$ be the bottleneck point of edge e_{ht} relative to vertex u. Then

$\xi_{ht} = (d(u, v_t) + l_{ht} - d(u, v_h))/2$, $\quad \xi_{ht}$ being independent of the location of $x \in S_j$ (Fig. 1a).

For each $x \in S_j$ the distance function $d(y, x)$ is piecewise linear and concave when y varies in $[v_h, v_t]$ and can be expressed as follows

$$d(x, y) = \begin{cases} x + d(v_h, u) + y, & \text{if} \quad 0 \leq y \leq \xi_{ht} \\ x + d(v_t, u) + l_{ht} - y, & \text{if} \quad \xi_{ht} < y \leq l_{ht}. \end{cases}$$
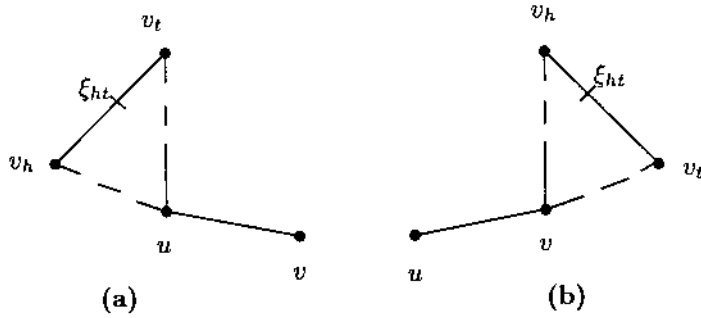

Figure 1.

Using these relationships and the definitions and properties associated with density function $f_{ht}(y)$ in Definition 3, leads to

$$D_{e_{ht}}(x) = p_{ht}x + p_{ht}H_D(u, c_{ht})$$

where $H_D(u, c_{ht}) = d(v_t, u) + l_{ht} - 2\xi_{ht}F_{ht}(\xi_{ht}) + 2\mu_{ht}(\xi_{ht}) - \mu_{ht}$.

The same procedure applied to $D_{e_{ht}}^{(2)}(x)$ gives

$$D_{e_{ht}}^{(2)}(x) = p_{ht}x^2 + 2p_{ht}H_\nu(u, c_{ht})x + p_{ht}R_\nu(u, c_{ht})$$

where

$$H_\nu(u, c_{ht}) = 2\mu_{ht}(\xi_{ht}) - \mu_{ht} + l_{ht} + d(v_t, u) - 2\xi_{ht}F_{ht}(\xi_{ht})$$

$$R_\nu(u, c_{ht}) = d(v_h, u)^2 F_{ht}(\xi_{ht}) + (l_{ht} + d(v_t, u))^2$$
$$\times (1 - F_{ht}(\xi_{ht})) + 4(d(v_h, u) + \xi_{ht})\mu_{ht}(\xi_{ht})$$
$$- 2(l_{ht} + d(v_t, u))\mu_{ht} + \mu_{ht}^{(2)}.$$

$\underline{c_{ht} = [v_h, v_t] \in E_v(S_j)}$

Now we have a symmetrical situation in which $\forall y \in c_{ht}$, $d(y, x)$ $= d(y, v) + l_{uv} - x$ and the bottleneck point in c_{ht} relative to vertex v

is given by $\xi_{ht} = (d(v, v_t) + l_{ht} - d(v, v_h))/2$ (Fig. 1b). Therefore, the distance function $d(y, x)$ yields

$$d(x, y) = \begin{cases} l_{uv} - x + d(v, v_h) + y, & \text{if } 0 \le y \le \xi_{ht} \\ l_{uv} - x + d(v, v_t) + l_{ht} - y, & \text{if } \xi_{ht} < y \le l_{ht} \end{cases}$$

and similar reasoning to the former case provides

$$D_{e_{ht}}(x) = -p_{ht}x + p_{ht}L_{\nu}(v, c_{ht})$$
$$D_{e_{ht}}^{(2)}(x) = p_{ht}x^2 - 2p_{ht}L_{\nu}(v, c_{ht})x + p_{ht}J_{\nu}(v, c_{ht})$$

where the coefficients are given by

$$L_{\nu}(v, c_{ht}) = d(v_t, v) + l_{ht} + l_{uv} - 2\xi_{ht}F_{ht}(\xi_{ht}) + 2\mu_{ht}(\xi_{ht}) - \mu_{ht}$$
$$L_{\nu}(v, c_{ht}) = 2\mu_{ht}(\xi_{ht}) - \mu_{ht} + l_{ht} + l_{uv} + d(v_t, v) - 2\xi_{ht}F_{ht}(\xi_{ht})$$
$$\begin{aligned} J_{\nu}(v, c_{ht}) = {} & (d(v_h, v) + l_{uv})^2 F_{ht}(\xi_{ht}) \\ & + (l_{ht} + l_{uv} + d(v_t, v))^2(1 - F_{ht}(\xi_{ht})) \\ & + 4(\xi_{ht} + d(v_h, v) + l_{uv})\mu_{ht}(\xi_{ht}) \\ & - 2(l_{ht} + l_{uv} + d(v_t, v))\mu_{ht} + \mu_{ht}^{(2)}. \end{aligned}$$

$\underline{c_{ht} = [v_h, v_t] \in E_{[u,v]}(S_j)}$

In this case a point $\xi_{ht}(x) \in c_{ht}$ is associated with each $x \in S_j$. Such a point assumes the role of the "bottleneck point" relative to x, that is $d(\xi_{ht}(x), v_h) + d(v_h, x) = d(\xi_{ht}(x), v_t) + d(v_t, x)$ (see Fig. 2).

According to the orientation of edge c_{ht} from vertex v_h to vertex v_t considered in (1), we obtain $\xi_{ht}(x) = (d(v_t, x) + l_{ht} - d(v_h, x))/2$. Since $v_h \in V_u(S_j)$ and $v_t \in V_v(S_j)$ we have $d(v_h, x) = d(v_h, u) + x$
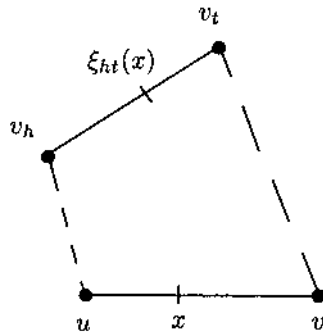

Figure 2.

and $d(v_t, x) = d(v_t, v) + l_{uv} - x$ which implies $\xi_{ht}(x) = (d(v_t, v) + l_{uv} + l_{ht} - d(v_h, u) - 2x)/2$. It is easy to see that the function $\xi_{ht}(x)$ is decreasing and varies in $[0, l_{ht}]$ when $x \in [\bar{x}_{j-1}, \bar{x}_j]$.

This dependence on x induces non-linearity in $D_{e_{ht}}(x)$. In effect, applying

$$d(x, y) = \begin{cases} x + d(u, v_h) + y, & \text{if} \quad 0 \le y \le \xi_{ht}(x) \\ l_{uv} - x + d(v, v_t) + l_{ht} - y, & \text{if} \quad \xi_{ht}(x) < y \le l_{ht} \end{cases}$$

leads to

$$\begin{aligned} D_{e_{ht}}(x) = p_{ht} \big[& x(2F_{ht}(\xi_{ht}(x)) - 1) + d(v_h, u)F_{ht}(\xi_{ht}(x)) \\ & + (d(v_t, v) + l_{ht} + l_{uv})(1 - F_{ht}(\xi_{ht}(x))) \\ & + 2\mu_{ht}(\xi_{ht}(x)) - \mu_{ht} \big]. \end{aligned}$$

The variation of $\xi_{ht}(x)$ with respect to x gives rise to a continuous change of demands on edge c_{ht} from "u–demands" to "v–demands", and contributes to indetermination of curvature of $z_m(x)$ within a primary region (see Chiu [5]). In a similar manner,

$$\begin{aligned} D_{e_{ht}}^{(2)}(x) = p_{ht} \big[& x^2 + F_{ht}(\xi_{ht}(x))(d(v_h, u)^2 + 2xd(v_h, u)) \\ & + (1 - F_{ht}(\xi_{ht}(x)))\Big((d(v_t, v) + l_{ht} + l_{uv})^2 \\ & - 2x(d(v_t, v) + l_{ht} + l_{uv})\Big) + 2\mu_{ht}(\xi_{ht}(x))(d(v_h, u) + d(v_t, v) \\ & + l_{ht} + l_{uv} + 2x) - 2\mu_{ht}(d(v_t, v) + l_{ht} + l_{uv} - x) + \mu_{ht}^{(2)} \big]. \end{aligned}$$

$\underline{c_0 = [u, v]}$

The orientation considered in this edge implies that, for $y \in c_0$,

$$d(x, y) = |x - y| = \begin{cases} x - y, & \text{if} \quad 0 \le y \le x \\ y - x, & \text{if} \quad x < y \le l_{uv}. \end{cases}$$

Therefore

$$\begin{aligned} D_{e_0}(x) &= p_0 \left[\int_0^x (x - y)f_0(y)\,dy + \int_x^{l_{uv}} (y - x)f_0(y)\,dy \right] \\ &= p_0[2xF_0(x) - x + \mu_0 - 2\mu_0(x)]. \end{aligned}$$

Since $d(x, y)^2 = (x - y)^2$, it follows that

$$D_{e_0}^{(2)}(x) = p_0 \int_0^{l_{uv}} (x^2 - 2xy + y^2)f_0(y)\,dy = p_0 \left[x^2 - 2x\mu_0 + \mu_0^{(2)} \right].$$

Using the expressions obtained in Proposition 4 for both functions in each case, yields

$$
\begin{aligned}
z_m(x; S_j) = a_j + g_j\, x + \Big(w_{E}(E_u(S_j)) - w_{E}(E_v(S_j)) \Big) x \\
+ \sum_{e_{ht} \in E_u(S_j)} p_{ht}\, H_{D}(u, c_{ht}) + \sum_{e_{ht} \in E_v(S_j)} p_{ht}\, L_{D}(v, c_{ht}) \\
+ \sum_{e_{ht} \in E_{[u,v]}(S_j)} D_{e_{ht}}(x) + p_0\left[2x F_0(x) - x + \mu_0 - 2\mu_0(x) \right] \quad (2)
\end{aligned}
$$

$$
\begin{aligned}
z_\nu(x; S_j) = \Big(w_{\cdot} + w_{E}(E_u(S_j)) + w_{E}(E_v(S_j)) + p_0 \Big) x^2 + 2c_j\, x \\
+ 2x \left(\sum_{e_{ht} \in E_u(S_j)} p_{ht}\, H_\nu(u, c_{ht}) \right. \\
\left. - \sum_{e_{ht} \in E_v(S_j)} p_{ht}\, L_\nu(v, c_{ht}) - p_0\mu_0 \right) \\
+ \sum_{e_{ht} \in E_{[u,v]}(S_j)} D^{(2)}_{e_{ht}}(x) + b_j + p_0\mu_0^{(2)} \\
+ \sum_{e_{ht} \in E_u(S_j)} p_{ht}\, R_\nu(u, c_{ht}) + \sum_{e_{ht} \in E_v(S_j)} p_{ht}\, J_\nu(v, c_{ht}) \\
- \Big(z_m(x; S_j) \Big)^2. \quad (3)
\end{aligned}
$$

We now analyse the behaviour of function $z_\nu(x)$ in a primary region. The continuous median $z_m(x)$ is convex over any edge $[u, v]$ in a tree network (see Chiu [5]), however, the same does not occur with the function $z_\nu(x)$. In effect, the function $z_\nu(x)$ is neither concave nor convex over any edge of a tree network, as can be observed in the counterexample presented in Figures 3–5. This behaviour can be extended to a general network, and therefore:

COROLLARY 5: *The continuous variance function $z_\nu(x)$ is neither concave nor convex over any primary region of a network.*

As stated earlier, in order to illustrate these results we present a tree network shown in Figure 3. In this example demand is assumed to be uniformly distributed along the edges, and so only the length and weight of each edge are shown. Node weights are written next to each node with small numbers, edge weights with round brackets, and the lengths are indicated
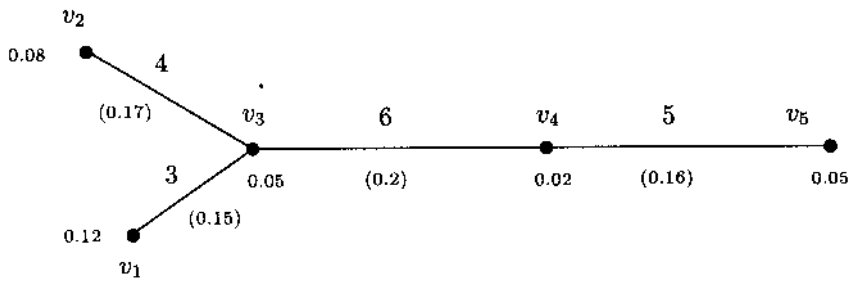
Figure 3.
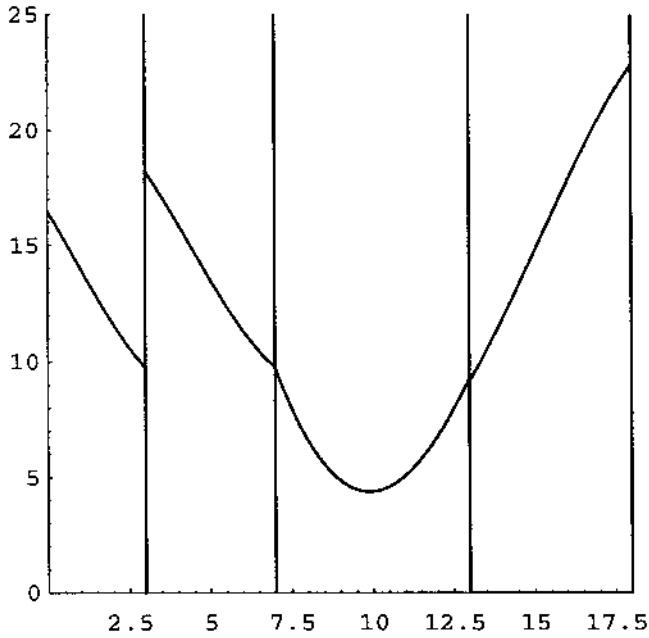
Figure 4.

above each edge. Figure 4 shows the continuous variance function along edges $[v_1, v_3]$, $[v_2, v_3]$, $[v_3, v_4]$, and $[v_4, v_5]$, in this order. Figure 5 shows the corresponding second derivatives.

4. RELATIONSHIPS BETWEEN PRIMARY REGIONS

From the aforementioned results it follows that for solving the continuous variance problem on a network it is necessary to find an efficient way to find all local minima, that is, it is required to solve the restricted problem

Figure 5.

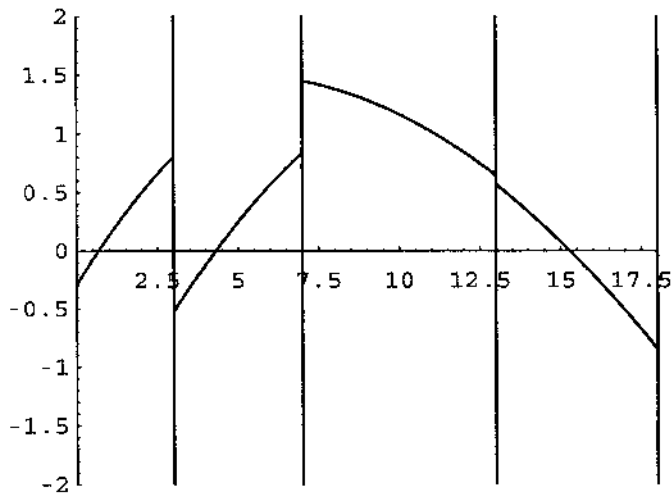$\min\{z_\nu(x; S_j), \ x \in [\bar{x}_{j-1}, \bar{x}_j]\}$ over each primary region (note that nm restricted problems must be solved at most). The search for the minimum of such problems involves a procedure to find the zeros of $d(z_\nu(x; S_j))/dx$, (and a similar procedure must be applied to solve the continuous median problem). This procedure is exhaustive in nature, and depends very much on the edge density functions dealt with.

When only uniform density functions are considered, both functions ($z_\nu(\cdot)$ and $z_m(\cdot)$) are polynomials over any primary region (of degree 4 and 2, respectively), and the search for the local minimum takes constant time. However, the identification of functions $z_\nu(x; S_j)$ or $z_m(x; S_j)$ over the j-th primary region requires examining all vertices and all edges in order to determine the corresponding partitions of V and E for obtaining all terms of expressions (2) and (3). Supposing that the shortest distance matrix is calculated already (in a preprocessing phase), the exhaustive procedure with uniform density functions provides the minimum in $O(mn(n+m))$ time.

The only algorithm that tackles the continuous median problem is heuristic [5]. In the following we present an exact procedure for solving the continuous variance problem (as well as the continuous median problem, as explained in the introduction). Even in the worst case, the complexity of this algorithm will be lower than that of the exhaustive procedure, and computational results will show that it runs in $O(m(m+n))$ time for the test networks used.

Now we will develop some recursive relationships between a primary region and its predecessor. To this end we introduce some definitions and properties associated with the concept of bottleneck point.

DEFINITION 6: *Let $\bar{x}_1 < \ldots < \bar{x}_k$ be the bottleneck points of edge $c_0 = [u,v]$ and let $\bar{x}_0 = 0$, $\bar{x}_{k+1} = l_{uv}$. For $j = 0 \ldots, k+1$, $B(\bar{x}_j)$ denotes the set of vertices relative to point $\bar{x}_j$, and is given by*

$$B(\bar{x}_j) = \{v_t \in V / \, d(v_t, u) + \bar{x}_j = d(v_t, v) + l_{uv} - \bar{x}_j\}.$$

Clearly $B(\bar{x}_j) \subset V_u(S_j)$. The following proposition is introduced by Hansen and Zheng [11], and its proof can be found in [16].

PROPOSITION 7: *The following statements hold:*
  (i) $B(\bar{x}_i) \cap B(\bar{x}_j) = \emptyset$, $\forall i, j = 0, \ldots, k+1$, $i \neq j$.
 (ii) $\displaystyle\sum_{j=0}^{k+1} |B(\bar{x}_j)| \leq n.$
(iii) $V_u(S_j) = V_u(S_{j-1}) \backslash B(\bar{x}_{j-1})$, *and* $V_v(S_j) = V_v(S_{j-1}) \cup B(\bar{x}_{j-1})$,
      $\forall j = 1, \ldots, k+1.$
As a consequence, $V_u(S_{j-1}) \supset V_u(S_j)$ and $V_v(S_{j-1}) \subset V_v(S_j)$, $j = 1, \ldots, k+1$.

DEFINITION 8: *For $j = 0, \ldots, k+1$, define the following subsets of edges:*

$$\begin{aligned}
B^{(1)}(E_u(S_j)) &= \{c_{ht} = [v_h, v_t] \in E_u(S_j) / \\
&\qquad \text{either} \quad v_h \in B(\bar{x}_j), v_t \notin B(\bar{x}_j), \\
&\qquad \text{or} \qquad v_h \notin B(\bar{x}_j), v_t \in B(\bar{x}_j)\} \\
B^{(2)}(E_u(S_j)) &= \{c_{ht} = [v_h, v_t] \in E_u(S_j) / v_h, v_t \in B(\bar{x}_j)\} \\
B^{(1)}(E_{[u,v]}(S_j)) &= \{c_{ht} = [v_h, v_t] \in E_{[u,v]}(S_j) / \\
&\qquad \text{either} \quad v_h \in B(\bar{x}_j), v_t \notin B(\bar{x}_j), \\
&\qquad \text{or} \qquad v_h \notin B(\bar{x}_j), v_t \in B(\bar{x}_j)\}.
\end{aligned}$$

Set
$$B(E_u(S_j)) = B^{(1)}(E_u(S_j)) \cup B^{(2)}(E_u(S_j)).$$

For $q = 1, 2$ the set $B^{(q)}(E_u(S_j))$ contains those edges of $E_u(S_j)$ such that either one of the end vertices (if $q = 1$) or both (if $q = 2$), belong to

$B(\bar{x}_j)$. Similarly the edges of $B^{(1)}(E_{[u,v]}(S_j))$ are edges of $E_{[u,v]}(S_j)$ with an unique end vertex in $B(\bar{x}_j)$. Therefore $B^{(q)}(E_u(S_j)) \subset E_u(S_j)$, $q = 1,2$, $B(E_u(S_j)) \subseteq E_u(S_j)$ and $B^{(1)}(E_{[u,v]}(S_j)) \subseteq E_{[u,v]}(S_j)$.

PROPOSITION 9: *For $j = 1,\ldots,k+1$,*

$$E_u(S_j) = E_u(S_{j-1}) \backslash B(E_u(S_{j-1})). \qquad (4)$$

PROPOSITION 10: *For $j = 1,\ldots,k+1$,*

$$E_v(S_j) = E_v(S_{j-1}) \cup B^{(1)}(E_{[u,v]}(S_{j-1})) \cup B^{(2)}(E_u(S_{j-1})). \qquad (5)$$

PROPOSITION 11: *For $j = 1,\ldots,k+1$,*

$$E_{[u,v]}(S_j) = \left(E_{[u,v]}(S_{j-1}) \backslash B^{(1)}(E_{[u,v]}(S_{j-1})) \right) \cup B^{(1)}(E_u(S_{j-1})). \qquad (6)$$

PROPOSITION 12: *Let $j,q \in \{1,\ldots,k+1\}$ be such that $j \neq q$. The following statements hold:*
1. $B(E_u(S_j)) \cap B(E_u(S_q)) = \emptyset$.
2. $B^{(1)}(E_{[u,v]}(S_j)) \cap B^{(1)}(E_{[u,v]}(S_q)) = \emptyset$.

As a consequence of this Proposition, we have the bound

$$\sum_{j=1}^{k+1} |B(E_u(S_j))| \leq m.$$

Furthermore, for $q \neq j$ the intersection $B(E_u(S_j)) \cap B^{(1)}(E_{[u,v]}(S_q))$ can be non–empty, since for $q > j$ the sets $B^{(1)}(E_u(S_j))$ and $B^{(1)}(E_{[u,v]}(S_q))$ are not necessarily disjointed. Thus, we can obtain the upper bound

$$\sum_{j=1}^{k+1} \left(|B^{(1)}(E_{[u,v]}(S_j))| + |B(E_u(S_j))| \right) \leq 2m.$$

On the other hand, although $\bigcup_{j=1}^{k+1} E_{[u,v]}(S_j) \subseteq E$, the sets $E_{[u,v]}(S_j)$, $j = 1,\ldots,k+1$ have no cause to be disjointed. Therefore, the quantity $\sum_{j=1}^{k+1} |E_{[u,v]}(S_j)|$ can not be bounded by m. However, from (6) each set

$E_{[u,v]}(S_j)$ is the union of two disjointed sets, and therefore we can write

$$\sum_{j=1}^{k+1} |E_{[u,v]}(S_j)| = \sum_{j=1}^{k+1} |E_{[u,v]}(S_{j-1}) \backslash B^{(1)}(E_{[u,v]}(S_{j-1}))|$$
$$+ \sum_{j=1}^{k+1} |B^{(1)}(E_u(S_{j-1}))|.$$

In order to estimate the value of this sum in computational terms, the following definition is introduced.

DEFINITION 13: *With each edge $[u,v]$ of N is associated a quantity $\tau(u,v)$, whose value is given by*

$$\tau(u,v) = \sum_{j=0}^{k} \tau_j(u,v)$$

where for $j = 0,\ldots,k$, $\tau_j(u,v) = |E_{[u,v]}(S_j) \backslash B^{(1)}(E_{[u,v]}(S_j))|$.
This definition yields

$$\sum_{j=1}^{k+1} |E_{[u,v]}(S_j)| = \tau(u,v) + \sum_{j=0}^{k} |B^{(1)}(E_u(S_j))|.$$

The sets $\{E_{[u,v]}(S_j) \backslash B^{(1)}(E_{[u,v]}(S_j))\}$ $j = 0,\ldots,k$ are not disjointed in general which supposes that, in the worst case, $\tau(u,v) \in O(nm)$. However, it is interesting to explore the average value of $\tau(u,v)$. In the computational experience it will be seen that for not excessively large values of m, (as happens, for example, in planar networks, where $m \leq 3n - 6$) the values of $\tau(u,v)$ are quite moderate. In fact we will obtain that $\tau(u,v) \in O(m)$ for the tested networks.

5. ALGORITHM AND COMPLEXITY

Before solving the restricted problem $\min\{z_\nu(x; S_j), \ x \in [\bar{x}_{j-1}, \bar{x}_j]\}$ over the j-th primary region, we need to compute all terms of $z_m(x; S_j)$ and $z_\nu(x; S_j)$, given in expressions (2) and (3). Now it will be shown that, except for the terms corresponding to the demand of edges $E_{[u,v]}(S_j)$, the remaining can be recursively obtained from the former primary region.

Nodal demands

The nodal contribution is obtained through coefficients a_j, g_j, b_j and c_j. They can be computed as follows:

$$a_j = a_{j-1} + z_.(v, B(\bar{x}_{j-1})) - z_.(u, B(\bar{x}_{j-1})) - l_{uv} w_.(B(\bar{x}_{j-1})) \quad (7)$$

$$g_j = g_{j-1} - 2w_.(B(\bar{x}_{j-1})) \quad (8)$$

$$b_j = b_{j-1} + \sum_{v_i \in B(\bar{x}_{j-1})} w_i [d(v_i, v) + l_{uv}]^2 - \sum_{v_i \in B(\bar{x}_{j-1})} w_i d(v_i, u)^2 \quad (9)$$

$$c_j = c_{j-1} - \Big(z_.(u, B(\bar{x}_{j-1})) + z_.(v, B(\bar{x}_{j-1})) + l_{uv} w_.(B(\bar{x}_{j-1})) \Big). \quad (10)$$

Edge demands

Using (4) and (5) with the terms of (2) and (3) relative to edge contribution, we find that when such demand is referred to edges belonging to $E_u(S_j)$ or to $E_v(S_j)$ we can write

$$w_.(E_u(S_j)) = w_.(E_u(S_{j-1})) - w_.(B(E_u(S_{j-1}))) \quad (11)$$

$$\sum_{e_{ht} \in E_u(S_j)} p_{ht} Q_{(.)}(u, c_{ht}) = \sum_{e_{ht} \in E_u(S_{j-1})} p_{ht} Q_{(.)}(u, c_{ht})$$
$$- \sum_{e_{ht} \in B(E_u(S_{j-1}))} p_{ht} Q_{(.)}(u, c_{ht}) \quad (12)$$

where $Q_{(.)}$ denotes H_p, H_ν or R_ν. Similarly, when $c_{ht} \in E_v(S_j)$,

$$\sum_{e_{ht} \in E_v(S_j)} p_{ht} Q_{(.)}(v, c_{ht}) = \sum_{e_{ht} \in E_v(S_{j-1})} p_{ht} Q_{(.)}(v, c_{ht})$$
$$+ \sum_{e_{ht} \in B^{(1)}(E_{[u,v]}(S_{j-1}))} p_{ht} Q_{(.)}(v, c_{ht})$$
$$+ \sum_{e_{ht} \in B^{(2)}(E_u(S_{j-1}))} p_{ht} Q_{(.)}(v, c_{ht}) \quad (13)$$

where again $Q_{(.)}$ denotes L_p, L_ν or J_ν. However, these relations cannot be used when demand arises on any edge $c_{ht} \in E_{[u,v]}(S_j)$. This demand is obtained by means of functions $D_{e_{ht}}(x)$ and $D_{e_{ht}}^{(2)}(x)$, that are $\xi_{ht}(x)$ dependent functions. Both functions change when x changes from one

primary region to another. Therefore, their calculation requires the checking of all edges of each set $E_{[u,v]}(S_j)$.

THEOREM 14: *For each $c_0 = [u,v] \in E$ and $x \in [u,v]$, the number of edges and vertices of N that must be examined to obtain $z_m(x)$ and $z_\nu(x)$ by means of the procedure described by (7) to (13) is not greater than $n + 2m + \tau(u,v)$.*

Proof: When $x \in [\bar{x}_{j-1}, \bar{x}_j]$ is $z_\nu(x) = z_\nu(x; S_j)$ and $z_m(x) = z_m(x; S_j)$. Suppose we know all necessary information to calculate $z_\nu(x; S_{j-1})$ and $z_m(x; S_{j-1})$. To obtain both functions over S_j requires computations of the terms corresponding to:

- nodal contribution. From relationships (7) to (10), their calculation requires examining all vertices in $B(\bar{x}_{j-1})$.
- Edge contribution. From relationships (11) to (13), for demand on edges in $E_u(S_j)$ or $E_v(S_j)$ one has to check the sets $B(E_u(S_{j-1}))$, $B^{(1)}(E_{[u,v]}(S_{j-1}))$ and $B^{(2)}(E_u(S_{j-1}))$. For demand on edges in $E_{[u,v]}(S_j)$ it is necessary to check the set $B^{(1)}(E_u(S_{j-1}))$ together with $\tau_{j-1}(u,v)$ edges of E.

Since $B^{(q)}(E_u(S_{j-1})) \subset B(E_u(S_{j-1}))$, $q = 1, 2$, the number of vertices and edges that must be examined to obtain $z_\nu(x; S_j)$ and $z_m(x; S_j)$ is given by

$$|B(\bar{x}_{j-1})| + |B^{(1)}(E_{[u,v]}(S_{j-1}))| + |B(E_u(S_{j-1}))| + \tau_{j-1}(u,v).$$

Therefore, to obtain $z_\nu(x)$ and $z_m(x)$ along edge $[u,v]$ it is sufficient to extend the former quantity to all primary regions on $[u,v]$. Hence, the total number of elements examined takes the following value

$$\sum_{j=1}^{k+1} |B(\bar{x}_{j-1})| + \sum_{j=1}^{k+1} \left(|B^{(1)}(E_{[u,v]}(S_{j-1}))| + |B(E_u(S_{j-1}))| \right)$$
$$+ \sum_{j=1}^{k+1} \tau_{j-1}(u,v) \leq n + 2m + \tau(u,v)$$

where Proposition 7, definition of $\tau(u,v)$ and consequences of Proposition 12 have been used.   □

Now we present the algorithm for finding the continuous variance point (or the continuous median point, as already pointed out). First, some preliminary calculations need to be executed:

(i) compute the shortest distance matrix $(D(u,v))$.

(ii) for each edge $[u, v]$, put the bottleneck points $\bar{x}_1, \ldots, \bar{x}_k$ in increasing order, and compute $B(\bar{x}_j)$, $j = 0, \ldots, k + 1$.

In planar networks these operations can be done in $O(mn \log n)$ time, using Dijkstra's algorithm n times, and that of Heapsort m times to rank the bottleneck points (see [1]). This complexity increases when the network is nonplanar and very dense, since the distance matrix is computed in $O(n^3)$ time.

In the following we describe the pseudocode of the algorithm to solve the continuous variance problem. The algorithm finds the local optimum in each edge, and chooses the best solution as the global optimum. To obtain the minimum in each edge the continuous variance function must be characterised in all primary regions of the edge, and this process is recursively executed by means of the relationships developed in the last section.

5.1. Main algorithm

Step 0.

   Let Z be a suitably large number; $x_\nu^* \longleftarrow \emptyset$; $[u^*, v^*] \longleftarrow \emptyset$.

Step 1.

   For $e_0 = [u, v] \in E$ **do**

   Let $\bar{x}_0 < \ldots < \bar{x}_k$ be the bottleneck points.

   $Z_{uv} \longleftarrow M$ (a large number); $x_{\nu_{(uv)}}^* \longleftarrow \emptyset$.

   Compute $V_u(S_0)$, $V_v(S_0)$, $E_u(S_0)$, $E_v(S_0)$, $E_{[u,v]}(S_0)$.

      while $j < k + 1$ **do**

      Compute $V_u(S_j)$, $V_v(S_j)$, $E_u(S_j)$, $E_v(S_j)$, $E_{[u,v]}(S_j)$, and $B^{(1)}(E_{[u,v]}(S_j))$, $B^q(E_u(S_j))$ $(q = 1, 2)$ by using (4), (5), (6).

      Compute $z_\nu(x; S_j)$ $(z_m(x; S_j))$ by using (7) to (13).

      Find the optimum $x_{\nu_j}^*$ of $\min\{z_\nu(x; S_j), \ x \in [\bar{x}_{j-1}, \bar{x}_j]\}$.

      If $z_\nu(x_{\nu_j}^*; S_j) < Z_{uv}$, then $Z_{uv} \longleftarrow z_\nu(x_{\nu_j}^*; S_j)$; $x_{\nu_{(uv)}}^* \longleftarrow x_{\nu_j}^*$.

      end while

   If $Z_{uv} < Z$, then $Z \longleftarrow Z_{uv}$; $x_\nu^* \longleftarrow x_{\nu_{(uv)}}^*$; $[u^*, v^*] \longleftarrow [u, v]$.

   end for

Step 2.

   Stop. The incumbent Z, x_ν^*, $[u^*, v^*]$ is the solution.

5.2. Complexity

We denote the complexity of the preprocessing phase by $O(\alpha(n, m))$ (where $\alpha(n, m)$ can be $mn \log n$ or n^3, according to the type of network). In order to analyse the complexity of the main procedure, we define

$$\tau_N = \max\{\tau(u, v) : [u, v] \in E\}.$$

On the assumption that all density functions are uniform, the main computational effort is given by Step 1, where the global minimum in each edge $[u, v]$ is calculated. From Theorem 14 this process requires a time of $O(m + n + \tau(u, v))$. Since Step 1 is executed with m edges, the main algorithm can be performed in $O(m(m + n + \tau_N))$. Therefore, the overall time to solve the problem is $O(\alpha(n, m) + m(m + n + \tau_N))$. However, in order to compare this algorithm with the exhaustive procedure, we will centre our interest on the complexity of the main procedure, since the computational effort in the preprocessing phase is the same for both algorithms.

Although the complexity of the main algorithm depends on τ_N (whose value is associated with each network), in the computational experience we will see that for all networks tested, the values of τ_N were markedly less than the corresponding worst case values.

We remark that including the worst case for τ_N (in which $\tau_N \in O(nm)$), the algorithm runs in less time than the exhaustive procedure. In effect, this fact can be deduced from the following.

PROPOSITION 15: *Even in the worst case, the number of vertices and edges visited by the exhaustive procedure is greater than in the main algorithm.*

Proof: The number of vertices and edges visited by the exhaustive procedure is $mn(n + m)$, and from Theorem 14, the number of visits executed by the main algorithm is not greater than

$$\sum_{[u,v] \in E} (n + 2m + \tau(u, v)).$$

However, this number is bounded by $m(n + 2m + nm)$, since $\tau(u, v) \leq nm$ for all $[u, v] \in E$. Therefore, it will be sufficient to prove that $2m + n + nm \leq n(n + m)$. In effect, this relationship holds because in a network without loops, $m \leq \dfrac{n(n - 1)}{2}$.

6. COMPUTATIONAL EXPERIENCE

In this section some computational results are presented to analyse the order of τ_N and so to analyse the complexity of the algorithm. Although we have seen that the complexity of the main algorithm is $O(m(m + n + \tau_N))$, with $\tau_N \in O(nm)$ in the worst case, it is quite likely that the computational time would be substantially reduced for most problems. This fact happens with

networks of moderate density, as can be observed in the results displayed in all tables of this section.

These tables show the behaviour of the main algorithm over a set of networks. In most types of network we have considered not excessively large values of m. It will be seen that for all networks observed, the values of τ_N were quite moderate; in fact we have found that $\tau_N \in O(m)$ in all cases examined, and therefore the complexity of the algorithm will remain reduced to $O(m(m + n))$ time.

With the aim of showing the performance of τ_N, we introduce a parameter a_N, with $0 \leq a_N \leq n$, and we study whether $\tau_N \leq a_N m$ occurs.

We have analysed randomly generated networks, with a variable number, n, of vertices, and for each input n we have considered a variable number, m, of edges. The method used to generate each network started with a randomly obtained generator tree whose number of edges was increased up to the desired quantity m. The resultant networks are not planar in general, since the edges have been randomly placed. Weights and lengths were generated uniformly in $[1, 20]$ and $[1, 50]$, respectively.

For each type of $N(n, m)$ network, 100 instances were tested. The values of a_N range from an initial to a final value, with a fixed increment (which is 0.5 in the first table, and 1 in the remaining tables). Tables 1, 2 and 3 show the results obtained. In each table, each element represents the number of times (or percentage) that $\tau_N \leq a_N m$ occurred.

TABLE 1

$n = 25$	a_N										
m	0.5	1	1.5	2	2.5	3	3.5	4	4.5	5	5.5
30	65	100									
40		1	32	84	100						
50				6	44	80	95	100			
60					2	33	81	97	100		
70						4	37	81	97	100	
80							15	61	90	99	100
90							3	48	92	98	100
100							1	41	76	89	100

TABLE 2

$n = 50$	a_N										
m	1	2	3	4	5	6	7	8	9	10	11
60	78	100									
80		27	66	100							
100			2	37	89	100					
120				19	55	99	100				
140						26	81	100			

TABLE 3

$n = 80$	a_N										
m	1	2	3	4	5	6	7	8	9	10	11
100	5	44	100								
140	1	34	87	100							
170				29	72	100					
200						12	67	100			
230							18	72	98	100	

In all tables shown it can be observed that initially small values of a_N are sufficient to trap all the instances, and these values increase when m increases. However, even with a very dense network, (as happens in Tab. 1, for $n = 25$ and $m = 100$), every single case is trapped with $a_N = 5.5$. This means that $\tau_N \leq a_N m = 550$, since $a_N m$ is an upper bound of τ_N. Likewise, a similar situation occurs in Tables 2 and 3: the upper bound of all types $N(50, 140)$ and $N(80, 230)$ is obtained for $a_N = 8$ and $a_N = 10$, respectively. It is highly probable that such values can be reduced when planar networks are considered.

We remark that the bound $a_N m$ was never reached, that is, the last "a_N-value" of each row provided a strict upper bound for 100 instances tested (for example, in the type $N(50, 120)$, the greatest value of τ_N was 742, and in $N(80, 230)$, such a value was 2109). Regarding these data, we can infer that the running time of τ_N was $O(m)$ in all networks tested in these computational experiments.

From these results we can expect that, when the underlying network is of moderate density, the value of $\tau_{_N}$ is bounded by a small value of $a_{_N}$, and therefore the computational effort of the main algorithm would be appreciably lower. When the network is very dense we cannot know the size of this upper bound. However, we have already seen that, in such a case, the number of visits made by the main algorithm is smaller than the visits executed by the exhaustive procedure.

In conclusion, we have shown in this paper how to reduce the effort to solve the continuous variance problem in a network by applying a recursive procedure to compute the objective function along each edge. In order to analyse the complexity we have considered uniform density functions, since the minimum in each subproblem is obtained in constant time, and this has no influence on the computational order. However, several questions remain to be studied, as we comment in the following. When arbitrary density functions are considered, a numerical procedure will be needed to compute the minimum in each primary region, the difficulty of which will depend on the type of density function involved. It would be interesting to know what density function provides the worst complexity. Another question deals with discretization of the problem by means of a number of artificial nodes, and comparison of the complexity of the resultant problem with the continuous case, together with determining the level of approximation of both solutions. As far as the authors are aware there is no theoretical analysis of the aggregation issue for continuous edge demand problems.

APPENDIX A

Proof of Proposition 1: Developing the first component of $z_\nu(x)$ relative to nodal contribution we obtain

$$\sum_{v_i \in V} w_i d(v_i,x)^2 + w_{_N} z_m(x)^2 - 2z_m(x)\sum_{v_i \in V} w_i d(v_i,x).$$

A similar procedure applied to the component of $z_\nu(x)$ relative to edge contribution yields

$$\sum_{e_{jk} \in E} p_{jk}\int_0^{l_{jk}} d(y,x)^2 f_{jk}(y)\,dy + w_{_E} z_m(x)^2$$

$$- 2z_m(x)\sum_{e_{jk} \in E} p_{jk}\int_0^{l_{jk}} d(y,x)f_{jk}(y)\,dy$$

where $F_{jk}(l_{jk}) = 1$ has been used. The Proposition holds by summing both expressions and using the fact that $w(N) = 1$. ▢

Proof of Proposition 9: By the above relations, we have $E_u(S_j) \subset E_u(S_{j-1})$. First, suppose that $c_{ht} \in E_u(S_j)$. If it were true that $c_{ht} = [v_h, v_t] \in B(E_u(S_{j-1}))$ then at least one of the end vertices (or both) would belong to $B(\bar{x}_{j-1})$. Let v_h be such a vertex. From Proposition 7 it follows that $v_h \notin V_u(S_j)$, which contradicts the fact that $[v_h, v_t] \in E_u(S_j)$. This proves that $E_u(S_j) \subseteq E_u(S_{j-1}) \backslash B(E_u(S_{j-1}))$.

Reciprocally if $c_{ht} \in E_u(S_{j-1}) \backslash B(E_u(S_{j-1}))$ then $v_h, v_t \in V_u(S_{j-1})$ and $v_h, v_t \notin B(\bar{x}_{j-1})$ (since $B(E_u(S_{j-1}))$ is the set of edges of $E_u(S_{j-1})$ such that at least one vertex belongs to $B(\bar{x}_{j-1})$). Then, from Proposition 7 it follows that $v_h, v_t \in V_u(S_j)$ and therefore $c_{ht} \in E_u(S_j)$, which proves the converse inclusion. Both inclusions complete the proof. ▢

Proof of Proposition 10: Since $V_v(S_{j-1}) \subset V_v(S_j)$, we have $E_v(S_{j-1}) \subset E_v(S_j)$. Consider $c_{ht} \in B^{(2)}(E_u(S_{j-1}))$. Then $v_h, v_t \in B(\bar{x}_{j-1})$, and therefore $v_h, v_t \in V_v(S_j)$ which implies $c_{ht} \in E_v(S_j)$. If $c_{ht} \in B^{(1)}(E_{[u,v]}(S_{j-1}))$, then an unique vertex of edge c_{ht} belongs to $B(\bar{x}_{j-1})$. Suppose this vertex is v_h, then $v_h \in V_v(S_j)$. But, as $B(\bar{x}_{j-1}) \subset V_u(S_{j-1})$ and $c_{ht} \in E_{[u,v]}(S_{j-1})$ then $v_h \in V_u(S_{j-1})$ which implies $v_t \in V_v(S_j)$, and therefore $c_{ht} \in E_v(S_j)$. This reasoning proves

$$E_v(S_j) \supseteq E_v(S_{j-1}) \cup B^{(1)}(E_{[u,v]}(S_{j-1})) \cup B^{(2)}(E_u(S_{j-1})).$$

On the other hand suppose $c_{ht} \in E_v(S_j)$. Then $v_h, v_t \in V_v(S_j) \cup B(\bar{x}_{j-1})$. If $v_h, v_t \in V_v(S_{j-1})$ or $v_h, v_t \in B(\bar{x}_{j-1})$, the result is trivial. If $v_h \in V_v(S_{j-1})$, $v_t \in B(\bar{x}_{j-1})$, then $v_t \in V_u(S_{j-1})$ which implies $c_{ht} \in E_{[u,v]}(S_{j-1})$. Since $v_h \notin B(\bar{x}_{j-1})$ (otherwise we return to the former case) it follows that $c_{ht} \in B^{(1)}(E_{[u,v]}(S_{j-1}))$. This proves the desired inclusion and concludes the proof. ▢

Proof of Proposition 11: This relationship will be proved by using the same procedure as in Proposition 10. First, suppose $c_{ht} \in E_{[u,v]}(S_j)$. Then $v_h \in V_u(S_{j-1}) \backslash B(\bar{x}_{j-1})$, $v_t \in V_v(S_{j-1}) \cup B(\bar{x}_{j-1})$. The assumption $v_t \in V_v(S_{j-1})$ implies $c_{ht} \in E_{[u,v]}(S_{j-1}) \backslash B^{(1)}(E_{[u,v]}(S_{j-1}))$ (otherwise from (5) it follows that $c_{ht} \in E_v(S_j)$, which is a contradiction). If $v_t \in B(\bar{x}_{j-1})$, $c_{ht} \in E_u(S_{j-1})$ must hold from Proposition 7. Furthermore, $c_{ht} \in B^{(1)}(E_u(S_{j-1}))$ since $v_h \notin B(\bar{x}_{j-1})$. This reasoning proves the relationship

$$E_{[u,v]}(S_j) \subseteq \left(E_{[u,v]}(S_{j-1}) \backslash B^{(1)}(E_{[u,v]}(S_{j-1})) \right) \cup B^{(1)}(E_u(S_{j-1})).$$

Conversely if $c_{ht} \in E_{[u,v]}(S_{j-1}) \backslash B^{(1)}(E_{[u,v]}(S_{j-1}))$ then $v_h \in V_u(S_{j-1})$, $v_t \in V_v(S_{j-1})$ and $v_h, v_t \notin B(\bar{x}_{j-1})$. In effect, the supposition $v_t \in B(\bar{x}_{j-1})$ implies $v_t \in V_u(S_{j-1})$, which is a contradiction. Furthermore $v_h \notin B(\bar{x}_{j-1})$ holds from the assumption $c_{ht} \notin B^{(1)}(E_{[u,v]}(S_{j-1}))$ together with $v_t \notin B(\bar{x}_{j-1})$. Therefore $c_{ht} \in E_{[u,v]}(S_j)$ since, from Proposition 7, $v_h \in V_u(S_j)$, $v_t \in V_v(S_j)$. A similar reasoning proves that $c_{ht} \in E_{[u,v]}(S_j)$ is also obtained when $c_{ht} \in B^{(1)}(E_u(S_{j-1}))$, and this completes the proof. $\square$

Proof of Proposition 12: Suppose that $\bar{x}_j < \bar{x}_q$. First we show that the Proposition holds for any two consecutive half-open sets S_j, S_{j+1}. The general case ($q > j + 1$) will be obtained by repeating this process.

If $q = j + 1$, the first relationship follows directly from (4) and from the fact that $B(E_u(S_{j+1})) \subseteq E_u(S_{j+1})$. Similarly the second expression is a consequence of (6) and the fact that $B^{(1)}(E_{[u,v]}(S_{j+1})) \subseteq E_{[u,v]}(S_{j+1})$.
$\square$

Glossary

$z_v(x)$	Continuous variance function.
$z_m(x)$	Continuous median function.
$\bar{x}_j, \, j = 0, \ldots, k+1$	Bottleneck points of the edge $[u,v]$.
$[\bar{x}_{j-1}, \bar{x}_j]$	j–th. primary region.
$S_j = (\bar{x}_{j-1}, \bar{x}_j]$	j–th. half open set.
$\{V_u(S_j), V_v(S_j)\}$	Partition of V associated to S_j.
$\{E_u(S_j), E_v(S_j), E_{[u,v]}(S_j), [u,v]\}$	Partition of E associated to S_j.
$B(\bar{x}_j)$	Set of vertices whose bottleneck point is $\bar{x}_j$.
$B^{(1)}(E_u(S_j))$	Edges of $E_u(S_j)$ with an unique end vertex in $B(\bar{x}_j)$.
$B^{(2)}(E_u(S_j))$	Edges of $E_u(S_j)$ with the two end vertex in $B(\bar{x}_j)$.
$B(E_u(S_j))$	The set $B^{(1)}(E_u(S_j)) \cup B^{(2)}(E_u(S_j))$.
$B^{(1)}(E_{[u,v]}(S_j))$	Edges of $E_{[u,v]}(S_j)$ with an unique end vertex in $B(\bar{x}_j)$.
$\tau_j(u,v)$	The cardinal of the set $E_{[u,v]}(S_j) \backslash B^{(1)}(E_{[u,v]}(S_j))$.
$\tau(u,v)$	The sum of $\tau_j(u,v)$ along the edge $[u,v]$.
τ_N	The maximum of $\tau(u,v)$ values over N.

REFERENCES

1. A. AHO, J. HOPCROFT and J. ULLMAN, The design and analysis of computer algorithms. Addison-Wesley (1974).

2. O. BERMAN, Mean-Variance location problems. *Transportation Science* **24** (1990) 287-293.

3. M.L. BRANDEAU, S.S. CHIU and R. BATTA, Locating the two-median of a tree network with continuous link demands. *Ann. Oper. Res.* **6** (1986) 223-253.

4. T.M. CAVALIER and H.D. SHERALI, Network location problems with continuous link demands: p-medians on a chain and 2-medians on a tree. *European J. Oper. Res.* **23** (1986) 246-255.

5. S.S. CHIU, The minisum location problem on an undirected network with continuous link demands. *Comput. Oper. Res.* **14** (1987) 369-383.

6. R.L. CHURCH and R.S. GARFINKEL, Locating an obnoxious facility on a network. *Transportation Sci.* **12** (1978) 107-118.

7. R.S. GARFINKEL, A.W. NEEBE and M.R. RAO, The m-center problem: Minimax facility location. *Management Sci.* **23** (1977) 1133-1142.

8. S.L. HAKIMI, Optimum locations of Switching Centers and the Absolute Centers and Medians of a graph. *Oper. Res.* **12** (1964) 450-459.

9. S.L. HAKIMI, Optimum distribution of switching centers in a communications network and some related graph-theoretic problems. *Oper. Res.* **13** (1965) 462-475.

10. J. HALPERN and O. MAIMON, Accord and conflict among several objectives in locational decisions on tree networks, in Locational Analysis of Public Facilities, edited by J.F. THISSE and H.G. HOLLER, Series of Studies in Mathematical and Managerial Economics. North-Holland, Amsterdam (1983) 301-314.

11. P. HANSEN and M. ZHENG, An algorithm for the minimum variance point of a network. *RAIRO Oper. Res.* **25** (1991) 119-126.

12. J.N. HOOKER, Solving nonlinear single–facility network location problems. *Oper. Res.* **34** (1986) 732-743.

13. R.K. KINCAID and O. MAIMON, Locating a point of minimum variance on triangular graphs. *Transportation Sci.* **23** (1989) 216-219.

14. R.K. KINCAID and O. MAIMON, A note on locating a central vertex of a 3-cactus graph. *Comput. Oper. Res.* **17** (1990) 315-320.

15. M. LABBÉ, The median location problem with continuous demand. *Cahiers Centre Étude Rech. Opér.* **34** (1992) 103-115.

16. M.C. LÓPEZ DE LOS MOZOS, Localización en redes con el criterio varianza, *Tesis doctoral*, Universidad de Sevilla (1997).

17. O. MAIMON, The variance equity measure in locational decision theory. *Ann. Oper. Res.* **6** (1986) 147-160.

18. P.T. NKANSAH and H.T. DAVID, Network median problems with continuously distributed demand. *Transportation Sci.* **20** (1986) 213-219.