Solving the Two-level Clustering Problem by Hyperbolic Smoothing Approach, and Design of Multicast Networks

Gentil Veloso Barbosa, Sergio Barbosa Villas-Boas, Adilson Elias Xavier
Federal University of Rio de Janeiro - Brazil
e-mail: gentil@cos.ufrj.br, sbvb@cos.ufrj.br, adilson@cos.ufrj.br

Abstract

The design of multicast networks constitute an important problem in phone networks, the public Internet, ad-hoc wireless networks, transportation and sensor networks.

We consider the problem of designing a multicast network by modeling it as a two-level clustering problem, that is, to find the location of multiple hubs and the location of one main hub (super hub), connected in a tree. The location of hubs is defined in continuous bi-dimensional space.

The optimum design is defined as the one that minimizes the cost function as the sum of distances of the observations to its nearest hub. There is a transportation problem with similar formulation.

The specification of the problem corresponds to a strongly non differentiable *min-sum-min* formulation. This is a NP-hard problem to be solved [6, 8]. We propose a method to overcome this difficulty by using hyperbolic smoothing strategy. The solution is ultimately obtained by solving a sequence of differentiable unconstrained low dimension optimization sub-problems.

The proposed method is tested with some computational experiments. The results confirm the robustness and accuracy of the proposed method.

Keywords: Two-level clustering problem, continuous space, Multicast Hub Location, Min-Sum-Min, Hyperbolic Smoothing

1 Introduction

The design of multicast networks constitute an important problem in various communication contexts such as phone networks, the public Internet, ad-hoc wireless networks, sensor networks and transportation.

In this work it is considered the problem of designing a multicast network by modeling it as a two-level clustering problem, that is, to find the location of multiple hubs and the location of one main hub, connected in a tree. The location of hubs is defined in continuous bi-dimensional space. The optimum design is defined as the one that minimizes the cost function as the sum of distances of the observations to its nearest hub. There is a transportation problem with similar formulation. The specification of the problem corresponds to a strongly non differentiable *min-sum-min* formulation. This is a NP-hard problem to be solved [6, 8].

In recent years, several studies concerning the hub location problem, being the hubs connected in hierarchical tree, have been published [1].

In [13] it is addressed the problem of determining the routes and the hubs to be used in order to send, at minimum cost, a set of commodities from sources to destinations in a network of given given capacity. In [4] it is presented the tree of hubs location problem. It is a network hub location problem with single assignment where a fixed number of hubs have to be located, with the particularity that it is required that the hubs are connected by means of a tree. In [10] it is discussed the hub location problem for weighted links, which is a variant of the classical hub problem. In [16] it is discussed the techniques for organizing hierarchies using clustering, which could be applied to design a multicast network. In [18] it is proposed the use of clustering algorithms to help in determining hierarchical multicast trees, considering mainly the proximity, by apply the k-means clustering algorithms iteratively to construct successive hierarchical levels. In [17] it is introduced a flexible new method of constructing hierarchical multicast structures suitable for supporting large-scale GRID applications. In [2] it is shown how to construct a class of hierarchical multicast trees and analyze their performances.

In this work it is proposes a method to solve the hierarchical two-level clustering problem, that is to find the location of hubs with one main super hub to that minimizes the cost of sum of links. The formulation adopted in this work is the same as [3, 14, 7, 4], where it is shown that the multicast network design problem is mapped to the same model of the hierarchical two-level clustering problem. This model is a NP-hard problem to be solved [6, 8]. There are other optimization problems that can also be mapped to the two-level clustering model.

The proposed solution uses hyperbolic smoothing approach. It assumes that it is a possibility to position the hubs continuous space, that is, minor adjustments can be made to the position of centroids in any position in $\mathbb{R}^2$.

This work is organized in the following way. The specification of the continuous hub location multicast problem is presented in the next section. The hyperbolic smoothing methodology is described in section 3. The illustrative computational results are presented in section 4, and some concluding remarks are drawn in a final section.

2 The Continuous Hubs Location Problem

The continuous hubs location problem consists in calculating the location of p hubs in a planar region, in order to define a multicast network that connects m places, also known as "consumer points" or "cities", while minimizing the network's cost. One of the hubs is known as a "super-hub", because all other hubs are connected to it.

To formulate this problem, we proceed as follows. Let $S = \{s_1, \dots, s_m\}$ denote a set of m cities in a planar region, $s_i \in \mathbb{R}^2$. Let x_i, $i = 1, \dots, p$ be the hubs, where each $x_i \in \mathbb{R}^2$. The set of these hubs is represented by $X \in \mathbb{R}^{2p}$. Each city s_i must be connected to one hub (might be the super-hub). Each hub must be connected to the one super-hub. The cost to be minimized is the sum of the distances of the cities to hubs plus the distances among hubs.

A simple configuration of this network is depicted in figure 1. The distances from cities (circles) to hubs (diamonds) are in blue. The distances from hubs to super-hub (pentagon) are in red.

The unitary cost associated to city s_j and hub x_i is equal to the Euclidean distance between them:

$$z_{ij} = \|s_j - x_i\|_2 \tag{1}$$

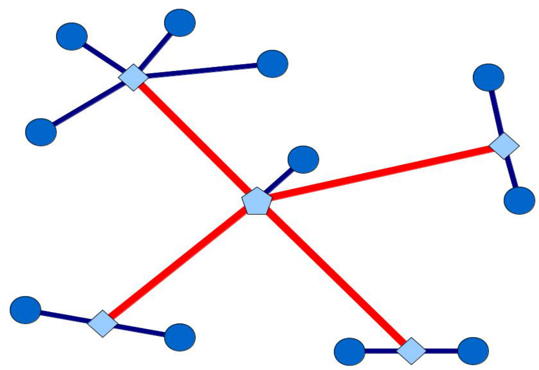


Figure 1: Super-hub (pentagon), hubs (diamonds) and cities (circles)

The unitary cost that actually is considered (added) to the cost function to be minimized is the distance of a city j to the nearest hub i , that is,

$$z_j = \min_{i=1,\dots,p} z_{ij}, \tag{2}$$

The problem is to calculate the hub locations x_i to minimize the total cost of the network, that is, the sum of all costs of each city to its nearest hub, plus the cost of connecting the hubs.

$$\begin{aligned} & \text{minimize} \left(\sum_{j=1}^{m} z_j \right) + v && (3) \\ \text{subject to} \quad & z_j = \min_{i=1,\dots,p} z_{ij}, \quad j = 1, \dots, m. \\ & v = \min_{i=1,\dots,p} \sum_{l=1}^{p} \|x_l - x_i\|_2 \end{aligned}$$

The v component of the cost function represent the cost produced by the connecting the hubs. For each hub i it is calculated the sum of distances of the all hubs to it. The *min* operator makes v the minimum of all possible sums. This formulation implicitly produces the concept of super-hub.

The super-hub is the hub to which all other hubs will be connected to. It is also allowed to cities to connect directly to the super-hub. The formulation of the problem will make one of the hubs x_i to be the super-hub. The index that identifies which of the hubs is the super-hub shall be defined as $\bar{\iota}$, that is $x_{\bar{\iota}}$ is the super-hub.

Whatever the index $\bar{\iota}$ be, the minimization process will produce the position of the hubs that minimizes the global cost, taking into account the portion of the cost of the interconnections among hubs.

After the minimization calculation is done, it can be determined the index of the super-hub using the equation (4).

$$\bar{\iota} = \arg\min_i \left(\sum_{l=1}^{p} ||x_l - x_i||_2 \right) \tag{4}$$

3 Smoothing the Problem

This problem has a *min-sum-min* structure with non-differentiable and non-convex characteristics, having a very large number of local minima. A series of transformations will be performed in order to obtain a continuous formulation. First, considering its definition, each z_{ij} must necessarily satisfy the following set of inequalities:

$$z_j - z_{ij} \le 0, \quad i = 1, \dots, p. \tag{5}$$

Substituting these inequalities for the equality constraints of problem (3), we perform a relaxation and the following problem is obtained:

$$\begin{aligned} &\text{minimize} \left(\sum_{j=1}^{m} z_j \right) + v \\ \text{subject to} \quad & z_j - z_{ij} \le 0, \quad i = 1, \dots, p; \quad j = 1, \dots, m. \\ & v = \min_{i=1,\dots,p} \sum_{l=1}^{p} \|x_l - x_i\|_2 \end{aligned} \tag{6}$$

The variables z_{ij}, from equation (2), must be non-negative. However, in the last formulation, the variables z_j are no longer bound from below. So, in order to obtain the desired equivalence with the original problem, we must modify problem (6) to create an adequate lower bound for variables z_j . We define function $\psi(y)$ as in equation (7).

$$\psi(y) = \max(0, y) \tag{7}$$

Applying (7) to (6), it follows that

$$\sum_{i=1}^{p} \psi(z_j - z_{ij}) = 0, \quad j = 1, \ldots, m. \tag{8}$$

For any point j assuming a strictly increasing sequence of the p distances z_{ij} denoted by $d_1 < d_2 < \cdots < d_p$, Figure 2 illustrates the first three summands of (8) as a function of z_j.

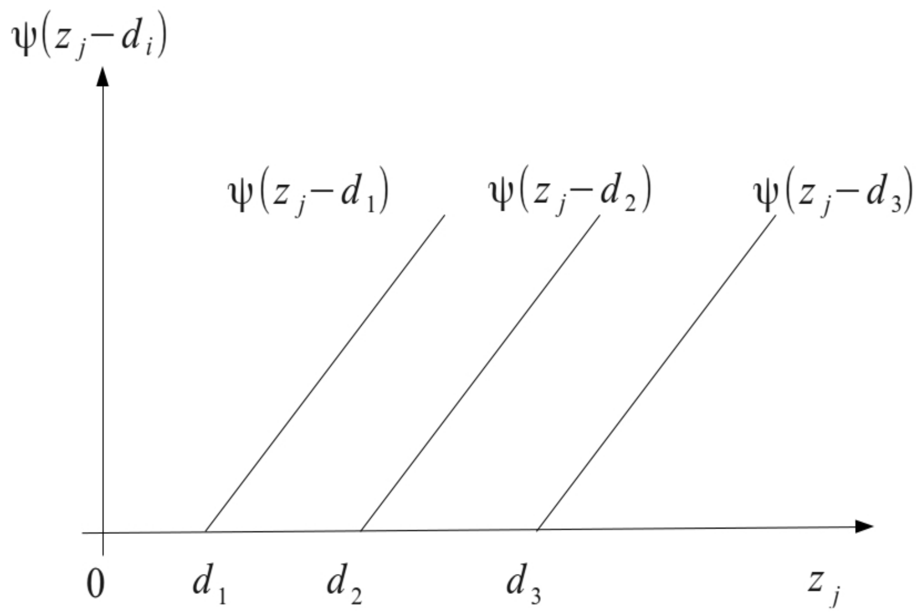


Figure 2: Summands in (8), $\psi(z_j - d_i)$

Using (8) in place of the set of inequality constraints in (6), we would obtain an equivalent problem maintaining the undesirable property that z_j still has no lower bound. Considering, however, that the objective function of problem (6) will force downward each z_j, $j = 1, \ldots, m$, we can think of bounding the latter variables from below by including an $\varepsilon > 0$ perturbation in (8). So, the following modified problem is obtained:

$$\begin{aligned} & \text{minimize} \left(\sum_{j=1}^{m} z_j \right) + v && (9) \\ \text{subject to} \quad & \sum_{i=1}^{p} \psi(z_j - z_{ij}) \ \geq \ \varepsilon , \quad j = 1, \ldots, m. \\ & v \ = \min_{i=1,\ldots,p} \sum_{l=1}^{p} \|x_l - x_i\|_2 \end{aligned}$$

Since the feasible set of problem (3) is the limit of that of (9) when $\epsilon \to 0_+$, we can consider solving (3) by solving a sequence of problems like (9) for a sequence of decreasing values of ε that approaches zero. However, the definition of function ψ endows problem (9) with an extremely rigid nondifferentiable structure, which makes its computational solution very hard. In view of this, the numerical method we adopt for solving problem takes a smoothing approach. From this perspective, let us define the function:

$$\phi(y, \tau) = \left(y + \sqrt{y^2 + \tau^2} \right) / \, 2 \, , \tag{10}$$

for $y \in \mathbb{R}$ and $\tau > 0$.

Function ϕ has the following properties:

(a) $\phi(y, \tau) \ > \ \psi(y), \quad \forall \, \tau > 0;$

(b) $\lim_{\tau \to 0} \ \phi(y, \tau) \ = \ \psi(y);$

(c) $\phi(y, \tau)$ is an increasing convex C^{∞} function in variable y.

Therefore, function ϕ constitutes an approximation of function ψ. Adopting the same assumptions used in figure 2, the first three summands

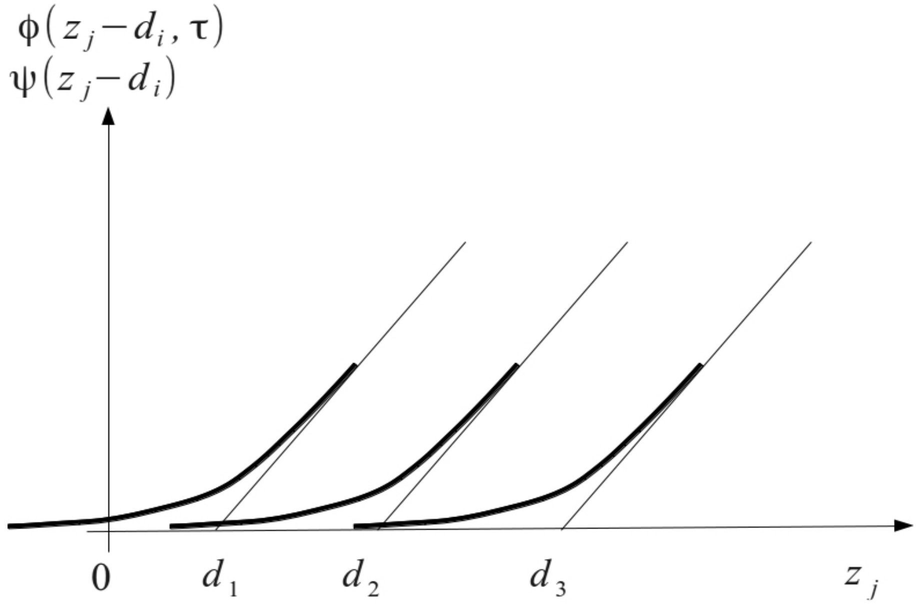


Figure 3: Original $[\psi(z_j - d_i)]$, smoothed $[\phi(z_j - d_i, \tau)]$ summands in (8)

of (8) and their corresponding smoothed approximations, given by (10), are depicted by the curves in figure 3.

By using function ϕ in place of function ψ in (9), we get to the problem:

$$\text{minimize} \left(\sum_{j=1}^{m} z_j \right) + v \tag{11}$$

$$\text{subject to} \quad \sum_{i=1}^{p} \phi(z_j - z_{ij}, \tau) \geq \varepsilon, \qquad j = 1, \ldots, m.$$

$$v = \min_{i=1,\ldots,p} \sum_{l=1}^{p} \| x_l - x_i \|_2$$

To obtain a differentiable problem, it is necessary further to smooth the Euclidean distances z_{ij}. For this purpose, let us define the function

$$\theta(c , d , \gamma) = \sqrt{(c_1 - d_1)^2 + (c_2 - d_2)^2 + \gamma^2} \, , \tag{12}$$

where $c, d \in \mathbb{R}^2$ and $\gamma > 0$.

Function θ has the following properties:

(a) $\lim_{\gamma \to 0} \theta(c, d, \gamma) = \|c - d\|_2$;

(b) θ is a C^∞ function.

By using function θ in place of the Euclidian distances, we obtain the completely differentiable problem:

$$\begin{aligned} &\text{minimize} \left(\sum_{j=1}^{m} z_j \right) + v \qquad (13) \\ &\text{subject to} \quad \sum_{i=1}^{p} \phi(z_j - \theta(s_j, x_i, \gamma), \tau) \geq \varepsilon, \quad j = 1, \dots, m. \\ &v = \min_{i=1,\dots,p} \sum_{l=1}^{p} \|x_l - x_i\|_2 \end{aligned}$$

Now, the properties of functions ϕ and θ allow us to seek a solution to problem (9) by solving a sequence of subproblems like problem (13), produced by decreasing the parameters $\gamma \to 0$, $\tau \to 0$, and $\varepsilon \to 0$.

The objective function minimization process will work towards reducing to the utmost the z_j, $j = 1, \dots, m$, values. On the other hand, given any set of hubs x_i, $i = 1, \dots, p$, due to property (c) of the hyperbolic smoothing function ϕ, the constraints of problem (13) are monotonically increasing functions in z_j. So, these constraints will certainly be active and problem (13) will ultimately be equivalent to the following problem:

$$\text{minimize} \left(\sum_{j=1}^{m} z_j \right) + v \tag{14}$$

$$\text{subject to} \quad h_j(z_j, x) \;=\; \sum_{i=1}^{p} \phi(z_j - \theta(s_j, x_i, \gamma),\, \tau\,) - \varepsilon = 0, \quad j = 1, \ldots, m.$$

$$v \;=\; \min_{i=1,\ldots,p} \sum_{l=1}^{p} \|x_l - x_i\|_2$$

The dimension of the variable domain space of problem (14) is $(2p+m)$. Since, in general, the value of parameter m, the cardinality of the set S of the cities s_j, is large, problem (14) has a large number of variables. However, it has a separable structure, because each variable z_j appears only in one equality constraint. Therefore, as the partial derivative of $h_j(z_j, x)$ with respect to z_j, $j = 1, \ldots, m$ is not equal to zero, it is possible to use the Implicit Function Theorem to calculate each component z_j, $j = 1, \ldots, m$ as a function of the hub location variables x_i, $i = 1, \ldots, p$. This way, the unconstrained problem:

$$\text{minimize } f(x) \;=\; \left(\sum_{j=1}^{m} z_j(x) \right) + v(x) \tag{15}$$

$$v \;=\; \min_{i=1,\ldots,p} \sum_{l=1}^{p} \|x_l - x_i\|_2$$

is obtained, where each $z_j(x)$ results from the calculation of a zero of each equation

$$h_j(z_j, x) \;=\; \sum_{i=1}^{p} \phi(z_j - \theta(s_j, x_i, \gamma),\, \tau\,) - \varepsilon = 0, \quad j = 1, \ldots, m \tag{16}$$

Due to property (c) of the hyperbolic smoothing function, each term ϕ above is strictly increasing with variable z_j and therefore the equation has a single zero.

Again, due to the Implicit Function Theorem, the functions $z_j(x)$ have all derivatives with respect to the variables x_i, $i = 1, \ldots, p$, and therefore it is possible to calculate the gradient of the objective function of problem (15),

$$\nabla f(x) = \left(\sum_{j=1}^{m} \nabla z_j(x) \right) + \nabla v(x), \tag{17}$$

where

$$\nabla z_j(x) = - \nabla h_j(z_j, x) \ / \ \frac{\partial h_j(z_j, x)}{\partial z_j}, \tag{18}$$

while $\nabla h_j(z_j, x)$ and $\partial h_j(z_j, x)/\partial z_j$ are directly obtained from equations (10), (12) and (16).

In the equation (17), the term $\nabla v(x)$ is indeed non-differentiable, however the super-hub index $\bar{\iota}$ rarely changes during the optimization procedure. In this condition, $\nabla v(x)$ is continuous and can be easily calculated by equation (19).

$$\nabla v(x) = \sum_{l=1}^{p} \frac{x_l - x_{\bar{\iota}}}{||x_l - x_{\bar{\iota}}||_2} \tag{19}$$

This way, it is easy to solve problem (15) by making use of any method based on first order derivative information. Finally, it must be emphasized that problem (15) is defined on a $(2p)-$dimensional space, so it is a small problem, since the number of hubs, p, is small, in general, for real world applications.

Simplified HSHS Algorithm

The solution of the original problem can thus be obtained by using what we call “Hyperbolic Smoothing Hub-and-Spoke Algorithm (HSHS)”, described below in a simplified way.

Initialization Step: Choose initial values: x^0, γ^1, τ^1, ε^1.

Choose values $0 < \rho_1 < 1$, $0 < \rho_2 < 1$, $0 < \rho_3 < 1$; let $k = 1$.

Main Step: Repeat until a stopping rule is attained

Solve problem (15) with $\gamma = \gamma^k$, $\tau = \tau^k$ and $\varepsilon = \varepsilon^k$, starting at the initial point x^{k-1} and let x^k be the solution obtained.

Let $\gamma^{k+1} = \rho_1 \gamma^k$, $\tau^{k+1} = \rho_2 \tau^k$, $\varepsilon^{k+1} = \rho_3 \varepsilon^k$, $k := k + 1$. ■

Notice that the algorithm causes τ and γ to approach 0, so the constraints of the subproblems it solves, given as in (15), tend to those of (9). In addition, the algorithm causes ε to approach 0, so, in a simultaneous movement, the problem (9) gradually approaches problem (3).

4 Computational Results

The numerical experiments have been carried out on a computer SGI-model C1104-2TY9, with 24G RAM, 2T of disk, CPU 24 cores (but only one core was actually used). The operating system was linux open Suse v 12.2. The compiler used was GNU's gfortran v 4.5.2. The unconstrained minimization tasks were carried out by means of a Quasi-Newton algorithm, employing the BFGS updating formula from the Harwell Library [11].

The computational results presented below were obtained from a first implementation of the algorithm, without any sort of pruning procedure and where the initial starting hubs $x_i^0, i = 1, \cdots, p$ were arranged, in a very simple manner, around the center of gravity of the set of cities, by making random perturbations proportional to the standard deviation of this set, following the *formulae*: $\bar{s} = \sum_{j=1}^{m} s_j / m$, $\sigma = (\sum_{j=1}^{m} \| s_j - \bar{s} \|_2^2 / m)^{1/2}$, $x_i = \bar{s} + a\,\sigma$, where the components of the vector a are uniform random variables in the interval $[-0.5, +0.5]$.

The value of τ^1 was taken as $1/100$ of this standard deviation. The following choices were made for the other parameters:
$\varepsilon^1 = 4\,\tau^1$, $\gamma^1 = \tau^1/100$, $\rho_1 = 1/4$, $\rho_2 = 1/4$ and $\rho_3 = 1/4$.

In order to illustrate the performance of the proposed algorithm, we show below the results obtained by processing 2 data-sets - one with 76 cities, and other with 1002 cities, both from [12]. For each data-set, we solve the problem for the range of 2 to 10 hubs. The obtained computational results are presented in tables 1 and 2. The images of the results for 5 to 10 hubs (only 6 images per data-set to save space) for each problem is depicted in figures 4 and 5.

As this is a global optimization problem, we use a multistart strategy with $T = 100$ different tentative randomly chosen starting points for each instance.

In tables 1 and 2 the first column presents the specified number of hubs (p). The second column presents the best objective function value (f_{HSHS}) produced by the HSHS algorithm. The next three columns present the number of occurrences of the best solution $(Occur.)$, the average percentage error of the T solutions (E_{Mean}) in relation to the best solution obtained (f_{HSHS}) and the CPU mean time given in seconds (T_{Mean}). By defining f^t as the optimum value of the objective function obtained from the tentative starting point t, the percentage error E_{Mean} observed in all T attempts is calculated by the expression:

$$E_{Mean} = \frac{100 \sum_{t=1}^{T} (f^t - f_{HSHS})}{T\ f_{HSHS}}. \tag{20}$$

The necessary effort for the resolution of (15) depends first on the dimension of the problem, equal to $2p$. However, each evaluation of the objective function involves the calculation of m distances. One at a time, each distance z_j is obtained through the calculation of a zero of equation (16), which has p terms. So, the resultant function evaluation complexity is $O(mp)$. By assuming that the number of iterations taken in the main step of the HSHS algorithm is proportional to the dimension of the variable domain space, the bulk problem resolution complexity would be $O(mp^2)$.

Table 1 presents the computational results obtained for the German Towns instance with 76 cities. For the cases of small p, the occurrences of the best solution, given by column $Occur.$ are relatively more common. The average errors presented by the 10 solutions, given by column E_{Mean}, have small values, which indicate a consistent performance of the algorithm.

From the results presented in tables 1 and 2 one sees that the first implementation of the HSHS Algorithm for solving large continuous hub-and-spoke instances confirms the consistency of the proposed methodology. The expressive number of occurrences of the best solution for small cases ($p \leq 4$) shows the algorithm's consistent performance. The low values of the average error of the 10 solutions (E_{Mean}), in relation to the best solution obtained, show the consistency and the numerical stability of the algorithm.

p	f_{HSHS}	$Occur.$	E_{Mean}	T_{Mean}
2	0.300797D+06	82	2.87	0.04
3	0.266465D+06	1	4.49	0.02
4	0.237632D+06	49	0.47	0.09
5	0.213695D+06	15	0.23	0.12
6	0.199957D+06	11	1.14	0.15
7	0.188349D+06	4	1.13	0.25
8	0.179200D+06	44	1.34	0.30
9	0.173214D+06	25	1.61	0.37
10	0.168442D+06	9	2.27	0.46

Table 1: Results for the 76-city problem (Padberg/Rinaldi)

p	f_{HSHS}	$Occur.$	E_{Mean}	T_{Mean}
2	0.340215D+07	100	0.00	0.25
3	0.281685D+07	24	0.13	0.43
4	0.232770D+07	90	0.37	0.61
5	0.194299D+07	81	1.10	0.94
6	0.170837D+07	50	0.80	1.32
7	0.158897D+07	31	0.23	1.66
8	0.147249D+07	38	0.90	2.23
9	0.136153D+07	70	1.56	2.67
10	0.130772D+07	40	1.19	3.34

Table 2: Results for the 1002-city problem (Padberg/Rinaldi)

Finally, the Hub Location Multicast Problem were suitably solved in adequate CPU times, as a consequence of both the low dimension of the nonlinear problem (15), defined on a $(2p)-$dimensional space, and of the use of a minimization algorithm that takes advantage of its C^∞ differentiability property.

5 Conclusions

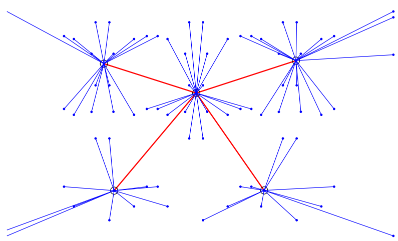
This work analyzed and proposed an algorithm to solve the two-level clustering problem, that is, to find the location of multiple hubs and the location of one main hub (super hub), connected in a tree. The location of hubs is defined in continuous bi-dimensional space. The designing of multicast network can be mapped to the considered problem.

The optimum design is defined as the one that minimizes the cost function as the sum of distances of the observations to its nearest hub. There is a transportation problem with similar formulation.

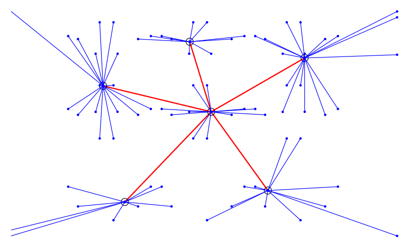
The specification of the problem corresponds to a strongly non differentiable *min-sum-min* formulation. This is a NP-hard problem to be solved [6, 8]. It was proposed a method to overcome this difficulty by using hyperbolic smoothing strategy. The original problem is altered by exogeneous variables that when tend to zero, make the altered problem tend to the original one. The solution is ultimately obtained by solving a sequence of differentiable unconstrained low dimension optimization sub-problems.

The proposed method does not guarantee that the solution is the global minimum. But due to the smooth effect of the method, it produces convexification that in turn produces a deep solution. The larger is the value of exogeneous variables introduced by the proposed method, the more smooth the problem becomes. By gradually reducing the exogeneous variables to zero, the altered problem tends to the original one. The final convergence of the algorithm is guaranteed, because by making the exogeneous variables enough close to zero, the altered problem can be as close the original one as it is desired. The final solution can be obtained with 8 significant digits or more.

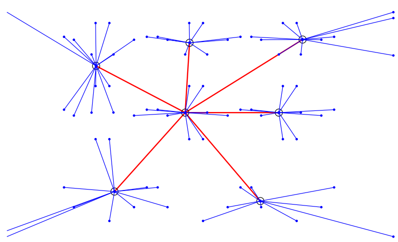
The proposed method is tested with some computational experiments. The results confirm the robustness and accuracy of the proposed method.



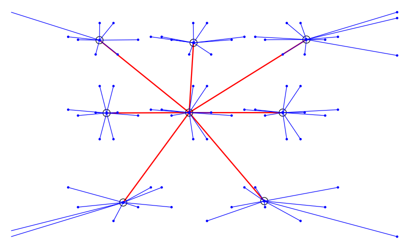
(a) 5 hubs



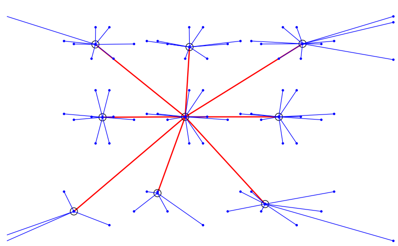
(b) 6 hubs



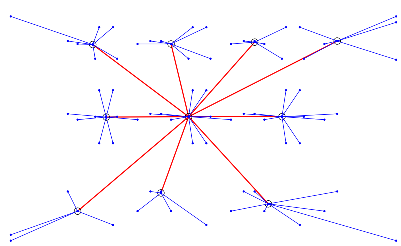
(c) 7 hubs



(d) 8 hubs

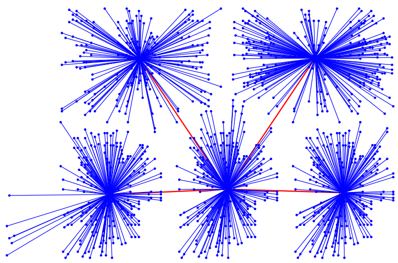


(e) 9 hubs

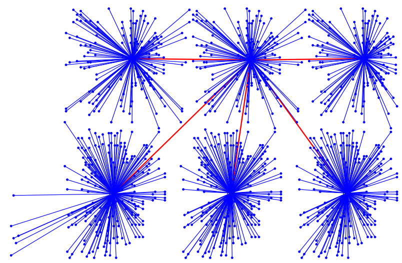


(f) 10 hubs

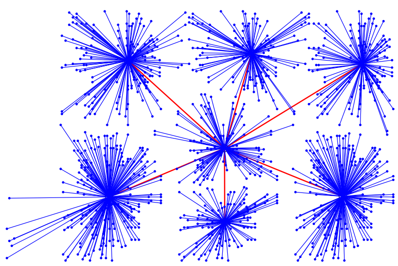
Figure 4: Results for 76 cities and 5 to 10 hubs



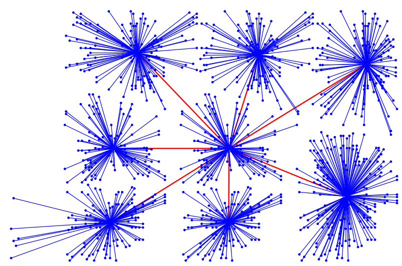
(a) 5 hubs



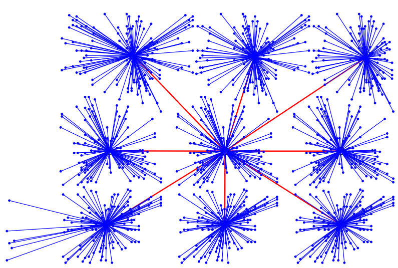
(b) 6 hubs



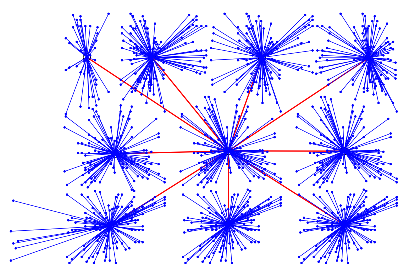
(c) 7 hubs



(d) 8 hubs



(e) 9 hubs



(f) 10 hubs

Figure 5: Results for 1002 cities and 5 to 10 hubs

Acknowledgments

The authors have research projects sponsored by CNPq, the Brazilian research council. They wish also to thank professors Nelson Maculan and Geraldo Xexeo of the Federal University of Rio de Janeiro for the helpful discussions and constructive comments. Gentil Veloso thanks CAPES - www.capes.gov.br for this doctorate scholarship.

References

[1] S. Alumur and B.Y. Kara. Network hub location problems: The state of the art. *European Journal of Operational Research*, 190:1–21, 2008.

[2] F. Baccelli and Rougier J.L. Self organizing hierarchical multicast trees and their optimization. *Proceedings of IEEE Inforcom*, 1999.

[3] A.M. Bagirov. Modified global k-means algorithm for minimum sum-of-squares clustering problems. *Pattern Recognition*, 41 (10):3192–3199, 2008.

[4] I. Contreras, E. Fernandez, and A. Marin. The tree of hubs location problem. *European Journal of Operational Research*, 202(2):390–400, April 2010.

[5] A. Demyanov. On the solution of min-sum-min problems. *Journal of Global Optimization*, 3 (31):437–453, 2005.

[6] Gonzalez T. F. On the computational complexity of clustering and related problems. *System Modeling and Optimization Lecture Notes in Control and Information Sciences*, 1982.

[7] L. Jia, A. Bagirov, I. Ouveysi, and A. M. Rubinov. Optimization based clustering algorithms in multicast group hierarchies. *Proceedings of the Australian Telecommunications, Networks and Applications Conference (ATNAC)*, 2003.

[8] D.S. Johnson, J.K. Lenstra, and A.H.G. Rinooy Kan. *The complexity of the network design problem.* 1978.

[9] D. Koutsonikolas, S. Das, H.Y. Charlie, and I. Stojmenovic. Hierarchical geographic multicast routing for wireless sensor networks. *Sensor Technologies and Applications, 2007. SensorComm 2007. International Conference on*, 2007.

[10] J. Kratica, Milanovic M., Stanimirovic Z., and D. Tosic. An evolutionary-based approach for solving a capacitated hub location problem. *Applied Soft Computing*, 2011.

[11] Harwell Library. http://www.cse.scitech.ac.uk/nag/hsl/ or http://www.hsl.rl.ac.uk/.

[12] G. Reinelt. Tsplib: a traveling salesman library. *ORSA Journal of Computing*, pages 376–384, 1991.

[13] I. Rodriguez-Martin and J.J Salazar-Gonzalez. Solving a capacitated hub location problem. *European Journal of Operational Research*, 2008.

[14] A. Rubinov. Methods for global optimization of nonsmooth functions with applications. *Applied and Computational Mathematics*, 1 (5):3–15, 2006.

[15] Z. Stanimirovic. Solving the capacitated single allocation hub location problem using genetic algorithm. *Wold Scientific Publishing*, 2007.

[16] G. Waters. Hierarchies for network evolution. *European Journal of Operational Research*, 2000.

[17] G. Waters. Optimising multicast structures for grid computing. *Computer Communications*, 2004.

[18] G. Waters and Lim S. G. Applying clustering algorithms to multicast group hierarchies. Computer Science Technical Report 4-03, University of Kent, Computing Laboratory, August 2003.

[19] A. E. Xavier. The hyperbolic smoothing clustering method. *Pattern Recognition*, 43 (3):731–737, 2010.

[20] A. E. Xavier and A. A. F. Oliveira. Optimal covering of plane domains by circles via hyperbolic smoothing. *Journal of Global Optimization*, 31 (3)(3):493–504, 2005.

[21] A. E. Xavier and V. L. Xavier. Solving the minimum sum-of-squares clustering problem by hyperbolic smoothing and partition into boundary and gravitational regions. *Pattern Recognition*, 44 (1):70–77, 2011.