ISSN 0717-9103
ISSN Online 0718-8307
Ingeniería Industrial - Año 9 N° 1: 87 - 94 , 2010

ALGORITMO BASADO EN DISCRIMINACION POR DISTANCIAS CON BUSQUEDA GLOBAL APLICADO AL PROBLEMA DE LA P-MEDIANA.

ALGORITHM BASED DISCRIMINATION BY DISTANCE TO GLOBAL SEARCH APPLIED TO THE P-MEDIUM PROBLEM*

ALEJANDRO BASTÍAS GAJARDO[1] , EDER PADILLA VILLALOBOS[1], RAFAEL ORTEGA BURGOS[1], CRISTIAN OLIVA SAN MARTÍN
1Departamento de Ingeniería Industrial, Facultad de Ingeniería, Universidad Católica de la Santísima Concepción. Campus San Andrés, Concepción. Chile

RESUMEN

Este artículo presenta una heurística para resolver problemas del tipo p-median. La idea principal es la ejecución de un preproceso, cuyo objetivo es la reducción del número de variables del problema facilitando computacionalmente su resolución.
Este pre-procesamiento se basa esencialmente en la determinación de una distancia máxima, que permite discriminar si dos nodos cualesquiera son atendidos por la misma instalación. Si la distancia entre estos dos nodos es mayor que la distancia máxima, entonces no serán atendidos por la misma localización. Esto conlleva a fijar variables a cero y, por lo tanto, a la reducción del tamaño del problema.

Palabras Clave: Heurística, p-media, diseño de redes.

ABSTRACT

This paper presents an heuristics for solving p-median problems. The main idea is the execution of a preprocess routine whose goal is to reduce the number of variables facilitating its computacional resolution.
This preprocessing is essentially based in the determination of a maximum distance that enables you to discriminate if two any nodes are served for the same facility. If the distance between these nodes is greater than the maximum distance, then these nodes will not be served for the same facility. As a consequence, variables need to be fixed to zero in order to reduce the size of the problem.

Keywords: Heuristics, p-median, networks design.

*Trabajo financiado parcialmente por el proyecto DIN/11-2008 UCSC. Programa de Magíster en Ingeniería Industrial UCSC y presentado en el VIII Congreso Chileno de Investigación Operativa, realizado por la Universidad del Bio-Bio en la ciudad de Chillán en Octubre del año 2009

Autor de correspondencia: coliva@ucsc.cl.

Recibido: 25.03.2009 Aceptado: 28.08.2009

INTRODUCCIÓN

El problema de p-median consiste en encontrar la mejor configuración de instalaciones para atender, de la mejor forma, la demanda de la población (Daskin, 1995). Dado un conjunto de coordenadas en el plano cartesiano de 1000 nodos, donde cada uno de estos nodos posee una cierta demanda que debe ser satisfecha, se deben instalar 10 servidores, de modo de satisfacer esta demanda al mínimo costo posible. El costo fue determinado de manera proporcional a la distancia entre cada par de nodos.

Dadas las características de este problema, la formulación que lo resuelve fue planteada por (*Hakimi, 1964*). El problema se denominó p-mediana. Este problema localiza p instalaciones, de modo de minimizar la distancia total entre las demandas y sus instalaciones más cercanas (Daskin, 1995). Para ello resuelve el problema de asignación y minimiza el costo dependiente de la distancia y de la demanda de los nodos.

Formulación de la P- mediana:

$$Min \sum_i \sum_j h_i d_{ij} y_{ij} \qquad (F.O.)$$

$$s/a \quad \sum_j y_{ij} = 1 \qquad \forall i \qquad (1)$$

$$\sum_j x_j = P \qquad (2)$$

$$y_{ij} \le x_j \qquad \forall i, j \qquad (3)$$

$$x_j, y_{ij} \in \{0,1\} \qquad \forall i, j \qquad (4)$$

Donde:

$$h_i = \text{Demanda del nodo i.}$$

$$d_{ij} = \text{Distancia minima de i a j.}$$

$$y_{ij} = \begin{cases} 1 & \text{Si la instalación j atiende a la demanda i.} \\ 0 & \text{Si no.} \end{cases}$$

$$x_j = \begin{cases} 1 & \text{Si se localiza una instalación en j.} \\ 0 & \text{Si no.} \end{cases}$$

La función objetivo representa el mínimo costo de satisfacer la demanda de todos los nodos de la red. La restricción (1) indica que para cada nodo de la red debe ser atendido por exactamente un servidor. La restricción (2) indica que se deben instalar exactamente *p* servidores (localizaciones). La restricción (3) es de contingencia; esto quiere decir que, cuando un servidor no se localice en el nodo *j*, ningún nodo puede ser atendido por dicho servidor. La restricción (4) nos indica el dominio de las variables, que en este caso son variables binarias.

MATERIAL Y MÉTODOS

El problema fue modelado en AMPL y resuelto utilizando CPLEX 11.0. Sin embargo, dado el gran número de variables de las instancias a resolver (en el orden de 10^6), éstas no pudieron ser resueltas por dos razones: a) falta de memoria y b) tiempos de CPU prohibitivos. Además, considerando que el problema de la p-mediana es NP_difícil, se hace necesaria la utilización de una heurística que permita obtener soluciones de buena calidad en tiempos de CPU razonables.

La idea principal de la heurística propuesta es reducir el número de variables (arcos). Esta reducción se basa en la determinación de una distancia máxima que permite discriminar si dos nodos cualesquiera son atendidos por la misma instalación. Si la distancia entre estos dos nodos es mayor que la distancia máxima (radio inicial de cobertura), entonces no serán atendidos por la misma localización. Esto conlleva a fijar variables a cero y, por lo tanto, a la reducción del tamaño del problema

Para determinar un radio de cobertura inicial, se analiza la red de nodos que se muestra en la figura 1. Para facilidad de explicación se asume que los nodos quedan distribuidos uniformemente en el plano en un rango de 100 x 100 unidades de distancia.

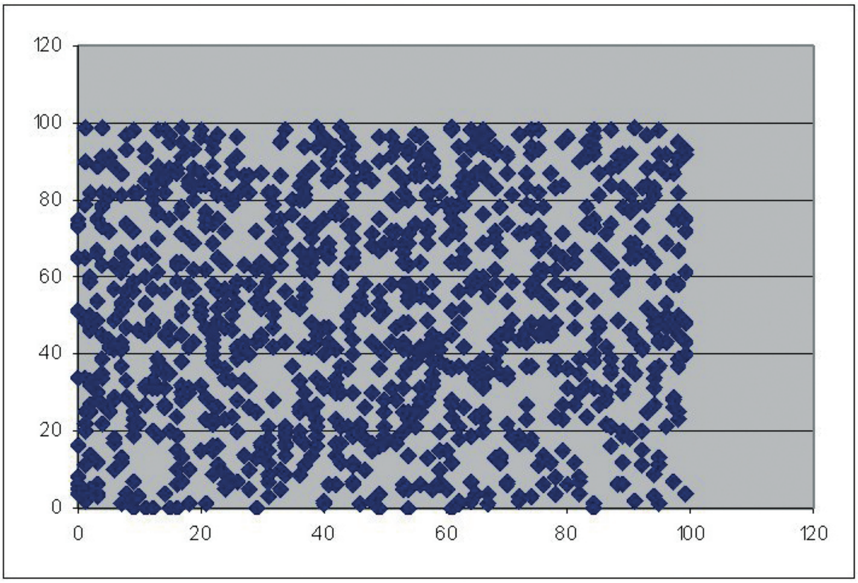


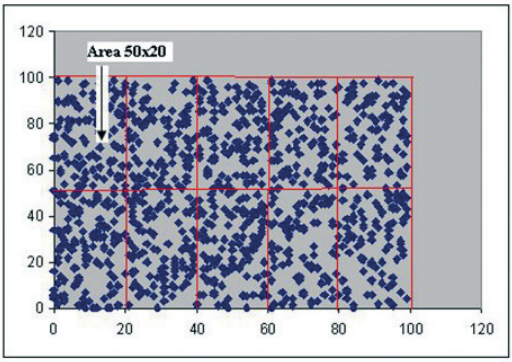
Figura 1. Distribución espacial de los nodos del problema[2].

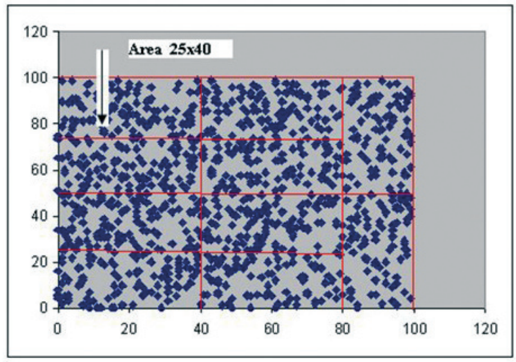
En la figura 1 se observa que hay que cubrir un área total de 10.000 unidades de distancia (GOCUP 2008). Dado que debemos instalar 10 servidores que atiendan la demanda de todos los nodos de la red, se divide la superficie en 10 regiones. En efecto, se tiene:

Área Total	= 10.000 (unidades de distancia)^2
N° Localizaciones	= 10 (servidores en la red)
Área a cubrir por cada servidor	= (Area Total / N° Localizaciones)
	= 1.000(unidades de distancia)^2/ (serv)

A continuación, se deben generar regiones con al menos el área a cubrir por cada servidor (lo más aproximado posible). A partir de cada una de ellas se obtiene la distancia entre los vértices más lejanos, de dicha figura. Esta distancia recibe el nombre de diámetro de cobertura. La mitad de éste equivale al radio inicial de cobertura (distancia máxima). Un problema que queda por resolver es cómo generar las regiones de 1000 unidades de distancia (GOCUP 2008).

[2]Instancia propuesta en el concurso GOCUP 2008



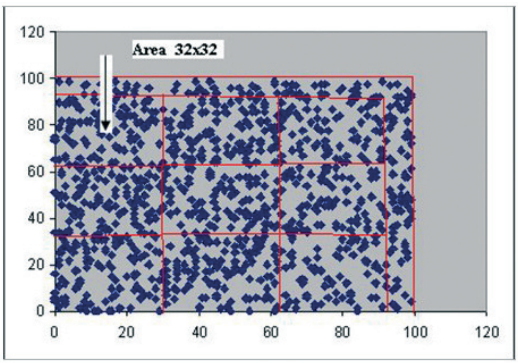


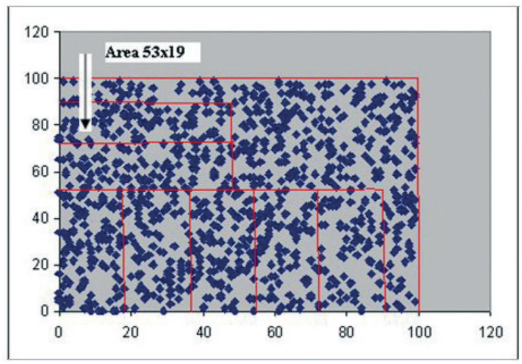
a

Área de cada rectángulo = 1.000 (u de área)
Lados del rectángulo = (50x20)
Diámetro = 53.85≈54
Radio de partida = (54/2)=27

b

Área de cada rectángulo = 1.000 (u de área)
Lados del rectángulo = (25x40)
Diámetro = 47.16≈48
Radio de partida = (484/2)=24





c

Área de cada rectángulo = 1.000 (u de área)
Lados del rectángulo = (32x32)
Diámetro = 45.25≈46
Radio de partida = (46/2)=23

d

Área de cada rectángulo = 1.000 (u de área)
Lados del rectángulo = (53x19)
Diámetro = 56.30≈57
Radio de partida = (57/2)=29

Figura 2: Divisiones del plano que se utilizaron para encontrar el radio de cobertura inicial.

En el caso de la figura 2a, cada región tiene dimensiones de 50x20, lo que hace que se generen 10 regiones del mismo tamaño. En este caso el radio inicial de cobertura es de 27 unidades de distancia[3]. La figura 2c genera 8 regiones de tamaño 32x32 y 2 regiones 7x93. El criterio para establecer el radio inicial de cobertura es: seleccionar el tamaño que presente el mayor número de regiones iguales; en este caso es el tamaño de 32x32. Por lo anterior, el radio inicial de cobertura es 23[4]. A continuación se muestra un resumen para cada una de las figuras anteriormente mencionadas.

Después de determinar el radio inicial de cobertura (distancia máxima), se obtienen dos radios auxiliares incrementando en 1 y disminuyendo en 1 el radio inicial de cobertura. A continuación se fijan las variables y_{ij} de la siguiente manera: si la distancia entre un nodo i y un nodo j es

[3] Radio de cobertura= $\left\lceil \sqrt{20^2+50^2} \big/ 2 \right\rceil = 27$

[4] Radio de cobertura= $\left\lceil \sqrt{32^2+32^2} \big/ 2 \right\rceil = 23$

mayor que el radio auxiliar, entonces la variable y_{ij} correspondiente se fija al valor cero. Con la información obtenida se resuelve (para cada radio auxiliar) de manera exacta el problema de p-median. Luego se comparan las funciones objetivos de estas dos soluciones, con el fin de determinar la dirección de búsqueda de una mejor solución. En el caso que la mejor solución se obtiene para el radio auxiliar aumentado en 1, entonces la dirección de búsqueda es ir aumentando en una unidad el radio auxiliar hasta que no se mejore la solución. En el caso contrario, si la mejor solución se obtiene para el radio auxiliar disminuido en 1, entonces la dirección de búsqueda es ir disminuyendo en uno el radio auxiliar hasta que no se mejore la solución. Este procedimiento constituye el motor de resolución del problema.

Pseudocódigo

A continuación se presenta el Pseudocódigo de la heurística propuesta.

```
// Inicialización //
S1 ← ∞ (cota superior inicial para el problema)
Radio_0 ← Radio Inicial de cobertura
k ← 0 contador de modificación del radio de cobertura

// Radio hacia abajo //
ArcosR ← Conjunto de arcos cuya matriz de distancia es menor o igual a "Radio_0 – 1"
Resolver P-Media (utilizando el conjunto de ArcosR )
OBJ_menos ← Solución de este P-media

// Radio hacia arriba //
ArcosR ← Conjunto de arcos cuya matriz de distancia es menor o igual a "Radio_0 + 1"
Resolver P-Media (utilizando el conjunto de ArcosR)
OBJ_mas ← Solución de este P-media

// Dirección hacia un solución factible //
If OBJ_menos > OBJ_mas Then
        S1 ← OBJ_mas
        k ← k+2
        Radio ← Radio_0 + 1
        Repeat until Radio_0 = 0
                ArcosR ← Conjunto de arcos cuya matriz de distancia es menor o igual a
                          "Radio_0 + k"
                Resolver P-Media (utilizando el conjunto de ArcosR)
                S2 ← Solución de este P-media
                If S2 < S1 Then
                        Localbest ← Almacenar Localizaciones
                        Radio ← Almacenar "Radio_0 + k"
                        S1 ← S2
                Else
                        // Mostrar la mejor Solución, Radio e Instalaciones //
                        Radio_0 ← 0
                End If
                k ← k+1
        End Repeat
Else
        S1 ← OBJ_menos
        k ← k+2
        Radio ← Radio_0 - 1
        Repeat until Radio_0 = 0
                ArcosR ← Conjunto de arcos cuya matriz de distancia es menor o igual a
                          "Radio_0 - k"
                Resolver P-Media (utilizando el conjunto de ArcosR)
                S2 ← Solución de este P-media
                If S2 < S1 Then
                        Localbest ← Almacenar Localizaciones
                        Radio ← Almacenar "Radio_0 - k"
                        S1 ← S2
                Else
                        // Mostrar la mejor Solución, Radio e Instalaciones //
                        Radio_0 ← 0
                End If
                k ← k+1
        End Repeat
End If
```

RESULTADOS

La Heurística para la P-median se programó en lenguaje AMPL y se resolvió de manera exacta cada subproblema con CPLEX 11.0, en un CPU Intel Core 2Duo 2.64 GHz y 1.99 RAM.

Las instancias de estudio se obtuvieron del concurso GoCup 2008. Son 4 instancias, cada una de 1000 nodos. Para cada nodo se especifican las coordenadas en el plano cartesiano y su respectiva demanda.

La figura 3 muestra las localizaciones obtenidas con la heurística que se menciona en este trabajo para cada instancia resuelta. Los puntos azules corresponden a las localizaciones.

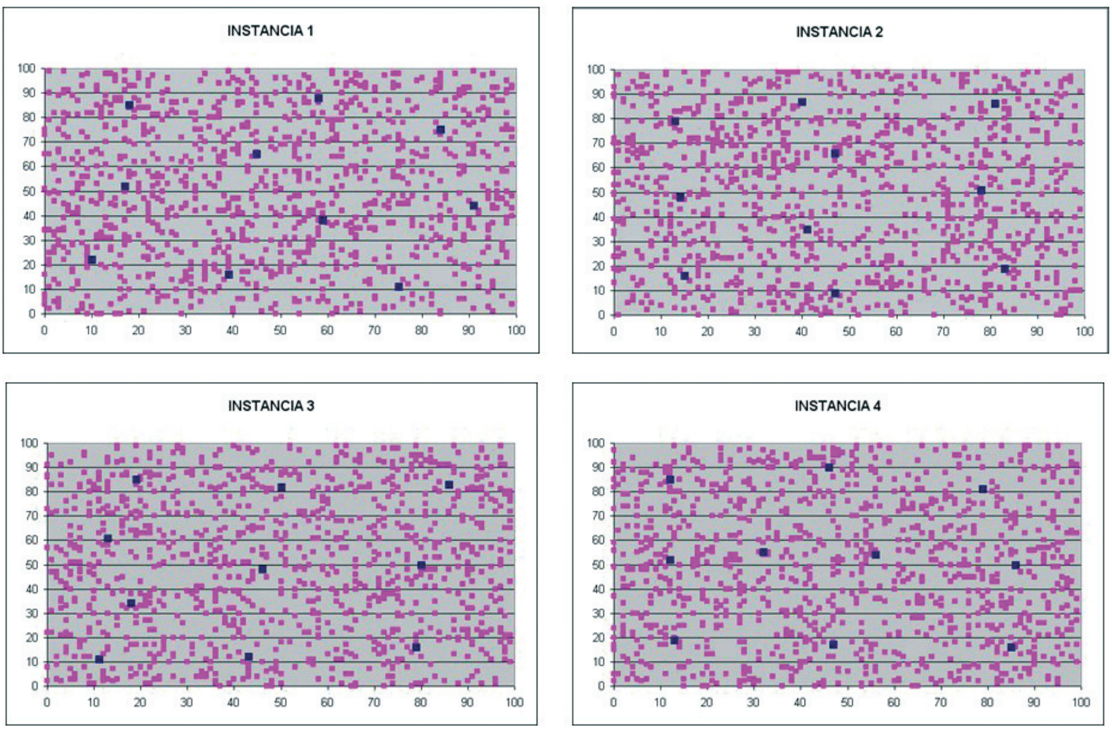


Figura 3. Resultados para las cuatro instancias estudiadas.

La tabla 1 ilustra la resolución del problema aplicando la heurística mencionada en el apartado anterior. Aquí se considera que las variables y_{ij} son binarias. La primera columna indica el radio inicial de cobertura obtenida para las 4 figuras. La segunda columna indica el radio óptimo obtenido por cada figura. La tercera columna muestra la función objetivo y, finalmente, la cuarta columna indica el tiempo CPU en segundos.

Se puede observar para la instancia 1 en la tabla 1, que el menor tiempo se obtuvo con un radio inicial de cobertura igual a 27, obtenido al crear figuras de 50x20. Para las otras instancias lo mejor fue comenzar con un radio de cobertura igual a 23. Estas se obtienen al formar figuras geométricas cuadradas (32x32).

La tabla 2 ilustra la resolución del problema aplicando la heurística mencionada en el apartado anterior. Aquí se considera que las variables y_{ij} son continuas en el intervalo [0,1]. La primera columna indica el radio inicial de cobertura obtenida para las cuatro instancias de la figura 3. La segunda columna indica el radio óptimo obtenido por cada figura. La tercera columna muestra la función objetivo y, finalmente, la cuarta columna indica el tiempo CPU en segundos.

Se puede observar, para la instancia 2 en la tabla 2, que el menor tiempo se obtuvo con un radio inicial de cobertura igual a 23, obtenido al crear figuras de 32x32. Para las otras instancias lo mejor fue comenzar con un radio de cobertura igual a 24. Estas se obtienen al formar figuras geométricas cuadradas (25x40).

Se observa que para el caso relajado, los tiempos de ejecución son mucho menores a los encontrados para el caso no relajado.

Tabla 1. Problema no relajado.

Instancia 1			
Radio inicial	Radio opt	OBJ	Tiempo (seg)
23	25	6203859	660,98
24	25	6203859	546,89
27	26	6203859	389,94
29	28	6203859	460,77

Instancia 2			
Radio inicial	Radio opt	OBJ	Tiempo (seg)
23	24	6137625	334,61
24	25	6137625	389,48
27	26	6137625	385,81
29	28	6137625	440,27

Instancia 3			
Radio inicial	Radio opt	OBJ	Tiempo (seg)
23	24	6273000	1.125,55
24	23	6273000	1.130,75
27	26	6273000	1.484,36
29	28	6273000	1.756,50

Instancia 4			
Radio inicial	Radio opt	OBJ	Tiempo (seg)
23	24	6182962	986,13
24	25	6182962	1.022,20
27	26	6182962	1.148,78
29	28	6182962	1.493,92

Tabla 2. Problema relajado.

Instancia 1			
Radio inicial	Radio opt	OBJ	Tiempo (seg)
23	25	6203859	439,52
24	25	6203859	342,38
27	26	6203859	380,50
29	28	6203859	435,63

Instancia 2			
Radio inicial	Radio opt	OBJ	Tiempo (seg)
23	24	6137625	268,59
24	25	6137625	293,48
27	26	6137625	353,25
29	28	6137625	398,50

Instancia 3			
Radio inicial	Radio opt	OBJ	Tiempo (seg)
23	24	6273000	895,09
24	23	6273000	811,05
27	28	6273000	1.451,81
29	28	6273000	1.469,06

Instancia 4			
Radio inicial	Radio opt	OBJ	Tiempo (seg)
23	24	6182962	927,42
24	25	6182962	867,83
27	26	6182962	988,47
29	28	6182962	1.178,64

CONCLUSIÓN

En este trabajo se ha desarrollado una heurística fácil de implementar y que permitió que se obtuvieran los mismos resultados con el ganador del GoCup 2008 [2], a pesar de que el tiempo computacional es mayor. Podemos agregar también que, relajando un conjunto de variables, los tiempos de ejecución disminuyen hasta en un 28%.

REFERENCIAS

Daskin, Mark S. 1995. Network and Discrete Location, Models, Algorithms, and Applications. John Wiley and Sons, Inc.

Maturana, D. 2009. Pontificia Universidad Católica de Chile Ganador del GoCup 2008. Presentación Congreso de Localización y Diseño de Redes. 22-25 de marzo 2009. Pucón-Chile.

Hakimi, S.L. 1964. Optimum locations of switching centres and the absolute centers and medians of a graph, Operations Research 12, 450-459,.