

COMPARACIÓN DEL DESEMPEÑO DEL ALGORITMO GENÉTICO DE CHU-BEASLEY Y EL ALGORITMO COLONIA DE HORMIGAS EN EL PROBLEMA DE P-MEDIANA.

Performance Comparison of Genetic Algorithm of Chu-Beasley and Ant Colony Algorithm for P-median Problem.

RESUMEN

Los problemas de localización son de gran importancia en la vida diaria, siendo el problema de P-mediana uno de ellos y sus aplicaciones constituyen una importante línea de investigación. En este trabajo se presenta el problema general de la P-mediana así como el modelo matemático que lo representa y la codificación empleada en su solución. También se describen las técnicas de optimización combinatorial implementadas (Algoritmo Genético de Chu-Beasley y Algoritmo Colonia de Hormigas). El desempeño de los métodos es estudiado a través de una comparación de varios casos de prueba de baja, media y alta complejidad matemática.

PALABRAS CLAVES: Algoritmo Colonia de Hormigas, Algoritmo Genético de Chu-Beasley, metaheurísticas, optimización combinatorial, P-mediana.

ABSTRACT

Location problems are of great importance in daily life and their applications are an important line of research. The problem with P-median is a location problem. In this paper the general problem of the P-median is presented, as well as its mathematical model and the encoding used for solving. It also describes the implemented combinatorial optimization techniques (Genetic Algorithm of Chu-Beasley and Ant Colony Algorithm) to solve this problem. The performance of the methods is studied through a comparison of several test cases of low, medium and high mathematical complexity.

KEYWORDS: *Ant Colony Algorithm, Genetic Algorithm of Chu-Beasley, combinatorial optimization, metaheuristics, P-median.*

CÉSAR ADRIÁN MUÑOZ B.

Ingeniero Electricista.
Profesor Catedrático
Programa de Ingeniería Eléctrica
Universidad Tecnológica de Pereira
ceadmuno@utp.edu.co

RAMÓN ALFONSO GALLEGO.

Ingeniero Electricista, Ph.D.
Profesor Titular
Programa de Ingeniería Eléctrica
Universidad Tecnológica de Pereira
ragr@utp.edu.co

ELIANA M. TORO O.

Ingeniera Industrial, M.Sc.
Profesor Asociado
Facultad de Ingeniería Industrial
Universidad Tecnológica de Pereira
elianam@utp.edu.co

1. INTRODUCCIÓN

En la presentación clásica y de forma general el problema de localización consiste en ubicar uno o varios centros de servicios para suplir las necesidades de una serie de puntos de consumo establecidos, usualmente denominados puntos de demanda, optimizando alguna medida de efectividad como por ejemplo la distancia. El problema de P-mediana consiste en ubicar un conjunto de facilidades llamadas medianas para suplir las necesidades de un grupo de usuarios de forma que la distancia media de los usuarios a las medianas sea la menor posible.

Dentro de las aplicaciones que se enmarcan en este tipo de problema están: la localización de entidades de servicio como bancos, escuelas, centros de recaudo, la localización de antenas repetidoras de telecomunicaciones, de centrales de alimentos, de oficinas de clasificación de correspondencia, y en redes eléctricas la localización de subestaciones nuevas en las

redes de subtransmisión y transformadores en la red de nivel de tensión II.

La formulación del problema en su versión clásica, atribuida a Hakimi, se realizó a mediados de los años 60's como una extensión a los problemas simples de localización de instalaciones. Algunos de los estudios más recientes muestran en 2006 un análisis computacional para problemas de P-mediana de gran tamaño [1]; en 2008 un estudio sobre algoritmos evolutivos aplicados a este mismo problema [2]. En 2009 y 2010 fueron usados Algoritmos Genéticos y Algoritmos Colonia de Hormigas para resolver este problema [3] y [4].

Para el problema de la P-mediana no se conoce un algoritmo de orden polinomial que lo resuelva, por lo que es considerado un problema NP-Duro. Por tal razón, en la literatura especializada, este problema ha sido enfrentado usualmente a través de técnicas heurísticas y metaheurísticas aprovechando la dimensionalidad y la

explosión combinatorial que se presenta al incluir una mediana adicional en el problema. Dentro de las técnicas metaheurísticas usadas están: Algoritmos Genéticos, Cúmulos de Partículas, Recocido Simulado, Colonia de Hormigas entre otros.

La dificultad de las técnicas de exploración exhaustiva radica en que para problemas de gran escala, el árbol de decisión es demasiado amplio y su evaluación completa requiere de demasiado esfuerzo computacional. En la literatura se han presentado alternativas para la relajación del problema, como se describe en [1].

En este trabajo fue resuelto el problema de P-mediana de forma general, usando los métodos de optimización combinatorial Algoritmo Genético de Chu-Beasley y Algoritmo Colonia de Hormigas, con el objetivo de comparar el desempeño de ambos. Esta comparación fue hecha con base en la calidad de la respuesta y tiempo de cómputo. Para esto fueron usados sistemas de prueba de baja, media y alta complejidad matemática.

2. DESCRIPCIÓN DEL PROBLEMA DE LOCALIZACIÓN A TRAVÉS DE LA P-MEDIANA

De forma general el problema de localización a través de la P-mediana consiste en asignar un grupo de clientes en exactamente P grupos. La cantidad de grupos es determinada a partir de un parámetro P donde cada uno representa un centro de servicio. Tanto los clientes como los centros de servicio están espacialmente distribuidos. Cada grupo está caracterizado por la ubicación de su mediana o centro de servicio, el grupo de clientes que lo conforman y por el costo asociado de acoger dicho grupo particular de clientes. En esta categoría de problemas no se especifican restricciones de capacidad para los centros de servicio, por lo que cada cliente es asignado a la mediana más cercana. Se puede utilizar una representación espacial discreta o de red. Se asume que los costos fijos para el emplazamiento de las instalaciones son idénticos y por lo tanto no son tomados en cuenta en la formulación del problema. Finalmente, el objetivo consiste en determinar los posibles centros de servicio que se requieren ubicar, habiendo sido asignado P, de manera que se minimice la suma total de las distancias entre cada cliente y su centro de servicio.

3. FORMULACIÓN MATEMÁTICA

La formulación más clásica del problema de la P-mediana es la siguiente:

Sea:

$N=\{1,\dots,n\}$ el conjunto de índices para los clientes.

$J=\{1,\dots,m\}$ el conjunto de índices para las localizaciones potenciales de las medianas.

Para cada (i, j) , $i \in N, j \in J$ sea C_{ij} el costo de asignación del cliente i a la mediana ubicada en la localización j .

Se definen las siguientes variables de decisión:

- Y_j . Tomará el valor de 1 si se ubica la mediana en la localización $j \in J$ y de 0 en otro caso.
- X_{ij} . Tomará el valor de 1 si el cliente $i \in N$ se asigna mediana ubicada en la localización $j \in J$ y de 0 en otro caso.

El problema de la P-mediana puede formularse de la siguiente manera.

$$\min \sum_{i \in N} \sum_{j \in J} C_{ij} X_{ij} \quad (1)$$

Sujeto a:

$$\sum_{j \in J} X_{ij} = 1, \quad \forall i \in N \quad (2)$$

$$\sum_{j \in J} Y_j = P \quad (3)$$

$$X_{ij} \leq Y_j, \quad \forall i \in N, j \in J \quad (4)$$

$$X_{ij} \in \{0,1\} \quad \forall i \in N, \quad Y_j \in \{0,1\} \quad \forall j \in J \quad (5)$$

Las restricciones (2) aseguran que cada cliente es asignado a una única mediana.

La restricción (3) garantiza que se seleccionen exactamente P localizaciones para las medianas.

Las restricciones (4) aseguran que los clientes se asignen a una mediana sólo si ésta ha sido seleccionada.

Finalmente, el conjunto de restricciones (5) especifica que todas las variables de decisión son binarias.

Esta formulación representa un total de n^2 variables y $n^2 + 1$ restricciones.

4. METODOLOGÍA DE SOLUCIÓN

4.1 Codificación

La codificación es un aspecto de gran importancia en la solución del problema. En esta se pueden incluir implícitamente algunas restricciones y facilita el manejo explícito de las demás. En la codificación de este problema el número máximo de medianas es controlado. Si el número de medianas no es controlado, la solución que minimiza la suma total de las distancias entre cada cliente y su centro de servicio es aquella con mayor cantidad de medianas.

La codificación adoptada es la siguiente:

P1	P2	P3	...	Pj
1	0	0	...	1

Figura 1. Codificación empleada

Donde P_j representa el centro de servicio o mediana, y su valor representa si la mediana asociada se ubica o no.

Esta codificación permite controlar el número de medianas a ubicar.

4.2 Cálculo de la Función Objetivo

Tal como se ve en la formulación matemática, la función objetivo consiste en calcular los costos asociados entre un grupo de clientes y una mediana, para finalmente calcular el costo total de todas las medianas y sus usuarios.

Diversos autores han planteado estrategias para el cálculo de la suma total de las distancias entre cada cliente y su centro de servicio asociado a partir de funciones de distancia usando norma Euclidiana y norma Manhattan. En este trabajo se uso la Euclidiana.

La estrategia planteada aquí, similar a la mostrada en [5], se basa en que la suma que menor costo genera es aquella en la que los clientes asociados a una mediana particular tienen el menor costo respecto a dicha mediana. Así resulta bastante intuitivo darse cuenta que si se asigna un cliente a otra mediana con la cual se tiene mayor costo este se incrementa.

De esta forma se tiene entonces que para cada mediana se asigna aquel usuario que tenga mayor cercanía o genere menor costo.

En la Figura 2 se da un ejemplo gráfico del problema, donde los puntos azules representan las medianas y los puntos rojos representan los usuarios. Cada uno de los usuarios debe ser asignado a una mediana y de acuerdo con lo expuesto, la forma de hacerlo y la cual genera menor costo es asociar los clientes a la mediana más cercana. Si a alguno de los puntos ya asociados se cambia de mediana el costo total aumenta, debido a que las distancias a cualquier otra mediana son mayores.

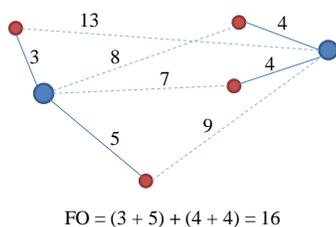


Figura 2. Cálculo de función objetivo.

5. ALGORITMOS IMPLEMENTADOS

Como ya se mencionó, este problema se considera como NP-Duro, por lo cual no es posible resolverlo usando numeración exhaustiva a través de una técnica exacta. Dada la característica combinatorial de este problema y la dificultad descrita, se propone resolverlo a través de técnicas metaheurísticas como el Algoritmo Genético Modificado de Chu-Beasley (AGCB) y el Algoritmo Colonia de Hormigas (ACH).

Tanto la estrategia del AGCB como del ACH para dar solución al problema de las P-medianas se basa en proponer alternativas de solución, y a través del proceso evolutivo ir encontrando soluciones de mejor calidad hasta encontrar la que optimice el problema tratado.

Para ambas metodologías, la generación de la población inicial es de carácter netamente aleatoria dado que no se cuenta con factores de sensibilidad que guíen este proceso y generen individuos de buena calidad. Para la generación de esta población inicial se controla la factibilidad respecto a la restricción del número máximo de medianas propuestas.

5.1 Algoritmo Genético de Chu-Beasley

El AGCB es una versión modificada del algoritmo genético básico el cual es basado en la genética presente en la naturaleza, y que además ha mostrado que su proceso evolutivo es altamente eficiente. La característica fundamental del AGCB consiste en mantener durante todo el proceso, la diversidad de los individuos que conforman la población. En cada generación, solo un individuo es reemplazado, siempre y cuando, se cumplan con las condiciones de optimalidad y factibilidad establecidas.

El proceso evolutivo del AGCB se desarrolla de la forma clásica a través de los procesos de selección, recombinación y mutación.

El proceso de *selección* se efectúa a través de dos torneos, en los cuales se determinan los padres que podrán pasar sus genes a la siguiente generación.

Se realiza *recombinación* de un punto a partir de la cual los cromosomas de los padres se comparten para generar dos nuevos individuos candidatos. En este proceso debe ser controlada la factibilidad respecto al máximo número de medianas candidatas. Al final del proceso sólo el nuevo individuo con mejor función objetivo sigue el proceso.

Para la *mutación* se emplea una tasa que es ajustable dentro de los parámetros generales del algoritmo. Si se cumple el criterio para realizar mutación esta se ejecuta alterando alguno de los alelos de forma aleatoria. En esta etapa se controla nuevamente la factibilidad respecto a la restricción del número máximo de medianas.

Finalmente se tienen dos *criterios de aceptación*. Si el nuevo individuo es mejor que el individuo de peor calidad este se reemplaza directamente, sino se emplea una tasa de aceptación en la que se le da una oportunidad de pasar a la siguiente generación. En esta etapa además se evalúa la diversidad, chequeando que no haya individuos con mucha similitud en la población.

5.2 Algoritmo Colonia de Hormigas

El ACH es basado en el comportamiento que muestran las hormigas en su entorno natural y en la capacidad que poseen para encontrar el alimento a través de la exploración de sus individuos y guiados, mediante la comunicación indirecta de estos, por los rastros de feromonas.

El proceso evolutivo del algoritmo se desarrolla de la forma clásica, donde la colonia es guiada por la feromona en el proceso de búsqueda. Esta feromona es generada a partir de la información suministrada por el individuo de mejor calidad durante la colonización, simulando un comportamiento de hormigas elitistas.

A partir de la población de hormigas (conjunto de alternativas de solución) se selecciona aquella de mejor calidad (mejor función objetivo) y mediante el proceso iterativo del algoritmo la feromona va siendo actualizada a partir de la información que suministra dicha hormiga. Para el proceso de *depósito de la feromona* se emplea una tasa la cual indica la cantidad de feromona que deja el mejor individuo en cada iteración. Para evitar el estancamiento del proceso en soluciones de baja calidad, se ejecuta el proceso de *evaporación de la feromona* de forma global, empleando una tasa de evaporación la cual indica la velocidad con la que se va perdiendo el rastro de feromona. Finalmente, para la obtención de la nueva población de hormigas, se acude a un proceso probabilístico y a la información contenida en la feromona ya actualizada. Los nuevos individuos caminan en la dirección dada por la feromona.

Al igual que en el caso ya descrito del AGCB, la factibilidad debe ser controlada respecto a la restricción del número máximo de medianas.

6. ANÁLISIS DE RESULTADOS

Para validar las metodologías implementadas y comparar su desempeño se evalúan varios casos de prueba en los cuales la información de las posibles medianas y los puntos de consumo son generados de forma aleatoria a partir de una distribución de probabilidad uniforme. Esta información describe la ubicación espacial de las P-medias candidatas y los centros de consumo.

Dentro de la descripción del problema se explicó que el objetivo es establecer la mejor ubicación de P centros de servicio para suplir una demanda o necesidad de tal forma que se minimice la suma total de las distancias entre cada cliente y su centro de servicio asociado.

De acuerdo con la formulación matemática y el planteamiento del problema, se presentan tres casos de análisis diferentes:

- a. Definiendo cuales son los posibles centros de servicio: de esta forma se tiene un número de posibilidades limitado. J (el número de medianas candidatas) es predeterminado.
- b. Cuando no son definidos cuáles son los posibles nodos de servicio, por lo cual las posibilidades son del mismo tamaño que la cantidad de clientes. J es igual a N.
- c. Teniendo fijos algunos nodos de servicio y tomando la decisión respecto a otras posibilidades las cuales se pueden plantear de las dos formas descritas anteriormente.

En cada análisis se plantearon 4 casos de prueba para los cuales el objetivo es encontrar la mejor ubicación de las P-medias a ser emplazadas.

Para cada uno de los casos de prueba se tienen diferentes cantidades de nodos de consumo y medianas candidatas, y de estas se asume que 3 serán finalmente las que podrán ser emplazadas. El número de medianas o centros de servicio a emplazar P es definido previamente y basado en algún criterio, por ejemplo económico.

6.1 Análisis asumiendo algunas medianas fijas y otras variables.

En este análisis se tienen preestablecidas algunas medianas existentes que son fijadas y deben ser tenidas en cuenta en el proceso de optimización, y las demás medianas son seleccionadas dentro de las posibles candidatas.

En este caso de análisis se establecieron las medianas fijas, mostradas en color rojo en la tabla 1, y el número de medianas candidatas para cada caso de prueba y se determinó que la cantidad de P medianas a ubicar sería 3. En la Tabla 1 se muestran los cuatro casos de prueba, donde las medianas resaltadas en color rojo corresponden a las medianas fijas. Además se presentan las soluciones encontradas, las cuales fueron las mismas con los dos métodos de solución.

Caso	N clientes	J candidatas	Solución	FO
BD1	70	10	5 1 6	430.790
BD2	200	20	10 9 18	1565.500
BD3	400	40	15 8 27	2969.400
BD4	1000	100	100 27 36	25294.000

Tabla 1. Casos de prueba y soluciones encontradas.

Estas soluciones encontradas son las que menor costo total generan, siendo la función objetivo la suma total de las distancias entre cada cliente y su mediana.

En la Tabla 2 se pueden ver los resultados de los tiempos computacionales de cada metodología para cada caso de prueba.

El tamaño del problema es resultado del espacio combinatorial que resulta de hacer combinaciones de J candidatas con P medianas a ubicar (JcP J combinado P).

Caso	Tamaño JcP	AGCB Tiempo [s]	ACH Tiempo [s]
BD1	36	1.221	0.558
BD2	171	2.454	1.433
BD3	741	22.148	12.702
BD4	4851	124.190	130.564

Tabla 2. Tiempos computacionales.

En la Figura 3 se ilustra el comportamiento de cada una de los métodos respecto al tamaño del problema.

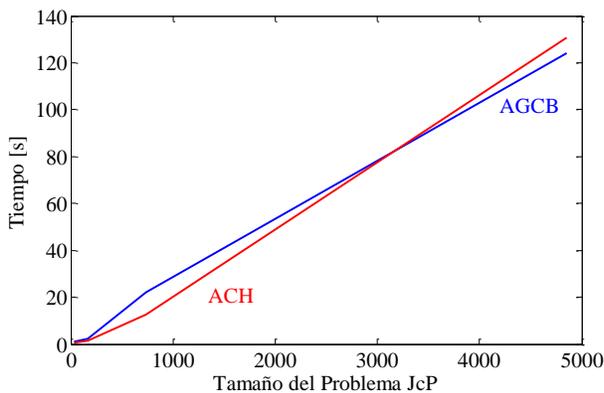


Figura 3. Comparación del desempeño computacional.

De la Figura 3 se puede notar que para este análisis de baja complejidad matemática, ambas metodologías presentan un desempeño similar.

6.2 Análisis donde J es preestablecido

Para este análisis se tiene que el número de medianas candidatas es preestablecido.

Los casos de prueba y las soluciones asociadas son mostrados en la tabla 3. En esta se muestra la mejor ubicación para un número de 3 medianas y el costo de la función objetivo.

Caso	N clientes	J candidatas	Solución	FO
BD1	70	10	3 4 7	420.70
BD2	200	20	4 9 14	1469.00
BD3	400	40	3 16 30	2748.80
BD4	1000	100	8 70 93	23541.00

Tabla 3. Casos de prueba y soluciones encontradas.

Las soluciones mostradas fueron obtenidas a través de los dos métodos, y ambos alcanzan la misma respuesta, sin embargo el tiempo de cómputo es diferente. En la tabla 4 se muestran los tiempos computacionales para cada caso de prueba.

Caso	Tamaño JcP	AGCB Tiempo [s]	ACH Tiempo [s]
BD1	120	1.733	1.156
BD2	1140	3.658	10.397
BD3	9880	22.109	120.963
BD4	161700	180.764	1158.158

Tabla 4. Tiempos computacionales.

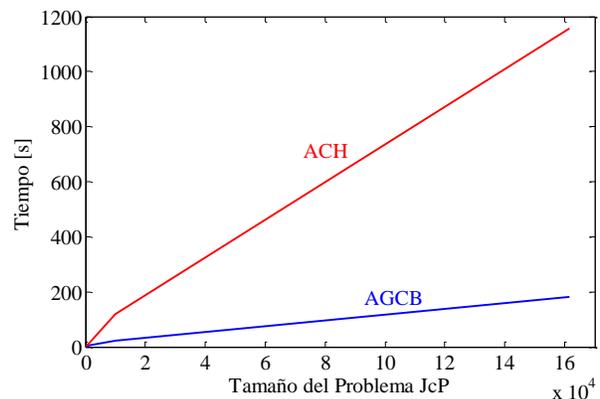


Figura 4. Comparación del desempeño computacional.

De los resultados mostrados en la Figura 4 se puede notar la fuerte dependencia del rendimiento de los algoritmos con el tamaño del problema, ya que el tiempo de cómputo aumenta con el número de medianas candidatas.

También se puede resaltar la diferencia que presentan los métodos en cuanto al desempeño ya que el AGCB muestra una tasa de crecimiento más baja que el ACH.

6.3 Análisis con J igual a N

Como ya se mencionó, en este análisis, se especifica el número de P medianas a ser ubicadas, sin embargo no se conoce ni el número, ni la localización de las medianas candidatas, por lo que se asume que la información para estas es la misma que la de los clientes.

Para este análisis se tienen los mismos 4 casos de prueba, la diferencia es que la cantidad de medianas candidatas es igual a la cantidad de clientes, haciendo de este un problema de gran complejidad matemática.

En la tabla 5 se muestran los casos de prueba y los resultados obtenidos para 3 medianas a emplazar.

Caso	N clientes	J candidatas	Solución	FO
BD1	70	70	13 37 57	395.608
BD2	200	200	10 134 183	1355.00
BD3	400	400	87 108 309	2695.40
BD4	1000	1000	250 604 795	23306.00

Tabla 5. Casos de prueba y soluciones encontradas.

Para este análisis las soluciones mostradas también fueron alcanzadas por ambos métodos, sin embargo el desempeño es diferente, como se observa en la tabla 6.

Caso	Tamaño JcP	AGCB Tiempo [s]	ACH Tiempo [s]
BD1	54740	12.986	227.133
BD2	1313400	52.709	2635.088
BD3	10586800	270.577	8652.496
BD4	166167000	758.877	36522.490

Tabla 6. Tiempos computacionales.

En la Figura 5 se ilustra el desempeño computacional de ambos métodos.

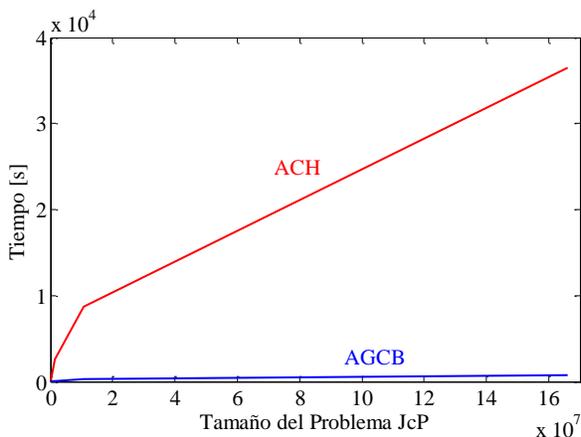


Figura 5. Comparación del desempeño computacional.

De acuerdo con lo mostrado en la Figura 5, el desempeño del AGCB es mejor que el del ACH.

7. CONCLUSIONES

Fue resuelto el problema de las P-medianas para los tres casos planteados: el primero donde se tienen preestablecidas algunas de las medianas, el segundo donde el número de medianas candidatas es definido previamente y el tercero en el que las posibles medianas a ubicar sea igual o menor que el número de clientes, usando algoritmos de optimización combinatorial AGCB y el ACH, obteniendo resultados que permiten establecer el desempeño de cada uno de los métodos.

Fue posible determinar de forma clara que el AGCB presenta mejor desempeño que el ACH, ya que sin importar el tamaño o la complejidad matemática del

problema, el tiempo computacional es menor, para la misma calidad de las respuestas.

La codificación usada para la solución del problema es eficiente, ya que simplifica su solución al incorporar de manera implícita algunas de las restricciones y facilitar el cálculo de la función objetivo.

El modelo estudiado es aplicable en muchas de las situaciones de la vida real como es el de ubicación de centros de salud, de centros de atención inmediata de la policía (CAI), localización de centros de consumo, escuelas, aulas, droguerías entre otros.

8. BIBLIOGRAFÍA

- [1] Avella, P., Sassano, A. e Vasil'ev, I., "Computational study of large-scale p-Median problems". Math. Program., pp. 89-114. 2006.
- [2] Sinaide N. Bezerra, Sérgio R. de Souza y João Francisco de A., "Um Estudo Sobre Algoritmos Evolutivos Aplicados Ao Problema Das P-Mediana". XIV Latin Ibero-American Congress on Operations Research (CLAIO). 2008.
- [3] K. Ghoseiri, S. F. Ghannadpour, "Solving Capacitated P-Median Problem using Genetic Algorithm". Industrial Engineering and Engineering Management, IEEE International Conference on, pp. 885-889. 2007.
- [4] Chen Liang, "An Ant Colony Algorithm for Text Clustering". Computing, Control and Industrial Engineering (CCIE), International Conference on, pp. 249-252. 2010.
- [5] R. Chen, G. Y. Handler, "The Conditional p-Center Problem in the Plane". Naval Research Logistic, Vol. 40, pp. 117-127. 1993.
- [6] Ramón A. Gallego, Antonio Escobar, Eliana M. Toro. "Técnicas Metaheurísticas de Optimización". Texto Universitario Universidad Tecnológica de Pereira. Segunda Edición. 2008.