



Universidad de las Ciencias Informáticas

Facultad 2

**Trabajo de diploma para optar por el título de
Ingeniero en Ciencias Informáticas.**

Título: Algoritmo MT_FDA aplicado al problema RND.

Autor: Isabel Maria Higuera Igarza

Tutor: Ing. Biasmey Morgado Guirola

Ciudad de la Habana, Junio 2009

*El futuro tiene muchos nombres. Para los débiles es lo inalcanzable.
Para los temerosos lo desconocido. Para los valientes es la oportunidad.*

Victor Hugo

Agradecimientos

A mi mamá y mi papá, por ser mi guía, mi luz, por enseñarme y educarme siempre por el mejor camino y por ayudarme a construir este sueño...al fin lo logramos.

A mi hermana y mi abuelita, que tanto me han ayudado y apoyado, las quiero muchísimo.

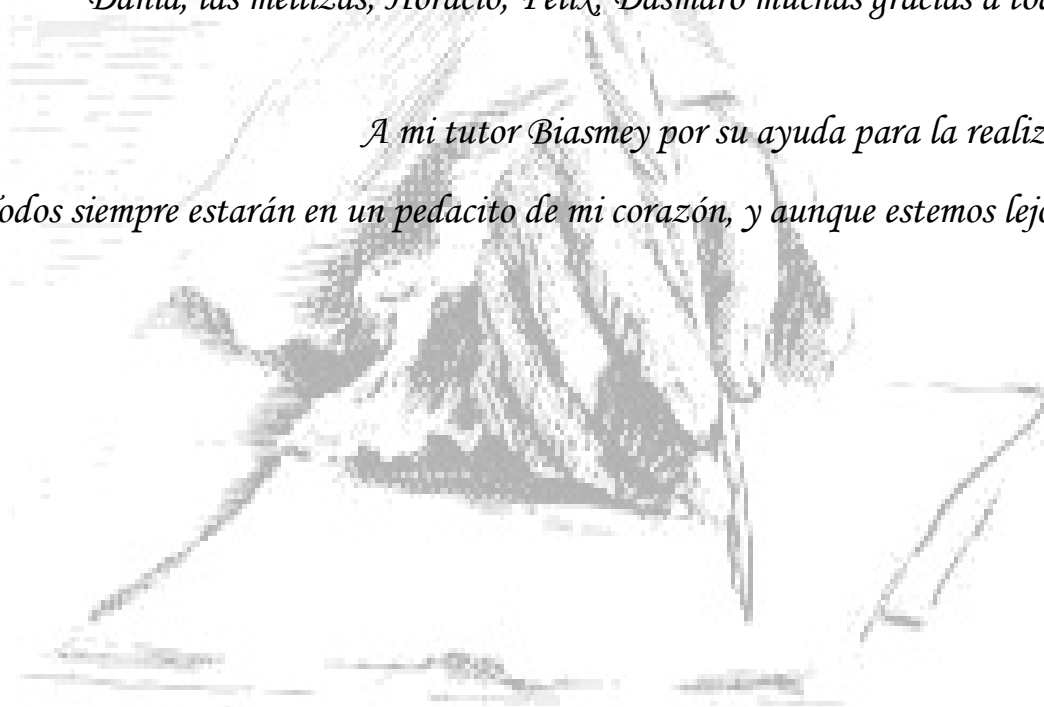
A Darian, por todo el amor, la comprensión y ayuda que me ha brindado todos estos años, gracias por formar parte de mi vida.

A mis amigos que fueron mi familia en estos 5 años, Suammy, Susana, Daymi, Yuneimy, Liset, Saskia, Yurima, Lisandra, Dennis, Alain, Moisés, Ronald, Dannier y a mis compañeros de grupo, a todos muchas gracias por ayudarme siempre que los necesité.

A Mileydis, mi abuelo Nené, Yosbani, Margarita, Ovidio, Digna, Yobania, Ideliza, Bertha, Dania, las mellizas, Horacio, Félix, Dasmáro muchas gracias a todos por apoyarme y ayudarme siempre.

A mi tutor Biasmey por su ayuda para la realización de esta tesis.

Todos siempre estarán en un pedacito de mi corazón, y aunque estemos lejos los recordaré con mucho amor.



Dedicatoria

A mi mamá y mi papá por toda su ayuda y amor.

Tabla Contenido

Resumen	9
Introducción	10
Resumen de cada capítulo:	12
Capítulo 1. Fundamentación Teórica.....	14
1. Introducción.....	14
2. Algoritmos Evolutivos.....	14
3. Algoritmo Evolutivo con Estimación de Distribuciones (EDAs).....	17
4. Algoritmo Evolutivo con Factorización de Distribuciones (FDAs).	22
5. Algoritmo Evolutivo con Factorización de distribuciones Basado en Mezcla de árboles.....	24
6. Mezclas de distribución.....	25
6.1. Mezcla de árboles.....	25
7. Problema a resolver	26
8. Algoritmos utilizados en otras investigaciones para resolver el problema RND.	27
9. Tecnologías de desarrollo.....	29
9.1. Entorno de Desarrollo Integrado.....	29
9.2. Sistema Operativo:	29
9.3. Lenguaje de Programación.....	30
9.4. Conclusiones	30
Capítulo 2. Descripción de la función Objetivo y el algoritmo Mt_FDA.	31
1. Introducción.....	31
2. Diseño de redes de Radio (Radio Network Design (RND)).	31

Tabla de Contenido

3. Descripción de la función objetivo.....	35
4. Algoritmo MT_FDA.....	38
5. Mezcla de árboles	39
5.1. Algoritmo Iterativo para la maximización de la verosimilitud en mezclas de árboles.	42
5.2. Muestreo de la mezcla de árboles.....	42
5.3. Pseudocódigo del algoritmo	42
5.4. Análisis de la complejidad del MT_FDA.	44
Capítulo 3. Resultados de la Investigación.....	46
1. Introducción.....	46
2. Experimentos	46
3. Métricas.....	47
3.1. Cantidad de evaluaciones.	47
3.2. Tiempo de ejecución.	47
3.3. Generación de parada.	48
4. Resultados	48
5. Conclusiones	52
Conclusiones.....	53
Recomendaciones.....	54
Anexos.....	55
Referencias Bibliográficas	67
Bibliografía	69

Tabla de Figuras

Figura 1. Algoritmo Evolutivo	17
Figura 2. Diagrama Ejemplificando la operativa de un EDAs	21
Figura 3. Coberturas ofrecidas por distintas antenas	33
Figura 4. Cobertura discretizada en el modelo de rejillas	34
Figura 5. Mezcla de árboles con 3 componentes Figura	41
Figura 6. Tiempo de Ejecución	48
Figura 7. Generación de Parada	49
Figura 8. Número de evaluaciones	50
Figura 9. Resultado de los algoritmos en cuanto a número de evaluación	51
Pseudocódigo 1. Algoritmo Estimación Distribuciones	19
Pseudocódigo 2. Función Objetivo	36
Pseudocódigo 3. Algoritmo MT_FDA	43
Tabla 1. Resultados NV: 149; NA: 1	55
Tabla 2. Resultados NV: 149; NA: 2	56
Tabla 3. Resultados NV: 149; NA: 4	57
Tabla 4. Resultados NV: 149; NA: 6	58
Tabla 5. Resultados NV: 199; NA: 1	59
Tabla 6. Resultados NV: 199; NA: 2	60
Tabla 7. Resultados NV: 199; NA: 4	61
Tabla 8. Resultados NV: 199; NA: 6	62
Tabla 9. Resultados: NV: 249; NA: 1	63

Tabla de Figuras

Tabla 10. Resultados NV: 249; NA: 2	64
Tabla 11. Resultados NV: 249; NA: 4	65
Tabla 12. Resultados NV: 249; NA: 6	66

RESUMEN

Esta investigación aborda el proceso de modelación probabilística en una clase de algoritmo de optimización basado en poblaciones utilizando la mezcla de árboles para modelar dicha distribución probabilística de la población, denominándose Algoritmo Evolutivo con Factorización de distribuciones basado en mezcla de árboles (MT_FDA).

El objetivo principal de la tesis es tratar de demostrar la eficiencia del MT_FDA en el diseño de redes de radio (RND), el cual es un problema importante en las telecomunicaciones móviles así como en el aumento del área de redes de sensores¹. El mismo consiste en darle cobertura de radio a una zona geográfica dada de manera que se cubra la mayor superficie del área con el menor número de antenas. Hasta la fecha varios modelos de antenas de radio han sido utilizados, entre ellos: plaza de cobertura de antenas que cubren un área cuadrada, las antenas omnidireccionales que cubren un área circular, etc. En el siguiente trabajo se utilizan las antenas cuadradas.

Para poder demostrar la eficiencia del algoritmo se compara con otros modelos utilizados en otras investigaciones para resolver este mismo problema.

¹ Sensor: Es un aparato capaz de transformar magnitudes físicas o químicas, llamadas variables de instrumentación, en magnitudes eléctricas. Las variables de instrumentación dependen del tipo de sensor y pueden ser ej: temperatura, intensidad lumínica, distancia, aceleración, inclinación, desplazamiento, presión, fuerza, torsión, humedad.

INTRODUCCIÓN

La Inteligencia Artificial aborda problemas poco estructurados, donde no se conoce de antemano cuál es el mejor método para resolverlo. Hay que descubrir, si acaso, alguna solución. De ahí que se utiliza la palabra heurística cuyo significado se asocia a búsqueda, pero una búsqueda diferente a la de los algoritmos, ya que es un camino para buscar lo nuevo, lo que no conocemos, y los algoritmos son un camino para realizar lo que ya conocemos bien. (1)

En los últimos años una nueva familia de algoritmos evolutivos conocida como Algoritmos con Estimación de Distribuciones (EDAs) ha sido objeto de gran atención por parte de la comunidad científica alrededor de la computación evolutiva y los modelos probabilísticos. Estos algoritmos han surgido como una alternativa a los métodos heurísticos estocásticos² donde es necesario ajustar un número elevado de parámetros, motivados además por la necesidad de identificar las interrelaciones entre las variables. (2)

Los EDAs al igual que los algoritmos genéticos (AG) están basados en poblaciones³ que evolucionan, con la diferencia que sustituyen los operadores de cruce y mutación por operadores que estiman o “aprenden” la distribución de probabilidad de los puntos seleccionados en la generación⁴ anterior. A diferencia de los algoritmos genéticos, donde las interrelaciones entre las variables se tienen en cuenta de manera implícita⁵, en los EDAs dichas interrelaciones se expresan de manera explícita⁶ a través de la distribución de probabilidad⁷ asociada a los puntos seleccionados en cada generación.

Calcular todos los parámetros necesarios para especificar el modelo probabilístico⁸ que será “aprendido” no es practicable. La solución a este problema ha sido la aproximación de la distribución de probabilidad por medio de distintas factorizaciones más o menos complejas de la misma, que ha traído consigo una amplia variedad de algoritmos en los dominios discretos y continuos.

² Estocástico: sistema que funciona, sobre todo, por el azar.

³ Población: es un conjunto de puntos que representan los individuos candidatos.

⁴ Generación: cada una de las poblaciones que evolucionan sucesivamente una tras otra.

⁵ Relación entre las variables implícita: cuando la relación no está clara, no se sabe que variable depende de cual.

⁶ Relación de las variables explícitas: cuando se tiene claro la dependencia entre las variables.

⁷ Distribución de probabilidad: Describe el rango de valores de una variable aleatoria así como la probabilidad de que el valor de la variable aleatoria esté dentro de un subconjunto de dicho rango.

⁸ Modelo Probabilístico: Es la forma que pueden tomar un conjunto de datos obtenidos de muestreos de datos con comportamiento que se supone aleatorio.

El algoritmo evolutivo con factorización de distribuciones basado en mezcla de árboles (MT_FDA) es un ejemplo de combinación de modelos probabilísticos en algoritmos de estimación de distribuciones, este algoritmo emplea mezcla de árboles para modelar la distribución probabilística de la población seleccionada. (3)

La clase de problemas NP completos son aquellos para los cuales todavía no se ha encontrado un algoritmo que los resuelva en un tiempo polinomial, siendo hoy en día un gran problema de optimización y consumo de recursos. Debido a esta **Situación Problémica**, se plantea como **Problema Científico** ¿Cómo demostrar la eficiencia del Algoritmo Evolutivo con Factorización de Distribuciones Basado en Mezcla de Árboles aplicado a un problema NP completo?

Para desarrollar la investigación y darle respuesta a dicho problema científico, se escogió un problema NP completo al cual se le aplicará el algoritmo MT_FDA, dicho problema es el conocido como Radio Network Design Problem (RND) o problema de Diseño de Redes de Radio, que consiste en situar un conjunto de antenas repetidoras en un área geográfica dada, de manera que se cubra la mayor superficie posible con el menor número de antenas.

Se parte de la **hipótesis** de que el algoritmo evolutivo con factorización de distribuciones basado en mezcla de árboles debe mejorar los resultados del problema en cuestión.

Se define como **Objeto de Estudio** los algoritmos evolutivos con estimación de distribuciones, y como **Campo de Acción** el algoritmo evolutivo con factorizaciones de distribuciones basado en mezcla de árboles.

Se plantea como **Objetivo General** de esta investigación determinar la eficiencia en el problema Diseño de Redes de Radio (RND) aplicando el algoritmo MT_FDA.

Para dar cumplimientos a este objetivo se han identificados los siguientes **Objetivos Específicos**:

1. Hacer un estudio del arte de los Algoritmos Evolutivos con Estimación de Distribuciones, Algoritmos Evolutivos de factorización de distribuciones y del Algoritmo Evolutivo con factorización de distribuciones basado en mezcla de árboles, así como de algunos algoritmos utilizados para resolver el problema RND.

2. Implementar la función objetiva para resolver el problema del Diseño de redes de Radio (RND).
3. Mostrar y comparar los resultados de la investigación con resultados obtenidos con otros algoritmos.

Se definen como **Tareas de la Investigación** para dar cumplimiento a los objetivos específicos:

- 1.1 Consultar diversas bibliografías especializadas en la temática en cuestión.
- 1.2 Buscar información de otros algoritmos utilizados para resolver este problema.
- 1.3 Arribar a conclusiones en cuanto a la temática en estudio.
- 1.4 Profundizar en el estudio del algoritmo evolutivo con factorización de distribuciones basado en mezcla de árboles (MT_FDA).
- 2.1 Describir en profundidad el problema RND.
- 2.2 Definir la función objetiva para resolver el problema.
- 2.3 Implementar mediante un lenguaje de programación la función objetiva definida para el problema abordado.
- 2.4 Hacer una descripción del algoritmo MT_FDA.
- 3.1 Aplicar el algoritmo MT_FDA al problema de la Diseño de redes de Radio.
- 3.2 Comparar los resultados obtenidos con los resultados arrojados por otros algoritmos utilizados en este mismo problema.
- 3.3 Llegar a conclusiones.

Resumen de cada capítulo:

Capítulo 1: Fundamentación Teórica

Se realizará una fundamentación teórica de los principales conceptos y algoritmos abordados en el trabajo, así como de las herramientas utilizadas para el desarrollo de la investigación.

Capítulo 2: Descripción de la Función Objetivo y del algoritmo MT-FDA

Se explicará detalladamente el problema al que se le aplicará el algoritmo. Se describe la implementación de la función objetiva del algoritmo evolutivo con factorización de distribuciones aplicada al problema de Diseño de redes de Radio.

Capítulo 3: Resultados de la investigación

Se arribarán a conclusiones teniendo en cuenta los resultados obtenidos. Y la comparación que se haga con otros algoritmos aplicado al mismo problema.

CAPÍTULO 1. FUNDAMENTACIÓN TEÓRICA

1. Introducción

Este capítulo hace referencia a conceptos y algoritmos que son de interés para el desarrollo de esta investigación, se tratarán algunos conceptos importante relacionados con los algoritmos evolutivos (AE), algoritmos evolutivos con estimación de distribuciones(EDAs), algoritmos evolutivos con factorización de distribuciones(FDAs), y el algoritmo evolutivo con factorización de distribuciones basado en mezcla de árboles(MT_FDA), además de hacer un estudio del problema al que se le aplicará el algoritmo y algunos algoritmos utilizados anteriormente para resolver el problema.

2. Algoritmos Evolutivos

Una rama de la inteligencia artificial son los algoritmos evolutivos (AE), estos son métodos de optimización y búsqueda estocásticos que se inspiran en la genética y en los principios establecidos por Darwin en 1859 sobre la evolución de las especies en el mundo biológico. Estos algoritmos se modelan computacionalmente simulando la selección natural y el entrecruzamiento de las especies por medio de la recombinación genética y la mutación. En ellos se mantienen un conjunto de entidades que representan posibles soluciones, las cuales se mezclan y compiten entre sí, de manera que las más aptas son capaces de prevalecer a lo largo del tiempo, evolucionando hacia mejores soluciones cada vez. (4)

En la naturaleza lo único que hay que optimizar⁹ es la supervivencia, y eso significa a su vez maximizar diversos factores y minimizar otros. Un algoritmo evolutivo, sin embargo, se usará habitualmente para optimizar sólo una función, no diversas funciones relacionadas entre sí simultáneamente. (5)

Se han aplicado con éxito en problemas de optimización, búsqueda, aprendizaje y simulación de sistemas dinámicos, en problemas con espacios de búsqueda extensos y no lineales donde otros métodos no son capaces de encontrar soluciones en un tiempo razonable.

Esta gran familia de algoritmos está compuesta por distintos métodos de búsqueda entre los que podemos mencionar, la Programación Genética, que consiste en la evolución automática de programas usando

⁹ Optimizar: Mejorar el rendimiento de algo.

Capítulo 1. Fundamentación Teórica

ideas basadas en la selección natural, se busca que poblaciones de programas evolucionen transmitiendo su herencia de manera que se adapten mejor al medio, los individuos tienen mayores probabilidades de reproducirse, la medida de la calidad del individuo dependerá del tipo de problema. (6)

La Programación Evolutiva que se basa en la optimización de funciones y es una abstracción de la evolución al nivel de las especies, por lo que no se requiere el uso de un operador de recombinación (diferentes especies no se pueden cruzar entre sí). Asimismo, usa selección probabilística. (7)

También se encuentra las estrategias evolutivas, son métodos computacionales que trabajan con una población de individuos que pertenecen al dominio de los números reales, que mediante los procesos de mutación y de recombinación evolucionan para alcanzar el óptimo de la función objetivo. (4)

Los Algoritmos Genéticos usados fundamentalmente en la optimización de problemas combinatorios, en ellos cada cromosoma es una estructura de datos que representa una de las posibles soluciones del espacio de búsqueda del problema. Los Cromosomas son sometidos a un proceso de evolución que envuelve evaluación, selección, crossover¹⁰ y mutación. Después de varios ciclos de evolución la población deberá contener individuos más aptos. (4)

También forman parte de los algoritmos evolutivos los algoritmos de estimación de distribuciones y otros métodos.

Un algoritmo evolutivo consiste en: hallar de qué parámetros depende el problema, codificarlos en un cromosoma¹¹, y aplicar los métodos de la evolución: selección y reproducción sexual con intercambio de información y alteraciones que generan diversidad. (5)

Un algoritmo evolutivo sigue generalmente los siguientes pasos:

1. Generar una POBLACIÓN aleatoria de **N** individuos
2. Evaluar los individuos de la POBLACIÓN de acuerdo a la función objetivo¹²
3. Repetir durante GENERACIONES iteraciones:
 - 3.1. Aplicar el operador de selección para elegir **k** individuos de la POBLACIÓN

¹⁰ Crossover, recombinación sexual.

¹¹ Cromosomas: Representan posibles soluciones al problemas.

¹² Función objetivo, se trata de una función evaluadora de la calidad de los individuos.

Capítulo 1. Fundamentación Teórica

- 3.2. Aplicar los operadores genéticos a esos k individuos para generar la descendencia
- 3.3. Evaluar los nuevos individuos de acuerdo a la función objetivo
- 3.4. Reemplazar los peores individuos en POBLACIÓN por los individuos recién creados.

Capítulo 1. Fundamentación Teórica

En la figura se muestra la representación de un algoritmo evolutivo.

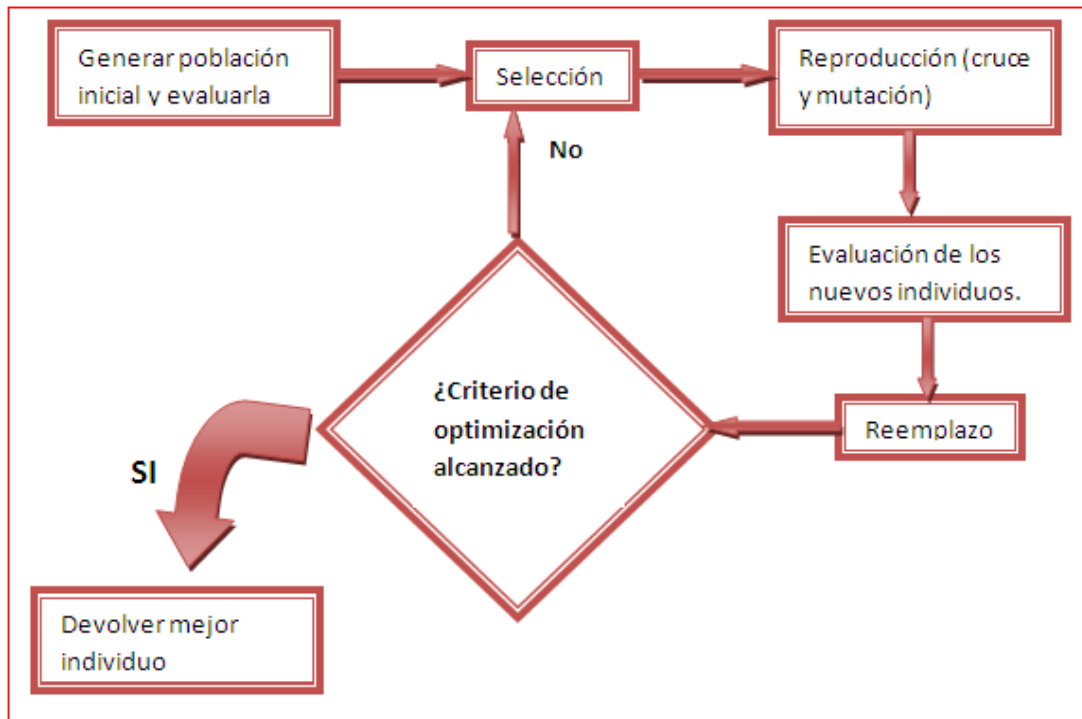


Figura 1. Algoritmo Evolutivo

3. Algoritmo Evolutivo con Estimación de Distribuciones (EDAs).

El comportamiento de los algoritmos genéticos depende en gran parte de los operadores de cruce¹³ y mutación¹⁴, las probabilidades de cruce y mutación, el tamaño de la población¹⁵, la tasa de reproducción generacional, el número de generaciones, etc. De hecho la determinación de valores adecuados para dichos parámetros constituye por sí mismo un verdadero problema de optimización. Por otra parte una mala elección de los valores de los parámetros puede llevar a que el algoritmo obtenga soluciones alejadas del óptimo. Este es uno de los motivos por los que desde hace varios años se han venido estudiando alternativas a los métodos heurísticos estocásticos existentes que no necesiten el ajustar un número alto de parámetros. (3)

¹³ Cruzamiento: Mezcla de la información de dos o más individuos.

¹⁴ Mutación: operador genético usado para mantener la diversidad genética de una población

¹⁵ Población: Conjunto de individuos que representan soluciones al problema.

Capítulo 1. Fundamentación Teórica

Otro motivo básico por el que se ha desarrollado la búsqueda de nuevos métodos heurísticos de optimización es por la necesidad de identificar las interrelaciones entre las variables utilizadas para representar a los individuos con la codificación utilizada. (3)

Todo ello ha motivado al nacimiento de una nueva familia de algoritmos evolutivos llamados algoritmos evolutivos con estimación de distribuciones (EDAs). Estos son algoritmos basados en poblaciones al igual que los algoritmos genéticos en los que la transición de una generación a otra ha sido modificada, y basan su búsqueda en el carácter estocástico de la misma (3), sin embargo a diferencia de los AGs¹⁶, en los EDAs no existen los operadores de cruce y mutación. En su lugar la población de individuos se regenera mediante una distribución de probabilidad, que se estima a partir de una base de datos formada por individuos de generaciones anteriores, es decir se sustituyen los operadores de cruce y mutación, por operadores que estiman o “aprenden” la distribución de probabilidad de los puntos seleccionados en la generación anterior. Se utilizan para resolver problemas de optimización en los cuales las variables de decisión son reales ($x_i \in \mathbb{R}$), en una generación la población se construye realizando un muestreo probabilístico de la población anterior y escogiendo los mejores individuos. (8)

Mientras que en los algoritmos genéticos las interrelaciones entre las variables representando a los individuos se tienen en cuenta de manera implícita, en los EDAs dichas interrelaciones se expresan de manera explícita a través de la distribución de probabilidad asociada con los individuos seleccionados en cada generación.

A continuación se muestra los pasos que siguen los algoritmos evolutivos de estimación de distribuciones:

1. Primeramente, se genera la población inicial formada por **N** puntos, que frecuentemente se realiza asumiendo una distribución uniforme de cada variable. Tras generar los puntos, estos se evalúan mediante la aplicación de la función de adaptabilidad o de costo (función objetivo).
2. Segundo, en cada generación se seleccionan **M** puntos (**M** \leq **N**) siguiendo un criterio de selección.
3. Tercero, se induce el modelo probabilístico n-dimensional que mejor representa las interdependencias entre las **n** variables. Este paso es conocido como el de aprendizaje, y es el más crucial en los EDAs debido a la importancia de tener en cuenta todas las dependencias entre las variables para asegurar mejoras sucesivas de los puntos.

¹⁶ AG: Algoritmo Genético.

Capítulo 1. Fundamentación Teórica

4. Finalmente, la nueva población se conforma con **N** nuevos puntos obtenidos tras simular o muestrear la distribución de probabilidad aprendida en el paso previo. Es frecuente en este paso utilizar una aproximación elitista, que mantiene el mejor o los mejores puntos de la población **t-1** en la nueva población. En este caso, se crean en cada generación un total de **N - E** puntos, donde **E** representa el número de puntos éliticos.

Los pasos 2, 3 y 4 se repiten mientras no se cumple una condición de parada en concreto. Ejemplos de criterios de parada son: llegar a un número máximo de generaciones, uniformidad de la población generada, no mejora con respecto al mejor de los puntos obtenido en las generaciones previas, etc. A continuación se muestra el pseudocódigo de un EDAs.

```
Poner  $t \leftarrow 1$ ;  
Generar  $N \gg 0$  puntos aleatoriamente;  
While no se cumpla el criterio de parada do  
Seleccionar  $M \leq N$  puntos de acuerdo a un método de selección;  
Estimar la distribución  $ps(x, t)$  del conjunto seleccionado;  
Generar  $N$  nuevos puntos de acuerdo a la distribución  $ps(x, t)$ ;  
Poner  $t \leftarrow t + 1$   
end While
```

Pseudocódigo 1. Algoritmo Estimación Distribuciones

Estructura general de un algoritmo evolutivo con estimación de distribuciones. (EDA)

Capítulo 1. Fundamentación Teórica

1. D_0 <- Generar la Población inicial (m individuos).
2. Evaluar la población inicial D_0 ; $k=1$.
3. Repetir hasta la condición de parada.
 - 3.1. $D_{\{k-1\}}$ <- seleccionar $n < m$ individuos de $D_{\{k-1\}}$.
 - 3.2. Estimar un nuevo modelo M a partir de $D_{\{k-1\}}$.
 - 3.3. $D_{\{k-1\}_m}$ <- Muestrear m individuos a partir de M .
 - 3.4. Evaluar $D_{\{k-1\}}$.
 - 3.5. D_k <- seleccionar m individuos de $D_{\{k-1\}}$ unión $D_{\{k-1\}_m}$.

Como se puede observar al igual que en la mayoría de los algoritmos generacionales, se parte de una población inicial con m individuos, generada (en la mayoría de los casos) aleatoriamente. En el segundo paso se selecciona un número n ($n \leq m$) de individuos, se selecciona (normalmente aquellos con los mejores valores en la función de evaluación) como base de datos para la estimación del modelo¹⁷. A continuación se introduce el modelo probabilístico n -dimensional que mejor refleja las interdependencias entre las n variables. A partir del modelo inducido se genera una población auxiliar mediante muestreo. Por último la nueva población D_k se obtiene a partir de la población anterior $D_{\{k-1\}}$ y de la población auxiliar. Normalmente esta selección se realiza de forma elitista. (7) A continuación se muestra como proceden los EDAs para arrojar las nuevas generaciones. (8)

¹⁷ Modelo: nos referimos como modelo a las relaciones existentes entre las variables.

Capítulo 1. Fundamentación Teórica

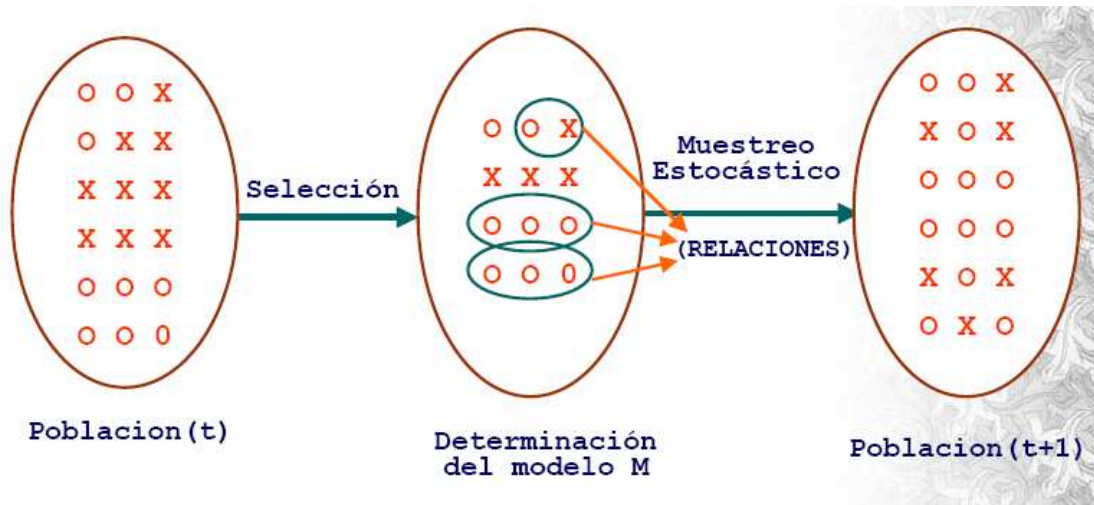


Figura 2. Diagrama Ejemplificando la operativa de un EDAs

Como se puede observar, partimos de una población inicial de individuos a la cual aplicamos un proceso de selección, en este paso decantamos aquellos individuos menos prometedores a mejorar nuestra función objetivo, posteriormente utilizando los individuos seleccionados determinamos el modelo que caracteriza las relaciones existentes entre las variables, este modelo permite luego aplicando un muestreo estocástico seleccionar la nueva generación.

El principal problema que se presenta es la estimación del modelo **M**, ya que cuanto más complejo sea el modelo mejor recogerá las dependencias entre las variables, pero más compleja/costosa será su estimación, puesto que en cada generación un nuevo modelo puede ser construido.

Existen tres formas fundamentales de realizar el modelado:

- **Análisis univariado:** este es el modelo más simple, no considera relaciones entre las variables y las probabilidades conjuntas se factorizan como producto de distribuciones (marginales) univariadas e independientes.
- **Análisis bivariados:** permite la dependencia entre las variables, sin embargo la dependencia se restringe a dos variables, ej: x_j depende de x_i y solo de x_i . En este caso es necesario realizar dos tareas: hallar el grafo de dependencia y calcular las probabilidades condicionales $P(x_j|x_i)$. Se puede decir que existen heurísticas específicas para determinar el grafo de dependencia.

Capítulo 1. Fundamentación Teórica

- Análisis multivariado: es el enfoque más completo, pero también el más complejo de implementar, los modelos estadísticos utilizados son de orden n , típicamente se modelan mediante redes bayesianas. Para obtener la población en $t+1$ se realiza el muestreo de \mathbf{M} , habitualmente mediante la utilización de modelos lógicos probabilísticos.

Las diferencias consisten en el tipo de relaciones que son consideradas para la construcción del modelo. (8)

4. Algoritmo Evolutivo con Factorización de Distribuciones (FDAs).

Para resolver problemas más complejos, como pueden ser aquellos con un número alto de variables, dentro del campo de los EDAs se encuentra una subclase especial que agrupa los algoritmos que usan factorizaciones de la distribución probabilística y a esta se le llama Algoritmos Evolutivos con Factorización de distribuciones (FDAs). (3)

Se tienen dos motivos fundamentales para el estudio de los FDAs, el incremento de la expresividad de los modelos probabilísticos, permitiendo la representación de interacciones de mayor orden; y la reducción de la complejidad de los métodos y algoritmos usados para la estimación y muestreo de probabilidades.

Una factorización de una distribución (\mathbf{p}) es el producto de un conjunto de marginales¹⁸ de (\mathbf{p}). El uso de las factorizaciones cobra una gran importancia en EDAs porque permite obtener una representación condensada de distribuciones de probabilidad que en otro caso serían muy difíciles de almacenar. Este hecho, y la existencia de métodos que permitan muestrear de manera eficiente las distribuciones de probabilidad almacenadas en forma de factorizaciones, han motivado su uso en los EDAs. (3)

La mayoría de los FDAs existentes hasta el momento utilizan modelos probabilísticos que son factorizaciones, pero para la estimación de las distribuciones hasta ahora no se propone un método práctico, por tal motivo en el presente trabajo se utilizará la mezcla de árboles para estimar la distribución.

Los FDA son clasificados de acuerdo a la complejidad de los modelos usados para capturar las dependencias entre las variables. Para el análisis de esta investigación se agruparán los FDA de acuerdo

¹⁸ Probabilidades marginales: son los valores en los márgenes de la tabla de probabilidades conjuntas, que proporcionan la probabilidad de cada evento por separado.

Capítulo 1. Fundamentación Teórica

a la manera que se realiza el aprendizaje en el modelo gráfico usado. Una clase agrupa a los algoritmos que realizan un aprendizaje paramétrico de las probabilidades, y la otra comprende aquellos algoritmos que hacen un aprendizaje estructural del modelo. (3)

A continuación mencionamos algunos algoritmos que son miembros de la familia FDAs.

Algoritmo Evolutivo con Distribución Marginal Univariada (UMDA) que usa marginales univariadas para aproximar las distribuciones conjuntas, este algoritmo transforma el paisaje original de una función en un nuevo paisaje definido para la media de la función calculada a partir de la probabilidad de un modelo. Esta transformación suaviza la superficie de la función de tal forma que si existe una tendencia al óptimo global el UMDA lo encontrará.

El Algoritmo Evolutivo con Distribución Factorizada basado en Árboles (Tree-FDA) utiliza como modelo probabilístico al árbol. Aunque el Tree-FDA puede usar otros métodos para construir el árbol, el algoritmo propuesto por Chow y Liu es utilizado en su implementación actual y este algoritmo lo que hace es que calcula el árbol de cubrimiento de peso máximo a partir de la matriz de información mutua entre las variables.

Algoritmo Evolutivo con Distribución Factorizada y modelo fijo de las dependencias (FDA*) construye un árbol de cliques para representar la factorización de la probabilidad conjunta. El árbol de cliques permite un aprendizaje paramétrico eficiente, y muestrear la próxima población. Este algoritmo no necesita una factorización exacta de la distribución, distribuciones aproximadas pueden ser usadas. Emplea factorizaciones basadas en el conocimiento previo sobre la estructura de la función objetivo. Los resultados muestran que el FDA* se desempeña mejor que otros algoritmos evolutivos en la optimización de funciones aditivas. (3)

La otra clase agrupa los algoritmos con un aprendizaje estructural, entre los que se encuentra los siguientes algoritmos:

Algoritmo de Maximización de la Información Mutua para la Clasificación de Entradas (Mutual Information Maximization for Input Clustering (MIMIC) algorithm (9) que usa una búsqueda glotona para generar el modelo probabilístico en forma de cadena donde cada variable está condicionada a la anterior en la cadena. MIMIC busca en cada generación el modelo en forma de cadena que está más próximo en la

Capítulo 1. Fundamentación Teórica

distancia de Kullback-Leibler a la distribución de los puntos seleccionados.

En el Algoritmo Evolutivo con Distribución Marginal Bivariada (Bivariate Marginal Distribution Algorithm (BMDA)) (10) las más importantes interacciones bivariadas, de acuerdo al test Chi-cuadrado de Pearson (11) son tomadas en consideración para modelar la distribución probabilística. El modelo probabilístico usado para representar estas dependencias es un bosque. El BMDA es capaz de optimizar una clase más amplia de funciones que el UMDA porque su modelo probabilístico cubre interacciones entre pares de variables. Sin embargo, este algoritmo puede no ser capaz de optimizar funciones con dependencias de mayor orden.

El Algoritmo Genético Compacto Extendido (Extended Compact Genetic Algorithm (ECGA)) (12) usa una factorización de la probabilidad donde las variables están separadas en grupos que no se solapan. Los grupos son encontrados minimizando la medida de Mínima Extensión de la Descripción (Minimum Description Length (MDL) metric) de los datos. En el modelo, se asume que cada grupo es independiente de los restantes. El modelo probabilístico de MIMIC fue extendido a una clase mayor de grafos de dependencias que pueden ser representados usando redes con forma de árbol.

El algoritmo, llamado Combinando Optimizadores Locales con Árboles de Información Mutua (Combining Optimizers with Mutual Information Trees (COMIT)) busca una red Bayesiana donde cada nodo puede tener un solo padre.

5. Algoritmo Evolutivo con Factorización de distribuciones Basado en Mezcla de árboles.

El algoritmo evolutivo con factorización de distribuciones basado en mezcla de árboles, es un algoritmo que utiliza la mezcla de árboles para modelar la distribución probabilística de la población. (3) Es un algoritmo relativamente novedoso, no teniéndose conocimiento de aplicaciones en el campo de las telecomunicaciones. En esta investigación se va a aplicar al problema del Diseño de redes de Radio (RND), buscando demostrar una solución eficiente.

Una de las alternativas para tratar con problemas que exhiben complejas interacciones entre sus variables es la combinación de modelos más simples. Es de esperar que la combinación de estos modelos pueda

Capítulo 1. Fundamentación Teórica

contribuir a una mejor representación de estas interacciones. Existen varias formas de combinar modelos probabilísticos. En esta tesis se investiga el uso de mezclas de árboles en los FDAs. La modelación probabilística usando mezclas de distribuciones finitas (10) consiste en la modelación de una distribución probabilística con una suma pesada de otras distribuciones.

6. Mezclas de distribución

Los enfoques usados para el aprendizaje¹⁹ de mezclas de distribuciones incluyen (13): métodos gráficos, métodos de mínima distancia, métodos de máxima verosimilitud, métodos de los momentos y enfoques Bayesianos. Cuando la variable de selección es conocida, el aprendizaje de la mezcla se limita al aprendizaje de la estructura y de los parámetros en cada componente, y posteriormente a la determinación de los coeficientes. Cuando la variable de selección no es conocida, una de las alternativas que pueden ser utilizadas es el algoritmo EM (14), que busca una mezcla que maximice la verosimilitud de los datos.

Aunque existen muchas variantes de este algoritmo según el tipo de mezclas, un problema común asociado con todas ellas es que pueden quedar atrapadas en máximos locales de la verosimilitud. Diversos métodos y heurísticas se han propuesto para permitir al EM escapar de los óptimos locales (15), sin embargo estas soluciones no son todavía lo suficientemente robustas y generales. En (16) los métodos de Monte Carlo basados en cadenas de Markov reversibles (Reversible Markov chain Monte Carlo Methods (17) se utilizan para encontrar mezclas normales univariadas con un número desconocido y variable de componentes. Estos métodos son capaces de saltar entre los subespacios de los parámetros que corresponden a un número diferente de componentes en la mezcla. En (18) proponen un método para aprender la mezcla de redes Bayesianas (MBNs) en las cuales cada red Bayesiana está definida por distribuciones Gaussianas.

6.1. Mezcla de árboles.

Las mezclas de árboles fueron introducidas en (19), donde su utilidad como algoritmo de estimación de densidades y como herramienta de clasificación fue demostrada en un conjunto de problemas.

¹⁹ Aprendizaje de un algoritmo: son los datos que se le pasan a la máquina para inducir el conocimiento.

Capítulo 1. Fundamentación Teórica

Presentamos primeramente un algoritmo para aprender una mezcla de árboles que se ajuste a un conjunto de datos observado. El algoritmo sigue el paradigma de la Máxima Verosimilitud a través del algoritmo EM. En nuestro conocimiento existe solamente otro algoritmo utilizado para aprender mezclas de árboles: en (20) los autores utilizan un algoritmo glotón que a partir de una mezcla inicial de árboles aplica perturbaciones locales hasta que no resulta posible mejorar la verosimilitud del modelo. Este algoritmo simple no es de interés para nuestra investigación pues se puede aplicar solamente a una mezcla de árboles con estructura compartida.

El algoritmo que se utilizó para el aprendizaje de las mezclas de árboles en el MT_FDA fue el algoritmo iterativo EM, este es un método numérico general y confiable para obtener estimadores de parámetros en contextos con datos incompletos.

7. Problema a resolver

La colocación de antenas es la principal tarea que se efectúa a la hora de establecer la infraestructura de una red de radio. La instalación de una antena de telefonía no es algo trivial, ya que se debe buscar un emplazamiento bien situado, que ofrezca una buena cobertura y esté alejado de lugares restringidos tales como pudieran ser un colegio o un hospital. Además tales instalaciones suelen tener asociado un elevado coste económico, no sólo por los equipos necesarios, sino por el coste de alquiler del emplazamiento, que suele ser un bloque de viviendas. Debido a esto, el problema de colocación de antenas, también conocido como diseño de red de radio (Radio Network Design, RND) constituye una tarea crítica que puede alcanzar una gran complejidad cuando se pretende realizar en un entorno extenso como puede ser una gran ciudad.

Esta tarea se puede plantear como un problema de optimización en el que se pretende alcanzar la mayor cobertura posible en un escenario a la vez que se emplea el menor número de antenas, las cuales sólo pueden colocarse en un conjunto de emplazamientos disponibles (generalmente azoteas de edificios). De esta manera se pueden aplicar con éxito técnicas de optimización combinatoria que permitan obtener configuraciones de antenas eficientes en un tiempo razonable. Para ello se emplea una función objetivo que calcula un valor de mérito para cada solución a partir de los objetivos.

Este tipo de problemas es de la clase NP completo y existen técnicas como los algoritmos evolutivos que

Capítulo 1. Fundamentación Teórica

han demostrado su valor como excelente estrategia de resolución de problemas de este tipo y a menudo producen soluciones que serían inabordables mediante cualquier otro método.

En esta tesis se investiga la aplicación del algoritmo MT_FDA a dicho problema, debido a que es relativamente novedoso en el campo de los EDAs y no se le conocen aplicaciones en la rama de las telecomunicaciones, además coincide con el perfil de la facultad 2, y puede ser interés de futuras investigaciones mayores.

8. Algoritmos utilizados en otras investigaciones para resolver el problema RND.

El diseño de redes de radio (RND) es uno de los problemas fundamentales dentro del diseño de redes celulares para telecomunicaciones, varios algoritmos se han utilizado para tratar de resolver este problema de la manera más eficiente, entre estos se encuentran:

- PBIL (Population-Based Incremental Learning) es un método que combina un algoritmo genético con el aprendizaje competitivo para la función objetivo. En lugar de aplicar operadores, este algoritmo deduce una distribución de probabilidad de la población actual, y muestrea de la nueva población para inferir la distribución. (21)
- DE (Differential Evolution) es un algoritmo que tiene como objetivo resolver problemas de optimización y se ha utilizado arrojando resultados satisfactorios. Es un método estocástico simple basado en poblaciones en función de maximizar y minimizar, utilizado en una amplia gama de problemas de optimización, incluyendo la optimización multiobjetivo. (21)
- SA (Simulated annealing) es un método de búsqueda local basado en la simulación Monte Carlo para hallar la configuración más estable de un sistema físico de n partículas. La generalización del método a problemas de optimización fue propuesta por Kirkpatrick en (22). SA no es una técnica basada en población, ya que mantiene una única solución candidata para el problema (análoga al estado actual de un sistema) con una función objetivo asociada (análoga a la función de energía) para la que se quiere hallar el mínimo global (análogo al estado más estable). El espacio de búsqueda se explora mediante un operador de transición y para controlar la probabilidad de

Capítulo 1. Fundamentación Teórica

aceptar soluciones peores que la actual, este utiliza un parámetro temperatura T que no tiene un análogo en el problema de optimización, por lo cual definir un esquema de enfriamiento que permita evitar mínimos locales es un arte. También los restantes parámetros del método (el valor inicial de T , el número de iteraciones realizadas en cada paso y el largo de la cadena de Markov utilizada) son usualmente determinados empíricamente. (23)

- GenGA (Algoritmo Genético Generacional) El algoritmo aplica operadores estocásticos tales como selección, recombinación y mutación sobre la población para calcular una generación completa de nuevos individuos. (24)
- SsGA (Algoritmo Genético de estado estacionario) el algoritmo utilizado es un algoritmo genético de estado estacionario estándar que utiliza selección por torneo binario. La codificación de soluciones, la recombinación y la mutación dependerán del problema concreto. Este se puede considerar como un algoritmo híbrido. (24)
- CHC puede entenderse como un tipo de algoritmo genético (AG) en el que no se emplea la mutación para producir soluciones nuevas sino únicamente un mecanismo de recombinación o cruce llamado HUX. La selección de los individuos para la siguiente generación se realiza siguiendo un criterio elitista²⁰ entre padres e hijos, y en el momento en que se detecta el estancamiento de la población se provoca un reinicio de la misma. La ausencia de mutación unida al criterio elitista de selección para la nueva población tiende a homogeneizar ésta. Con el fin de retrasar la convergencia de la población. Llegado el umbral de incesto a cero (la distancia mínima entre dos individuos para que se puedan recombinar es 0), si transcurren k iteraciones sin que ninguna solución nueva haya ingresado en la población, se considera que la población ha convergido y la ejecución del algoritmo se ha estancado. Entonces se provoca el reinicio de la población: todos los individuos excepto los mejores se modifican drásticamente mediante una mutación por inversión de bit con probabilidad de mutación de bit muy elevada. (25)

²⁰ Elitista: los mejores individuos de la población.

9. Tecnologías de desarrollo

9.1. Entorno de Desarrollo Integrado

Eclipse: La plataforma Eclipse está diseñada para la integración de IDEs²¹ que pueden ser utilizados para diversos propósitos, como desarrollar programas en Java, C++, PHP, etc. Además, para integrarse y ejecutar módulos llamados plug-ins²².

Eclipse fue desarrollado originalmente por IBM como el sucesor de su familia de herramientas para VisualAge. Eclipse es ahora desarrollado por la Fundación Eclipse, una organización independiente sin ánimo de lucro que fomenta una comunidad de código abierto. (4)

9.2. Sistema Operativo:

Linux (Debian): Debian o Proyecto Debian es una comunidad conformada por desarrolladores y usuarios, que mantiene un sistema operativo GNU²³ basado en software libre precompilado y empaquetado, en un formato sencillo en múltiples arquitecturas de computador y en varios núcleos.

Debian nace como una apuesta por separar en sus versiones el software libre del software no libre. El modelo de desarrollo del proyecto es ajeno a motivos empresariales o comerciales, siendo llevado adelante por los propios usuarios, aunque cuenta con el apoyo de varias empresas en forma de infraestructuras. Debian no vende directamente su software, lo pone a disposición de cualquiera en Internet, aunque sí permite a personas o empresas distribuir comercialmente este software mientras se respete su licencia.

La comunidad de desarrolladores de Debian cuenta con la representación de **Software in the Public Interest**, una organización sin ánimo de lucro que da cobertura legal a varios proyectos de software libre. (4)

²¹ IDE: Entorno de desarrollo integrado(IDE) por sus siglas en ingles

²² Plug-in: es una aplicación que se relaciona con otra para aportarle una funcionalidad nueva y generalmente muy específica.

²³ GNU: sistema operativo completamente libre, acrónimo recursivo que significa, NO es Libre.

9.3. Lenguaje de Programación

C++: En la actualidad existe una gran diversidad de lenguajes de programación, cada uno con sus características ventajosas y otras menos deseadas. Esto hace dificultoso a los desarrolladores escoger el lenguaje en el cual van a implementar alguna aplicación que necesiten desarrollar. En esta investigación se hace uso del lenguaje de programación c++, este constituye una extensión orientada a objetos del lenguaje c, conservando su eficiencia, poder y elegancia, al mismo tiempo que provee otras características avanzadas de programación y el soporte para programar siguiendo el paradigma orientado a objetos. En el siguiente trabajo se hace uso del lenguaje C++ debido a que el algoritmo utiliza librerías programadas en esta tecnología. (26)

9.4. Conclusiones

En este capítulo se abordaron los conceptos fundamentales para la comprensión de la investigación, se argumentaron las diferentes técnicas de optimización tenidas en cuenta, además de explicar el problema RND en el cual se pretende testear el algoritmo en estudio.

Capítulo 2. Descripción de la Función Objetivo y el Algoritmo MT_FDA

CAPÍTULO 2. DESCRIPCIÓN DE LA FUNCIÓN OBJETIVO Y EL ALGORITMO MT_FDA.

1. Introducción

En este capítulo se hace una explicación más detallada del problema que se va a analizar, además de explicar el algoritmo MT_FDA y la función objetivo para el problema que se estará analizando.

2. Diseño de redes de Radio (Radio Network Design (RND)).

La Diseño de redes de Radio (RND) es uno de los problemas fundamentales dentro del diseño de redes celulares para telecomunicaciones. La división del espacio en celdas de cobertura asociadas cada una a una antena persigue un doble objetivo: maximizar el tamaño del área cubierta por la señal de radio y reducir el número de antenas empleadas. (25)

En la literatura este problema ha sido resuelto con anterioridad aplicando técnicas monoobjetivo y un valor numérico obtenido combinando los valores de los distintos objetivos. Estas técnicas han permitido encontrar soluciones óptimas en base a la combinación de valores objetivos definidas, pero adolecen de una falta de visión de la naturaleza multiobjetivo del problema. En esta investigación se presenta un algoritmo que debido a sus características, se pretende demostrar, que dará una solución más eficiente que los resultados obtenidos anteriormente con otros algoritmos utilizados.

La determinación de las localizaciones²⁴ de las antenas es una tarea fundamental dentro del diseño de redes celulares como las empleadas en telefonía móvil. El emplazamiento de una antena determina la zona de cobertura de la misma también llamada celda, por lo que al decidir los emplazamientos de todas las antenas de una red se está estableciendo la zona de cobertura, el grado de cobertura (número de antenas que cubren una determinada zona) y el número de antenas que se colocan. Por ello, una buena o mala decisión puede tener consecuencias directas sobre la calidad del servicio ofrecido y el coste

²⁴ Localizaciones de antenas: lugar donde se ubica una antena

Capítulo 2. Descripción de la Función Objetivo y el Algoritmo MT_FDA

incurrido.

Consideremos un escenario urbano al que se quiere ofrecer cobertura de señal de radio. El área cubierta es un factor esencial; idealmente se quiere cubrir toda la zona, pero los recursos disponibles son limitados, por lo que hay que intentar colocar el menor número de antenas disponibles. La empresa encargada realiza un estudio previo que determina una serie de emplazamientos aptos para la colocación de antenas: techos de edificios elevados y lugares con buena visibilidad (sin obstáculos cercanos) y alejados de zonas no permitidas (como hospitales o colegios). En total existen un centenar de puntos, repartidos de manera más o menos uniforme por el área que se quiere cubrir, entre los cuales escoger dónde colocar las antenas.

El problema de Diseño de Redes de Radio tal y como se considera en los trabajos previos y en este trabajo, consiste en seleccionar un subconjunto de localizaciones de entre un conjunto de localizaciones disponibles para colocar antenas en ellas. El objetivo que se persigue es doble, ya que por un lado se quiere maximizar la cantidad de área cubierta por las antenas, pero por otro se espera que el número de antenas sea mínimo. Ambos objetivos están claramente contrapuestos. (25)

Para manejar la información de este problema se empleó una discretización del problema que empleaba un modelo de rejilla cuadrada para representar el terreno. Este modelo es el mismo que se empleó en trabajos anteriores y el mismo que se utilizará en este trabajo. (25)

Capítulo 2. Descripción de la Función Objetivo y el Algoritmo MT_FDA

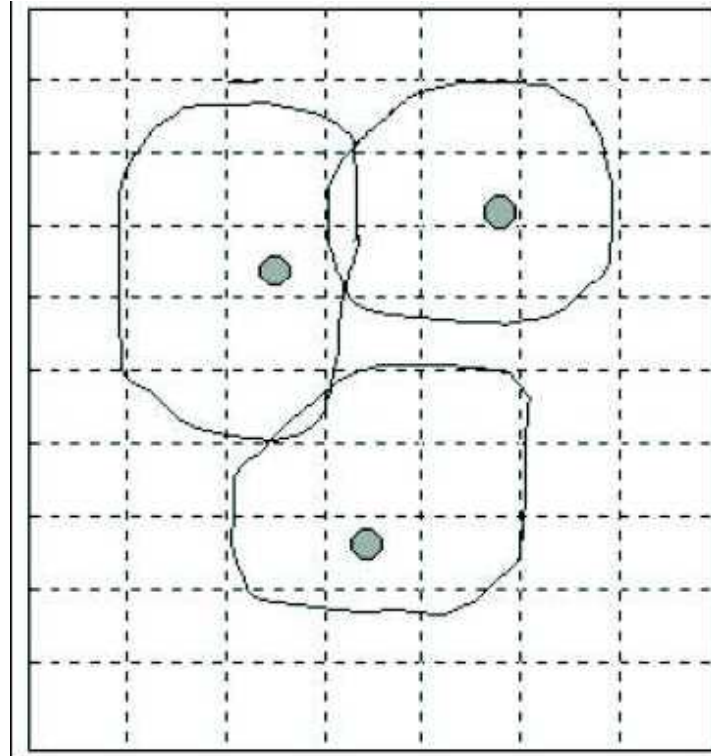


Figura 3. Coberturas ofrecidas por distintas antenas

Capítulo 2. Descripción de la Función Objetivo y el Algoritmo MT_FDA

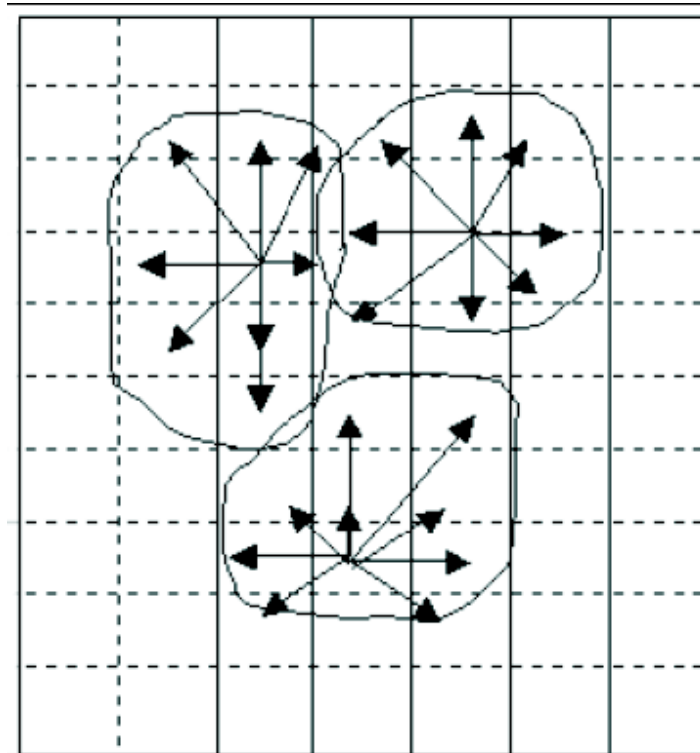


Figura 4. Cobertura discretizada en el modelo de rejillas

La rejilla empleada para modelar el terreno es cuadrada y contiene 287×287 cuadros, cada uno de los cuales representan un sector atómico de terreno que o se cubre en su totalidad o no se cubre en absoluto. La Figura 4 ilustra un modelo simplificado de discretización del terreno con una rejilla 10×7 .

El conjunto de emplazamientos disponibles (ALS) para la colocación de antenas se representa mediante una lista de coordenadas de la rejilla, indicando los sectores del terreno en los que se encuentran esos puntos. Se considera que la discretización es suficientemente precisa como para que cualquier posición de la antena dentro del sector resulte en una misma celda de cobertura.

Finalmente, un diseño de red consiste en un subconjunto de las coordenadas del ALS²⁵. Cualquier subconjunto válido constituye una solución del problema; el objetivo de este trabajo es encontrar la

²⁵ ALS: Localizaciones disponibles.

Capítulo 2. Descripción de la Función Objetivo y el Algoritmo MT_FDA

solución óptima. Las soluciones se codifican mediante cadenas binarias cuya longitud iguala a la cardinalidad del ALS, y en las que cada bit se corresponde unívocamente con una localización de la lista. Un “1” indica la colocación de antena en las coordenadas correspondientes, mientras que un “0” indica la no colocación.

El espacio de soluciones del problema es el total de valores que puede adoptar la cadena, y su tamaño, que depende del número de localizaciones disponibles, es $2^{|\text{ALS}|}$. Se considera que el tamaño de una instancia se corresponde con el tamaño del ALS empleado en esa instancia.

Para este trabajo se emplea el mismo modelo de cobertura que en los trabajos anteriores: una antena produce la misma celda independientemente de donde se coloque, y esta celda es una zona cuadrada de 41 x 41 sectores. La forma de la celda se eligió de manera que pudiera conseguirse la cobertura completa (100% del área) de manera óptima con 49 antenas formando una rejilla 7 x 7 ($7 \times 41 = 287$). (25)

3. Descripción de la función objetivo

Diseñar una función objetivo es uno de los pasos más importantes a la hora de aplicar un algoritmo evolutivo a un problema. Esa clase de representación que se utiliza en los algoritmos evolutivos influye en su uso en la función objetivo. En esta sección se procederá a la definición de la función objetivo, esta explica todos los detalles del conocimiento del problema ya que escanea la cadena binaria proporcionada por el algoritmo, de esta forma calcula la zona cubierta así como el tipo de medida que podrían ser necesarias para evaluar la calidad del individuo seleccionado como una solución tentativa del problema.

La función objetivo que se va a introducir en esta sección es adecuada para la resolución del problema RND ya que por un lado se quiere maximizar el área cubierta y por otro lado se pretende minimizar el número de transmisores.

A continuación se muestra un pseudocódigo de la misma:

Capítulo 2. Descripción de la Función Objetivo y el Algoritmo MT_FDA

```
Inicialización
Total Transmisores TT <- 349
Cantidad sectores horizontales del terreno SH <- 287
Cantidad sectores vertical del terreno SV <- 287
Total sectores del terreno TS <- 82369
Celda cobertura CC <- 41
Sea V conjunto de localizaciones disponibles (ALS)

Función objetivo (Individuo M)
Begin
Sea puntos_cobertura <- 0
Sea cantidad_transitores <- 0
Sea G conjunto de posiciones de cobertura de TT
Inicializar G <- 0
Sea B conjunto de puntos a los que da convergencia un transmisor T ∈ TT
Recorro los puntos de M y para cada punto k ∈ M
  Begin
  Si M [k] existe presencia de antena
    Begin
    Incrementar cantidad_transitores en 1
    Capturar coordenadas x, y
    G [x, y] <- ubicar antena
    Inicializar B <- 0
    B <- llenar celdas cobertura (B, x, y)
    Para cada punto j ∈ B
      Begin
      Capturar coordenadas x1, y1
      Si x1 ∈ SH && y1 ∈ SV
        Begin
        Si G [x1, y1] no existe presencia de antena
          Begin
          G [x1, y1] <- incrementar la cobertura en 1
          Si G [x1, y1] existe presencia de una antena
            puntos_cobertura <- incrementa en 1
          End
        End
      End
    End
  End
End
Sea cobertura <- 100 * puntos_cobertura / TS
Retornar (cobertura * cobertura) / cantidad_transitores
End
```

Pseudocódigo 2. Función Objetivo

Capítulo 2. Descripción de la Función Objetivo y el Algoritmo MT_FDA

En el diseño de una red celular es preferible el incremento del área cubierta frente al ahorro en el número de antenas (nunca es deseable la solución sin antenas). En (24) se propuso una función de medición de calidad (Ecuación 1) para las soluciones *función objetivo*) que permitía combinar ambos factores para asignar un único valor numérico a cada solución y poder emplear una técnica monoobjetivo. (27)

$$f(\vec{x}) = \frac{\text{Cobertura}(\vec{x})^\alpha}{\text{Número de antenas}(\vec{x})}$$

Ecuación 1. Función Objetivo

En esta formulación el parámetro α ajusta la relación de importancia entre cobertura y número de antenas. El valor que se emplea para α es 2. (28)

En el **Pseudocódigo 2** se puede observar que se parte de inicializar algunos parámetros que son puntos clave para poder resolver la función. El conjunto **G** indica las posiciones de cobertura de todas las antenas que serán ubicadas, este se inicializa a cero indicando que no existe ninguna antena ubicada en el terreno. Tenemos además el conjunto **B** el cual es utilizado para guardar los puntos a los que ofrece cobertura un determinado transmisor.

Cada individuo está formado por un vector que contiene las variables, o sea contiene los puntos que ofrece la empresa donde se pueden colocar las antenas (ALS), de ahí que se recorra todos los puntos del individuo y para cada uno se verifica si existe presencia de antenas (si en esa posición hay un **1**, indica que hay antenas, si hay un **0** indica que no hay antenas), si existe presencia de antenas se aumentará la cantidad de transmisores, y se capturan las coordenadas de donde se encuentra ésta. Luego en el conjunto de todas las posiciones de cobertura (**G**) se ubica una antena, y se inicializa el conjunto de puntos (**B**) a los que esa antena le da cobertura en **0**.

Se llena **B** con las posiciones de cobertura de la antena llamando al método “llenar celdas cobertura” al cuál se le pasa como parámetro el vector que se va a llenar (**B**) y las coordenadas de la antena (**x, y**). Este método se mueve por la celda de cobertura y encuentra todas las posiciones a las que le da cobertura esa antena.

Capítulo 2. Descripción de la Función Objetivo y el Algoritmo MT_FDA

Luego para cada punto de **B** se capturan las coordenadas y se verifica que pertenezcan al terreno (**SH**, **SV**), si pertenecen al terreno y en esa coordenada no existe una antena, se incrementa la cobertura. Luego si la cobertura en esa coordenada es **1** (lo que indica que hay una sola antena ofreciéndole cobertura a esa área) incremento los puntos de cobertura.

Por último, cuando tengo todos los parámetros aplico la **ecuación 1**.

Donde

$$Cobertura = (100 * puntos_cobertura) / total\ sectores\ terreno$$

Ecuación 2. Cobertura

4. Algoritmo MT_FDA

Existen diferentes formas de usar una distribución de probabilidad entre las que se encuentra:

- **Estimadores:** Cuando un modelo gráfico es usado como estimador (o evaluador) se asocia un valor de probabilidades a cada punto **x**. La forma tradicional de usar los modelos gráficos como estimadores es buscando en el modelo las probabilidades marginales y condicionales correspondientes a la solución **x**. Estos marginales son sustituidos en la factorización representada por el modelo.
- **Generadores:** Si queremos usar **p** como un generador esperamos que el modelo gráfico no de un valor de probabilidad sino uno o más puntos que forman una muestra y donde cada punto **x** tiene aproximadamente la frecuencia **p(x)**.

En los EDAs, los modelos gráficos se usan como estimadores para representar el modelo probabilístico²⁶ de los individuos seleccionados. La estructura del modelo probabilístico se puede fijar de acuerdo a la

²⁶ Modelo probabilístico: es la forma que pueden tomar un conjunto de datos obtenidos de muestreos de datos con comportamiento que se supone aleatorio.

Capítulo 2. Descripción de la Función Objetivo y el Algoritmo MT_FDA

información existente sobre el problema. Usados como generadores, los modelos gráficos sirven para generar nuevas soluciones. Esto se puede lograr con los algoritmos de muestreo o utilizando el algoritmo de las K configuraciones más probables. (3)

5. Mezcla de árboles

Definición 1. Una mezcla de distribuciones se define con una distribución de la siguiente forma:

$$Q(x) = \sum_{j=1}^m \lambda_j \cdot f_j(x)$$

Ecuación 3. Mezcla de Distribución

con $\lambda_j \geq 0$, $j = 1, \dots, m$, $\sum_{j=1}^m \lambda_j = 1$.

A las funciones f se les llama densidades o componentes de la mezcla, y a las λ se les denomina proporciones o coeficientes de la mezcla. m es el número de componentes de la mezcla.

Una mezcla de distribuciones puede ser vista como un modelo contentivo de una variable no observada \mathbf{z} , que toma valores $j \in \{1, \dots, m\}$ con probabilidad λ_j . En algunos problemas la variable \mathbf{z} puede ser conocida u observada. Ejemplos de mezclas de distribuciones incluyen:

- Mezclas de distribuciones Gaussianas.
- Multiredes Bayesianas.
- Mezclas de árboles.

En las mezclas de distribuciones Gaussianas²⁷ cada componente es una distribución multivariada

²⁷ Distribución Gaussianas: es la distribución de probabilidad que con más frecuencia aparece en estadística y teoría de probabilidades. Esto se debe a dos razones fundamentalmente: Su función de densidad es simétrica y con forma de campana, lo que favorece su aplicación como modelo a gran número de variables estadísticas. Es, además, límite de otras distribuciones y aparece relacionada con multitud de resultados ligados a la teoría de las probabilidades gracias a sus propiedades matemáticas.

Capítulo 2. Descripción de la Función Objetivo y el Algoritmo MT_FDA

Gaussianas. Las mezclas de distribuciones Gaussianas han demostrado ser útiles para representar interacciones no lineales. (3)

Una Multired Bayesiana puede ser vista como una red donde las aristas pueden aparecer o desaparecer dependiendo de los valores de ciertos nodos en el modelo gráfico, una noción que se ha llamado aserciones de independencia asimétricas. Las Multiredes Bayesianas constituyen una generalización de las redes Bayesianas. La familia de las distribuciones de probabilidad representables por las redes Bayesianas y por las Multiredes Bayesianas es la misma. Sin embargo, una Multired puede dar lugar en la práctica a un ahorro substancial de la complejidad necesaria para aprender y muestrear una red Bayesiana equivalente.

Las redes bayesianas son un grafo acíclico y dirigido donde los nodos representan las variables y las aristas representan las relaciones que se establecen entre las variables, y una multired bayesiana es una red bayesiana donde cada nodo es una red.

La mezcla de árboles es un caso particular de Multired Bayesiana, donde cada componente es un árbol. (3)

Definición 2: Una distribución de probabilidad²⁸ basada en un árbol se define como:

$$T(x) = \prod_{i=1}^n p(x_i | pa(x_i))$$

Ecuación 4. Distribución de probabilidad basada en un árbol

Donde $pa(x_i)$ es el padre de x_i en el grafo, y $p(x_i | pa(x_i)) = p(x_i)$ cuando $pa(x_i) = \Phi$ es la raíz del árbol. La distribución T será también llamada árbol cuando no exista confusión posible.

El grafo (V, E) representa la estructura de la distribución. Para todos los árboles se define en el dominio V al conjunto de aristas E , el cual define únicamente la estructura del árbol.

²⁸ La distribución de probabilidad: describe el rango de valores de una variable así como la probabilidad de que el valor de la variable esté dentro de un subconjunto de dicho rango.

Capítulo 2. Descripción de la Función Objetivo y el Algoritmo MT_FDA

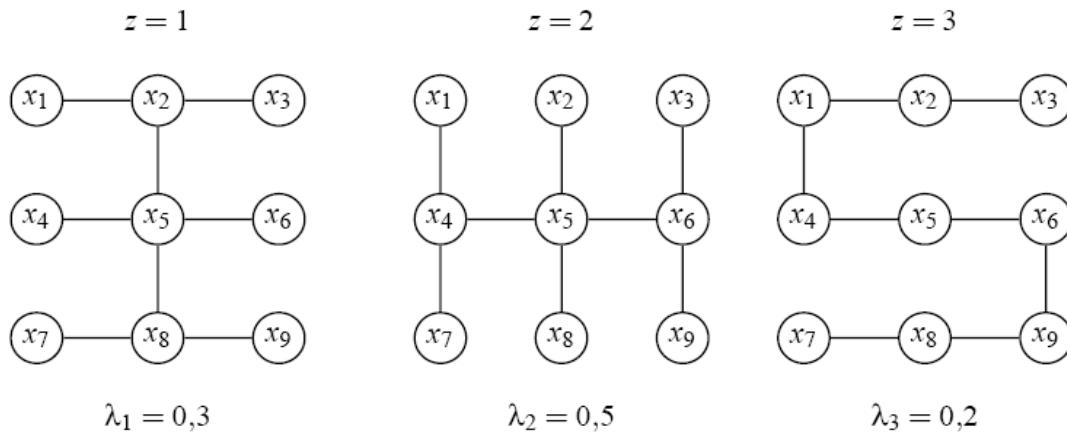


Figura 5. Mezcla de árboles con 3 componentes Figura

Definición 3: Una mezcla de árboles se define como una distribución de la forma:

$$Q(x) = \sum_{j=1}^m \lambda_j T^j(x)$$

Ecuación 5: Mezcla de árbol

con $\lambda_j \geq 0$, $j = 1, \dots, m$, $\sum_{j=1}^m \lambda_j = 1$.

En este caso las distribuciones basadas en cada árbol son las componentes de la mezcla. Los m árboles pueden tener diversas estructuras y diversos parámetros. Cuando todos los árboles tienen la misma estructura pero los parámetros de los árboles son diferentes, el modelo se llama una mezcla de árboles con estructura compartida. (3)

Capítulo 2. Descripción de la Función Objetivo y el Algoritmo MT_FDA

5.1. Algoritmo Iterativo para la maximización de la verosimilitud en mezclas de árboles.

El algoritmo Esperanza-Maximización (EM), se usa en estadística para encontrar estimadores de máxima verosimilitud²⁹ de parámetros en modelos probabilísticos que dependen de variables no observables. El algoritmo EM alterna pasos de esperanza³⁰ (paso E), donde se computa la esperanza de la verosimilitud mediante la inclusión de variables latentes como si fueran observables, y un paso de maximización (paso M), donde se computan estimadores de máxima verosimilitud de los parámetros mediante la maximización de la verosimilitud esperada del paso E. Los parámetros que se encuentran en el paso M se usan para comenzar el paso E siguiente, repitiéndose el proceso. (4)

El algoritmo EM se utiliza frecuentemente para algoritmos de agrupamiento en aprendizaje automático³¹ y visión artificial. El algoritmo de aprendizaje utilizado por los EDAs es una variación del EM como algoritmo de aprendizaje automático y se le nombra algoritmo EM iterativo (IEM). (3)

5.2. Muestreo de la mezcla de árboles

El método clásico para muestrear mezclas de árboles tiene dos pasos. En el primero se decide qué árbol será muestreado. Este paso se realiza escogiendo entre las componentes proporcionalmente a sus coeficientes en la mezcla. Primeramente la variable en la raíz es instanciada a partir de las probabilidades marginales, después el árbol es recorrido siguiendo una estrategia de primero a lo ancho, instanciando cada variable dado el valor de su padre en el árbol y utilizando las probabilidades condicionales. (3)

5.3. Pseudocódigo del algoritmo

A continuación se plantea el pseudocódigo del algoritmo

²⁹ Máxima Verosimilitud: es una técnica para estimar los valores esperados dado una muestra finita de datos.

³⁰ Esperanza matemática: también llamada valor esperado o media poblacional es el número que formaliza la idea de valor medio de un fenómeno aleatorio.

³¹ Aprendizaje automático: proceso de inducción del conocimiento.

Capítulo 2. Descripción de la Función Objetivo y el Algoritmo MT_FDA

```
t <= 0. Generar N puntos aleatoriamente.  
  
Do {  
  
    Evaluar los puntos en la función objetivo.  
  
    Seleccionar un conjunto S de  $k \leq N$  puntos de acuerdo a un método de selección.  
  
    Aprender una mezcla de árboles Q a partir del conjunto seleccionado.  
  
    Generar N nuevos puntos muestreando Q.  
  
    t <= t + 1  
  
} Until Algún criterio de terminación ha sido satisfecho.
```

Pseudocódigo 3. Algoritmo MT_FDA

Como se muestra anteriormente, se parte de una población inicial de N puntos generados aleatoriamente, en el segundo paso se evalúan los puntos en la función objetivo y se selecciona un conjunto S de k individuos ($k \leq N$), normalmente los que mejores resultados arrojen de la evaluación en la función objetivo y se mantienen como base de datos para la estimación del modelo. Luego se introduce el modelo probabilístico que mejor refleja las interdependencias entre las variables, en este caso el modelo probabilístico que se está investigando es la “mezcla de árboles” (Q). A partir de este modelo se genera una población auxiliar muestreando, por último la nueva población se genera a partir de la población anterior y de la población auxiliar, realizándose esto de forma elitista.

En el caso del problema de Diseño de redes de Radio las variables son los puntos que la empresa encargada detecta que están aptos para colocar antenas, teniendo en cuenta las escuelas, techos de edificios altos y todas las restricciones que se deben de tener en cuenta para colocar una antena y esos puntos se brindan en forma de una lista de coordenadas del terreno donde estaría el lugar que se pueden colocar las antenas. Esos puntos se evalúan en la función objetiva mencionada anteriormente y posteriormente se selecciona un subconjunto de puntos del conjunto inicial que por el método de selección usado “Truncamiento” se seleccionan un porcentaje (30 %) de la población, de manera elitista. A

Capítulo 2. Descripción de la Función Objetivo y el Algoritmo MT_FDA

continuación se realiza una distribución de probabilidades basada en mezcla de árboles generándose el modelo (relaciones existentes entre las variables). A partir del modelo obtenido se muestrea una generación auxiliar, tomando de esta, junto con la generación anterior los nuevos individuos que conforman la nueva generación. Para nuestro problema estamos interesados en obtener una generación de puntos en los cuáles se pueda ubicar antenas, tratando de que se cubra toda el área con el menor número de antenas que serían 49 antenas para el 100% del área a cubrir.

5.4. Análisis de la complejidad del MT_FDA.

5.4.1. Iniciación:

El paso de iniciación consiste en asignar valores a todos los individuos en la población inicial. Tiene complejidad nN .

5.4.2. Evaluación

La complejidad computacional de este paso depende del problema. Algunas funciones pueden demandar un tiempo exponencial en el número de variables. Sea $cost_f$ el tiempo asociado a la evaluación de la función f , la complejidad algorítmica de este paso es $O(Ncost_f)$.

5.4.3. Selección:

La complejidad del paso de selección depende del método de selección usado. En esta tesis usamos principalmente el método de selección por truncamiento. El algoritmo consiste en seleccionar N de los mejores individuos de la población. En el caso peor, la complejidad de este paso es $N\log(N)$.

5.4.4. Algoritmo IEM

La complejidad del IEM para el aprendizaje de mezclas de árboles fue calculada en (3), donde la complejidad total es $O(mn^2N + mn r_{MAX}^2)$ y $r_{MAX} = \max_{i \in \{1, \dots, n\}} |X_i|$. La complejidad de cada paso del algoritmo IEM se muestra en el cuadro siguiente:

Capítulo 2. Descripción de la Función Objetivo y el Algoritmo MT_FDA

Pasos	5	7	8	9	10
Complejidad	$O(mnN)$	$O(mn^2N)$	$O(mnr^2_{MAX})$	$O(mn^2)$	$O(mnr^2_{MAX})$

Complejidad del algoritmo IEM

5.4.5. Muestreo

La complejidad del muestreo de un árbol tiene orden $O(n)$, la complejidad total del paso de muestreo es $O(Nn)$.

5.4.6. Complejidad Total

La complejidad total del MT-FDA es $O(Gm(n^2N + nr^2_{MAX}))$, donde el tamaño de la población N y la cantidad de generaciones G varían según la dificultad del problema. Los resultados presentados en esta tesis sólo son para variables binarias ($r^2_{MAX} = 2$), por consiguiente la complejidad en nuestro caso es $O(Gmn^2N)$. (3)

Capítulo 3. Resultados de la Investigación

CAPÍTULO 3. RESULTADOS DE LA INVESTIGACIÓN

1. Introducción

En este capítulo se hará referencia a los resultados obtenidos en esta investigación además de comparar los resultados con los arrojados por otros algoritmos, y llegar a conclusiones.

2. Experimentos

Las instancias resueltas en este trabajo fueron las mismas que se utilizaron en investigaciones anteriores (25) Son todas instancias del problema RND en las que se emplean antenas con cobertura cuadrada, tal y como se describe en el capítulo 2, aplicando el enfoque monoobjetivo descrito anteriormente. Los tamaños de instancia varían desde 149 localizaciones disponibles (ALS), aumentando de 50 en 50 hasta llegar a 249.

En cada caso se realizan 25 ejecuciones independientes para asegurar la significación estadística. El criterio de parada es encontrar la solución óptima descrita anteriormente (100% de cobertura con 49 antenas) o alcanzar la generación máxima definida en este caso 40, ya que de esta manera cabe establecer comparaciones entre los esfuerzos realizados por los algoritmos en este trabajo y aquellos mostrados en (25).

Los parámetros que se le deben de pasar al algoritmo son:

- **Población:** es un conjunto de individuos y cada individuo está compuesto por un vector que representa el número de variables.
- **Truncamiento:** se refiere al método de selección utilizado por el algoritmo, en este caso se está truncando al 30 % de la población inicial.
- **Generación Máxima:** el algoritmo debe de regenerar varias generaciones hasta lograr la convergencia, en caso que no se logre es utilizada la generación máxima (en este caso definimos 40) para dar finalidad a la ejecución (como criterio de parada).
- **Número de árboles:** significa la cantidad de árboles que conforman la mezcla, en este caso se

Capítulo 3. Resultados de la Investigación

prueba con 1, 2, 4 y 6 árboles.

- **Número de variables:** es un parámetro que se define con 149, 199, y 249 y significa los puntos en que pueden ubicarse las antenas, es decir es un subconjunto del ALS, que un 1 indica la colocación y un 0 la no colocación.

En todas las instancias del problema se incluyen las coordenadas de las 49 localizaciones que permiten dar cobertura completa (100% del terreno) con el mínimo número de antenas. Esa configuración, definida como óptima en los trabajos anteriores, determina que el problema ha sido resuelto. La función objetivo descrita en la ecuación 1 alcanza su mejor valor de configuración en 203,839.

3. Métricas

3.1. Cantidad de evaluaciones.

La primera métrica empleada que servirá para establecer comparaciones entre algunos algoritmos aplicados en otras investigaciones a este mismo problema para obtener soluciones, dado que las ejecuciones se realizan hasta alcanzar la solución óptima, es el esfuerzo computacional. Este esfuerzo se define como el número de puntos (soluciones) que es necesario evaluar para resolver el problema (cantidad de evaluaciones). La interpretación de esta métrica es directa: el algoritmo que resuelve el problema de manera más eficiente es el que requiere menos esfuerzo, y por tanto el que tenga un valor más bajo de esfuerzo computacional. Aunque los enfoques dados a los distintos tipos de algoritmo no son iguales, puede establecerse esta comparación debido a que todos los algoritmos miden su esfuerzo hasta alcanzar la configuración óptima definida anteriormente.

3.2. Tiempo de ejecución.

El tiempo de ejecución que se demora cada algoritmo no es una métrica que se va a tener en cuenta a la hora de comparar entre los diferentes algoritmos, debido a que, los resultados que arrojan cada algoritmo se obtuvieron de otras investigaciones y no se muestra el tipo, ni la cantidad de máquinas que se usaron para cada corrida de los diferentes algoritmos, por tanto no es un parámetro que se pueda tener en cuenta a la hora de comparar con los demás algoritmos, pero si a la hora de arrojar conclusiones entre los valores de las diferentes variables.

Capítulo 3. Resultados de la Investigación

3.3. Generación de parada.

La generación de parada es la cantidad de veces que es necesario regenerar la población para llegar a un resultado final, esta métrica es directamente proporcional al esfuerzo computacional debido a que la cantidad de evaluaciones es la multiplicación de la generación de parada por la población. Este es un parámetro que vamos a tener en cuenta para arrojar conclusiones entre los resultados de las diferentes variables, pero no se comparará con los resultados de los demás algoritmos debido a que su resultado está implícito en el número de evaluaciones que arroja cada algoritmo.

4. Resultados

A continuación se muestran los resultados obtenidos en las diferentes corridas que se le realizaron al algoritmo. Están descritos los resultados obtenidos con 149, 199, y 249 variables, aplicando la mezcla de árboles con 1, 2, 4, y 6 árboles. Primeramente se comparará los resultados obtenidos entre las diferentes variables con las cuales se testeó el algoritmo.

La siguiente grafica muestra cómo se comporta el tiempo de ejecución para las diferentes variables.



Figura 6. Tiempo de Ejecución

Como se puede observar el tiempo de ejecución aumenta a medida que aumenta el número de árboles de

Capítulo 3. Resultados de la Investigación

la mezcla, al igual que con el aumento del número de variables, aunque el tiempo no debe de influir en los resultados arrojados por el algoritmo.

La siguiente gráfica muestra los resultados obtenidos en la generación de parada, debido a que la generación de parada como se mencionó anteriormente, es la cantidad de veces que el algoritmo tiene que regenerar la población para llegar a resultados satisfactorios, entonces la mejor generación de parada sería el que menor valor de generación de parada arroje.



Figura 7. Generación de Parada

En el gráfico anterior se puede observar que con 149 variables la generación de parada aumenta a medida que aumenta el número de árboles, significando esto que la mezcla de árboles no ha arrojado los resultados esperados, ya que tendría que arrojar valores más pequeños a medida que aumentara el número de árboles.

Con un número de variables de 199 se puede observar la mejora que se obtiene con 6 árboles con respecto a los resultados con 2 y 4 árboles, pero no supera los resultados de 1 árbol, lo que prueba que la mezcla de árboles no está arrojando mejores resultados. Tomando 249 variables se está teniendo mejores resultados utilizando 4 y 6 árboles en la mezcla, manifestándose un comportamiento similar que con 199

Capítulo 3. Resultados de la Investigación

variables. El experimento refleja que a medida que se aumenta el número de variables de la población y el número de árboles en la mezcla, se mejoran los resultados.

El número de evaluaciones o esfuerzo computacional es proporcional a la generación de parada debido a que el algoritmo que mejores resultados arroje es el que menor valor tenga, ya que el esfuerzo computacional se puede ver como el número de individuos que es necesario evaluar para resolver el problema.

A continuación se muestra una gráfica donde se compara los resultados obtenidos de cada variable.



Figura 8. Número de evaluaciones

Se puede observar que con 149 variables el número de evaluaciones aumenta a medida que aumenta el número de árboles, lo que demuestra que la mezcla de árboles para esta cantidad de variables no mejora los resultados.

Con un número de variables de 199 la mejora se obtiene cuando la mezcla de árboles tiene 6 árboles, pero no llega a ser mejor que con 1 árbol, y ya con 249 variables con 4 y 6 árboles se obtienen mejores resultados lo que prueba que la mezcla de árboles está arrojando mejores resultados con mayores valores

Capítulo 3. Resultados de la Investigación

de variables y número de árboles.

Como habíamos dicho anteriormente se hará una comparación con otros algoritmos para ver los resultados arrojados por el MT_FDA, se comparará con los algoritmos CHC, SA, ssGA, genGA, donde anteriormente se hizo una breve descripción de ellos. Solo compararemos con el parámetro “número de evaluaciones” por las razones antes expuestas.

La gráfica siguiente muestra los resultados obtenidos por el MT_FDA y comparado con los demás.



Figura 9. Resultado de los algoritmos en cuanto a número de evaluación

Capítulo 3. Resultados de la Investigación

Como se puede observar antes de comparar con el MT_FDA, de todos los algoritmos el CHC es el que arroja mejores resultados que los demás, por tanto el MT_FDA para lograr ser el mejor de todos debe de superar los resultados del CHC. Con el número de variables 149 el MT_FDA es mejor que el CHC por muy pocos valores de diferencia, y por lo tanto mejor que los demás también, es decir con 149 variables el MT_FDA es el algoritmo que arroja mejores resultados, no siendo así cuando aumentamos a número de variables 199 y 249 donde el CHC arroja los mejores resultados, aunque el MT_FDA tiene mejores resultados que el SA, ssGA, genGA, por lo que podemos llegar a la conclusión que con variables pequeñas el MT_FDA es mejor que los demás algoritmo, no siendo así a la hora que se incrementen las variables, donde el CHC sigue siendo el mejor de todos y el MT_FDA pasa a ocupar un segundo lugar.

5. Conclusiones

En este capítulo se mostraron los resultados del algoritmo MT_FDA, y se compararon con otros algoritmos.

CONCLUSIONES

En esta investigación se tuvo como principal objetivo demostrar la eficiencia del algoritmo MT_FDA, donde se hicieron algunas comparaciones con otros algoritmos utilizados anteriormente para la resolución del problema y se pudo llegar a las siguientes conclusiones:

- Las mezclas de árboles no influyeron a la hora de mejorar los resultados del algoritmo para el número de variables 149 y 199.
- Cuando se aumenta el número de variables, y la mezcla contiene varios árboles, entonces se ve un ligero mejoramiento de los resultados.
- Con variables pequeñas el algoritmo MT_FDA llega a ser mejor que los demás algoritmos comparados, pero cuando se aumentan las variables no mejora los resultados del CHC que es el que mejores resultados arroja, pero si los resultados de los demás algoritmo.
- El tiempo de ejecución aumenta considerablemente a medida que se aumenta el número de árboles de la mezcla y el número de variables, cosa que no influye en el resultado del algoritmo.

RECOMENDACIONES

- Incrementar el número de variables en el experimento hasta llegar a 349.
- Paralelizar el algoritmo MT_FDA debido al costo computacional que presenta el mismo.
- Desarrollar una aplicación informática la cual permita gestionar la ubicación de antenas obtenidas luego de la utilización del algoritmo MT_FDA.
- Testear la eficiencia del algoritmo MT_FDA en otros problemas complejos de optimización.

ANEXOS

Variable 149				
Poblacion 1400				
# Arboles 1				
# Corridas	Tiempo (Segundos)	Mejor Fitness	Generación de parada	Cantidad de Evaluaciones
1	143	203,839	19	26600
2	166	203,839	23	32200
3	166	203,839	23	32200
4	155	203,839	21	29400
5	140	203,839	19	26600
6	147	203,839	20	28000
7	141	203,839	19	26600
8	147	203,839	20	28000
9	148	203,839	20	28000
10	135	203,839	18	25200
11	216	203,839	19	26600
12	141	203,839	19	26600
13	142	203,839	19	26600
14	186	203,839	19	26600
15	147	203,839	20	28000
16	159	203,839	22	30800
17	148	203,839	20	28000
18	208	203,839	19	26600
19	148	203,839	20	28000
20	148	203,839	20	28000
21	140	203,839	19	26600
22	147	203,839	20	28000
23	156	203,839	21	29400
24	140	203,839	19	26600
25	162	203,839	22	30800
Medias	155,04	203,839	20	28000

Tabla 1. Resultados NV: 149; NA: 1

Variable 149 Población 1400 # Árboles 2				
# Corridas	Tiempo de Ejecución	Mejor Fitness	Generación de parada	Cantidad de Evaluaciones
1	694	203,839	20	28000
2	715	203,839	24	33600
3	636	203,839	20	28000
4	679	203,839	20	28000
5	595	203,839	19	26600
6	643	203,839	20	28000
7	588	203,839	19	26600
8	612	203,839	20	28000
9	632	203,839	19	26600
10	700	203,839	24	33600
11	767	203,839	26	36400
12	700	203,839	21	29400
13	685	203,839	21	29400
14	628	203,839	19	26600
15	653	203,839	19	26600
16	659	203,839	19	26600
17	742	203,839	22	30800
18	644	203,839	19	26600
19	685	203,839	20	28000
20	459	203,839	21	29400
21	857	203,839	19	26600
22	542	203,839	21	29400
23	465	203,839	20	28000
24	513	203,839	19	26600
25	626	203,839	22	30800
Medias	644,76	203,839	20,52	28728

Tabla 2. Resultados NV: 149; NA: 2

Variable 149				
Poblacion 1400				
# Arboles 4				
# Corridas	Tiempo de Ejecución	Mejor Fitness	Generación de parada	Cantidad de Evaluaciones
1	1051	203,839	19	26600
2	1007	203,839	20	28000
3	1054	203,839	20	28000
4	1153	203,839	21	29400
5	1124	203,839	21	29400
6	1202	203,839	19	26600
7	966	203,839	20	28000
8	1046	203,839	20	28000
9	1090	203,839	19	26600
10	1061	203,839	20	28000
11	1118	203,839	21	29400
12	1162	203,839	21	29400
13	1168	203,839	21	29400
14	1233	203,839	21	29400
15	1021	203,839	20	28000
16	1149	203,839	22	30800
17	1087	203,839	21	29400
18	1075	203,839	21	29400
19	1250	203,839	23	32200
20	1128	203,839	22	30800
21	1164	203,839	21	29400
22	1002	203,839	20	28000
23	1031	203,839	20	28000
24	1088	203,839	21	29400
25	1029	203,839	21	29400
Medias	1098,36	203,839	20,6	28840

Tabla 3. Resultados NV: 149; NA: 4

Variable 149 Poblacion 1400 # Arboles 6					
# Corridas	Tiempo de Ejecución	Mejor Fitness	Generación de parada	Cantidad de Evaluaciones	
1	1581	203,839	23	32200	
2	1188	203,839	20	28000	
3	1653	203,839	23	32200	
4	1373	203,839	22	30800	
5	1293	203,839	20	28000	
6	1238	203,839	20	28000	
7	1400	203,839	24	33600	
8	1372	203,839	21	29400	
9	1245	203,839	20	28000	
10	1370	203,839	21	29400	
11	1487	203,839	25	35000	
12	1386	203,839	23	32200	
13	1299	203,839	21	29400	
14	1326	203,839	21	29400	
15	1463	203,839	28	39200	
16	1229	203,839	20	28000	
17	1390	203,839	24	33600	
18	1161	203,839	20	28000	
19	1618	203,839	23	32200	
20	1187	203,839	20	28000	
21	1256	203,839	20	28000	
22	1011	203,839	20	28000	
23	1142	203,839	20	28000	
24	1283	203,839	20	28000	
25	1243	203,839	21	29400	
Medias	1327,76	203,839	21,6	30240	

Tabla 4. Resultados NV: 149; NA: 6

Variable 199 Poblacion: 3700 # Arboles 1				
# Corridas	Tiempo (Segundos)	Mejor Fitness	Generación de parada	Cantidad de Evaluaciones
1	538	203,839	23	85100
2	558	203,839	23	85100
3	538	203,839	23	85100
4	542	203,839	23	85100
5	511	203,839	21	77700
6	522	203,839	22	81400
7	538	203,839	23	85100
8	537	203,839	23	85100
9	539	203,839	23	85100
10	617	203,839	27	99900
11	537	203,839	23	85100
12	612	203,839	25	92500
13	617	203,839	23	85100
14	550	203,839	23	85100
15	629	203,839	23	85100
16	669	203,839	24	88800
17	623	203,839	25	92500
18	558	203,839	23	85100
19	564	203,839	23	85100
20	583	203,839	24	88800
21	554	203,839	21	77700
22	538	203,839	23	85100
23	602	203,839	24	88800
24	566	203,839	23	85100
25	588	203,839	24	88800
Medias	569,2	203,839	23,28	86136

Tabla 5. Resultados NV: 199; NA: 1

Variable 199 Poblacion 5000 # Arboles 2				
# Corridas	Tiempo (Segundos)	Mejor Fitness	Generación de parada	Cantidad de Evaluaciones
1	4670	203,839	25	125000
2	3898	203,839	23	115000
3	3873	203,839	24	120000
4	3463	203,839	22	110000
5	3704	203,839	22	110000
6	4504	203,839	23	115000
7	3742	203,839	23	115000
8	4311	203,839	25	125000
9	4163	203,839	24	120000
10	3794	203,839	23	115000
11	3826	203,839	22	110000
12	3814	203,839	23	115000
13	3287	203,839	23	115000
14	3596	203,839	23	115000
15	3624	203,839	24	120000
16	3820	203,839	22	110000
17	3879	203,839	23	115000
18	4171	203,839	22	110000
19	4336	203,839	22	110000
20	5098	203,839	27	135000
21	3864	203,839	25	125000
22	4804	203,839	24	120000
23	3099	203,839	23	115000
24	3617	203,839	24	120000
25	3183	203,839	24	120000
Medias	3925,6	203,839	23,4	117000

Tabla 6. Resultados NV: 199; NA: 2

Variable 199 Poblacion 5000 # Arboles 4				
# Corridas	Tiempo (Segundos)	Mejor Fitness	Generación de parada	Cantidad de Evaluaciones
1	9034	203,839	24	120000
2	8548	203,839	23	115000
3	8584	203,839	23	115000
4	8445	203,839	23	115000
5	8093	203,839	22	110000
6	8101	203,839	22	110000
7	8122	203,839	22	110000
8	8422	203,839	23	115000
9	8599	203,839	23	115000
10	8159	203,839	22	110000
11	10511	203,839	23	115000
12	8398	203,839	23	115000
13	8769	203,839	24	120000
14	8487	203,839	23	115000
15	10209	203,839	40	200000
16	8204	203,839	22	110000
17	8465	203,839	23	115000
18	8001	203,839	22	110000
19	8367	203,839	23	115000
20	8227	203,839	23	115000
21	8893	203,839	24	120000
22	8557	203,839	23	115000
23	9063	203,839	24	120000
24	7847	203,839	21	105000
25	8580	203,839	23	115000
Medias	8587,4	203,839	23,52	117600

Tabla 7. Resultados NV: 199; NA: 4

Variable 199 Poblacion 5000 # Arboles 6				
# Corridas	Tiempo (Segundos)	Mejor Fitness	Generación de parada	Cantidad de Evaluaciones
1	14599	203,839	23	115000
2	14649	203,839	23	115000
3	16413	203,839	24	120000
4	14612	203,839	23	115000
5	14508	203,839	23	115000
6	13886	203,839	22	110000
7	14460	203,839	23	115000
8	18638	203,839	23	115000
9	31486	203,839	25	125000
10	14948	203,839	23	115000
11	33620	203,839	23	115000
12	15608	203,839	23	115000
13	15085	203,839	23	115000
14	14689	203,839	22	110000
15	15284	203,839	24	120000
16	14568	203,839	23	115000
17	15194	203,839	23	115000
18	15248	203,839	24	120000
19	16020	203,839	25	125000
20	15609	203,839	24	120000
21	15648	203,839	23	115000
22	14469	203,839	23	115000
23	14497	203,839	23	115000
24	14726	203,839	23	115000
25	16005	203,839	25	125000
Medias	16578,76	203,839	23,32	116600

Tabla 8. Resultados NV: 199; NA: 6

Variable 249					
Poblacion 12000					
# Arboles 1					
# Corridas	Tiempo (Segundos)	Mejor Fitness	Generación de parada	Cantidad de Evaluaciones	
1	2304	203,839	27	324000	
2	2227	203,839	26	312000	
3	2219	203,839	26	312000	
4	2147	203,839	25	300000	
5	2194	203,839	25	300000	
6	2215	203,839	26	312000	
7	2411	203,839	28	336000	
8	2261	203,839	26	312000	
9	2140	203,839	25	300000	
10	2315	203,839	27	324000	
11	2527	203,839	31	372000	
12	2061	203,839	24	288000	
13	2327	203,839	27	324000	
14	2151	203,839	25	300000	
15	2172	203,839	25	300000	
16	2315	203,839	27	324000	
17	2237	203,839	26	312000	
18	2381	203,839	28	336000	
19	2228	203,839	26	312000	
20	2421	203,839	29	348000	
21	2381	203,839	28	336000	
22	2172	203,839	25	300000	
23	2333	203,839	27	324000	
24	2177	203,839	25	300000	
25	2239	203,839	26	312000	
Medias	2262,2	203,839	26,4	316800	

Tabla 9. Resultados: NV: 249; NA: 1

Variable 249				
Poblacion 12000				
# Arboles 2				
# Corridas	Tiempo (Segundos)	Mejor Fitness	Generación de parada	Cantidad de Evaluaciones
1	15276	203,839	30	360000
2	11067	203,839	25	300000
3	11544	203,839	29	348000
4	13549	203,839	25	300000
5	11145	203,839	27	324000
6	11986	203,839	25	300000
7	14800	203,839	28	336000
8	13185	203,839	29	348000
9	14180	203,839	28	336000
10	14537	203,839	26	312000
11	12971	203,839	27	324000
12	16038	203,839	28	336000
13	14267	203,839	27	324000
14	13640	203,839	29	348000
15	12256	203,839	27	324000
16	10885	203,839	27	324000
17	11632	203,839	28	336000
18	14450	203,839	27	324000
19	11000	203,839	26	312000
20	11931	203,839	28	336000
21	11132	203,839	27	324000
22	10589	203,839	25	300000
23	13908	203,839	26	312000
24	11431	203,839	26	312000
25	11972	203,839	27	324000
Medias	12774,84	203,839	27,08	324960

Tabla 10. Resultados NV: 249; NA: 2

Variable 249				
Poblacion 12000				
# Arboles 4				
# Corridas	Tiempo (Segundos)	Mejor Fitness	Generación de parada	Cantidad de Evaluaciones
1	36412	203,839	25	300000
2	43113	203,839	27	324000
3	42826	203,839	27	324000
4	44979	203,839	27	324000
5	47996	203,839	25	300000
6	36566	203,839	26	312000
7	63520	203,839	26	312000
8	34110	203,839	26	312000
9	67550	203,839	27	324000
10	39258	203,839	27	324000
11	38472	203,839	25	300000
12	40293	203,839	28	336000
13	66206	203,839	25	300000
14	38229	203,839	25	300000
15	39582	203,839	26	312000
16	41161	203,839	28	336000
17	38032	203,839	25	300000
18	65172	203,839	26	312000
19	35738	203,839	24	288000
20	40915	203,839	27	324000
21	70527	203,839	27	324000
22	32681	203,839	24	288000
23	39640	203,839	26	312000
24	34404	203,839	25	300000
25	35575	203,839	26	312000
Medias	44518,28	203,839	26	312000

Tabla 11. Resultados NV: 249; NA: 4

Variable 249 Poblacion 12000 # Arboles 6				
# Corridas	Tiempo (Segundos)	Mejor Fitness	Generación de parada	Cantidad de Evaluaciones
1	53359	203.839	27	324000
2	53295	203.839	25	300000
3	53576	203.839	27	324000
4	53361	203.839	26	312000
5	60176	203.839	26	312000
6	56170	203.839	25	300000
7	54054	203.839	27	324000
8	49261	203.839	25	300000
9	50813	203.839	26	312000
10	52692	203.839	27	324000
11	57781	203.839	27	324000
12	52985	203.839	27	324000
13	50963	203.839	26	312000
14	57670	203.839	25	300000
15	61546	203.839	26	312000
16	55729	203.839	25	300000
17	51863	203.839	26	312000
18	40261	203.839	25	300000
19	40361	203.839	25	300000
20	51261	203.839	26	312000
21	30271	203.839	25	300000
22	61261	203.839	27	324000
23	40261	203.839	25	300000
24	10561	203.839	24	288000
25	61761	203.839	27	324000
Medias	50451.68	203.839	25.88	310560

Tabla 12. Resultados NV: 249; NA: 6

REFERENCIAS BIBLIOGRÁFICAS

1. **Rich, Elaine.** *Introducción a la Inteligencia Artificial*. 1ra.
2. **Valls, Maykel Diaz.** *Algoritmo Evolutivo con estimación de distribuciones paralelo basado en el aprendizaje de poliárboles*. Camaguey : s.n., 2005.
3. **Hermidas, Roberto Santana.** *Modelacion probabilisticas basada en modelos graficos no dirigidos en EDAs*. 2004.
4. [En línea] [Citado el: 28 de 01 de 2009.] www2.ing.puc.cl.
5. [En línea] [Citado el: 12 de 02 de 2009.] www.upo.es/RevMetCuant/art4.pdf.
6. Taller de Algoritmos Evolutivos. [En línea] sci2s.ugr.es/seminars/5/taller_ae_documentacion__lunes13.pdf.
7. **Puerta, P.Larragaña y Jose M.** [En línea] 05 de 12 de 2003. [Citado el: 16 de 02 de 2009.] heur.uv.es/metodos/EDA.pdf.
8. **D.Hand, B.Everitt y.** *Mixture Models: Inference and Applications to Clustering*. London : s.n., 1981.
9. **Meila, M.** *Learning Mixtures of Tress*. Massachusetts Institute of Technology. 1999.
10. **D.Peel, G. Mclachlan y.** *Finite Mixture Models*. 2000.
11. **A.P.Dempster, N.M.Laird y D.B.Rubin.** *Maximum likelihood from incomplete data via the EM algorithm*. Journal of the Royal Statistical Society. 1977. págs. 1-38.
12. **G.Celeux, D.Chauveau y J.Diebolt.** *On stochastic versions of the EM algorithm*. 1995. tecnico.
13. **P.j.Green, S.Richardson y.** *On Bayesian analysis of mixtures with an unknown number of component*. Journal of the Royal Statistical Society. 1997. págs. 731-792.
14. **P.Green.** *Reversible jump markov chain Monte Carlo computation and Bayesian model determination*. 1995. págs. 711-732.

Referencias Bibliográficas

15. **B.Thiesson, C.Meek, D.Chickering y D.Hecherman.** *learning mixtures of bayesian networks.* advance twchnology division, Microsoft Research. 1997. tecnico.
16. **D.Forsyth, S.Ioffe y.** *Human tracking with mixtures of trees.* 2001. págs. 690-695.
17. *finding optima by estimating probability densities.* **J.S.Debonet, C.Isbell y P.Viola.** [ed.] M.Jordan y T.Peetsche M.Mozer. Cambridge : s.n., 1997, Advances in neural information processing systems, Vol. 9, pág. 424.
18. *The Bivariate Marginal Distribution Algorithm.* **H.Muhlenbein, M.Pelikan y.** [ed.] T.Furuhashi y P.Chawdhry R.Roy. London : s.n., 1999, Advances in soft computing-Engineering desing and manufacturing , págs. 521-535.
19. **S.E.Fienberg.** *The analysis of cross-classified categorical data.* 1981.
20. **G.Harik.** *Linkage learning via probabilistic modeling in the ECGA.* University of Illinois at urbana-champaign. 1999.
21. **Enrique ALba, Guillermo Molina y Antonio Nebro.** Disposicion Optima de antenas usando CHC. [En línea] [Citado el: 1 de 03 de 2009.] neo.lcc.uma.es/staff/guillermo/index_files/files/RND_con_CHC_multiobjetivo.pdf.
22. convocatoria de ayudas de proyectos de investigacion cientifica y tecnologica. 2002.
23. **Enrique Alba, Guillermo Molina y Antonio J.Nebro.** *Dispocision optima de antenas usando CHC multiobjetivo.*
24. **Calegari, P.Guidic, F.Kuonen, P.Kobler, D.Parallel.** *genetic algorithm for radio network design.* Joirnal of Parallel and Distributed Computing. 1997.
25. [En línea] [Citado el: 13 de 03 de 2009.] oplink.lcc.uma.es/problems/rnd.html.

BIBLIOGRAFÍA

1. **Pacheco, Marco A. Cavalcanti.** *Algoritmos Genéticos: Principios y Aplicaciones.* Departamento de ingeniería eléctrica, Universidad Católica de Rio de Janeiro. Brasil : s.n. Documento.
2. **E. Alba, C. Cotta, F. Chicano, and A.J. Nebro.** *Parallel Evolutionary Algorithms in Telecommunications: Two Case Studies.* lenguajes y ciencias de la computación, universidad de malaga.
3. **Elaine, Rich.** *Inteligencia Artificial.* 1ra.
4. **Bello, Rafael.** *Modelos Computacionales Avanzado.*
5. **Alba, Enrique.** *Parallel evolutionary algorithms can achieve superlinear performance.* Information Processing Letters. 2002. págs. 7-13.
6. **Goldberg, D.** *Genetic Algorithms in Search.* 1989.
7. **Pacheco, Marco Aurélio.** Notas de Aula em Computação Evolucionária. [En línea] www.ica.ele.pucrio.pucrio.br.
8. **P. Larrañaga, José M. Puerta.** *Algoritmos de Estimación de Distribución.* 2003.
9. **Julio Madera Quintana, Lorver Duarte Puig.** *Biblioteca para el rápido desarrollo de algoritmos evolutivos que estiman distribuciones.* Departamento de Computación, Facultad de Informática, Universidad de Camagüey. Camaguey : s.n.
10. *Algoritmos Evolutivos.* Centro de Cálculo, Instituto de Computación, Facultad de Ingeniería, Universidad de la República. Uruguay : s.n., 2007.
11. **Torres, Enrique Alba.** *Convocatoria de ayudas de Proyectos de Investigación Científica y Desarrollo.* 2002.
12. *Estado del Arte. Investigación en Navegación Robótica Móvil Cooperativa.*
13. **Holland, J.** *Adaptation in Natural and Artificial Systems.* 1995.
14. **Ansel Y. Rodríguez González, Adanay Martín Pérez.** *Un estudio sobre los Algoritmos con Estimación de las Distribuciones basados en modelos simplemente conectados.*
15. **Valero, Francisco Luna.** *Metaheurísticas avanzadas para problemas reales en redes de telecomunicaciones.* Lenguajes y Ciencias de la Computación, UNIVERSIDAD DE MÁLAGA. Malaga : s.n., 2008.
16. **Alba Enrique, F.Chicano.** *On the Behavior of Parallel Genetic Algorithms for Optimal Placement of Antennae in Telecommunications.* Departamento de Lenguajes y Ciencias de la Computación, University of Málaga. 2004.

17. **Abdelmalik Moujahid, Iñaki Inza y Pedro Larrañaga.** *Algoritmos de Estimación de Distribuciones.* Departamento de Ciencias de la Computación e Inteligencia Artificial, Universidad del País Vasco.
18. **Enrique Alba, Guillermo Molina, Francisco Chicano.** *Optimal Placement of Antennae using Metaheuristics.* Departamento de Lenguajes y Ciencias de la Computación, University of Málaga.
19. **E. Alba, J. F. Chicano, B. Dorronsoro, G. Luque.** *Diseño de Códigos Correctores de Errores con Algoritmos Genéticos.*
20. **Sergio Nesmachnow, Héctor Cancela, Enrique Alba.** *Técnicas evolutivas aplicadas al diseño de redes de comunicaciones confiables.*
21. *Una introducción a la computación evolutiva y algunas de sus aplicaciones en Economía y Finanzas.* **Luis Vicente Santana Quintero, Carlos A. Coello Coello.** diciembre de 2006, Revista de métodos cuantitativos para la economía y la empresa, págs. 3-26.
22. *Programación Genética.* **Aranda, Jorge Baier.**
23. *Practica 1. Diseño Redes de Radio RND.* universidad carlos III . Madrid : s.n., 2008.
24. *Redes Bayesianas.*
25. **Andrés Cano Utrera, Javier García Castellano, Andrés R. Masegosa Arredondo.** *Métodos para Determinar el Atributo Distinguido en Multiredes Bayesianas.* Department of Computer Science and Artificial Intelligence, Granada University . españa : s.n.
26. **Reina, José L. Ruiz.** *Introducción a las redes bayesianas.* Dpto. de Ciencias de la Computación e Inteligencia Artificial, Universidad de Sevilla. 2006.
27. **Enrique Alba, Guillermo Molina y Antonio J. Nebro.** *Disposición óptima de antenas usando CHC multiobjetivo.*
28. *TALLER DE ALGORITMOS EVOLUTIVOS.*
29. **Informática, Facultad de.** *Algoritmos Evolutivos con Estimación de Distribuciones Basados en Pruebas de Independencia.* Facultad de Informática, Universidad de Camagüey. Tesis de maestría.
30. **Hermida, Roberto Santana.** *Modelación probabilística basada en modelos gráficos no dirigidos en Algoritmos Evolutivos con Estimación de Distribuciones.* Instituto de Cibernética, Matemática y Física. Ciudad Habana : s.n., 2004. tesis doctorado.
31. **Valls, Maykel Díaz.** *Algoritmo evolutivo con estimación de distribución paralelo basado en el aprendizaje de poliárboles.* Facultad de Informática, Universidad de Camagüey. Camagüey : s.n., 2005. Tesis de Pregrado.

32. **Jose A. Lozano, Pedro Larrañaga.** *A Gentle Introduction to Estimation of Distribution Algorithms.* Intelligent Systems Group, University of the Basque Country . España : s.n., 2005.
33. **MEILA., M.** *Learning Mixtures of Trees.* Massachusetts Institute of Technology. 1999. Tesis Doctora.
34. **M. MEILA, M. JORDAN.** *Learning with mixtures of trees.* Journal of Machine Learning.
35. **BALUJA., S.** *Population-based incremental learning: A method for integrating genetic search based function optimization and competitive learning.* Carnegie Mellon University. 1994. págs. 94-163.
36. wikipedia. [En línea]
37. **THIERENS, P. A. BOSMAN Y D.** *Ideas based on the normal kernels probability density function.* Utrecht University. 2002. informe tecnico.
38. **HECKERMAN, D. GEIGER Y D.** *Knowledge representation and inference in similarity networks and Bayesian multinets.* Artificial Intelligence. 1996. págs. 45-72.
39. *Dynamic Bayesian multinets Uncertainty.* **Uncertainty, J. BILMES.** 2000, Proceedings of the 16th conference on Uncertainty, págs. 38-45.
40. **D. F. BROWN, A. GARMENDIA-DOAL Y J. A. W. MCCALL.** *Markov random field modelling of royal road genetic algorithms.* 2002. págs. 65-76.
41. **GOLDBERG, K. DEB Y D. E.** *Sufficient conditions for deceptive and easy binary functions.* Annals of Mathematics and Artificial Intelligence. 1994. págs. 385-408.
42. **LAURITZEN.(1996).** *Graphical Models.* Oxford:Clarendon Press .
43. **MÜHLENBEIN, M. PELIKAN Y H.** *The Bivariate Marginal Distribution Algorithm.* london : s.n., 1999. págs. 521–535.
44. **S.E.FIENBERG.** *The analysis of cross-classified categorical data.* 1981.