

## CAPITULO 2: MARCO TEÓRICO

En el desarrollo de este capítulo se presentan descripciones generales, definiciones y métodos, que nos pueden ayudar a entender con claridad el método que desarrolló en esta tesis. También incluye una explicación detallada de lo que se hace en Algoritmos Genéticos, esto para facilitar la explicación del funcionamiento de SSPGA catalogado como Algoritmo Evolutivo.

### 2.1 Problemas de Ruta más Corta

El problema de la ruta más corta surge cuando en algún proceso es necesario enviar algún material, vehículo, señal, etc. entre dos puntos dentro de una red lo más rápido, económico o rentable posible.

En el problema de la ruta más corta se cuenta con una red dirigida (que no tiene ciclos, los arcos son unidireccionales), es decir, se presentan una serie de nodos, y las conexiones entre ellos (arcos); éstas conexiones tienen un costo establecido el cual puede representar tiempo, distancia, etc. Si tomamos cualquier ruta para llegar de un nodo a otro consideraremos el costo de la ruta como la suma de los costos de cada uno de los arcos que se utilizan para completar la ruta. Este problema consiste en determinar la ruta con el menor costo posible para llegar de un nodo a otro.

Para explicar más a fondo este problema, es primero necesario familiarizarnos con la notación correspondiente, así como con los supuestos necesarios para el caso.

Consideremos una gráfica dirigida  $G = (N, A)$  con una longitud entre nodos (Costo de arco)  $c_{ij}$  asociado con el arco  $(i, j) \in A$ . Esta red debe tener un nodo  $s$ ,

el cual denominaremos como *fuentes*. Además  $A(i)$  representa la lista de arcos adyacentes al nodo  $i$ , es decir, los nodos con los que está conectado el nodo  $i$ . También definiremos la *longitud de una ruta definida* como la suma de los costos de los arcos en la ruta. Bajo este esquema, el problema de la ruta más corta consiste en determinar para cada nodo  $i \in N, i \neq s$  la ruta dirigida de menor costo de el nodo  $s$  al nodo  $i$ .

Desde otro enfoque, se puede ver como el problema de mandar una unidad de flujo de la manera más económica posible de cada nodo  $s$  hacia cada uno de los nodos  $N - \{s\}$  en una red “incapacitada” (No restringida en capacidad). Este punto de vista nos permite formular el problema de la ruta más corta como un problema de programación lineal [1].

$$\text{Min} \quad \sum_{(i,j) \in A} c_{ij} x_{ij}$$

Con las siguientes restricciones:

$$\sum_{\{j:(i,j) \in A\}} x_{ij} - \sum_{\{j:(j,i) \in A\}} x_{ij} = \begin{cases} -1 \\ 0 \\ 1 \end{cases}$$

$$x_{ij} \geq 0 \text{ para todo } (i,j) \in A$$

Existen muchos tipos de redes que, por sus características, necesitan diferentes métodos para ser resueltas. Para esta tesis en particular se tomaron en cuenta solo las redes que cumplan con las siguientes características:

1. La red no debe de contener ciclos.
2. Sólo un arco por cada par de nodos en la red
3. Debe de ser una red dirigida

## 2.2 Algoritmos Genéticos

El avance tecnológico y las nuevas técnicas para resolver problemas de optimización han permitido incorporar más variables a los problemas clásicos de optimización. Con estos avances, un nuevo modelo puede contener mucha más información, y representar con mayor precisión la realidad [2].

Con esta incorporación de variables, los problemas que se intentan resolver en el área de Investigación de Operaciones se vuelven más complejos y surge la necesidad de desarrollar nuevos métodos y técnicas eficaces para encontrar mejores soluciones.

Antes de comenzar a analizar la manera de resolver el problema de la ruta más corta con Algoritmos Genéticos y Evolutivos, es conveniente conocer el esquema general de los primeros para después aplicarlo al problema que se intenta resolver. Cuando se tienen problemas con espacios de soluciones de gran tamaño, y no es posible utilizar enumeración exhaustiva (analizar una por una todas las soluciones posibles) como método de búsqueda, es necesario buscar maneras alternativas de búsqueda para una buena solución (aunque no sea la mejor). A este tipo de métodos de búsqueda se les denomina heurísticas y dentro de estas podemos encontrar Algoritmos Genéticos.

El uso de Algoritmos Genéticos es una técnica relativamente nueva. La primera vez que se habló del uso de este método en búsquedas matemáticas fue en 1975 en el libro *Adaptation in Natural and Artificial Systems* [3]. Sin embargo, fue hasta 1985 cuando por primera vez se utilizó ésta técnica en la optimización: *Genetic Algorithms in Search and Machine Learning* [4].

Basados en la teoría de la selección natural, los Algoritmos Genéticos son métodos de búsqueda (Heurística) en los cuales, a partir de un conjunto de soluciones, se van eliminando las más pobres para ser sustituidas por nuevas, mientras que las mejores soluciones se mantienen y se combinan para mejorarse y llegar a una solución final.

En general, los problemas de optimización tienen una función que se quiere minimizar o maximizar. Cada vez que se obtiene una solución cualquiera, esta tiene un valor de dicha función, con este valor se pueden comparar las soluciones y determinar cual es mejor. A esta función se le llama función objetivo, a su valor lo denominaremos rendimiento (Costo en el caso del problema de la ruta más corta).

A grandes rasgos, un algoritmo genético hace lo siguiente: genera una población inicial de soluciones, calcula su rendimiento, selecciona padres y reproduce soluciones utilizando información de los padres este proceso se repite hasta que se cumpla una condición de terminación previamente establecida.

### 2.2.1 Población inicial

Las soluciones del problema son denominadas cromosomas, que se componen de genes. La población es un conjunto de estos cromosomas, cada uno con su respectivo rendimiento. Existen varias maneras de representar cromosomas, la más común es la

representación binaria en la cual se representa la solución como un vector en el cual cada gen se representa con ceros y unos.

El primer paso en un Algoritmo Genético consiste en generar una población inicial de cromosomas. Usualmente esto se hace generando cromosomas aleatoriamente, pero dependiendo del problema, también se pueden utilizar otros métodos para obtener esta solución inicial.

El tamaño de esta población depende de varios factores, uno de estos es el tipo de representación que se tiene para los cromosomas. Usualmente, con la representación binaria, las poblaciones suelen ser grandes mientras que con otro tipo de representaciones se reduce el tamaño de la población.

El método de selección de la población inicial, para la mayoría de los problemas, pasa a segundo término ya que lo que se debe de llevar a una buena solución es la estructura y la calidad del algoritmo, y no la solución de la que parte. Muchas veces se utilizan métodos para generar soluciones iniciales con la finalidad de reducir el tiempo de búsqueda del algoritmo, ya que se parte de una población con buenos cromosomas [5].

### 2.2.2 Selección de los padres

Una vez obtenida la población inicial, es necesario seleccionar dos cromosomas considerados “padres”, estos se combinan para generar nuevos cromosomas que, dependiendo de su rendimiento, pueden sustituir a otros cromosomas de la población.

El proceso de selección de los padres se puede realizar de varias maneras. Puede ser totalmente aleatorio, es decir, se escogen al azar dos padres de la población. Sin embargo, también se puede realizar este proceso dando preferencia a los mejores

cromosomas, entonces, las soluciones que presenten mejor valor de la función objetivo tienen mayor probabilidad de ser seleccionadas como padres. Otra manera de realizar esta selección es cortar la población en dos partes utilizando un criterio previamente determinado (usualmente se utiliza un porcentaje de la población), de esta manera se utiliza una sección de este corte como padres para las siguientes generaciones [5].

### 2.2.3 Reproducción de cromosomas

Una vez que se tienen seleccionados los padres es necesario combinarlos para generar nuevos cromosomas que nos sirvan para ir mejorando los rendimientos y a partir de ahí llegar a una solución final.

Una manera de hacer esta reproducción es tomar información de cada uno de los padres y combinarla, es decir, se seleccionan algunos genes de los padres y se mezclan para crear un nuevo cromosoma. A este proceso se le determina recombinación (crossover).

También se pueden reproducir los cromosomas mediante un proceso llamado mutación, por medio del cual se toma un cromosoma, y se modifican algunos de sus genes (De acuerdo a un porcentaje llamado porcentaje de mutación) para producir un nuevo cromosoma.

Dependiendo de cada las características del problema que se quiere resolver se pueden utilizar diferentes métodos para la reproducción de cromosomas, incluso se pueden combinar métodos. En algunas ocasiones parte de la población es reproducida por medio de recombinación y el resto por medio de mutación [5].

### 2.2.4 Esquema General

Una vez que se generó la población inicial, se seleccionaron los padres y se reprodujo la población, el proceso se repite hasta que se cumple un criterio previamente establecido que puede variar dependiendo del problema, pero usualmente el procedimiento se realiza un determinado número de veces denominadas Generaciones. En resumen (Diagrama de flujo en Figura 2.1):

1. Se crea una población inicial (Existen muchos métodos, pero usualmente se hace aleatoriamente).
2. Se buscan las mejores soluciones de acuerdo con criterios que dependen del problema que se quiere resolver o del método que se está utilizando
3. Se aplican los métodos de mutación y recombinación para obtener una nueva población en la búsqueda de mejores soluciones.
4. Se decide si ya se cumplió el criterio, previamente establecido, para terminar el algoritmo. Si no se ha cumplido, repetimos desde el paso 2, de lo contrario termina el algoritmo.

La idea en la cual se basan los algoritmos genéticos es la selección natural, las mejores soluciones “sobreviven” a lo largo de las generaciones y las peores van desapareciendo y son sustituidas por nuevas.

La parte más representativa de un Algoritmo Genético es la recombinación (reproducción de cromosomas), la idea principal de esta parte es que se obtenga información de dos o más cromosomas para crear nuevos y mejores.

Existen diferentes maneras de realizar cada uno de los pasos representados en la figura 2.1, dependiendo del problema y su complejidad un algoritmo genético realiza

los pasos de diferentes maneras. Más adelante se mencionara una manera de aplicar este método en el problema de la ruta más corta [5].

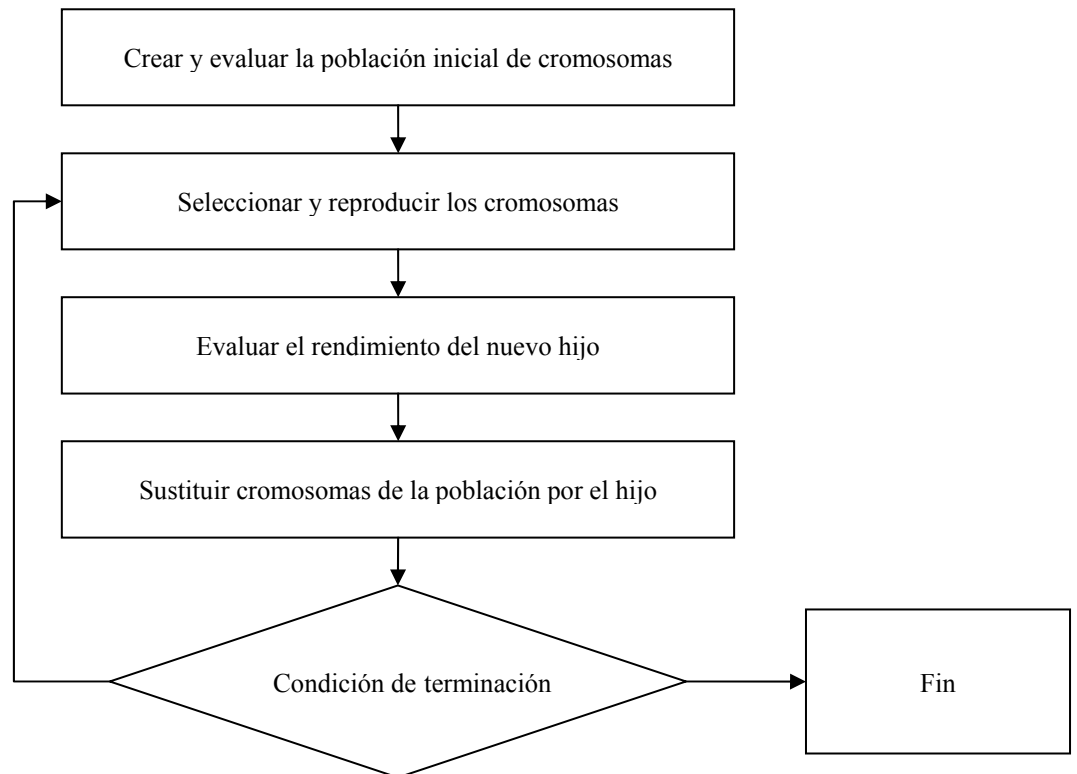


Figura 2.1 Diagrama de flujo para Algoritmos Genéticos

Para el caso del problema de la Ruta más Corta, la aplicación de Algoritmos Genéticos es una de las maneras más eficaces cuando se quiere buscar la Ruta más Corta en una red de gran tamaño ya que puede llegar a soluciones muy cercanas al óptimo en un tiempo relativamente pequeño. Otros métodos pueden llegar a la solución óptima pero el tiempo de búsqueda es considerablemente mayor.

Como se mencionaba anteriormente, poco a poco es más fácil adaptar los modelos tradicionales a la realidad, pero a medida que esto ocurre, también es necesario adaptar los métodos tradicionales a las nuevas necesidades o, en su caso, construir



nuevos. Como consecuencia de esta adaptación surge el método SSPGA para resolver problemas de redes.

Los Algoritmos Evolutivos son considerados una variación de Algoritmos Genéticos, los cuales presentan la misma idea que los primeros, sin embargo hay diferencias en la metodología que se explicaran posteriormente. Por lo anterior, sería útil conocer la representación del problema de la ruta más corta con Algoritmos Genéticos Puros, para después comprender correctamente su aplicación con Evolutivos.

La base de esta heurística son los cromosomas (Soluciones), que a su vez se componen de genes (nodos), los cuales se combinan para renovarse y mejorarse. En el caso particular del Problema de la Ruta más Corta los cromosomas son rutas, de tal manera que un cromosoma (ruta) puede ser representado por un vector de tamaño igual al número de nodos de la red, y se representa con 1 ó 0 dependiendo si el nodo está presente o no en la ruta, así cualquier ruta puede ser representada por un vector del mismo tamaño. Así mismo se cuenta con una función  $f(t)$  la cual determina el costo de un cromosoma en particular

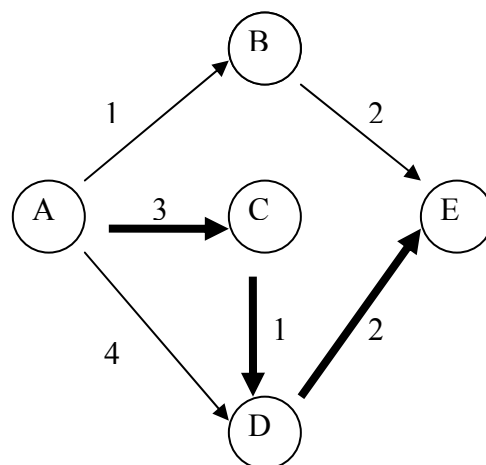


Figura 2.2 Ejemplo de Red: Ruta ACDE

Supongamos que tenemos la red representada en la figura 3.1. De esta obtenemos una la ruta ACDE (representada con líneas oscuras en la figura anterior).

Este cromosoma se puede representar con el siguiente vector:

1	0	1	1	1
A	B	C	D	E

Bajo este esquema, cada una de las combinaciones con las cuales se puede crear este vector es una ruta diferente (sea o no factible en la red). Con este vector, y con los costos de cada uno de los arcos podemos calcular el resultado de la función  $f(t)$  la cual nos determinará el valor total de la ruta, es decir, la suma de los costos de los arcos por los que pasa la ruta, en este caso:

$$f(t) = c_{AC} + c_{CD} + c_{DE} = 3 + 1 + 2 = 6$$

Previamente se definió la representación vectorial de un cromosoma en Algoritmos Genéticos para el problema de la ruta más corta. Es común que, bajo dicha representación, un vector pueda hacer referencia a una ruta no factible en la red. Por esta razón en algunos algoritmos se requiere verificar la factibilidad de cada cromosoma creado.

### 2.3 El Método SSPGA

Hasta ahora solo hemos discutidos redes con costos constantes en los arcos, sin embargo, esto puede causar complicaciones cuando se requiere llevar el problema a la realidad, donde los costos no necesariamente tienen que ser constantes. Estos costos pueden ser constantes e incluso variables aleatorias con distribuciones de probabilidad. Al hacer esto, los métodos tradicionales para resolver el problema de redes dejan de ser funcionales, es ahí donde se puede aplicar el método SSPGA (Stochastic Shortest Path Genetic Algorithm).

Desarrollado en 2001 por el Dr. Miguel Ángel Gómez, el método SSPGA tiene dos innovaciones importantes. La primera, y la más relevante, es la introducción de distribuciones de probabilidad en los costos de los Arcos. La segunda, es el uso de estructuras dinámicas para la codificación de la red. Anteriormente, para representar una ruta (cromosoma) se utilizaba un vector de tamaño definido, ahora con este método, el vector se reduce simplemente al número de nodos que utiliza la ruta. Esto hace más fácil el costeo y más eficiente el tiempo de búsqueda. Esta nueva representación vectorial soluciona el problema de la factibilidad del cromosoma, ya que no se puede crear un vector que represente a una ruta no factible.

Hay que recordar que aunque su nombre lo indica, SSPGA no es un algoritmo genético puro, sino evolutivo. Por esta razón el método de SSPGA que a partir del capítulo siguiente llamaremos SSPEA (Stochastic Shortest Path Evolutionary Algorithm) no sigue el esquema que se describió cuando hablamos de Algoritmos Genéticos.

Si bien existen muchas similitudes cuando se comparan ambos métodos, algoritmos Evolutivos es mucho más flexible que Genéticos en su estructura, en el caso

de SSPEA, no se considera un Algoritmo Genético por el hecho de que, como se explicará más adelante, para crear nuevos cromosomas solo utiliza la información de un solo padre que se recombina a sí mismo. En otras palabras, en SSPEA no existe la recombinación característico de Algoritmos Genéticos [2].