

**INSTITUTO TECNOLÓGICO Y DE ESTUDIOS SUPERIORES DE MONTERREY
CAMPUS ESTADO DE MÉXICO**

2000

PROGRAMA DE GRADUADOS EN INGENIERÍA Y CIENCIAS

MAESTRÍA EN CIENCIAS COMPUTACIONALES

ESPECIALIDAD EN REDES COMPUTACIONALES

BIBLIOTECA



TESIS

**ALGORITMOS GENÉTICOS APLICADOS AL DISEÑO
DE REDES DE TELECOMUNICACIONES**

Aspirante: Manuel Jaime Lima Romero

Asesor: Dr. Edgar E. Vallejo Clemente

**Comité de Tesis: Dr. Carlos Rodríguez Lucatero
Dr. Roberto Gómez Cárdenas**

Matrícula: 471000
Especialidad y plan: MCC '95
Campus: Estado de México

Atizapán de Zaragoza, Edo. Méx., Diciembre de 2000.

RESUMEN.

El diseño de redes es materia de estudio formal por parte de la investigación de operaciones, aunque cualquier otra área que haga uso de redes intentará dar solución al problema de obtener la red óptima para resolver algún problema específico.

Existen procedimientos formales para encontrar solución a los problemas de minimización de costos o de longitud de enlaces en una red. Estos procedimientos muestran que los árboles de mínima expansión, o de mínima cobertura, son parte del espacio de soluciones y entre ellos se encuentra la solución óptima al problema de minimización.

El uso de algoritmos genéticos (AG) es una técnica que ayuda a aproximarse a las soluciones óptimas cuando alguna o varias de las siguientes condiciones está presente:

1. La definición del problema lo hace intratable para usar procedimientos de búsqueda que operen en tiempo polinomial.
2. No se necesita la solución exacta sino que es suficiente una buena aproximación al máximo global.
3. No se conoce con detalle el espacio de búsqueda.
4. No tiene una solución unimodal.

Este trabajo tuvo como principio, la intención de desarrollar un procedimiento de optimización de la topología de una red de telecomunicaciones, que es mejor conocida como una red de área amplia, *WAN*.

El propósito de usar los AG en la búsqueda de soluciones parte de la necesidad de obtener las mejores propuestas para el diseño de una red *WAN*. Conocidas las características de capacidad de los enlaces que conectarán a los nodos de la red, el tráfico entre los usuarios, y la pertenencia de cada usuario a uno de los nodos definidos en la red, es lógico suponer que se desea obtener el máximo global, en el espacio de búsqueda de que se trate. Y aunque la obtención del éste máximo es deseable, al conocer las alternativas con las mejores aptitudes, es viable que el diseñador opte por usar una de ellas.

Los resultados muestran que es posible obtener soluciones con un buen nivel de aptitud, permitiendo seleccionar la topología de la red de mínimo costo entre las soluciones de mayor aptitud obtenidas en el proceso.

El modelo utilizado, así como los procedimientos desarrollados durante su implementación tienen una referencia en la literatura de AG, donde ha dominado el uso de cromosomas binarios, en tanto este AG usa no sólo un alfabeto que contiene un número de símbolos proporcional al número de objetos a representar, sino que también utiliza cromosomas en duplas no homogéneas, lo que introduce nuevamente estos elementos poco usados en el tratamiento de problemas con AG.

Adicionalmente, ha sido una oportunidad de ampliar el conocimiento de la aplicación de los AG en la búsqueda de soluciones en redes, permitiendo seguir afinando los parámetros del AG y de los operadores genéticos que intervienen en la selección y modificación de las poblaciones durante el proceso evolutivo de la población de soluciones.

CONTENIDO

Resumen	iii
Contenido	v
Lista de figuras	vii
Lista de tablas	viii
1. Introducción	10
2. Antecedentes	16
2.1. Las redes y su uso	16
2.1.1. Redes de flujo	17
2.1.2. Problema del árbol de expansión mínima (MST)	19
2.2. Las redes de comunicaciones	19
2.3. Integración de redes y computadoras	20
2.3.1. Redes locales	21
2.3.2. Redes de área amplia	23
2.3.2.1. Conmutación de circuitos	24
2.3.2.2. Conmutación de paquetes	25
2.3.3. Redes propietarias	26
2.3.3.1. Redes de área local	26
2.3.3.2. Redes de área amplia	27
2.3.4. Redes de protocolo abierto	28
2.4. Diseño de redes	29
2.4.1. Notación y definiciones	29
2.4.1.1. Árboles ordenados	35
2.4.1.2. Árboles posicionales	36
2.4.2. Representación de una red	37
2.4.3. Selección del modelo de red	39
2.4.4. Criterios de diseño	41
2.4.5. Árbol de expansión mínima como solución a problemas de minimización de costos	42
2.5. Algoritmos genéticos	45
2.5.1. Algoritmos evolutivos	46

2.5.2.	Terminología	47
2.5.3.	Espacio de búsqueda	48
2.5.4.	Implementación de un algoritmo genético	50
2.5.4.1.	Codificación	51
2.5.4.2.	Métodos de selección	52
2.5.4.3.	Operadores genéticos	55
2.5.4.4.	Parámetros de un algoritmo genético	56
3.	Planteamiento del problema	58
3.1.	El diseño de red de telecomunicaciones y el proceso de optimización	58
3.1.1.	La organización	58
3.1.2.	Los usuarios y la red	59
3.1.3.	Equipos de interconexión	60
3.1.4.	Topología de la red	61
3.2.	Árbol de expansión mínima (MST)	63
3.3.	Criterios de diseño de un árbol de expansión mínima	64
3.3.1.	Componentes	64
3.3.2.	Medición del desempeño de la red	66
3.3.3.	Puente transparente	67
3.3.4.	Retraso de señal en la red	68
3.3.5.	Definiciones	70
3.4.	Alcance de la solución	71
3.5.	Trabajos previos	72
4.	Modelo Propuesto	75
4.1.	Creando la representación del problema de diseño de la topología de la red	75
4.1.1.	Árbol de Huffman	76
4.1.2.	Reparación de etiquetas de red	77
4.1.3.	Representación del árbol como cromosoma	77
4.1.4.	Reparación del cromosoma	78
4.1.5.	Desarrollo del algoritmo	79
4.1.5.1.	Generación de cromosomas	80
4.1.5.2.	Matrices de decisión	81
4.1.5.3.	Selección y cruzamiento	86
4.1.5.4.	Parametrización	88
5.	Experimentos	90
5.1.	Datos de entrada	90
5.2.	Resultados	92
5.2.1.	Grupos de pruebas	93
6.	Conclusiones y trabajo futuro	107
7.	Referencias bibliográficas	109
8.	Glosario	113
	Apéndice A	117

LISTA DE FIGURAS

2.3.1.1.	Conmutación de circuitos y conmutación de paquetes	22
2.3.2.1.1	Una central telefónica es una red de conmutación de circuitos	24
2.3.2.2.1.	Red de conmutación de paquetes	25
2.4.1.1.	Ejemplo de grafo dirigido o red dirigida	30
2.4.1.2.	Ejemplo de grafo no dirigido	30
2.4.1.3.	Ejemplo de camino y de trayectoria	31
2.4.1.4.	Ejemplo de ciclos dirigidos	32
2.4.1.5.	Ejemplos de grafos conectados, y no conectados	32
2.4.1.6.	Ejemplo de arcos que están en el corte	33
2.4.1.7.	Ejemplo de árboles	33
2.4.1.8.	Ejemplo de bosque	34
2.4.1.9.	Ejemplo de árbol enraizado	34
2.4.1.10.	Ejemplo de árboles de expansión	35
2.4.1.2.1.	Ejemplo de un árbol 2-posicional	37
2.4.3.1.	Ejemplos de topologías de árboles	39
2.4.4.1.	Descripción de árboles de expansión	41
2.5.3.1.	Un mapa de aptitud	49
3.3.3.1.	Diagrama de decisión de un puente transparente	68
4.1.0.1.	Con el árbol de Huffman se puede construir cualquier configuración de red	75
4.1.1.1.	Árbol completo de Huffman donde $P=5$.	76
4.1.2.1.	Un árbol puede recibir cualquier etiqueta aleatoriamente	77
4.1.5.2.	Lista de prefijos y representación de matriz de listas	84
5.1.1.	Matriz de distribución del tráfico entre usuarios	91
5.1.2.	Matriz de distribución de usuarios por nodo	92
5.2.1.1.	Gráfico del ejemplo 1	95
5.2.1.2.	Representación de la red generada en el ejemplo 1	96
5.2.1.3.	Gráfico del ejemplo 2	97
5.2.1.4.	Representación de la red generada en el ejemplo 2	98
5.2.1.5.	Gráfico del ejemplo 3	99
5.2.1.6.	Representación de la red generada en el ejemplo 3	100

LISTA DE FIGURAS

2.3.1.1.	Conmutación de circuitos y conmutación de paquetes	22
2.3.2.1.1	Una central telefónica es una red de conmutación de circuitos	24
2.3.2.2.1.	Red de conmutación de paquetes	25
2.4.1.1.	Ejemplo de grafo dirigido o red dirigida	30
2.4.1.2.	Ejemplo de grafo no dirigido	30
2.4.1.3.	Ejemplo de camino y de trayectoria	31
2.4.1.4.	Ejemplo de ciclos dirigidos	32
2.4.1.5.	Ejemplos de grafos conectados, y no conectados	32
2.4.1.6.	Ejemplo de arcos que están en el corte	33
2.4.1.7.	Ejemplo de árboles	33
2.4.1.8.	Ejemplo de bosque	34
2.4.1.9.	Ejemplo de árbol enraizado	34
2.4.1.10.	Ejemplo de árboles de expansión	35
2.4.1.2.1.	Ejemplo de un árbol 2-posicional	37
2.4.3.1.	Ejemplos de topologías	39
2.4.4.1.	Descripción de árboles de expansión	41
2.5.3.1.	Un mapa de aptitud	49
3.3.3.1.	Diagrama de decisión de un puente transparente	68
4.1.0.1.	Con el árbol de Huffman se puede construir cualquier configuración de red	75
4.1.1.1.	Árbol completo de Huffman donde $P=5$.	76
4.1.2.1.	Un árbol puede recibir cualquier etiqueta aleatoriamente	77
4.1.5.2.	Lista de prefijos y representación de matriz de listas	84
5.1.1.	Matriz de distribución del tráfico entre usuarios	91
5.1.2.	Matriz de distribución de usuarios por nodo	92
5.2.1.1.	Gráfico del ejemplo 1	95
5.2.1.2.	Representación de la red generada en el ejemplo 1	96
5.2.1.3.	Gráfico del ejemplo 2	97
5.2.1.4.	Representación de la red generada en el ejemplo 2	98
5.2.1.5.	Gráfico del ejemplo 3	99
5.2.1.6.	Representación de la red generada en el ejemplo 3	100

5.2.1.7.	Gráfico del ejemplo 4	102
5.2.1.8.	Representación de la red generada en el ejemplo 4	102
5.2.1.9.	Gráfico del ejemplo 5	105
5.2.1.10.	Representación de la red generada en el ejemplo 5	105

LISTA DE TABLAS

2.1.1.1.	Características comunes a las redes físicas	18
2.4.5.1.	Crecimiento de número de nodos en un árbol de Huffman	44

1.0 INTRODUCCIÓN.

En plena era de la información, existe un interés creciente por el diseño eficiente de las redes de comunicaciones. El problema de diseño de topologías de red ocurre en muchos campos, pero particularmente en el campo de las comunicaciones, en donde la evolución de los muy conocidos sistemas telefónicos analógicos hasta los actuales sistemas digitales de redes locales *LAN*, y de área amplia *WAN*, forman una amplia gama de posibilidades. En esta parte final del espectro podemos incluir también aplicaciones como las redes ISDN y las digitales inalámbricas para uso de telefonía móvil.

La optimización del diseño de redes de comunicaciones ha recibido atención y es tema de interés en diferentes áreas de investigación [WL91] como en la investigación de operaciones [Gav92], [Ahu93] y en la industria [TDW97].

El principal reto en el diseño de redes, estriba en la selección de la topología adecuada al problema que se plantea. En muchos casos, se hace uso de topologías especiales[GD90] para el diseño de redes. Pero en general, se buscan métodos que permitan la optimización de la red, con el propósito de disminuir el costo o la longitud de los enlaces que unen los nodos de la red, o incrementar el desempeño de la red al ofrecer menores tiempos de retardo en el tráfico entre nodos. Están documentados los avances obtenidos para mejorar el diseño de las redes, buscando mejores modelos y métodos de solución[Min89]. También existen trabajos que con una combinación de técnicas[BMS98a] logran mejores aproximaciones con respecto al uso de procedimientos heurísticos.

Los métodos heurísticos tienen sus limitaciones, sean éstas por causa del tiempo que emplean en las operaciones para la determinación de las rutas, como por las limitaciones que presentan al momento de soportar las restricciones que presentan algunos problemas de diseño de redes.

Cuando se plantea el problema de minimización de costos o de longitudes de los enlaces de una red, la topología que resuelve este problema es la de árbol de mínima expansión, o de mínima cobertura [Ahu93], para lo cual han sido diseñados procedimientos heurísticos mejorados, estos son los algoritmos de Kruskal, Prim y Sollin [Ahu93], entre otros. Estos algoritmos son conocidos como *avaros*, por la forma en que operan al hacer la búsqueda de

las rutas para el flujo de menor costo, o para la elección de las rutas más cortas que permitan unir todos los nodos de una red, una de sus características es que resuelven el problema en tiempo polinomial. Estos algoritmos no siempre proporcionan resultados satisfactorios debido a que pueden convergen en óptimos locales. Con el uso de otras técnicas, como los algoritmos genéticos (AG), es posible obtener mejores resultados [BMS98a].

El uso de árboles de mínima expansión como solución a problemas de minimización está no sólo documentado, sino también demostrado [Ahu93] [BMS95]. Los encontramos aplicados en una gran diversidad de problemas normalmente relacionados con el flujo, sea en sistemas de transporte, autopista, o telecomunicaciones; desarrollándose en forma particular de acuerdo al ambiente de aplicación.

Algunos de estos problemas presentan restricciones y en consecuencia un mayor grado de dificultad para ser resueltos. En caso de que se limite el número de nodos que pueden estar conectados a un nodo de nivel superior en el árbol, el problema se llama de restricción en el grado, y es un problema NP-Hard, y si la restricción está en la capacidad de los enlaces al momento de permitir el paso de tráfico será un problema NP-Completo [Gen97] [ACG99] [GJ79]. Estos problemas no pueden ser resueltos en tiempo polinomial, por lo que se hace atractivo el uso de métodos de alternativas de optimización combinatoria, como es el caso de los mencionados algoritmos genéticos.

Los algoritmos genéticos están incluidos dentro de los llamados algoritmos evolutivos. Lo que pretende esta técnica de búsqueda, es la determinación de las mejores soluciones a un problema, por medio del perfeccionamiento paulatino de una población de soluciones. A esta población de soluciones, llamadas cromosomas, se les aplican operadores que asemejan los procesos y modificaciones que ocurren en la evolución de las especies vivas. Ocurre que a través del cambio evolutivo de las soluciones, las que obtienen mejor aptitud logran obtener mejores oportunidades de transmitir sus características a los individuos de la siguiente generación; y en consecuencia sobrevivir ellos y las características que les permitieron obtener esa mayor aptitud.

Una de las razones que invitan a usar los algoritmos genéticos es su habilidad para evitar detenerse durante el proceso en óptimos locales, lo cual significa que en el proceso de búsqueda no hará uso de la monotonía decreciente, que pudiese existir en las tendencias hacia alguna solución óptima de costo mínimo. Esta posibilidad resulta más en una fortaleza que en una debilidad.

El problema del diseño de la topología de la red, para el cual se desea encontrar solución, está determinado por el agrupamiento que se hace de los usuarios de la red, así como del enrutamiento del tráfico generado por estos usuarios al enviar mensajes, o cualquier tipo de información entre ellos. Estas características del modelo llevan a determinar un número de grupos o nodos, a los cuales debe pertenecer cada uno de los usuarios, y cada usuario sólo podrá pertenecer a un nodo.

Una característica más del modelo, es que todos los enlaces entre nodos tienen una capacidad determinada al inicio del proceso, lo cual establece una limitación al flujo y determina el tiempo de retraso del flujo de bits entre cualesquiera dos nodos que mantienen un tráfico a través de la red. Esta capacidad está referida, a la capacidad de los nodos recibir y transmitir una señal que los atraviesa, al *ancho de banda* que es asignado a cada enlace y paralelamente, a la capacidad de los *puentes*, que son los equipos de comunicación encargados de la administración del flujo de bits de entrada y de salida en cada una de las conexiones entre nodos.

Lo que se desea es minimizar el tiempo de retraso que en promedio puede existir en la red por causa de tráfico que se establece entre los diferentes nodos, y éste por razón de los usuarios conectados a él.

Los árboles de expansión mínima, *MST*, han sido utilizados, junto con los AG, en el diseño de topologías de red. El trabajo de Gen, reportado en GECCO 2000 y recopilado en el Workshop Program [Gen00], muestra el estado del arte en la investigación de estos procedimientos. Gen hace la enumeración y muestra de los diferentes métodos de diseño de redes usando AG y los modelos que presenta se basan en el uso de árboles de expansión mínima con restricciones, donde revisa siete tipos de restricciones, incluidas la restricción en el grado y la del número de hojas, las cuales son muy semejantes al modelo representado por el árbol de Huffman, en este trabajo; también reporta el uso de cuatro tipos diferentes de representación, usadas en el planteamiento de los problemas antes mencionados, las cuatro representaciones usan alfabetos binarios.

Los objetivos que plantea alcanzar este trabajo son:

- Crear una aproximación al problema de diseño de la topología de una red de telecomunicaciones, a través una representación de datos no binaria, que permita modelar la representación de la red y los datos de modo flexible.
- Usar los algoritmos genéticos AG, como herramienta eficaz para encontrar buenas soluciones, aunque no necesariamente representando el óptimo global, que permitan decidir sobre el diseño de una red de telecomunicaciones.
- Conocer y ampliar el conocimiento del uso de los diversos operadores genéticos, probando y reconociendo su utilidad en la búsqueda de soluciones dentro del proceso de evolución de poblaciones.
- Ampliar el espectro de uso del modelo utilizado, acoplando el modelo a redes de área amplia, WAN.

Se propone como modelo para la red, un árbol de expansión mínima, al igual que los métodos de minimización arriba mencionados, debido a que, como se mencionó antes, son una solución a los problemas de minimización de costos, de longitudes y de capacidades en una red [BMS95] [ACG99].

Una de las diferencias básicas en el tratamiento del problema, está en la elección de un modelo de árbol, conocido como árbol ordenado, el cual permite la identificación única de cada uno de sus nodos, a través de una etiqueta, en la cual se almacena información que localiza e identifica individualmente a todos los nodos [Eve79].

El árbol que se elige para describir la topología de la red es denominado árbol de Huffman, que es un árbol ordenado, ya que este modelo posee la flexibilidad necesaria que permite obtener cualesquiera de las formas de árbol para la red, que van de la colocación de todos los nodos en línea a ubicarlos en forma de estrella[ES96] [Eve79].

Además, se hace uso de los AG para la búsqueda de la mejor solución al problema propuesto, pensando en que la solución requiere métodos de búsqueda que no trabajen en tiempo polinomial, debido a que el problema es NP-Completo dadas las restricciones a la capacidad de los puentes para el enrutamiento del tráfico, y la distribución de los usuarios en los diferentes nodos.

Es deseable la obtención de un óptimo global, pero no es parte del proceso el garantizar la obtención de éste, sino obtener soluciones de la mayor aptitud posible, que representen una alternativa viable al diseño de una red de telecomunicaciones, o red WAN.

La forma en que está diseñado el cromosoma no toma la opción de usar un alfabeto binario, particularmente porque el uso de un alfabeto de 10 dígitos para la etapa contemplada en éste trabajo, junto con el uso de etiquetas de identificación de los nodos del árbol de Huffman [ES96], facilita las operaciones de búsqueda de rutas entre nodos, localización de nodos intermedios en las rutas y la reparación del cromosomas, los cuales se encuentran constituidos por duplas que representan a los nodos y al orden relativo de dichos nodos en la red.

Se usa mucho del trabajo desarrollado por Elbaum y Sidi, para el diseño de una red LAN [ES96], pero se cambian los objetivos dejando a un lado la distribución de los usuarios dentro de la red, que se da como parte de las condiciones iniciales, y se hace énfasis en la búsqueda de la mejor distribución de nodos y el mejor desempeño de la red.

En los resultados preliminares se observa la formación de redes con una aptitud similar a la obtenida en el trabajo de Elbaum y Sidi, en estos ensayos se usaron 6 nodos como parte de la red. Y en los resultados reportados, donde se incluyen 10 nodos, se observa la formación de redes cercanas al óptimo global, las cuales ofrecen un desempeño notable si las comparamos con el promedio y con las redes de menor aptitud que aparecen dentro de las población final.

Este trabajo no pudo ser comparado con algunos otros que usen la misma técnica de diseño de topologías de red, ya que esta se encuentra reportada únicamente en el trabajo de Elbaum y Sidi [ES96].

Los resultados reflejan un comportamiento del AG de acuerdo a lo esperado, es decir, convergiendo hacia soluciones con una aptitud creciente y un comportamiento estable de los resultados que se van obteniendo a lo largo del proceso. Considero que el camino por recorrer con este modelo es amplio, y es posible profundizar en su estudio y aplicación en la optimización de topologías de red, no sólo en el área de telecomunicaciones, sino en general en cualquier área que requiera la minimización de costos, longitud, o algún otro atributo de una red, a través del uso de árboles de expansión.

El documento está organizado de la siguiente manera:

En este capítulo 1 *Introducción*, Se resume el contenido y propósitos del trabajo de tesis, exponiendo los antecedentes y los alcances que se plantean para el mismo, definiendo las metas que se pretenden alcanzar, con el desarrollo y aplicación del AG al problema de obtener un diseño de red que tenga el menor retraso en el tráfico a través de la misma.

El conocimiento de las características de una red sea ésta de comunicaciones o no, hace necesario conocer de manera general el trabajo que se ha hecho para encontrar solución al mismo problema, presente en diferentes campos de aplicación. Y para el caso de las redes de comunicaciones, el conocer las limitaciones que tienen por causa de la tecnología que poseen, por diseño que la hace corresponder con alguna función particular, por los protocolos que usa en sus diferentes capas hace necesario un recorrido por los antecedentes inmediatos de la red. Esto se trata en el capítulo 2 *Antecedentes*; además, en el mismo, también se dan los antecedentes de los algoritmos genéticos, sus características y los operadores que de común se usan en este tipo de procesos de búsqueda.

A continuación, en el capítulo 3 *Planteamiento del problema*, se hace una revisión de las partes que integran el sistema en el cual estará inmersa la red de comunicaciones; así como una revisión del modelo que se usará, y de las condiciones de diseño que están involucradas, finalmente se define el alcance que con este modelo se puede tener.

El capítulo 4 *Modelo propuesto*, incluye la creación de la representación del problema así como la descripción de las características especiales del algoritmo, tales como el proceso de cruzamiento y reparación de cromosomas, los mecanismos que se utilizaron para encontrar los nodos intermedios entre los diferentes usuarios que mantiene intercambio de información y de las rutas que unen a estos nodos intermedios. También se describe la forma en que se fue parametrizando el algoritmo, hasta alcanzar las condiciones adecuadas de estabilidad en la generación de soluciones, así como poseer una convergencia no muy rápida para permitir la evolución de las soluciones.

El capítulo 5 *Experimentos*, incluye una descripción de los datos de entrada, describiendo su uso y las características que resultan relevantes para la conformación de las matrices de entrada.

Se describe, a través de ejemplos, la forma en que se fue modificando el algoritmo en sus parámetros y en su operación, pensando siempre en obtener el mejor comportamiento de los cromosomas en su evolución. A lo largo de los ejemplos, se van expresando las conclusiones obtenidas, tanto del comportamiento del AG, como de la convergencia de los resultados hacia la solución óptima global ya conocida.

El capítulo 6 *Conclusiones y trabajo futuro* incluye una evaluación de las ventajas que representa el trabajo con este tipo de representación, y lo que en el futuro podría hacerse haciendo uso de la misma técnica.

Al final del trabajo se incluyen las *Referencias Bibliográficas*, y un *Glosario* de términos para asistir en la comprensión del texto, estas secciones están numeradas como 7 y 8 respectivamente.

Hacia el final del trabajo, en el *Apéndice A*, se resume todo el proceso de pruebas, anotando en cada caso, las posibles alternativas de modificación para las pruebas subsecuentes. Es posible observar aquí las alternativas que se tomaron al analizar los resultados obtenidos en cada etapa conforme se hacía las modificaciones de parámetros, o se incluían nuevos operadores genéticos.

2.0 ANTECEDENTES.

2.1 LAS REDES Y SU USO.

En nuestra vida diaria las redes son ubicuas; tales son los casos de las redes de suministro eléctrico y las redes de distribución de agua, que forman parte fundamental en las ciudades proporcionando luz, entretenimiento y vida a los hogares. De igual modo ocurre con las redes telefónicas, donde damos casi por hecho que existen y que nos permiten comunicarnos con mucha gente en nuestras comunidades y con casi cualquier lugar en nuestro país o el mundo. Algunas redes más son: redes de carreteras, de ferrocarriles, de transporte aéreo, las que pertenecen a empresas y están a cargo de la distribución de bienes de consumo; todas ellas permiten la ocurrencia de nuestra vida tal y como la conocemos permitiéndonos cumplir con nuestros compromisos de trabajo o de placer, haciéndonos llegar hasta personas y lugares distantes, con lo cual, nuestro mundo se ensancha y a la vez se convierte cada vez con mayor intensidad en una sola comunidad, permitiéndonos disfrutar de nuevas experiencias y lugares.

La actividad de una línea aérea, actualmente es inconcebible sin la red de computadoras que integra su sistema de reservaciones, y los procesos de manufactura a gran escala, sólo son posibles debido a las redes de distribución que permiten el arribo de los insumos al menor costo, y hacen llegar los bienes que se producen a los centros de consumo con la mayor eficiencia posible.

Todas las redes se establecen con un objetivo común, que es el hacer llegar a su destino algún tipo de bien necesario, pudiendo nombrar la electricidad, el agua, un bien de consumo, un insumo, una persona o un mensaje, de entre dos lugares distantes, a través de un medio que permita el tránsito de una manera eficiente, otorgando a los usuarios un medio confiable, eficiente, rápido y en la mayoría de los casos, costoso.

El diseño de una red es un problema que involucra diversos campos de investigación tales como, matemáticas aplicadas, ciencias computacionales, ingeniería, administración,

investigación de operaciones, y aquellas materias específicas que corresponden con la naturaleza de la red misma, como las ingenierías, civil, eléctrica, hidráulica, etc.

Para hablar del inicio de la investigación sobre redes nos tendremos que remontar a los trabajos de Gustav Kirchhof y de sus contemporáneos, en el análisis sistemático de los circuitos eléctricos [Evd78] [RM72]. Estos trabajos sirven de fundamento a lo que actualmente conocemos como teoría de redes de flujo, y también establece las redes o grafos como un objeto matemático útil, para representar muchos sistemas físicos. Sabemos que como en cualquier trabajo pionero que intenta describir o modelar una parte de un sistema, lo que pretende es obtener respuestas que describan la naturaleza del sistema, respondiendo preguntas tales como: Si se aplican voltajes a la red en algunos de sus ramales ¿Cuál será la corriente resultante?

2.1.1 REDES DE FLUJO.

Una parte importante del estudio de las redes, considera como su objetivo el flujo de alguna entidad a través la red y enfoca el estudio del mismo en tres cuestiones básicas [Ahu93]:

1. El problema de la trayectoria más corta. Lo que se desea es encontrar el modo de recorrer la red entre dos puntos del modo más económico posible, basando la economía en el tiempo, la distancia o algún valor asociado a los canales por los que se va a transitar.
2. El problema de la maximización del flujo. Dada una red en donde cada uno de los lados o arcos que unen sus nodos tiene una capacidad establecida, se desea determinar el flujo máximo que se puede establecer entre dos puntos, de modo que sean respetadas las capacidades de los enlaces.
3. El problema de la minimización del costo del flujo. Si se incurre en un costo por unidad de flujo a través de una red cuyos arcos tienen capacidades y costos por unidad de flujo definidos, y deseamos enviar algún bien entre dos puntos dentro de la red, ¿Qué ruta utilizaremos de modo que podamos enviar el material con el menor costo posible?

Estos problemas han sido tratados extensivamente y podemos encontrar muchos modos de plantearlos, de modo que su tratamiento puede parecer trivial. La dificultad no estriba en encontrar el método para resolver el problema, sino en que debemos de restringir nuestro análisis a un número finito de alternativas en cada caso. Ya que desde un punto de vista puramente teórico, la solución está en evaluar cada una de las posibles soluciones y compararlas, con lo cual podremos obtener la solución óptima. No es posible lograr esto de modo práctico, ya que el conjunto de las posibles soluciones puede llegar a ser mayor que el número de átomos el universo, y esto para problemas muy cercanos a nuestra cotidianidad.

De modo que, aunque contemos con algoritmos bien diseñados, que permitan obtener resultados en un mínimo de tiempo computacional, el tiempo utilizado por una computadora para evaluar cada una de la soluciones depende del tamaño del problema y de la complejidad del algoritmo que calcula la solución. Una condición adicional, que es necesario incluir, es el número de instancias del problema que es necesario evaluar antes de encontrar una solución satisfactoria. Y cuando las instancias crecen en número en forma exponencial, o aún más

rápido, la evaluación y comparación de las diferentes soluciones es imposible ya que el tiempo necesario es tan grande que resulta inaceptable.

Para obtener soluciones que aunque no se garantice que serán óptimas, son aceptables para los fines de la evaluación del problema, se utilizan algoritmos donde el tiempo de ejecución no crece del mismo modo que las instancias del problema que se analiza.

El uso de redes esta muy difundido; su alcance llega a muchas aplicaciones y de diversas maneras. El establecimiento de una red física, que es aquella integrada por entidades físicas, es quizás la más común y más identificable clase de red. A continuación se resumen algunas de sus características[Ahu93].

Tabla 2.1.1.1 Características comunes a las redes físicas.

Aplicaciones	Analogía física entre nodos	Analogía física de las conexiones	Flujo
Sistemas de comunicaciones	Centrales y conmutadores telefónicos, computadoras, equipos de comunicaciones, satélites	Cableado de cobre, fibra óptica, microondas, ondas de radio	Voz, video, datos
Sistemas hidráulicos	Estaciones de bombeo, represas, lagos, tanques de almacenamiento	Tuberías	Agua, gas, petróleo, fluidos hidráulicos
Circuitos integrados de computadora	Compuertas, registros, procesadores	Alambrado	Corriente eléctrica
Sistemas mecánicos	Juntas	Varillas, viguetas, resortes	Calor, energía
Sistemas de transporte	Intersecciones, aeropuertos, patios de ferrocarril	Autopistas, vías férreas, rutas aéreas	Pasajeros, carga, vehículos, operadores

Y entre las redes físicas, las redes de transporte son quizás las más visibles y que influyen más en nuestras vidas. Frecuentemente, los modelos de este tipo de redes son homogéneos, tal como en el caso de las vías de ferrocarril y las autopistas. Pero en otras ocasiones, ellas corresponden a modelos compuesto de una diversidad de entidades, como en el caso de procesos de distribución complejos que involucran decisiones en la logística. El problema tradicional de transporte en la investigación de operaciones es un buen ejemplo. En este problema, un proveedor cuenta con un inventario de bienes en sus almacenes y deberá entregar esos bienes a un conjunto geográficamente disperso de centros de distribución al

menudeo, cada uno con una demanda diferente para los distintos artículos, y el proveedor desea cubrir esa demanda con el mínimo posible de costos de transporte. En este caso, la parte de transporte en la infraestructura de la red podría corresponder con un conjunto complejo de canales de distribución que incluya, por ejemplo, el uso de transportes por carretera del almacén a los centros de embarque por ferrocarril, el uso de la red ferroviaria, y la transferencia al centro de distribución al menudeo a través de transportes por carretera desde el punto de destino del ferrocarril.

Las redes físicas van más allá de las redes de transporte, surgen en varias disciplinas tanto de ciencias aplicadas como en la ingeniería, tales como matemáticas, química, ingeniería eléctrica, comunicaciones, mecánica, e ingeniería civil. Cuando una red de este tipo ocurre en cualquiera de estas disciplinas sus nodos, arcos y flujos modelan muy diferentes tipos de entidades físicas. Por ejemplo, en una red de comunicaciones típica, los nodos pueden representar conmutadores telefónicos y centrales de transferencia, los arcos serán las líneas de cobre o los enlaces de fibra óptica, y el flujo puede significar la transmisión de datos o de voz. En nuestro caso particular la red de comunicaciones se restringe al uso de datos.

2.1.2 PROBLEMA DEL ÁRBOL DE EXPANSIÓN MÍNIMA (MST).

Un modelo de red importante es el Árbol de Mínima Expansión, cuya definición y representación veremos más adelante; y de este modelo es importante hacer notar, que sin pertenecer a los modelos de redes de flujo estrictamente, es usado en algunos de ellos, como en el caso de los problemas de flujo de costo mínimo, trayectoria mínima y flujo máximo.

Un árbol de expansión, se define sintácticamente como un grafo conectado acíclico, que al extenderse toca todos los nodos de una red. El costo de un árbol de expansión es la suma de los costos, o de las longitudes, de sus arcos. En el problema del *MST*, lo que se desea es identificar el árbol de expansión de mínimo costo, o longitud.

Las aplicaciones del problema del *MST* son variadas e incluyen:

1. Construcción de autopistas o vías de ferrocarril que conecten varias ciudades.
2. Diseño de tuberías que conectan plataformas petroleras, refinerías y centros de consumo.
3. Diseño de redes tanto de área local, *LAN*; como de área amplia, *WAN*.
4. Alambrado de conexiones eléctricas en tableros de control.

2.2 LAS REDES DE COMUNICACIONES.

Las razones que se pueden aducir para el establecimiento de redes de comunicaciones son varias, ya que el uso de una red de comunicaciones está presente en la mente de quienes deben resolver los problemas de intercambio de información, tanto de datos, como de voz, vídeo y documentos; siendo actualmente una solución al alcance de casi cualquier organización.

La primera razón que está a nuestro alcance es la de compartir recursos, haciendo que los programas, datos y equipos estén disponibles para aquellos que se encuentran conectados a la red y tienen permitido su acceso, sin importar la localización geográfica del recurso y del usuario. Esto significa que para un usuario el recurso está disponible, no importando que esté ubicado en otra ciudad, país o continente.

Un segundo objetivo de una red podrá ser el ofrecer una mayor disponibilidad de recursos de que se trate, sean estos datos o capacidad de procesamiento de información, y contar con fuentes alternativas de suministro. De este modo, la información puede estar depositada en bancos de información que operen como *espejos* del banco de datos principal y que en caso de falla de estos últimos, sean estos bancos alternativos los que suministren la información sin que ninguno de los usuarios perciba la falla; o contar con grupos de computadoras donde operen en modo de *clustering* sus sistemas operativos y compartan la carga de procesamiento entre sus CPU's, esto permite distribuir la carga de procesamiento y mantener los procesos ocurriendo aún en el caso de daño en alguna de las computadoras. El uso de éste tipo de arreglos está muy difundido en sistemas de misión crítica donde el sistema debe permanecer operando, como es el caso de aplicaciones bancarias, de control de tráfico aéreo, y actualmente en Internet, donde los portales deben de permanecer operando las 24 horas del día.

Un objetivo adicional está en el ahorro económico que representa el mantener un sistema con procesamiento distribuido en comparación con los centros de procesamiento centralizado en computadoras de alta capacidad. También representa un ahorro a los usuarios de los sistemas el disponer de capacidad de procesamiento sin ellos contar con un centro de cómputo propio, al alquilar tiempo de procesamiento de centros de cómputo remotos, o al rentar aplicaciones que podrán ser utilizadas desde las terminales de la red del usuario, siendo que las aplicaciones se encuentran residentes y ejecutándose en centros de procesamiento remoto.

Las alternativas de contar con un centro de procesamiento centralizado o distribuido, dependen más del modelo de la operación de que se trate que de las ventajas *per se* de alguno de los dos modelos de procesamiento. En el caso del procesamiento centralizado la restricción de acceso a la información por razones de seguridad es una ventaja, y el control de acceso se facilita cuando existe una ventana única para este propósito. El procesamiento distribuido, por su parte, es flexible en el uso de los recursos de procesamiento.

134918 BIBLIOTECA

2.3 INTEGRACIÓN DE REDES Y COMPUTADORAS.



Pertenece ya a la historia la forma en que se fueron integrando las computadoras en redes, para el caso que nos ocupa, únicamente es necesario describir los diferentes niveles de integración que han ocurrido, muchos de los cuales aún se encuentran en uso, así también es importante

conocer las condiciones necesarias de cada uno de estos tipos de conexión entre computadoras.

El medio de conexión para transferir datos entre computadoras debe ser función del número de computadoras involucradas, la naturaleza de las aplicaciones y la separación física entre las computadoras.

Si la comunicación involucra a dos computadoras y ambas se encuentran en el misma oficina, entonces el medio utilizado podrá ser de un simple cable de conexión de punto a punto. Por otra parte, si las computadoras se encuentran en diferentes sitios separados cientos de metros o kilómetros, el medio será un *carrier* público. Este puede proporcionar dos alternativas: Un enlace punto a punto o un enlace conmutado, en ambos casos se requerirá un equipo llamado módem en cada uno de los sitios, el local y el remoto, para efectuar la transmisión de los datos.

Si se enlazan más de dos computadoras, se requiere un medio que pueda conmutar la comunicación entre ellas asignando diferentes intervalos de tiempo para la transferencia de mensaje entre las mismas, esto ya es una red. Si las computadoras se encuentran distribuidas en un mismo piso o edificio lo adecuado es instalar una red conocida como Red de Área Local o LAN, por sus siglas en inglés.

En el caso de que el conjunto de computadoras no ocupe una sola localidad, sino que se encuentren distribuidas en diferentes localidades, la conexión a través de un *carrier* se hace necesaria. La red resultante se conoce como WAN o Red de Área Amplia. El *carrier* podrá facilitar el medio de comunicaciones de más bajo nivel a través de la renta de enlaces dedicados de ancho de banda fijo, los cuales pueden ser analógicos, prácticamente en desuso, o digitales; estos últimos pueden ser canales de 64 Kbps., o múltiplos de esta cantidad. Algunos *carriers* proveen el servicio de Redes de Valor Agregado *VAN*, las cuales proporcionan la conexión en un nivel superior permitiendo la asignación de ancho de banda sobre demanda. En ambos casos, el uso de la red no queda restringido a la transmisión de datos ya que la transmisión de voz también puede integrarse dentro de la red.

El medio de interconexión de computadoras que actualmente observa mayor crecimiento es Internet, el cual permite unir a través de una red pública, donde interviene un número indeterminado de *carriers*, las computadoras de una misma empresa. Hay problemas no resueltos relativos a la confidencialidad de la información y al control de acceso, pero esto no impide que su uso se vaya extendiendo y permita la transferencia de información, sobre todo datos, también vídeo y voz, y estos dos últimos no en tiempo real, para su uso en localidades muy alejadas entre sí.

2.3.1 REDES LOCALES.

Hay dos tipos esenciales de redes locales: la basada en la conmutación de circuitos y aquella cuya tecnología que hace uso de la difusión de paquetes en el medio.

El ejemplo con el que estamos familiarizados es el de un conmutador conocido también como *PBX*.

El *PBX* se desarrolla originalmente para proporcionar un sistema telefónico local. Un *PBX* de voz proporciona un punto de interconexión a las extensiones telefónicas dentro de una oficina, y también la salida a la red telefónica pública a través de troncales telefónicas.

La tecnología digital actualmente domina el mercado de los *PBX*, en estos equipos todas las señales se manejan como señales digitales y hacen uso de la conmutación de circuitos al efectuar los enlaces. El uso de esta tecnología digital permite su uso tanto al establecer conexiones de voz como de datos.

Una solución diferente para el establecimiento de una red local se tiene cuando se hace uso de un medio compartido por un conjunto de equipos.

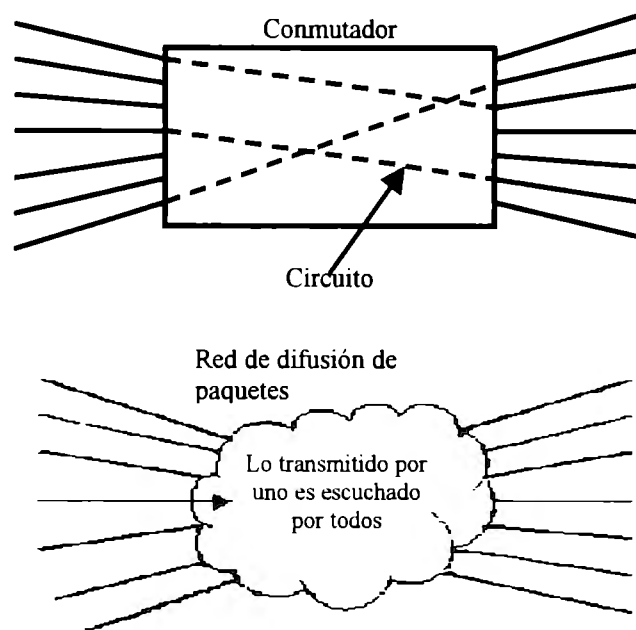


Figura 2.3.1.1 Conmutación de circuitos y conmutación de paquetes. En el dibujo superior, se desea ilustrar la conmutación de circuitos; en este ejemplo, se establecen circuitos fijos en tanto se realiza la transferencia de mensajes entre los usuarios de los enlaces. En el dibujo inferior, se muestra lo que ocurre con la conmutación de paquetes, en donde lo transmitido por un usuario es susceptible de ser escuchado por otros usuarios de la red.

Una de las primeras formas de compartir un medio fue el uso de enlaces multipunto. El multipunto sin embargo, es usado para comunicar un equipo llamado estación primaria, normalmente un *mainframe*, y una cierta cantidad de estaciones secundarias, usualmente terminales. La comunicación es controlada por la estación primaria y las estaciones secundarias no se pueden comunicar entre ellas.

Una red local real, requiere que la comunicación entre cualesquiera, dos estaciones o equipos conectados a ella sea permitida. Este tipo de red, normalmente llamamos Red de Área Local, LAN, y tiene las siguientes características básicas:

- El medio de transmisión se comparte entre los equipos conectados.
- Se transmite en forma de paquetes (de modo semejante a la conmutación de paquetes de las redes de área amplia).
- La transmisión desde cualquier estación es recibida por todas las otras estaciones (de aquí el término de difusión de paquetes).
- No hay una estación maestra, es decir, existe cooperación entre las estaciones para asegurar ordenadamente el uso del medio de transmisión.

2.3.2 REDES DE ÁREA AMPLIA.

La red de área amplia WAN, puede pertenecer por completo a la organización que la usa (red privada) o ser proporcionada por un proveedor especializado en este servicio (red pública), en este último caso serán varias las organizaciones que compartirán el recurso.

Las redes WAN podemos designarlas como redes de comunicación basadas en la conmutación. Una red está integrada por un conjunto de nodos, en la cual la información transmitida entre una estación fuente a una estación destino es conducida a través de los nodos de la red. Los nodos están unidos por líneas de transmisión. Una señal que proviene de una estación entra a la red y es conducida a su destino conmutándola de nodo a nodo.

Algunas características básicas de una red de conmutación son:

- Algunos nodos sólo conectan con otros nodos. Su tarea consiste en la conmutación de información. Otros nodos tienen una o más estaciones conectadas, en este caso tienen funciones de asignación de rutas además de las funciones de conmutación, lo cual les permite aceptar y entregar información a las estaciones conectadas.
- Los enlaces entre nodos son generalmente enlaces dedicados. Estos enlaces, entre nodos, son generalmente multiplexados, ya sea por frecuencia como *FDM*, o por tiempo *TDM*.
- Normalmente la red no está totalmente conectada, es decir, no hay un enlace directo entre cada posible par de nodos. Sin embargo, en algunos casos es deseable contar con alguna trayectoria alternativa entre pares de nodos, considerando el nivel de disponibilidad deseado en la red.

Son dos las tecnologías que se usan en una red de área amplia conmutada: la conmutación de circuitos y la conmutación de paquetes. La diferencia reside en la forma en que los nodos conmutan la información de un enlace a otro en la ruta que va de una estación fuente a una destino.

2.3.2.1 Conmutación de circuitos.

La conmutación de circuitos involucra tres fases:

- Establecimiento del circuito. Antes de que una señal se transmita, un circuito de punta a punta (de estación a estación) debe establecerse. Lo que procede en estos caso es la selección de un canal libre entre cada par de nodos que se encuentran en la ruta que se va estableciendo entre las dos estaciones. Esto ocurre cuando la estación fuente indica, al nodo al que se encuentra conectada, la identificación del nodo destino, con lo cual éste primer nodo inicia la construcción de la trayectoria, transfiriendo la información necesaria a los nodos subsecuentes, hasta lograr el establecimiento de la ruta completa.
- Transferencia de información. Una vez establecido el circuito completo será posible la transmisión de información. Generalmente el circuito opera como *full duplex*, con lo cual ambas estaciones pueden transmitir su señal simultáneamente.
- Circuito desconectado. Una vez que se ha transferido la información entre las estaciones se inicia el proceso de desconexión, usualmente por iniciativa de alguna de las dos estaciones. Este proceso libera los canales ocupados a través de la trayectoria establecida, con lo cual quedan disponibles para su uso.

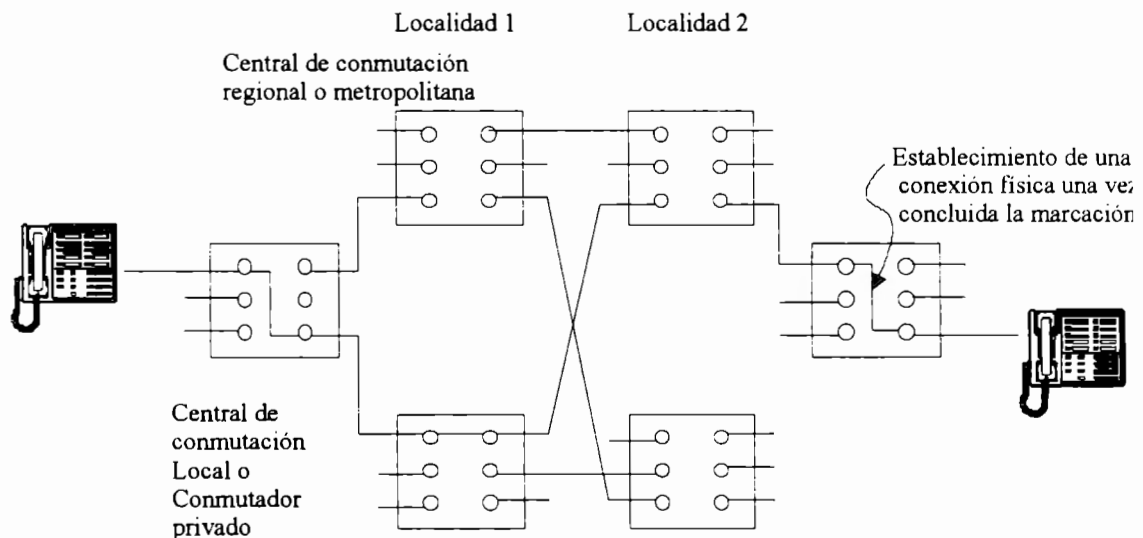


Figura.2.3.2.1.1. Una central telefónica es una red de conmutación de circuitos.

Algunas aplicaciones para este tipo de redes son:

- Red telefónica pública. Provee la interconexión a través de un enlace de voz de dos teléfonos. Las llamadas pueden ser establecidas entre cualesquiera dos subscriptores en la red local, nacional, o internacional. Esta red maneja las conexiones de datos a través módems. Es también posible la conexión digital entre dos computadoras a través de la tecnología *DSL*, por enlaces digitales conmutados, aunque no ha llegado a ser práctica común.

- *PBX*. Conmutadores telefónicos. Proporcionan la capacidad de interconexión de voz y datos en un edificio o en un grupo de ellos. Las llamadas se puede establecer entre cualquier par de teléfonos en la localidad. Naturalmente estos equipos están interconectados a las redes publicas a través de troncales digitales o analógicas.
- Red de área amplia privada. Proporciona la interconexión entre varios sitios de una misma organización. Generalmente se usa para interconectar los *PBX* de las diversas localidades.

2.3.2.2 Conmutación de paquetes.

Alrededor de 1970, se inició la investigación de una nueva arquitectura para la comunicación digital a larga distancia: la conmutación de paquetes.

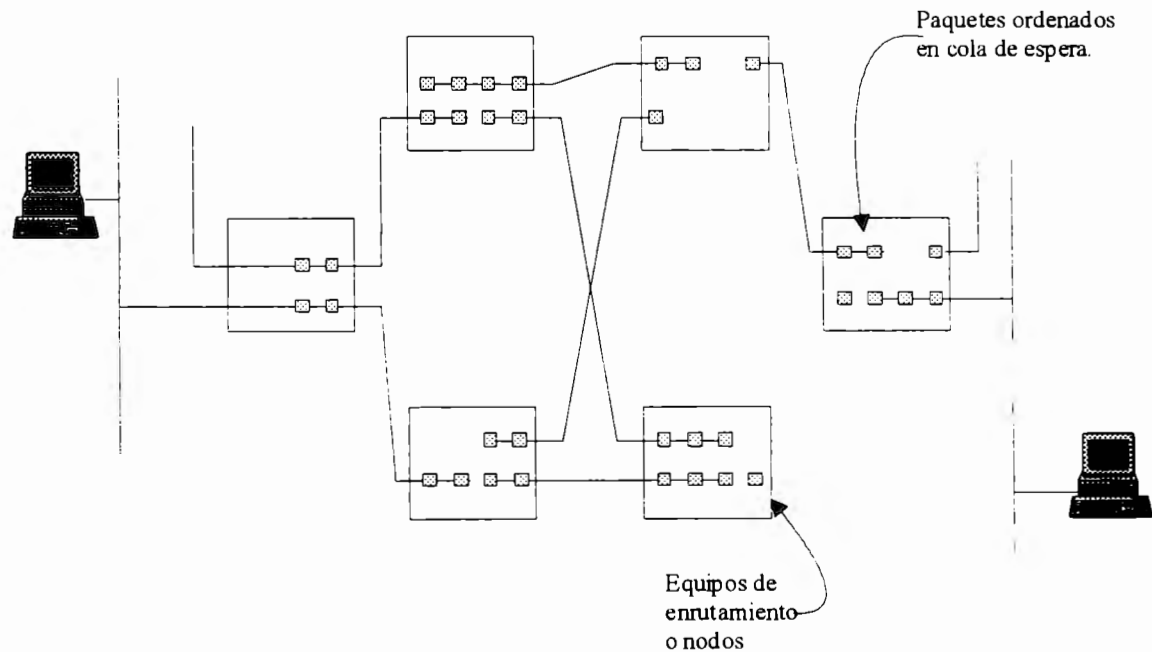


Figura 2.3.2.2.1. Red de conmutación de paquetes.

El proceso para el envío de información por este método lo podemos resumir de la siguiente manera:

- El proceso de conexión no existe para este método de transmisión. Es necesario que la estación sea dada de alta en la red, es decir, en el nodo local correspondiente, para que esté habilitada para transmitir o recibir mensajes. Esto no es equivalente al proceso de conexión en la conmutación de circuitos ya que sólo ocurre una vez, en el momento en que la estación se enciende e inicia su proceso de alta en la red.

- La estación transmisora, una computadora u otro equipo, envía un mensaje como una secuencia de paquetes. Cada paquete incluye la información de control indicando la estación de destino.
- Cada paquete es enviado inicialmente al nodo al cual está conectado el equipo.
- El paquete arriba al nodo y éste determina por cuál de los enlaces, a los cuales está conectado, debe dirigirse el paquete. Pone en fila al paquete por esa ruta.
- Una vez que el enlace está disponible, el paquete se transmite al siguiente nodo.
- El proceso se repite trazando una ruta hasta alcanzar su destino.

Es posible que los paquetes enviados a través de la red tengan una longitud mayor a la que disponible en alguno de los nodos. Este nodo se hará cargo de fraccionar el paquete construyendo nuevos paquetes con la longitud requerida.

Hay dos modos en que se maneja el flujo de paquetes generados durante el proceso de transmisión de información entre dos estaciones: por *datagrama* y por *circuito virtual*.

Cuando se usa el modo de *datagrama* cada paquete es tratado independientemente sin asumir ninguna referencia con respecto a los paquetes que lo antecedieron.

En el caso del *circuito virtual* una ruta planeada con anticipación se establece antes de que los paquetes con información sean enviados.

2.3.3 REDES PROPIETARIAS.

Aunque las redes protegidas por patentes, de algunas compañías como IBM, son cada vez menos usadas, es necesario mencionarlas como parte de la evolución hacia las redes actuales.

2.3.3.1 Redes de Área Local.

Las redes de LAN podemos dividir las por sus características:

- **Técnica de Transmisión**
 - *Broadband*. Usa elementos de *CATV*, tal como el cable coaxial y módems. Usa varias frecuencias activas concurrentes. Soporta canales de vídeo. Un ejemplo es la WangNet de Wang.
 - *Baseband*. Posee una sola frecuencia portadora. Usa cable coaxial o par trenzado. Transmite entre 1 a 100 Mbps. Ejemplos: Ethernet, Token Ring, etc.
- **Topología:**
 - *Estrella*. Usa concentradores, *hubs*. Al tener un punto común de transferencia de datos se facilita la detección de fallas. Al fallar el concentrador puede fallar la red completa. Ejemplo, *StarLan* de AT&T: 1Mbps. Sobre par trenzado.

- *Anillo*. Usa la técnica de *Token Passing*. Es predecible la degradación de la respuesta en la red conforme se incrementa la carga. Se puede simular la topología de estrella al usar Unidades de Acceso Múltiple, *MAU*. Token Ring de IBM: 4-16 Mbps. Par trenzado blindado y sin blindaje.
 - *Bus*. Un cable es compartido por todos los equipos conectados a la red. Posee terminadores en ambos extremos. Todos los equipos escuchan todas las transmisiones. Sólo un equipo usa el medio para transmitir. Ethernet de Xerox, o su equivalente en IEEE como 802.3, este último protocolo es un estándar público.
 - *Árbol*. Esta topología la encontramos en dos casos:
 - Redes jerárquicas con protocolo de control de acceso al medio propietario, como Poll-Select de Unisys. La derivación en cascada de *nodos tipo 2* de SNA.
 - Derivada de la topología de bus, al conectar en cascada redes de éste tipo. Las conexiones se hacen a través de puentes, ya en desuso, o sus equivalentes avanzados, switches.
- **Medio de Transmisión:**
 - Par trenzado.
 - Cable coaxial.
 - Fibra óptica.
 - Inalámbrico (radio FM, luz infrarroja).
 - **Por el método de acceso al medio:**
 - Revisión secuencial o selectiva de estaciones. *Polled*. En uso por redes jerárquicas en árbol y de algunas redes de estrella. Un controlador de comunicaciones “censa” o inquiriere a una estación a la vez para permitirle la transmisión. Al fallar el controlador de comunicaciones la red queda fuera de servicio. *Poll-Select* de Unisys.
 - Detección de portadora, acceso múltiple con detección de colisiones. *CSMA/CD*. Se usa en ethernet y en el estándar 802.3 de IEEE. El acceso al medio es probabilístico. Las colisiones se manejan por retransmisión después de un tiempo de espera aleatorio.
 - Acceso al medio por transferencia de testigo. *Token Passing*. Usado por IBM en Token Ring, que se convirtió en estándar como 802.5 de IEEE. El tiempo de respuesta depende del número de estaciones y de la carga de la red. Una estación podrá transmitir sólo al recibir el token.

2.3.3.2 Redes de Área Amplia.

Para el caso de redes de área amplia, *WAN*, se hicieron uso intensivo de algunas tecnologías propietarias, las cuales en algunos casos se popularizaron y fueron adoptadas por los usuarios como “estándar”. El caso más notable es el de *SNA*, que es la propuesta de IBM para la interconexión de redes. Esta arquitectura apareció en los años setentas permitiendo la comunicación entre diferentes localidades, haciendo uso de líneas dedicadas que conectaban a las mismas. En su primera versión de 1974, sólo permitía la conexión centralizada de nodos. La forma de la red era de estrella. Dos años después, la topología de la red era de árbol al permitir la interconexión de los nodos tipo 5, llamados también procesadores de

comunicaciones frontales, *FEP*, Front End Processor. En 1979 se elimina esta restricción. Para mediados de los ochentas la comunicación se hace más flexible.

Otros fabricantes de computadoras crearon arquitecturas paralelas a *SNA*, tal es el caso de *BNA*, de Unisys, que ofreció la interconexión a las redes *SNA*, más la interconexión a su propia red propietaria *Poll-Select* de arquitectura muy restringida.

2.3.4 REDES DE PROTOCOLO ABIERTO.

Hasta hace pocos años la industria de computadoras hizo uso de estándares relacionados principalmente con la operación interna de los equipos o con la conexión de equipos de manera local. Como resultado de esta carencia de normas, fue que se crearon los primeros subsistemas de comunicaciones por parte de los fabricantes de computadoras, los cuales posibilitaban la comunicación únicamente entre sus propios equipos. Tales sistemas ofrecían la posibilidad de conectar de inmediato un equipo, siempre y cuando fuese de la misma marca. Tales sistemas son conocidos como sistemas cerrados, dado que las computadoras de otros fabricantes no pueden intercambiar información a menos que se utilicen los estándares, necesariamente protegidos por una patente, de algún fabricante en particular.

Por otra parte los *carriers*, operadores de redes telefónicas públicas, tenía muchos años formulado acuerdos internacionales que estandarizaban la conexión entre sus redes. Algunos ejemplos son:

- La serie V que es relativa a la conexión de un equipo, designado normalmente como *DTE*, a un módem que se conecta a la red telefónica pública *PSTN*.
- Las recomendaciones de la serie X para conectar un *DTE* a una red pública de datos.
- Las recomendaciones de la serie I para conectar un *DTE* a la red digital de servicios integrados, *ISDN*.

Como es de esperarse, el uso de las recomendaciones han resultado en la compatibilidad entre los equipos de diferentes proveedores, haciendo posible que el comprador pueda seleccionar el equipo que requiere entre un amplio rango de fabricantes.

En un principio, los servicios que proporcionaban los *carriers* públicos, fueron básicamente de transmisión, de aquí que se hiciera uso de estándares para la interfaz de los equipos con las redes públicas. Conforme los *carriers* públicos han iniciado su participación en otras áreas se ha hecho necesario la elaboración de estándares en otros niveles, los cuales son relativos al formato y control del intercambio de información.

El resultado de esto es la creación sistemas abiertos. Un resultado de este esfuerzo de normalización es el modelo de referencia *OSI*, ofrecido por la Organización Internacional de Normas, *ISO*.

Por otra parte, la creación de una red con un estándar abierto conocida como *ARPANET* tuvo su origen en una agencia del Departamento de Defensa de los Estados Unidos, esto ocurrió a finales de los sesentas. Esta primera red experimental se dio a conocer a finales de 1969.

Éste es el origen de la red de redes Internet, la cual es un sistema abierto porque mantiene públicas sus especificaciones, posibilitando entonces la participación de cualquier organización en la creación de equipos y software para integrarlo a la red.

Los estándares que usa Internet no corresponden con los de *OSI*, ya que los últimos son posteriores a la aparición de esta primera red; sin embargo, los protocolos de *ARPANET* cubren las mismas funciones que los protocolos de red y transporte de *OSI*. El protocolo *IP*, protocolo entre redes, es un protocolo sin conexión y fue diseñado para manejar un número muy grande de redes *WAN* y *LAN*. En tanto *IP* tiene como parte primordial la interconexión de redes, *OSI* lo toma como una idea posterior.

TCP, protocolo de control de transmisión, es el protocolo de transporte de *ARPANET*, el cual es un protocolo orientado a conexión, es decir establece un circuito virtual entre las estaciones que se comunican por la red, y si bien cumple con funciones parecidas a las del protocolo de transporte de *OSI*, difiere por completo en cuanto a sus formatos y procedimientos. La ventaja de *TCP* radica en que este protocolo es parte del sistema operativo *UNIX*, lo cual ha hecho que se difunda amplia y rápidamente.

2.4 DISEÑO DE REDES.

Algunos elementos necesarios dentro del diseño de la topología una red son la notación, la descripción de la representación de la que se hará uso, el modelo de red, y los criterios de diseño. Los elementos anteriores forman la base a partir de la se realiza la selección de la topología de una red.

2.4.1 NOTACIÓN Y DEFINICIONES.

Grafo dirigido y red dirigida.

Un grafo dirigido $G=(N,A)$ consiste de un conjunto N de nodos y un conjunto A de arcos cuyos elementos son pares ordenados de nodos donde el primer y segundo elemento del par son diferentes, n denota el número de nodos y m denota el número de arcos.

$$N=\{a_1, a_2, \dots, a_n\}, A=\{(a_i, a_j) \mid (a_i, a_j) \neq (a_j, a_i), a_i, a_j \in N, i, j \leq n, i \neq j\}$$

Una red dirigida es un grafo dirigido cuyos nodos y arcos tienen asociados valores numéricos, usualmente esto son costos, capacidades, suministros y/o demandas. En general pueden usarse como sinónimos los términos grafo y red a lo largo de este trabajo.

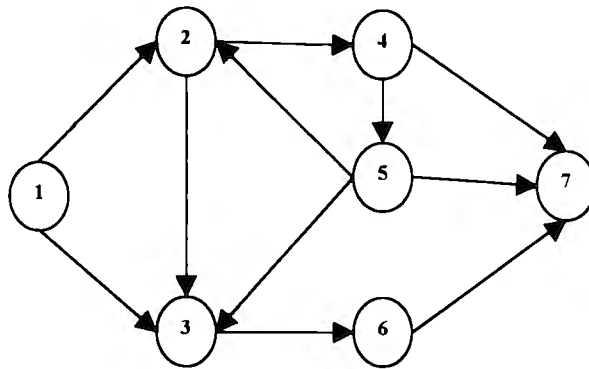


Figura 2.4.1.1 Ejemplo de grafo dirigido o red dirigida.

Grafo no dirigido y red no dirigida.

Grafo no dirigido, tiene una definición similar a la del grafo dirigido excepto que los arcos son pares no ordenados, donde el primer y segundo elemento del par son diferentes.

En tanto un grafo dirigido es similar a una calle de dos sentidos, el grafo dirigido se comporta como una calle de un sólo sentido.

$$N = \{a_1, a_2, \dots, a_n\}, A = \{(a_i, a_j) \mid (a_i, a_j) = (a_j, a_i), a_i, a_j \in N, i, j \leq n, i \neq j\}$$

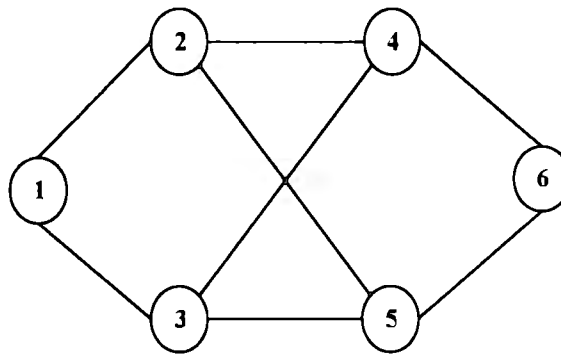


Figura 2.4.1.2. Ejemplo de grafo no dirigido.

Cabeza y cola.

Un arco (a_i, a_j) de un grafo dirigido tiene dos *extremos* a_i, a_j . El nodo a_i es la *cola*, y el nodo a_j es la *cabeza*. Decimos que el arco (a_i, a_j) *emana* de el nodo a_i y *termina* en el nodo a_j .

Un arco (a_i, a_j) es *incidente* en los nodos a_i y a_j . El arco (a_i, a_j) es un *arco saliente* del nodo a_i y es un *arco entrante* en el nodo a_j . Para todo arco $(a_i, a_j) \in A$, se dice que el nodo a_i es *adyacente* al nodo a_j .

Grado.

El grado de entrada de un nodo es el número de arcos entrantes a ese nodo, y su grado de salida es el número de arcos que salen.

Para el caso de los grafos que se usarán en este trabajo, se considerará simplemente el grado de un nodo como el número de arcos salientes.

Lista de adyacencia.

La lista de arcos adyacentes $A(i)$ de un nodo a_i , es el conjunto de arcos que emanan de ese nodo, esto es,

$$A(i) = \{(a_i, a_j) \mid (a_i, a_j) \in A, a_j \in N\}$$

Hay que señalar que $|A(i)|$ es igual al grado de salida de el nodo a_i . Debido a que la suma de todos los grados de salida es igual a m , se obtiene la siguiente propiedad:

Propiedad 1.1 $\sum_{a_i \in N} |A(i)| = m$

Arcos múltiples y rizos.

Arcos múltiples son dos o más arcos con los mismos nodos como cola y cabeza. Un rizo es un arco cuyo nodo cola es el mismo que el nodo cabeza.

En el presente trabajo no se incluye a ninguno de estos dos tipos de arco.

Subgrafo y subgrafo de expansión.

Un grafo $G' = (N', A')$ es un subgrafo de $G = (N, A)$ si $N' \subseteq N$ y $A' \subseteq A$.

Un grafo $G' = (N', A')$ es un grafo de expansión de $G = (N, A)$ si $N' = N$ y $A' \subseteq A$.

Camino.

Un camino en un grafo dirigido $G = (N, A)$ es un subgrafo de G que consiste de una secuencia de nodos y arcos $a_1 - a_2 - a_3 - \dots - a_{r-1} - a_r$, donde cada par de nodos adyacentes en la secuencia define un arco, denominado el arco como l_k , satisface la siguiente propiedad:

$$\forall 1 \leq k \leq r-1, \text{ donde } l_k = (a_k, a_{k+1}) \in A \text{ o } l_k = (a_{k+1}, a_k) \in A.$$

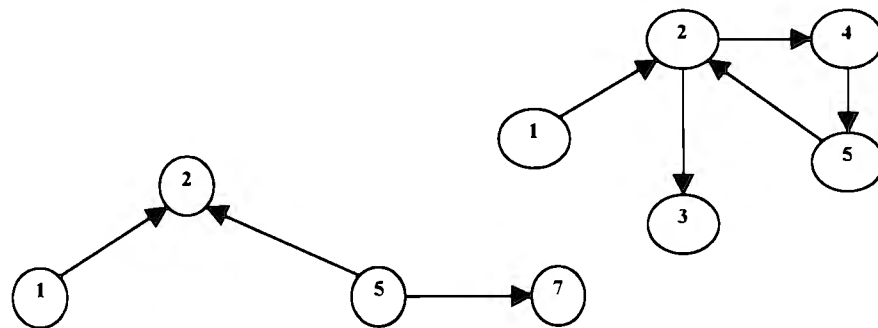


Figura 2.4.1.3 Ejemplo de camino, parte superior derecha; y de trayectoria, parte inferior izquierda.

Trayectoria.

Una trayectoria es un camino en el cual no hay repetición de nodos. Se pueden dividir los arcos de una trayectoria en dos grupos: los arcos hacia delante, y los arcos de regreso. Un arco (a_i, a_j) en una trayectoria, es un arco hacia delante, si la trayectoria visita el nodo a_i antes de visitar el nodo a_j , y es un arco de regreso en caso de que no se cumpla la condición.

Trayectoria dirigida.

Es un camino dirigido en el cual no hay repetición de nodos. En otros términos, una trayectoria dirigida no tiene arcos de regreso.

Ciclo.

Un ciclo es la trayectoria $a_1 - a_2 - \dots - a_r$ junto con el arco (a_r, a_1) o (a_1, a_r) .

Ciclo dirigido.

Un ciclo dirigido es una trayectoria dirigida $a_1 - a_2 - \dots - a_r$ junto con el arco (a_r, a_1) .

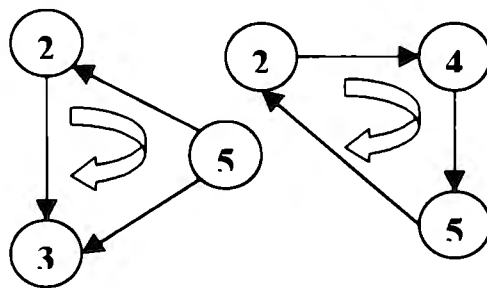


Figura 2.4.1.4. Ejemplo de ciclos dirigidos.

Red acíclica.

Un grafo es acíclico si el no contiene ciclos dirigidos.

Conectividad.

Podemos decir que dos nodos están conectados si el grafo contiene al menos una trayectoria del nodo a_i al nodo a_j . Un grafo está conectado si cada par de sus nodos está conectado, de otro modo el grafo está desconectado.

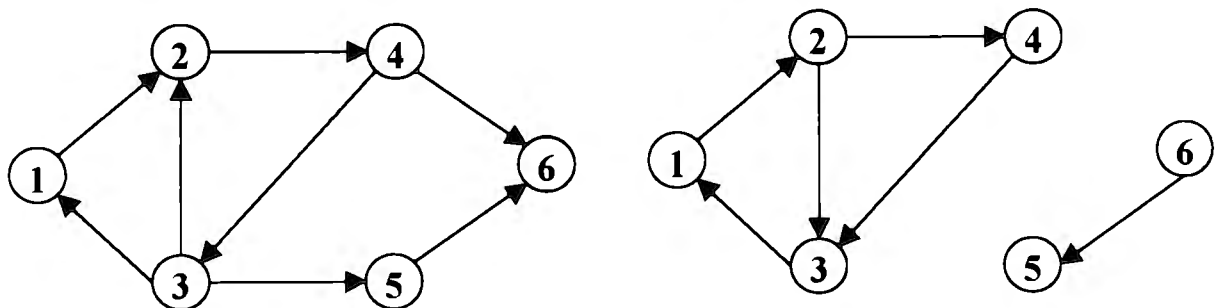


Figura 2.4.1.5. Ejemplos de grafo conectado, lado izquierdo; y grafo no conectado, lado derecho.

Grafo fuertemente conectado.

Un grafo está fuertemente conectado si contiene al menos una trayectoria dirigida de cada nodo a cualquier otro nodo.

Corte.

Un corte es la partición de un conjunto N de nodos en dos partes, S y $\bar{S} = N - S$. Cada corte define un conjunto de arcos, el cual consiste de los arcos que tienen un extremo en S y su otro extremo en \bar{S} . Este conjunto de arcos se denomina como un corte y se representa por $[S, \bar{S}]$.

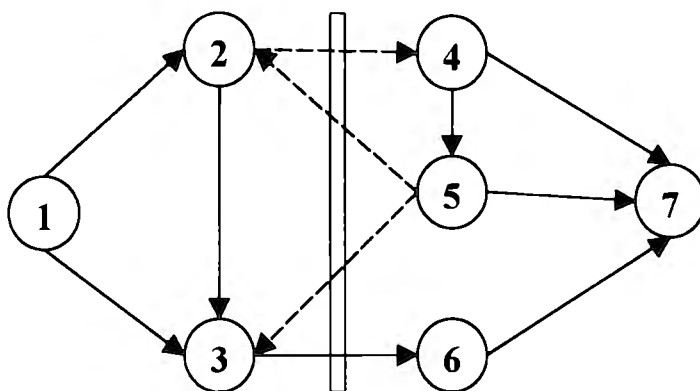


Figura 2.4.1.6. Ejemplo de arcos que están en el corte.

Arboles.

Un árbol es un grafo conectado que no contiene ciclos. Algunas de sus propiedades básicas son:

Propiedad 1.2.

1. Un árbol de n nodos contiene exactamente $n-1$ arcos.
2. Un árbol tiene al menos dos nodos como hojas (v.g., nodos con grado 1).
3. Cada par de nodos de un árbol está conectado por una trayectoria única.

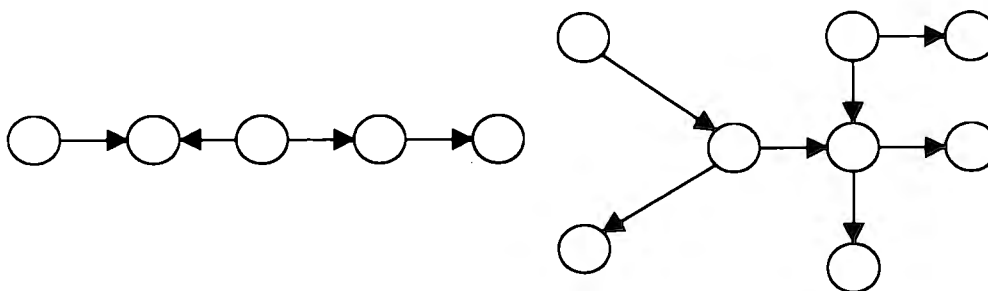


Figura 2.4.1.7. Ejemplo de árboles. A la izquierda, todos los nodos en fila. A la derecha, un ejemplo adicional de árbol.

Bosque.

Un grafo que no contiene ciclos es un bosque. También podemos decir que, un bosque es una colección de árboles.

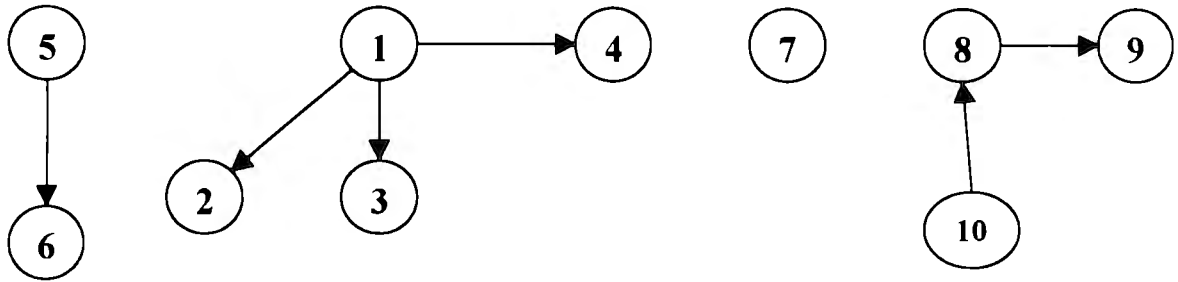


Figura 2.4.1.8. Ejemplo de bosque.

Subárbol.

Un subgrafo conectado de un árbol es un subárbol.

Árbol enraizado.

Un árbol enraizado es un árbol con un nodo que tiene una designación especial, se le llama raíz. Se visualiza el árbol enraizado pensando que se le tiene suspendido por la raíz.

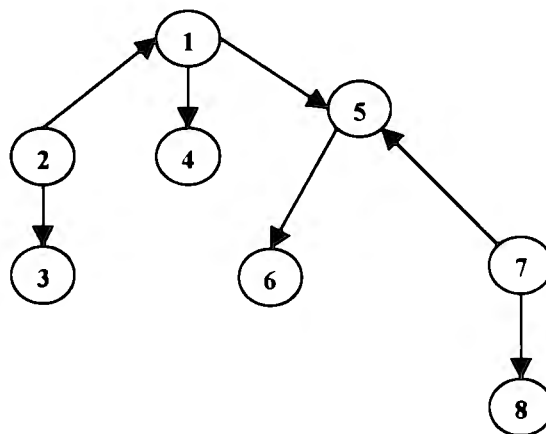


Figura 2.4.1.9 Ejemplo de árbol enraizado.

Árbol de expansión.

Un árbol T es un árbol de expansión del árbol G , si T es un subgrafo de expansión de G .

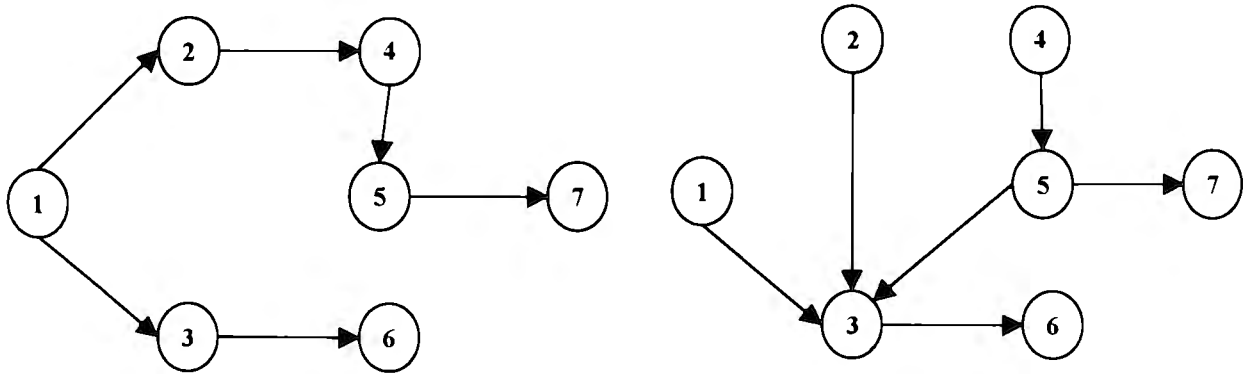


Figura 2.4.1.10. Dos árboles de expansión extraídos del grafo dirigido de la figura 2.4.1.1.

2.4.1.1 Árboles ordenados.

Un árbol ordenado es un árbol dirigido, tal que, para cada nodo hay un orden definido para sus nodos hijos.

Es a partir de la definición de los árboles ordenados que se crea el subconjunto de los árboles de Huffman.

Definiciones.

Alfabeto, letras y palabras.

Sea $\Sigma = \{0, 1, \dots, \sigma-1\}$

donde

Σ es un *alfabeto* y sus elementos son llamados *letras*, y el número de letras en el alfabeto es σ .

Una secuencia finita $a_1, a_2, a_3, \dots, a_l$,

Donde

a_i es una letra, es llamada una *palabra*, donde sus longitud es l . La longitud de una palabra w se expresa como $l(w)$. Una palabra es no vacía si $l(w) > 0$.

Código.

Un conjunto de palabras no vacías y diferentes entre ellas se llama *código*.

Mensaje.

Sean $c_1, c_2, c_3, \dots, c_k$, palabras codificadas. El *mensaje* $c_1 c_2 c_3 \dots c_k$ es la palabra resultante de la concatenación de las palabras codificadas c_1 con c_2 , con c_3 , etc.

Código descifrable de manera única.

Un código de C sobre Σ , es decir, el conjunto de palabras codificadas que están en C las cuales fueron construidas con letras de Σ , se dice que es *descifrable de manera única (UD)* si todos los mensajes construidos a partir de palabras codificadas de C pueden ser divididos de manera única en palabras que están contenidas en C .

Sufijo.

Si s, p y w son palabras y $ps = w$, entonces p es llamado un prefijo de w y s es llamado un sufijo de w .

Cola.

Una palabra no vacía t es llamada *cola* si existen dos mensajes c_1, c_2, \dots, c_m , y c_1', c_2', \dots, c_n' , con las siguientes propiedades:

1. $c_i, 1 \leq i \leq m$, y $c_j', 1 \leq j \leq n$ son palabras codificadas y $c_1 \neq c_1'$;
2. t es un sufijo de c_n' ;
3. $c_1 c_2 \dots c_m t = c_1' c_2' \dots c_n'$.

Lema 2.4.1.1.1 Un código C es *UD* si y solo si no hay cola que sea una palabra codificada.

Prefijo.

Un código C se dice que es *prefijo* si no contiene palabras codificadas que sean prefijo de otras.

2.4.1.2 Árboles posicionales.

Un σ -árbol posicional, o simplemente árbol posicional cuando conocemos σ , es un árbol dirigido con las siguientes propiedades:

1. Cada arco saliente de un nodo v esta asociado con una de las letras de un alfabeto $\Sigma = \{0, 1, \dots, \sigma-1\}$;
2. A diferentes lados se asocian diferentes letras. De aquí que, el número de lados que emergen de un nodo v es a lo más σ , pero podrían ser menos, de hecho una hoja no tiene ninguno.
3. Se asocia a cada nodo v una palabra, que consiste de la secuencia de letras asociadas con los arcos en la trayectoria de la raíz a v . Un ejemplo se observa en la figura 2.4.1.2.1, donde se considera un árbol binario, 2-árbol posicional, donde cada nodo tiene una palabra asociada.

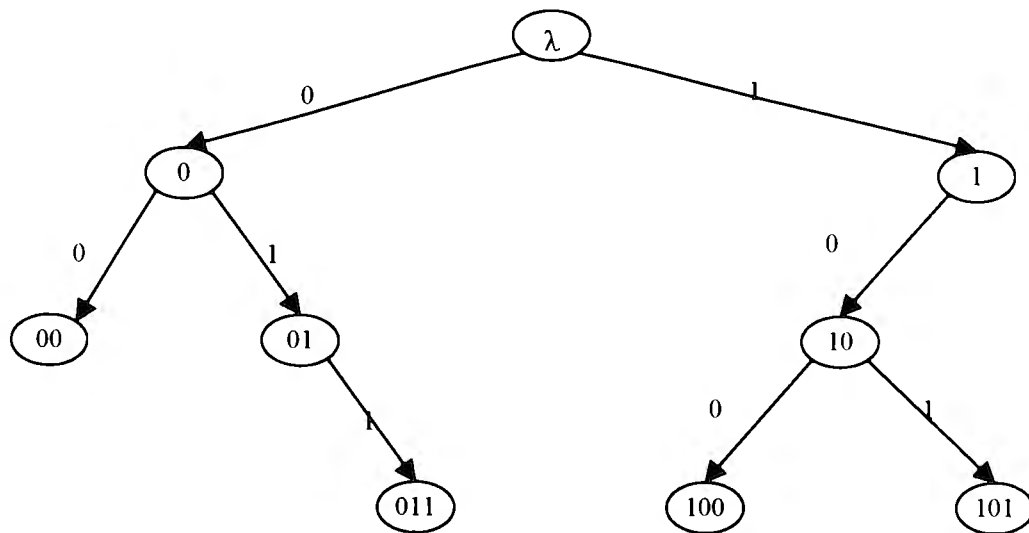


Figura 2.4.1.2.1. Muestra un ejemplo de un 2-árbol posicional.

El conjunto de palabras asociadas con las hojas de un árbol posicional es un código prefijo. Por consiguiente, todos los códigos prefijos se pueden describir con un árbol posicional de esta manera.

De lo anterior podemos concluir que, si un código de Huffman es un código prefijo, entonces podremos construir un árbol posicional con este código. Este árbol lo llamaremos árbol de Huffman.

2.4.2 REPRESENTACIÓN DE UNA RED.

La implementación de un algoritmo de red depende no sólo del desarrollo del algoritmo en sí, sino también de la manera en que se representa la red dentro de la computadora y el esquema de almacenamiento usado para mantener y actualizar los resultados intermedios. Si se representa una red con habilidad y usando estructuras de datos adecuadas, es posible mejorar el tiempo de ejecución de un algoritmo. Al representar una red, usualmente se requieren dos tipos de datos:

1. La topología de la red, es decir, la estructura de la red representada por sus nodos y arcos.
2. Los datos, tales como costos, capacidades, suministros/demandas, asociados con los nodos y arcos de la red.

Matriz de incidencia de nodo a arco.

La representación de esta matriz, también llamada matriz de incidencia, simplemente. Se usa para representar una red como la matriz de restricciones de un problema de flujo de mínimo

costo. Esta representación almacena la red como una matriz N de $n \times m$, la cual contiene una fila para cada nodo y una columna para cada arco. La columna correspondiente al arco (α_i, α_j) tiene únicamente dos elementos: un $+1$ en el renglón correspondiente al nodo α_i y un -1 en la fila correspondiente a α_j .

La matriz de incidencia tiene una estructura particular: únicamente $2m$ de los nm datos posibles son diferentes de cero. Todos los datos, diferentes de cero, son $+1$ o -1 , y cada columna tiene exactamente un $+1$ y un -1 . Además, el número de $+1$'s en un renglón representa el grado de salida del nodo y el número de -1 's es el grado de entrada del nodo. Si bien se entiende que ésta forma de almacenamiento no es eficiente en el uso del espacio, es importante ya que representa la restricción del problema de flujo de costo mínimo y por las propiedades teóricas que permiten un uso eficiente de la información que posee.

Matriz de adyacencia de nodo a nodo.

Llamada simplemente matriz de adyacencia, almacena a la red como una matriz H de $n \times n$, donde $H = \{h_{ij}\}$. La matriz tiene un renglón y una columna para cada nodo, y cumple con

$$h_{ij} = \begin{cases} 1 & \text{sí existe enlace entre } \alpha_i \text{ y } \alpha_j \\ 0 & \text{en cualquier otro caso} \end{cases}$$

Para toda $(\alpha_i, \alpha_j) \in A$.

Si es necesario almacenar los costos de los arcos, sus capacidades, etc. así como la topología de la red, entonces será necesaria una matriz para cada una de estas características.

La matriz de adyacencia tiene n^2 elementos, y únicamente m son diferentes de cero. En consecuencia, esta representación de la red es eficiente en el uso de espacio, si la red es suficientemente densa; para redes dispersas esta representación desperdicia, proporcionalmente, una gran cantidad de espacio. Sin embargo, la simplicidad de la representación de una red por medio de la matriz de adyacencia, permite usarla para implementar algoritmos de red de manera muy sencilla.

Para redes densas, el uso de esta representación hace eficiente el cálculo de los arcos de entrada y salida, pero para redes dispersas esta parte del proceso puede resultar en un cuello de botella para el algoritmo.

Hay representaciones que permiten identificar el conjunto de arcos de salida $A(i)$ de cualquier nodo en un tiempo proporcional a $|A(i)|$. Una de ellas es la siguiente:

Listas de adyacencia.

Ya se definió una lista de arcos adyacentes $A(i)$ de un nodo α_i como el conjunto de arcos que emanan de ese nodo, es decir, el conjunto de arcos $(\alpha_i, \alpha_j) \in A$ obtenidos al variar α_j sobre los

nodos de la red. De igual forma se define la lista de nodos adyacentes a un nodo α_i como los nodos α_j para los cuales $(\alpha_i, \alpha_j) \in A$. La representación en listas de adyacencia almacena la lista de nodos adyacentes de cada nodo como una lista ligada simple. Una lista ligada es una colección de celdas de uno o más campos.

La lista de nodos adyacentes al nodo α_i será una lista ligada que contendrá $|A(i)|$ celdas y cada una de estas corresponderá a un arco $(\alpha_i, \alpha_j) \in A$. La celda correspondiente al arco (α_i, α_j) tendrá tantos campos como la cantidad de información que se desee almacenar. Uno de esos campos almacenará el nodo α_j y podrán ser utilizados más campos para almacenar el costo del arco, así como su capacidad. Cada celda contendrá un campo adicional llamado enlace, el cual almacenará un apuntador a la siguiente celda en la lista de adyacencia. En el caso de que la celda sea la última de la lista, el valor del campo de enlace puede ser nulo o cero, dependiendo de la forma en que opere la lista.

Es necesario almacenar un acceso a n listas ligadas, una por nodo, de aquí que es necesario un arreglo de apuntadores que vayan señalando a la primera celda de cada lista ligada.

2.4.3 SELECCIÓN DEL MODELO DE RED.

La topología de la red define el modo en el que los nodos de la red se interconectan, además de las diferentes formas en que estos enlaces se establecen entre cada par de nodos. Lo normal es tener una jerarquía que asocie a nodos y/o a enlaces. Como vimos anteriormente, son seis los tipos básicos en que una topología de red puede clasificarse:

1. *Malla (mesh)*. Es la que representa a una red parcial o totalmente conectada.
2. *Derivación múltiple (multidrop)*. Se construye con base en árboles de expansión mínima.
3. *Enlaces direccionados*. Es semejante a una representación hecha por grafos dirigidos.
4. *Estrella*. Basándose en el modelo de una o más estrellas interconectadas.
5. *Colector (bus)*. Empleando un colector (bus) compartido por muchas estaciones o nodos.
6. *Anillo (ring)*. Un anillo que se comparte entre muchas estaciones.

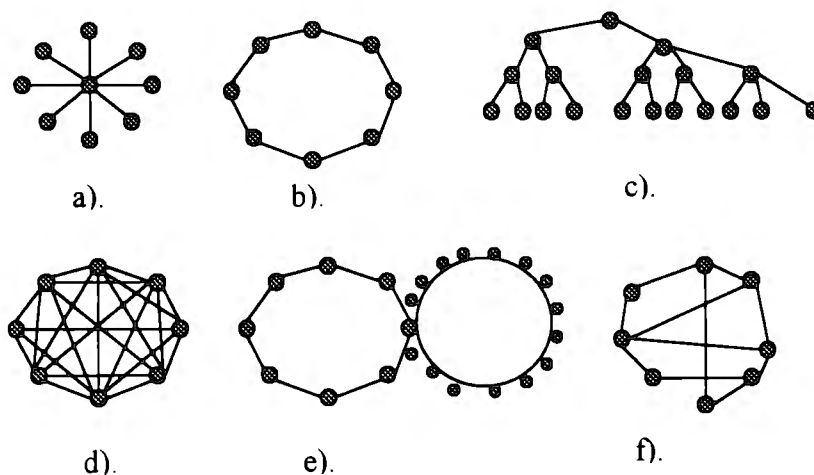


Figura 2.4.3.1. De entre las topologías posibles podemos mencionar a).Estrella. b).Anillo. c).Árbol. d).Malla completamente conectada. e).Anillos en intersección. f).Malla parcialmente conectada.

Estas formas de asociar a los nodos y enlaces pueden combinarse para representar un mayor número de redes, pero finalmente podemos reducir a las anteriores las formas básicas de representación.

El diseño de una red o de algún otro sistema nos lleva al problema del árbol de mínima expansión de manera directa o indirecta. Nos conduce de manera directa, cuando vemos un conjunto de puntos a conectar y debe obtenerse el resultado con los menores costos o la menor suma de las longitudes de los arcos. Estos puntos comúnmente representan entidades físicas que son desde componentes de computadora, a usuarios que requieren conectarse con otros o a un sistema de cómputo centralizado tal como un servidor o algún otro servicio.

Nos conduce de manera indirecta, a un árbol de mínima expansión, cuando:

- Queremos conectar un conjunto de puntos usando criterios de mérito que no se asemejan una función objetivo de un árbol de mínima expansión (suma de longitud de arcos, suma de costos de los arcos).
- El problema no se parece al problema de optimización un árbol de expansión. En este caso el modelo que podríamos ofrecer como alternativa debe acercarse al modelo de un árbol de mínima expansión, si es que nos decidimos por tal planteamiento.

El diseño de una red puede resultar una tarea compleja que involucre la interacción entre los objetivos de desempeño, como rendimiento y disponibilidad, incluir en el diseño el costo y la economía de su operación, y la tecnología disponible [Sha82] [Ahu93]. El principal criterio es simple: Se requiere diseñar una red que conecte un sistema de componentes dispersos geográficamente o que proporcione la infraestructura necesaria para que un conjunto de usuarios se comuniquen entre sí. En muchas de las posibles soluciones el sistema no necesita redundancia, de modo que el interés está en crear la interconexión más simple, un árbol de expansión.

Este tipo de aplicaciones surgen en la construcción de numerosos sistemas físicos: Autopistas, redes de computadoras, redes telefónicas, redes ferroviarias, líneas de televisión por cable, líneas de transmisión de alto voltaje. Los escenarios en donde surgen este tipo de problemas son, por ejemplo:

- Conectar terminales en los paneles de cableado de un equipo eléctrico. ¿Cómo deberán alambrarse las terminales para usar la menor longitud de cable?
- Construcción de la red de ductos para conectar poblaciones de modo que se utilice la menor cantidad de tubos.
- Enlazar pequeñas poblaciones en lugares alejados, las cuales están conectadas por caminos pero no por teléfono. En este caso se desearía conocer en cuáles caminos de harían los tendidos de cables telefónicos, usando la menor cantidad de kilómetros de cable, pero de modo que todos las poblaciones estuvieran enlazadas.
- La construcción de sistemas digitales de computadora, donde se usan circuitos de alta frecuencia y donde es importante minimizar la longitud de las rutas entre componentes, de manera que se reduzca tanto la capacitancia como los efectos de retraso en la línea.

- La construcción de redes de computadoras usando líneas de alta velocidad. Donde cada línea tiene un costo de renta, dependiendo de su capacidad, y deseamos determinar la configuración que conecte todas las localidades con el menor costo posible.

Este último es similar al problema que deseamos resolver, con la diferencia de que lo que se minimizará será el retraso en la red.

2.4.4 CRITERIOS DE DISEÑO.

Hay condiciones optimización que juegan un papel central en el diseño de algoritmos que resuelven el problema de los árboles de expansión y permiten establecer su validez.

Para los árboles de expansión, se pueden formular las condiciones de optimización de dos modos diferentes: Condiciones de optimización de corte, y condiciones de optimización de trayectoria. Ambas condiciones son equivalentes.

Algunos conceptos básicos relacionados con estas condiciones se pueden ilustrar de la siguiente manera.

El grafo (a) de la figura 2.4.4.1 tiene, entre otros posibles, dos árboles de expansión ilustrados en (b) y (c), en tanto el subgrafo (d) no es árbol de expansión ya que no está conectado y el subgrafo (e) no es un árbol debido a que tiene un ciclo en 1-3-4-1. Aquellos arcos que están incluidos en el árbol se denominan arcos del árbol, y los que no están incluidos serán arcos que no pertenecen al árbol. Estas dos definiciones tan simples son utilizadas más adelante.

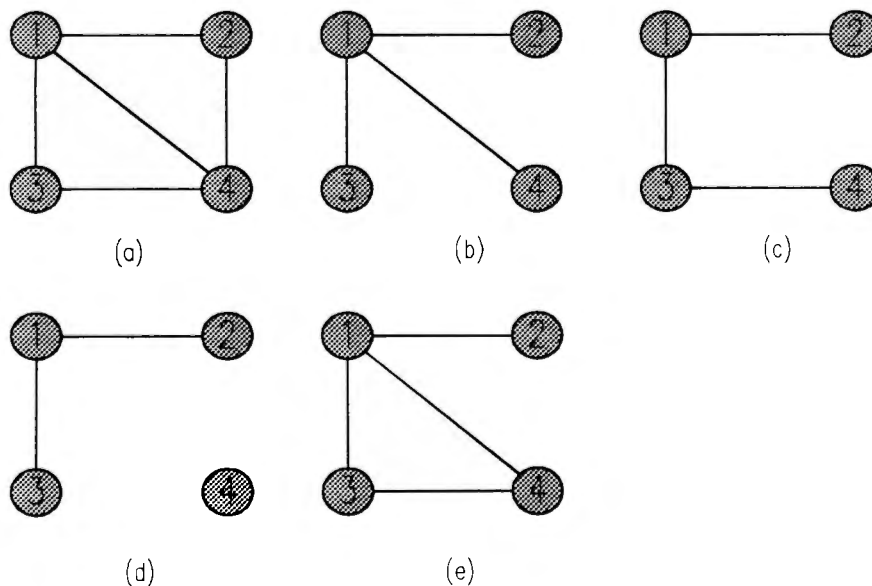


Figura 2.4.4.1. Descripción de árboles de expansión: (a) grafo básico; (b) y (c) son árboles de expansión; (d) nos es árbol de expansión ya que no incluye a todos los nodos; (e) tampoco es árbol de expansión ya que incluye un ciclo.

Es necesario incluir un par de observaciones básicas:

- Para todo arco (a_k, a_l) que no pertenece al árbol de expansión T , existe una única trayectoria del nodo a_k al nodo a_l , tal que el arco (a_k, a_l) junto con esa única trayectoria define un ciclo.
- Si se borra algún arco del árbol (a_i, a_j) del árbol de expansión, el grafo resultante parte el conjunto de nodos en dos subconjuntos. Los arcos del grafo básico G cuyas terminaciones pertenecen a los dos diferentes subconjuntos constituyen un corte.

Con lo anterior podemos establecer las siguientes definiciones:

Condiciones de optimización de corte.

Un árbol de expansión T^* es un árbol de mínima expansión si y sólo si satisface las siguientes condiciones de optimización de corte:

Para todo arco del árbol $(a_i, a_j) \in T^*$, $c_{ij} \leq c_{kl}$ ¹ para todo arco (a_k, a_l) contenido en el corte formado al borrar el arco (a_i, a_j) de T^* .

Condiciones de optimización de trayectoria.

Un árbol de expansión T^* es un árbol de mínima expansión si y sólo si satisface las siguientes condiciones de optimización de trayectoria:

Para todo arco que no pertenece al árbol $(a_k, a_l) \in G$, $c_{ij} \leq c_{kl}$ para todo arco (a_i, a_j) contenido en la trayectoria en T^* que conecta a los nodos a_k y a_l .

2.4.5 ÁRBOL DE EXPANSIÓN MÍNIMA COMO SOLUCIÓN A PROBLEMAS DE MINIMIZACIÓN DE COSTOS.

Dentro del campo de las redes de flujo los árboles de expansión juegan un papel importante. Los problemas de flujo de costos mínimo siempre tienen un árbol de expansión mínima *MST* como solución; de donde, en principio, para encontrar la solución de cualquier problema de minimización de costos de flujo de red, que incluye problemas de trayectoria más corta, y de flujo máximo, siempre podemos restringir nuestra atención a soluciones con árboles de expansión [Ahu93].

Desde el punto de vista de la investigación de operaciones, al existir en una red un número finito de árboles de expansión, el problema de flujo en la red se puede ver como un modelo de optimización discreta y resolverlo en un número finito de iteraciones. Para el caso que tratamos, es teóricamente posible resolver un número finito de operaciones; pero conforme crece el número de nodos y considerando la distribución de los mismos en la red, el número de posibles soluciones crece exponencialmente [ACG99], esto lo podemos observar en la tabla 2.4.5.1, donde A es el número de nodos que contiene un árbol de Huffman completo, que es el

¹ c_{ij} y c_{kl} son los costos o longitudes asociados a los arcos (a_i, a_j) y (a_k, a_l) respectivamente.

modelo que sirve de base a al tratamiento del problema en este trabajo. En dicha tabla, C es el número de instancias de solución al problema y tiene un crecimiento exponencial.

A diferencia de los problemas donde se requiere encontrar la ruta de costo mínimo, en el problema de *MST* se desea encontrar el árbol que une todos los nodos y que la suma de sus trayectorias, o la suma de los costos asociados, sea el mínimo. Como vimos, un árbol de expansión T de G es un subgrafo conectado acíclico que toca a todos los nodos. Cada árbol de G tiene $n - 1$ arcos. Dado un grafo no dirigido $G = (N, A)$ con $n = |N|$ nodos y $m = |A|$ arcos y con una longitud o costo c_{ij} asociado con cada arco $(\alpha_i, \alpha_j) \in A$, lo que se espera es encontrar un árbol de expansión, llamado árbol de expansión mínima, *MST*, que tenga el menor costo total, o la menor longitud, de la suma de los arcos que constituyen el árbol de expansión. Es posible ver que los árboles de mínima expansión difieren de los árboles de trayectoria mínima en dos aspectos importantes:

1. Para el problema del *MST* los arcos son no dirigidos. Y dada esta característica, los arcos entre los pares de nodos α_i y α_j pueden considerarse como (α_i, α_j) o (α_j, α_i) . Para los problemas de trayectoria mínima lo normal es que se consideren la red como una red dirigida. Pero obtener un árbol de mínima expansión en una red dirigida, donde las trayectorias emanan de un nodo raíz, es con mucho un problema más difícil que el problema actual del árbol de mínima expansión [Ahu93]. Por cierto, que la estructura referida donde las trayectorias emanan del nodo raíz se llama raíz arborescente.
2. Las funciones objetivo para el problema de los árboles de expansión mínima y para el problema del árbol de trayectoria más corta son muy diferentes. Para el problema del *MST* se cuenta el costo de cada arco exactamente una vez; para el árbol de trayectoria más corta, el costo de algunos arcos se llega a contar más de una vez: esto es igual al número de trayectorias del nodo raíz que pasan por tal arco, es decir, el número de trayectorias más cortas en el árbol que contienen tal arco.

El problema de los árboles de expansión mínima se encuentra en una gran cantidad de aplicaciones, tanto en un problema aislado, como en un subproblema de un problema más complejo.

Dos son las consideraciones de optimización que podemos usar para asegurar que un árbol es de mínima expansión. La primera condición se basa en comparar los costos de cada arco con todos los otros arcos del árbol que estén contenidos en el corte que se define al remover ese arco del árbol. La otra condición se basa en comparar el costo de un arco que no pertenece al árbol, con el costo de los arcos del árbol en la trayectoria que conecta las terminaciones del arco que no pertenece al árbol. Ambas condiciones son fáciles de establecer y evaluar, de modo que existen varios algoritmos para resolver el problema del *MST*.

Son tres los algoritmos que podemos mencionar por ser fáciles de implementar: Kruskal, Prim y Sollin [Gen97]. Ellos comparten una característica, son algoritmos avaros en el sentido de que en cada etapa adicionan un arco de costo mínimo de una lista de candidatos, considerando que este arco no forme ciclos con los arcos que ya se adicionaron anteriormente. En los tres tipos de algoritmos se conserva el bosque con los arcos que ya se eligieron y entonces se irán adicionando arcos para incrementar el tamaño del bosque.

Para el algoritmo de Kruskal, la lista de arcos que son candidatos la integran todos los definidos para el grafo [Sed92]. Para el algoritmo de Prim, el bosque es un único árbol más un conjunto de nodos aislados y la lista de candidatos contiene todos los arcos entre este árbol y los nodos que no pertenecen al árbol [Sed92]. Finalmente, el algoritmo de Sollin es un híbrido que mantiene varios componentes en el árbol, tal como lo hace Kruskal, pero puede adicionar varios arcos en cada iteración, eligiendo al modo de Prim, los arcos de costo mínimo que conectan cada componente del bosque con los nodos que no están en él [Ahu93].

Para el caso que nos ocupa, el uso de estas técnicas permite la solución del problema, las características asociadas al problema son las siguientes:

- El costo del enlace está asociado al flujo máximo de bits que por él puede circular y por tanto a la capacidad del enlace.
- Cada enlace tiene asociada una capacidad y en función del flujo tiene asociado un retraso en el flujo de bits.
- El flujo de bits por un enlace depende de la posición que cada nodo tiene asignada en la red.
- Al cambiar la posición de cualesquiera dos nodos, cambiará también el flujo asociado a ese enlace.
- El propósito del diseño de la red es encontrar el retraso mínimo dentro de la red, o retraso mínimo promedio de la red. El cual se elige por considerar que es la característica del diseño que calificar directamente a una red, debido a que cuestiones como costos o longitudes de los enlaces, son menos apreciadas cuando el desempeño de la red es inaceptable.

Además de considerar las características anteriores, es necesario observar que las instancias de solución que se pueden ofrecer crecen tan rápido como exponencialmente, conforme crece el tamaño del problema. Como lo muestra la tabla 2.4.5.1.

Tabla 2.4.5.1. Crecimiento del número de nodos en un árbol de Huffman completo y su relación con el número de combinaciones posibles de P nodos.

	A	B	A*B	C	C*B
P	$((P - 1) * ((P/2) + 1) - 1) / (P - 2)$	P!		A Combinaciones en P	
3	3	6	18	1	6
4	13	24	312	715	17,160
5	21	120	2,520	20,349	2,441,880
6	156	720	112,320	18,161,699,556	13,076,423,680,320

En la tabla anterior, A representa el número de nodos que contiene un árbol de Huffman en relación a P, que es el número de nodos del modelo a diseñar. C es el número de combinaciones posibles de los A nodos en grupos de P, que son finalmente, el número de posibles soluciones al problema.

En el caso que nos ocupa, para dar solución al diseño de una red, el espacio de búsqueda es suficientemente grande, por lo que es muy difícil considerar la realización de una búsqueda exhaustiva aún en el caso de una cantidad pequeña de nodos, del orden de 10.

Como hemos vistos, el problema de diseño de la topología de una red se puede resolver con las herramientas heurísticas disponibles, como los algoritmos arriba mencionados, pero algunos problemas de minimización presentan restricciones y en consecuencia un mayor grado de dificultad para ser resueltos. Si existe un límite al número de nodos que pueden conectar a un nodo de nivel superior en el árbol, el problema se llama de restricción en el grado, y es un problema NP-Hard, y si la restricción está en la capacidad de los enlaces al momento de permitir el paso de tráfico será un problema NP-Completo [Gen97] [ACG99] [GJ79]. Estos problemas no pueden ser resueltos en tiempo polinomial, por lo que se necesita utilizar un procedimiento alternativo de solución al problema, y en la literatura se reporta el uso con éxito de métodos de optimización combinatoria, como es el caso de los AG.

En razón de los argumentos anteriores, de tamaño del espacio de búsqueda y de restricciones al modelo, es que se usará un AG como alternativa de búsqueda de soluciones al problema.

2.5 ALGORITMOS GENÉTICOS.

En los años cincuentas y sesentas se inició el estudio de los sistemas evolutivos. La idea de estos sistemas fue hacer evolucionar una población de soluciones candidatas para un determinado problema, usando operadores inspirados por la genética y la selección natural.

En esta época, Ingo Rechenberg(1965, 1973) introdujo las que llamó *Estrategias evolutivas* (*Evolutionsstrategie*, en el original), que fueron usadas durante unas pruebas de cuerpos en un flujo, en un túnel de viento, en donde usó este método, en la optimización de los parámetros que definían la forma de los cuerpos. También se desarrollaron en esa misma época algoritmos que haciendo uso del modelo de la evolución, se aplicaron en optimización y máquinas de aprendizaje. La programación evolutiva de Fogel es una técnica de búsqueda a través de un espacio usando pequeñas máquinas de estados finitos.

Los algoritmos genéticos (AG) fueron creados por John Holland, en la misma década de los sesentas; los desarrolló Holland, sus estudiantes y colaboradores en la Universidad de Michigan entre los sesentas y setentas. A principios de los noventas, Koza propone uno sistema basado en la evolución, la llamada programación genética, para la búsqueda del programa de computadora con mejor desempeño [Mic92].

Se usa el término algoritmos evolutivos, para nombrar los sistemas basados en la evolución, tal es el caso de los mencionados arriba. La estructura de un algoritmo evolutivo es la siguiente:

Algoritmo evolutivo

Inicia

Para $t=0$

Genera Población(t)

Evalúa Población(t)

En tanto no se cumple una condición

Inicia

$t=t+1$

Selecciona Población(t) de Población($t-1$)

Modifica Población(t)

Evalúa Población(t)

Termina

Termina

El algoritmo consta de las siguientes partes básicas:

- Genera población. Cuenta con un método para generar de manera aleatoria la población inicial de cromosomas, o individuos.
- Evalúa población. A cada uno de los individuos se les expone al ambiente, que está representado por la función objetivo del modelo. De esa exposición al ambiente se obtiene una medida de la aptitud de cada uno de los individuos.
- Selecciona población. Con base a la evaluación y de acuerdo a algún método de selección, se determina los miembros susceptibles de reproducirse o de pasar a la nueva población sin ser afectados.
- Modifica población. En esta etapa al individuo o el par de individuos seleccionados, se les aplica operadores genéticos de cruzamiento, mutación, reordenamiento, y pasan a integrar la nueva población. En esta etapa ocurre la reparación de los cromosomas, si es que los individuos perdieron alguna(s) característica(s) que los hace válidos.

2.5.1 ALGORITMOS EVOLUTIVOS.

Los algoritmos evolutivos son algoritmos probabilísticos, que mantienen una población de individuos, $P(t)=\{x_1^t, \dots, x_n^t\}$ para t iteraciones. Cada individuo representa una solución potencial al problema de que se trate, y se representa a este individuo como una estructura de datos S . Cada solución x_i^t es evaluada para dar alguna medida de su aptitud o desempeño. A partir de este punto se genera una nueva población, en la iteración $t+1$, al seleccionar los individuos más aptos. Algunos miembros de la nueva población son alterados, por medio de un operador genético, con el propósito de formar nuevas soluciones. Estas alteraciones pueden ser transformaciones binarias, cruzamientos, que creen nuevos individuos a partir de la combinación de dos o más individuos de la población previa; o pueden ser transformaciones unarias, como una mutación, que crea un nuevo individuo a partir de la alteración de una pequeña parte de un individuo de la población previa. Después de un cierto número de

generaciones el programa converge y se espera que los mejores individuos representen soluciones razonablemente cercanas a la óptima.

Los *AG* usan generalmente cadenas binarias de longitud fija y dos operadores genéticos básicos.

2.5.2 TERMINOLOGÍA.

Un algoritmo genético, al igual que cualquier otro programa evolutivo, tiene los siguientes componentes para resolver un problema:

- Una representación interna para las soluciones potenciales del problema.
- Un modo para crear una población inicial de soluciones potenciales.
- Una función de evaluación que toma el papel del ambiente, midiendo las soluciones de acuerdo a su aptitud frente a este medio.
- Proceso de selección de individuos con la aptitud adecuada para los procesos de cruzamiento y los efectuados por los operadores genéticos.
- Operadores genéticos que alteran la composición de las nuevas generaciones.
- Valores para los parámetros que los *AG* utilizan. Algunos de estos son:
 - Tamaño de población.
 - Probabilidad para aplicar los operadores genéticos.
 - Número de generaciones.
 - Etc.

Parte de la terminología se toma directamente de la biología. En el contexto de los *AG* estos términos se usan como analogía con lo que ocurre en la verdadera biología, considerando naturalmente que las entidades a las que se hace referencia son mucho más simples que sus contrapartes biológicas.

Los organismos consisten de células, y cada célula contiene el mismo conjunto de uno o más cromosomas, que son cadenas de ADN, cuyo propósito es conservar y en su momento transmitir las características del individuo. Un cromosoma se divide en genes, los cuales son los bloques que conforman el ADN, cada uno de los cuales codifica una proteína en particular.

De modo simple, uno puede pensar que un gen codifica un rasgo o característica, como el color del cabello. Las diferentes opciones para un rasgo (rubio, castaño, negro, etc.) son llamados alelos. Cada gen está localizado en una posición particular dentro del cromosoma.

Los organismos en general poseen múltiples cromosomas en cada célula. El conjunto completo de material genético, es decir el grupo de cromosomas completo, se conoce como genoma del organismo. El término genotipo se refiere a un conjunto particular de genes contenidos en un genoma. Si dos individuos poseen idénticos genomas, entonces tendrán el mismo genotipo. El genotipo proporciona al embrión y después es el motor de desarrollo del fenotipo del organismo, el cual consiste de las características físicas y mentales, tales como color de ojos, pelo, estatura, complexión, inteligencia, etc.

Los organismos cuyos cromosomas se encuentran como pares se llaman diploides, y aquellos que no son de este modo se llaman haploides.

Durante la reproducción, la combinación (o cruzamiento) del material genético ocurre: cada padre proporciona una parte de material genético, los genes se fraccionan dividiendo el par de cromosomas para formar un gameto, o cromosoma simple, ocurre que con el gameto proporcionado por cada padre puede crearse el conjunto completo de genes para formar un cromosoma diploide. La descendencia está sujeta a la mutación en la cual nucleótidos individuales, los elementos básicos del ADN, son alterados de padres a hijos, estos cambios frecuentemente son consecuencia de errores en el copiado de la información genética.

La aptitud de un organismo frecuentemente es definido como la probabilidad de un organismo de vivir hasta reproducirse, su viabilidad, o como función del número de descendientes que llega a tener.

En AG el término cromosoma se refiere a una solución potencial al problema de que se trate, el cual está codificado como una cadena de bits. Los genes son bits individuales o pequeños bloques de bits adyacentes que codifican una elemento que integra una solución potencial. Un alelo en una cadena de bits es 0 o 1; para alfabetos mayores una mayor variedad de alelos es posible para cada posición.

El cruzamiento comúnmente consiste del intercambio de material genético entre dos de los cromosomas de padres haploides.

La mutación resulta de la alteración de un bit de manera aleatoria en alguna posición elegida; para el caso de mayores alfabetos, el remplazo de un símbolo, en alguna posición seleccionada aleatoriamente, con un nuevo símbolo seleccionado de modo aleatorio.

La mayoría de las aplicaciones de algoritmos genéticos utilizan individuos haploides, en particular con un sólo cromosoma por individuo. El genotipo de un individuo en un algoritmo genético que usa cadenas de bits, es la configuración de los bit del cromosoma del individuo. Ocurre con frecuencia que no existe la noción de fenotipo en el contexto de los AG, aunque en los últimos tiempos hay quienes están trabajando tanto en el nivel del genotipo como en el del fenotipo.

2.5.3 ESPACIO DE BÚSQUEDA.

La búsqueda de entre una colección de posibles candidatos para encontrar la solución deseada es común en la investigación de operaciones y en las ciencias computacionales, por lo que tiene un nombre definido: la búsqueda se hace en un espacio de búsqueda.

El espacio de búsqueda contiene la noción del conjunto de posibles soluciones, así como de la distancia que existe entre ellas.

Un concepto también importante es el mapa o topología de aptitud. Esto fue originalmente definido por el biólogo Sewell Wright, en el contexto de la genética de las poblaciones, donde el mapa de aptitud es una representación del espacio de todos los posibles genotipos y de sus aptitudes [Mit96].

Supongamos, del modo más simple, que un genotipo es una cadena de bits de longitud l , y que la distancia entre dos genotipos es su distancia de Hamming, el número de posiciones en las cuales los correspondientes bits pueden ser diferentes. Supongamos también, que a cada genotipo se le puede asignar un valor de aptitud verdadero. Un mapa de aptitud puede ser delineado como un diagrama de $(l+1)$ dimensiones en el cual cada genotipo es un punto en las l dimensiones y su aptitud es colocada en a lo largo del eje $(l+1)$. El mapa para un genotipo $l = 2$ es sencillo de dibujar.

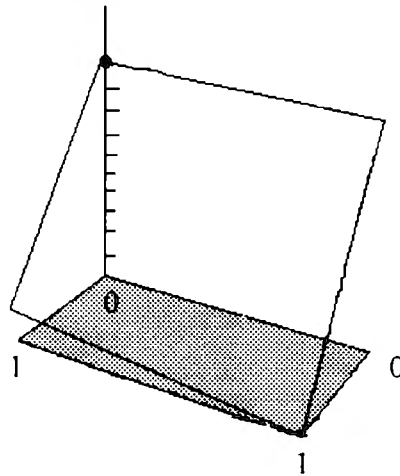


Figura 2.5.3.1 Un mapa de aptitud sencillo para $l = 2$. En este gráfico se representa el valor de la función de aptitud en el eje vertical, y el plano delimitado por los valores del cromosoma $f(00)$, $f(01)$, $f(10)$, $f(11)$ es el mapa de aptitud de este espacio.

Tales diagramas son llamados mapas porque los valores de la aptitud pueden formar valles, colinas y picos, y otras formaciones equivalentes a las que se observan en los paisajes físicos. Bajo la formulación de Wright, la evolución provoca que las poblaciones se muevan a lo largo del paisaje en caminos particulares, la adaptación puede verse como el movimiento de escalar un pico local. En este sentido, en los AG los operadores de cruzamiento y mutación se pueden ver como mecanismo que mueven a una población a través del mapa, definido por la función de aptitud.

Lo anterior no ocurre con poblaciones en el contexto de la biología, ya que no se le puede asignar una aptitud a un organismo independientemente de otros organismos en el mismo ambiente; es decir, las poblaciones cambian y las aptitudes de los genotipos también cambian,

esto significa que el mapa de aptitud esta ligado a los individuos que habitan el medio. Sin embargo, esta noción de mapa de aptitud es útil y de interés en el estudio de los AG.

2.5.4 IMPLEMENTACIÓN DE UN ALGORITMO GENÉTICO.

Aunque los AG han sido utilizados con éxito en un gran número de aplicaciones, también hay otros casos en donde su comportamiento es pobre [Mit96]. Al tener una aplicación que potencialmente puede hacer uso de ellos, ¿Cómo saber si los AG son un método a utilizarse?

Aún que no hay una respuesta rigurosa, hay algunas ideas útiles para saber si al usar AG hemos elegido el camino correcto:

- Acerca del espacio de búsqueda:
 - Es grande.
 - No está muy suavizado en su mapa de aptitud.
 - No es unimodal.
 - No es un espacio sobre el que se tenga un conocimiento detallado previo.
- Acerca de la función de aptitud:
 - Si está afectada por ruido.
 - Si la tarea no requiere obtener un óptimo global. Es decir, si al obtener una solución suficientemente buena se cumple el objetivo.

En el caso de que el espacio no sea muy grande, entonces se puede optar por una búsqueda exhaustiva, asegurando con esto que se tiene la solución óptima. Considerando que un AG podría converger en un óptimo local en lugar de hacerlo en la mejor solución global.

Si el espacio está suavizado o es unimodal, un algoritmo de gradiente ascendente tal como un *hill climbing*, el cual es mucho más eficiente que un AG en explotar los contornos del espacio de búsqueda.

Si del espacio de búsqueda se tiene un entendimiento previo, los métodos de búsqueda que usan dominios heurísticos específicos, pueden ser diseñados para ofrecer un mejor desempeño que los AG.

Si la función de aptitud contiene ruido, es decir, si involucra mediciones del mundo real en donde la función que describe la aptitud se ve afectada por errores presentes en la medición, aunque el problema fuese candidato a usar un algoritmo de hill climbing, este método podría desviarse del camino de manera irrecuperable. Pero los AG al trabajar por acumulación de estadísticas del desempeño sobre muchas generaciones, son lo suficientemente robustos como para soportar la presencia de una pequeña cantidad de ruido.

El problema que se desea resolver, de dar solución al diseño la topología de una red, el espacio de búsqueda es grande, y el problema esta planteado como NP completo, por lo que no se puede considerar la realización de una búsqueda exhaustiva aún en el caso de una cantidad pequeña de nodos, de 10 o menos.

2.5.4.1 Codificación.

La codificación binaria es un método común de codificación, algunas razones son:

- La razón histórica. Holland y sus estudiantes se concentraron en este tipo de codificación, en la práctica esto sigue siendo el método más utilizado.
- Una gran parte de la teoría actual de los AG esta basada en el hecho de considerar longitudes fijas en los cromosomas y codificación binaria. Naturalmente que la teoría puede extenderse al uso de codificaciones no binarias, pero esto no ha quedado plenamente desarrollado como en la teoría original.
- La heurística acerca de la parametrización adecuada, como en el caso de cruzamiento y mutación, ha sido desarrollada considerando la codificación binaria, aunque existen generalizaciones.

Existen métodos que son extensiones al uso directo de la codificación binaria, por ejemplo el uso del *código gray* [Mic92], donde lo que se pretende es contar con una distancia de Hamming mínima, entre elementos adyacentes; la codificación diploide [Mic92], que hace diferencia entre elementos del alfabeto para designar algunos alelos como dominantes o recesivos.

Holland da una justificación teórica al uso de la codificación binaria [Hol92], él indica que al comparar dos codificaciones con la misma capacidad de contenido de información, una con un pequeño número de alelos y cadenas de caracteres de gran longitud y la otra con un mayor número de alelos y menor longitud en la cadena de caracteres, la primera trae implícito un mayor paralelismo que la segunda; ya que si la primera es de longitud 100 y la segunda de longitud 30, la primera tendrá un mayor número de patrones de similitud, llamados esquemas, que la segunda. Será del orden de 2^{100} para la primera cadena y 2^{30} para la segunda. De acuerdo al argumento de Holland, la codificación binaria permite un mayor grado de paralelismo implícito que el uso de alfabetos de mayor cantidad de caracteres, debido a que una instancia de una cadena binaria contiene más esquemas que una instancia del segundo tipo. Por otra parte Mitchell [Mit96] indica que el argumento de Holland del *conteo de esquemas* es relevante en AG únicamente si el análisis de esquemas es relevante.

No obstante las ventajas mencionadas para el uso de la codificación binaria, este tipo de codificación, no resulta natural ni manejable para muchos problemas, por ejemplo, la representación con caracteres múltiples para la gramática de generación de grafos de Kitano, la representación con valores reales para los pesos de redes neuronales de Montana y Davis [Mit96].

El problema, en su planteamiento, requiere utilizar un alfabeto dependiente del número de nodos que se desean incluir en la red, por lo que crecerá conforme crezca la red. Lo anterior permite modelar adecuadamente los identificadores de los nodos de la red. Podemos identificar a los esquemas con los alelos, que son las etiquetas que identificarán a cada uno de los nodos de la red.

El propósito último de la codificación es elegir un alfabeto y un método para traducir cada instancia del problema en un cromosoma, y de este modo hacer uso de esta representación a través de los operadores genéticos que se hayan seleccionado.

2.5.4.2 Métodos de selección.

Una vez determinado el método de codificación, la segunda decisión está en la determinación del proceso de selección, es decir, resolver el cómo elegir a los individuos que crearán la descendencia para la siguiente generación, y cuanta descendencia podrá generar cada uno de estos individuos.

El fin de la selección es enfatizar la mejor aptitud de la población actual, esperando que este proceso origine una población con una mayor aptitud promedio. La selección tendrá que balancear los procesos de cruzamiento y mutación, ya que cada uno de ellos representa un modo de búsqueda de una mayor aptitud en la población. En tanto el cruzamiento desea explotar las características adquiridas por los individuos de la población, a través del intercambio de características que heredará la nueva población; la mutación provoca la búsqueda de características diferentes a las que actualmente posee la población, esperando que de algunas de estas alteraciones se genere alguna característica que mejore la aptitud de la nueva población.

Si la selección hace énfasis en la inclusión de los mejores individuos de la población evaluada, cuyos miembros seguramente poseen una aptitud que no es aún cercana a la óptima, se reducirá rápidamente la diversidad de la población la cual es necesaria para lograr los cambios y un mayor progreso en la población, a este problema se le denomina convergencia prematura.

Por el contrario, si se disminuye más de lo debido la participación de estos miembros que fueron calificados con la mejor aptitud, resultará en una muy lenta evolución de la población.

Algunos de los métodos más comunes son [Mit96] [Gol89a]:

- Selección proporcional a la aptitud o método de la ruleta.
- Muestreo universal estocástico.
- Muestreo estocástico con y sin remplazo.
- Muestreo estocástico con remanente y con y sin remplazo.
- Torneo estocástico o método de asignación de rango.
- Elitismo.
- Selección en base a métodos de escalamiento.

A continuación se describen brevemente algunos de los métodos mencionados.

Selección proporcional a la aptitud.

En el trabajo original de Holland se hizo uso de un método de selección proporcional a la aptitud, en el cual el número de veces que un individuo podía ser seleccionado para reproducirse, está dado por su aptitud individual dividida entre el promedio de aptitud de la

población. El método más común para implementar esta forma de selección se conoce como el método de la ruleta, que en resumen consiste en:

- A cada individuo se le asigna una porción o rebanada de una ruleta.
- El tamaño de la rebanada es proporcional a su aptitud.
- La rueda se hace girar n veces, donde n es el tamaño de la población.
- En cada oportunidad en que se hace girar la ruleta el individuo indicado por la misma es seleccionado para pertenecer al grupo de padres para la siguiente generación.

Este método fue modificado por James Baker [Mit96], quien propone un método de muestreo alternativo llamado muestreo universal estocástico, el cual es un variante del anterior, en el cual se garantiza un mínimo de oportunidad de reproducción a cada individuo y se limita la oportunidad de los más aptos.

Los métodos anteriores junto con otros de muestreos estocásticos, con remplazo y sin remplazo, en los que se mantiene la proporcionalidad entre la probabilidad de selección de un individuo y su aptitud presenta un problema al inicio del proceso de evolución. Con las primeras poblaciones ocurre que la varianza de la aptitud de la población es alta y un número reducido de individuos son mucho más aptos que la mayoría. Bajo estas circunstancias, y con estos métodos, estos individuos de mayor aptitud y sus descendientes se multiplicarán rápidamente en la población, teniendo como consecuencia la disminución de las posibilidades del AG de hacer exploración por la reducida variedad de individuos. A esto se le llama convergencia prematura.

Elitismo.

Fue introducido por Kenneth De Jong, y es una adición a muchos métodos de selección, obligando al AG a retener algunos de sus mejores individuos en cada generación. Tales individuos si no llegaran a ser seleccionados para la reproducción, o si son destruidos por causa del cruzamiento o mutación. Muchos investigadores han confirmado que el elitismo mejora significativamente el desempeño del AG [Gol89a] [ZF97].

Selección en base a métodos de escalamiento.

Para evitar los problemas causados por la selección proporcional, se hace uso de algún método que permita transformar en una escala los resultados de la función objetivo, que evalúa a cada individuo, con el propósito de transformar los valores mencionados en una escala que atenúe el efecto de los valores altos de aptitud y haga al proceso menos susceptible de caer en una convergencia prematura. Estos métodos son ampliamente aceptados en la práctica, conservando niveles apropiados de competencia a lo largo de una simulación. Estos métodos incluyen a los siguientes [Mic92]:

- Lineal.
- Sigma o truncamiento sigma.
- La regla de la potencia.

Escala lineal.

Se calcula la f' de los valores de la aptitud, los cuales se generaron al evaluar la función objetivo para cada individuo, usando una función lineal de la forma:

$$f' = af + b$$

En esta función los coeficientes se determinan para cumplir con:

- Mantener la proporcionalidad entre los valores de la aptitud, obtenidos de la función objetivo.
- Lograr que el valor máximo reflejado en la escala sea un múltiplo específico, comúnmente dos, de la aptitud promedio.

Estas dos condiciones aseguran que los miembros promedio de la población reciban una oportunidad de tener descendencia y los mejores reciban el múltiplo especificado, como número de oportunidades de obtener descendencia. Un cuidado especial se debe tener con el cálculo del escalamiento para prevenir la aparición de valores negativos en la escala.

Este procedimiento trabaja bien excepto cuando se generan aptitudes negativas. Esto ocurre normalmente cuando ya está avanzado el proceso de evolución y muchos de los individuos tienen una alta aptitud, en tanto existe todavía individuos con aptitudes muy bajas.

Truncamiento sigma.

Para evitar el problema descrito para la escala lineal, Forrest [Mit96] sugiere el uso de la varianza poblacional para realizar un procesamiento inicial a los valores obtenidos directamente de la función objetivo, antes de hacer la transformación en la escala. Este procedimiento, al cual se le llama truncamiento sigma, usa la información de la desviación estándar poblacional para abstraer un valor constante de los valores de la aptitud:

$$f' = f - (\bar{f} - c \cdot \sigma)$$

En esta ecuación la constante c se elige como un múltiplo adecuado de la desviación estándar (entre 1 y 3) y los resultados negativos son colocados arbitrariamente como 0. A continuación del truncamiento sigma, el escalamiento lineal de la aptitud puede proceder sin el peligro de resultados negativos.

Escala de la regla de la potencia.

Gillies [Mit96] sugirió la regla de la potencia como una forma de escala, donde la escala de aptitud se obtiene al aplicar una potencia especificada a la aptitud que se obtiene de la función objetivo [Mic92]:

$$f' = f^k$$

En algunos estudios del autor referido, él obtuvo valores para $k = 1.005$; sin embargo, en general k es dependiente del problema y podría requerir ajusten al momento de estar ejecutando el algoritmo, ya sea para estrechar o ampliar el rango.

2.5.4.3 Operadores genéticos.

La tercera decisión a tomar en el proceso de implementar un AG es la elección de los operadores genéticos a usar. Esta elección depende principalmente de la estrategia de codificación que se haya elegido.

Los operadores básicos son los de cruzamiento y mutación, aunque algunos investigadores [Mic92] sugieren otros para el tratamiento de características específicas de alguna población.

- Cruzamiento.
 - Con un único corte.
 - Con múltiples cortes.
 - Cruzamiento uniforme.
- Inversión y reordenamiento.
- Mutación.

Cruzamiento.

Puede decirse que la principal característica que distingue a un AG es el uso del cruzamiento. El modo más sencillo de hacerlo es con un solo corte sobre alguna posición seleccionada de manera aleatoria y con las partes obtenidas de los padres, intercambiarlas para formar dos descendientes. Esta tarea va dirigida a combinar los esquemas de diferente manera en la cadena que compone al cromosoma.

En el cruzamiento de 2 cortes, y en cualquier otro cruzamiento múltiple, los cromosomas se pueden visualizar como una cadena de caracteres circular, en donde los extremos se unen para formar un círculo. El segmento que se intercambia es el que se localiza entre los dos puntos de corte.

El cruzamiento uniforme es diferente al que hace uso de cortes, ya que utiliza una máscara o plantilla que indica cuáles alelos han de transferirse del padre y cuáles de la madre, esta plantilla es generada aleatoriamente para cada pareja seleccionada.

Algunos otros tipos de cruzamiento como, el de comparación parcial o PMX (por partially matched crossover), descrito por Goldberg [Gol89a] son usados en problemas basados en el orden, en donde la posición de los alelos es determinante al momento de calcular la aptitud de los cromosomas.

Inversión y reordenamiento.

Un modo alternativo y natural de obtener un mayor potencial de la información contenida en un cromosoma, considera la alteración de la posición de los alelos en las cadenas de caracteres que componen al cromosoma. Goldberg [Gol89a] hace una descripción de estos mecanismos, llamados operadores de reordenamiento.

Uno de ellos es la inversión, con este operador se elige dos puntos de corte en el cromosoma, y los alelos contenidos entre esos puntos de corte, invierten su posición intercambiando el primer alelo con el último, el segundo con el penúltimo, y así sucesivamente.

Otra forma de ejecución del reordenamiento es asignando al extremo izquierdo o derecho del cromosoma el primer punto de corte y el otro punto se determina a una distancia menor a la mitad del cromosoma, a partir del primer corte. Este método de inversión, descrito por Franz, D. R. (1972), intenta minimizar la tendencia a afectar mayormente a los alelos localizados cerca del centro con respecto a los que se encuentran en los extremos.

Mutación.

Se considera al cruzamiento como el mecanismo que permite la variación e innovación dentro de los AG [Hol75], viendo a la mutación como el proceso que asegura a la población contra la fijación en alguna posición de cualquier alelo, trabajando no como mecanismo de exploración sino como motivador de cambio.

Como ya fue comentado en otra sección, la mutación consiste en la alteración de la codificación de un alelo, cambiándolo por algún otro valor seleccionado aleatoriamente entre los definidos dentro del alfabeto.

2.5.4.4 Parámetros de un algoritmo genético.

La última decisión que se toma en la implementación de un AG es la definición de los valores de los varios parámetros, tales como:

- Tamaño de población.
- Probabilidad de cruzamiento.
- Probabilidad de mutación.

La interacción de estos parámetros no es lineal, de modo que no es posible optimizarlos uno por uno.

No hay acuerdo general acerca de la magnitud ideal de estos parámetros, ya que los valores que se aplican en cada caso, son el resultado de las pruebas realizadas.

De Jong [Mit96] desarrollo un estudio sistemático de cómo la variación de los parámetros afectaba al AG, en el desempeño de la búsqueda en línea y fuera de línea. Entendiendo que el desempeño en línea en el tiempo t es el promedio de aptitud de todos los individuos que han sido evaluados en t etapas de evaluación, digamos en la t generación. En tanto el desempeño fuera de línea en el tiempo t es el valor promedio, en las t generaciones, de las mejores aptitudes que han ocurrido en cada generación. Los experimentos de De Jong indican que:

- El mejor tamaño de población estuvo entre 50 y 100 individuos.
- La mejor probabilidad de cruzamiento fue 0.6.
- La mejor probabilidad de mutación fue 0.001.

Estos valores han sido ampliamente usados en la parametrización de AG, aún pensando que no hay garantía de que son los más adecuados en problemas que no son parecidos a los usados por De Jong en sus pruebas.

Sin embargo y considerando la experiencia generada a través de las pruebas, es posible que los parámetros sugeridos por este autor y otros, no sean los adecuados para lograr los mejores resultados en un determinado problema. Algunas condiciones del ambiente que afectan la parametrización del AG son:

- El tamaño del espacio de búsqueda.
- Las características del mapa de aptitud.
- La longitud el cromosoma.
- La composición del cromosoma.
- El uso de un alfabeto diferente al binario.

Como vemos, es posible que los valores sugeridos anteriormente sean los adecuados para un cierto entorno, y que naturalmente en problemas similares, o de entorno parecido, sirvan de referencia y proporcionen un punto de arranque para la correcta ejecución del AG.

Por lo cual los parámetros mencionados, y otros relacionados con la estrategia de selección (elitismo y su porcentaje con respecto a la población), el método de escalamiento, deban ser modificados de acuerdo a los resultados obtenidos durante las pruebas del AG de que se trate.

3.0 PLANTEAMIENTO DEL PROBLEMA.

Dentro de la definición del problema es necesario considerar los objetivos de la organización que hará uso del recurso, y de las partes del sistema que deberán de ser analizadas. El impacto que una red tiene en una organización, puede ser observado al notar las modificaciones que ocurren entre las entidades que componen la misma, incluyendo las nuevas dinámicas que se establecen conforme se asimila el uso de esta herramienta. Existen organizaciones que han hecho de la red de comunicaciones la herramienta básica de su negocio, en tanto algunas otras aún la consideran como una nueva tecnología que beneficiará a su negocio con su sola implantación sin visualizar las modificaciones, que a las operaciones actuales de la compañía, ocurrirán una vez que se intensifique su uso.

3.1 EL DISEÑO DE RED DE TELECOMUNICACIONES Y EL PROCESO DE OPTIMIZACIÓN.

3.1.1 LA ORGANIZACIÓN.

Las redes se establecen, crecen y se desarrollan en función de las necesidades de los usuarios. Naturalmente, estas necesidades corresponden con las demandas de información que surgen en las diferentes entidades de una institución.

El trabajo de crear y construir los canales de comunicación que servirán de medio para transferir información dentro y fuera de la institución se efectúa, en muchos de los casos en los que he participado, pensando en el costo y la función básica del equipo y las instalaciones, y poco o nada se piensa en el tráfico que estará presente en la red. Y por supuesto, el retraso en la red y la saturación de la misma quedan fuera de toda consideración.

En general, creo que este panorama es común en el medio de diseño e instalación de redes y esto puede tener alguna de las siguientes causas:

1. Ignorar que deben de considerarse parámetros como el tráfico y el retraso de señal en la red.
2. Desconocer que, bien dimensionados, estos parámetros pueden incrementar el valor futuro de la red al requerir un menor mantenimiento y posiblemente un periodo mayor de vida para la instalación original, permitiendo un ahorro en el mediano y largo plazo para los propietarios de la red.
3. Considerar estos parámetros durante la etapa de diseño agrega complejidad a este proceso y deja de ser algo sencillo de llevar a la práctica. Para estar en condiciones de incluir estos parámetros es necesario dimensionar la red, considerando el número de usuarios en el corto y mediano plazo, las cargas de trabajo prevista, en función de las actividades hacia el interior del departamento de que se trate e interdepartamentales.

Y en buena medida es bueno pensar que el proceso de diseño e implantación debe ser un proceso simple, ya que para organizaciones muy pequeñas, quizás menos de 50 estaciones, o 3 o 4 *LAN'es* interconectadas, los equipos estándar con software estándar (redes Ethernet 100baseT, con aplicaciones de oficina) no alcanzarán el límite de la capacidad de una LAN o no se creará saturación en la interconexión entre las mismas.

En el caso de organizaciones que incluyen varios sitios interconectados, donde en cada localidad hay varias *LAN'es*, en número mayor a esas 3 o 4 mencionadas arriba, la necesidad de interconectar eficientemente estas localidades, además de considerar el flujo eficiente de información entre los usuarios en una misma localidad, hace necesario prever la posibilidad de que esta organización llegue a tener problemas de saturación en las redes, y no sólo en la interconexión, sino hacia el interior de una o más de las localidades.

3.1.2 LOS USUARIOS Y LA RED.

Tenemos dos tipos de entidades que conforman el ambiente de una red. Una de ellas es la red misma, incluyendo sus componentes de software y hardware; y la otra son los usuarios, que en general los identificamos como quienes utilizan la red para efectuar parte o la totalidad de sus actividades.

De las características de los usuarios en general se puede decir poco, ya que su composición es característica de la organización de que se trate. Y aún dentro de organizaciones complejas donde se han estratificado y delimitado las actividades de cada componente del sistema, su composición resulta muy heterogénea. Si hablamos de una compañía automotriz, encontraremos diferentes niveles de uso de la red, involucrando también características que no son materia de este trabajo, pero que formarán parte del diseño de una red, tales como la seguridad, confiabilidad y nivel de especialización de los grupos (*LAN'es*) y de las estaciones de trabajo.

Podemos dividir las redes de telecomunicaciones en dos tipos básicos: redes conmutadas y no conmutadas. Las redes no conmutadas permiten el envío de información a través de enlaces establecidos de manera permanente, con ancho de banda definido. También son no

conmutadas las redes que sirven a sistemas centralizados que distribuyen información a partir de un nodo central. Este tipo de red no se contemplan en el presente trabajo.

Aún entre las redes conmutadas, las que usan el proceso de conmutación de circuitos no son consideradas como parte del problema a tratar en este documento, debido a que su forma de operar considera una asignación diferente de la capacidad de los enlaces, creando circuitos que tendrán asignaciones fijas de ancho de banda, en tanto la comunicación entre usuarios de esta red permanezcan conectados. El uso de estas redes, como ya se mencionó en los antecedentes, está dirigido a voz, los equipos de conmutación e interconexión son *PBX*.

El tipo de red que es materia de este trabajo es la conmutada que procesa paquetes, la cual permite a cualquier usuario comunicarse con algún otro a través de rutas establecidas para este propósito, y que ocupan temporalmente parte del ancho de banda de los enlaces entre grupos o nodos de la red.

El presente trabajo será para modelar la topología de una red típica, integrada por un cierto número de LAN'es, como grupos o nodos, y por estaciones de trabajo, que requieren del servicio de la red para transportar información entre ellas, dentro de una sola LAN o a otras estaciones que pertenecen a grupos o nodos remotos.

Cuando se desea instalar una red dentro de una organización, la primera clase de red que se establece es una LAN. Las topologías de red de área local más comunes son las de anillo, como el de una red token ring, 802.5 de acuerdo al IEEE; o la de bus, como es el caso de la red ethernet 802.3 de acuerdo al IEEE. El siguiente paso es hacer crecer la red o establecer nuevas redes en las diferentes instalaciones de la institución.

Con el crecimiento de las empresas y sus necesidades de intercambio de información, la infraestructura de comunicaciones tenderá a desarrollarse. Lo que lógicamente ocurre es unir las diferentes redes.

En vez de simplemente unir las redes existentes, lo deseable es distribuir los recursos de cómputo de modo que los enlaces que se establezcan entre las diferentes LAN'es sean correspondientes con las expectativas de una red funcional en cuanto:

- La obtención del mínimo retardo en la señal entre cualesquiera dos nodos.
- La optimización del ancho de banda contratado para los enlaces, sin afectar la capacidad de los enlaces para soportar el volumen de datos que se desea transferir.
- El disminuir el costo de los enlaces y de la instalación total, de modo que el diseño de la red pueda soportar la operación proyectada del negocio en el mediano, en el largo plazo, o cuando menos en un plazo que resulte rentable para el negocio.

3.1.3 EQUIPOS DE INTERCONEXIÓN.

La conexión entre las diferentes LAN'es se hace comúnmente con puentes transparentes (transparent bridges) [Sta97]. Aunque su uso a disminuido hasta prácticamente desaparecer, han sido substituidos por equipos que cumplen con la misma función, más otras relacionadas con la administración de la red. Asumo para este trabajo, que el proceso de interconexión

realizado por un enrutador es equivalente al de un puente transparente, durante el proceso de transferencia de información entre nodos. Por otra parte, la operación de un puente transparente es diferente a la de un repetidor, el cual opera en la capa Física del modelo de OSI y se usa para extender enlaces entre subredes [Sta97].

Los puentes operan en la Capa de Enlace, en la subcapa MAC (Media Access Control, Control de Acceso al Medio) [Tan91], de modo que los protocolos de las capas superiores son transparentes para la operación de estos equipos. Además, estos equipos permiten restringir las cargas de tráfico, el acceso a los usuarios, ayudan a limitar el problema en caso de fallas, y permiten extender las redes más allá de las limitaciones técnicas existentes. Una condición a considerar es que la red debe de construirse sin rutas múltiples entre nodos, esto para evitar la incertidumbre que se crea al intercalar segmentos que arriban al mismo destino por las diferentes trayectorias, ocasionando duplicación indefinida de paquetes y degradando el rendimiento de la red. Esto limita la topología de la red a una topología de árbol.

Naturalmente hay interconexiones que salen del entorno de este trabajo pero que con una modificación adecuada al modelo pudieran incluirse, tal es el caso de las redes con conexión *frame relay* entre nodos. En este caso la asignación de capacidad en la red es dinámica y tiene un comportamiento diferente al que se considera en el modelo a tratar, sobre todo porque la interconexión a la red pública se hace a equipos que conforman una nube de red VAN, en la cual los supuestos de interconexión son diferentes.

3.1.4 TOPOLOGÍA DE LA RED.

Al estar integrada la red por computadoras agrupadas en *LAN'es* (cluster, racimo o grupo) y estos grupos conectados por enlaces, el problema del diseño de la topología de la red esta determinado por:

- Agrupamiento (clustering).
- Enrutamiento (routing).

El problema de agrupamiento parte de las siguientes preguntas:

- ¿En cuántos segmentos se dividirá la red?
Esto significa determinar el número de grupos, LAN'es, en que se dividirá la red en el proceso de diseño.
- ¿Quiénes integrarán cada grupo?
Lo que implica conocer los criterios que determinarán el agrupamiento de las computadoras (o usuarios de la red) en los diferentes segmentos en que se dividirá el total de computadoras, aceptando que cada computadora puede pertenecer a un sólo segmento.

Los factores de mayor importancia en el diseño de una red son:

- El tráfico esperado entre las diferente localidades que conectará la red.
- El costo de los enlaces y su instalación.
- El costo de los equipos de comunicaciones.
- El retraso ocurrido entre la transmisión y recepción de los mensajes en la red.

El primero de los factores nos indica tanto el volumen como tipo de mensajes a transmitir, sean estos, simples datos, o voz, o video. Cada uno de los cuales presenta requerimientos especiales a la red. Yendo desde los procesos en lotes (*batch*) a los procesos de transmisión de imágenes en tiempo real.

La determinación del primer factor nos lleva en cascada, tanto a la determinación del tipo y ancho de banda de los enlaces que han de usarse. Como a determinar los equipos de comunicaciones que se instalarán.

El último de los factores no siempre es valuado por la dificultad inherente a determinar un valor esperado del retraso de señal para una red que no ha sido instalada; y cuando se hace una estimación, es usual que sea basada en la experiencia obtenida de instalaciones previas. Solamente he observado un esfuerzo por valuar el desempeño de redes, cuando su instalación será muy costosa y de elevado desempeño[Ers93] [Syk92] [Sin96].

Si consideramos que al diseñar una red que cumpla con los requerimientos establecido para el retraso de la señal y su diseño también considera el crecimiento futuro del tráfico y su relación con el retraso de señal esperado, entonces la inversión en equipo y enlaces puede hacerse sobre una base más objetiva y pensando en el mediano y largo plazos.

Si nuestro interés está centrado en proporcionar una red con el mínimo retraso posible en la información que fluye por ella, entonces el problema del enrutamiento queda delimitado a la determinación de los enlaces de interconexión entre los diferentes segmentos en una topología de árbol.

Como lo comente antes, es posible cambiar el parámetro de la red en función de la cual se desea optimizar el diseño de la misma, para esto es necesario definir dentro de la función objetivo, la forma en que dicho parámetro afecta la operación general de la red y el modo en que su variación determinaría el costo de operación de la red propuesta.

Lo que se propone como objetivo de diseño es minimizar el retraso promedio entre computadoras al considerar la capacidad, y por las consideraciones hechas, disminuir el costo presente y futuro de los enlaces y equipos instalados.

El problema de optimización a tratar en este trabajo tiene dos vertientes: Una es resolver el problema de agrupamiento y de las rutas de menor costo, que se establecen entre los grupos formados al reunir las estaciones y la segunda es el enrutamiento del flujo de datos, de acuerdo a la carga prevista.

El primer caso es de agrupamiento de estaciones, y lo que se desea es determinar la forma más conveniente de dividir el total de estaciones de modo que el flujo de datos hacia el exterior de cada grupo sea minimizado, este problema es conocido como la "división de un grafo"(graph partition) y es un problema NP- Completo [Fed99].

La división de un grafo dentro de k partes A_1, A_2, \dots, A_k con un patrón fijo de requerimientos como el que A_i es disperso o denso, y los pares A_i, A_j son completamente no adyacentes o completamente adyacentes. Y estos requerimientos pueden ser codificados en una matriz M de $k \times k$, y colocamos en cada posición ij un valor que represente los requerimientos. A esta partición la podemos llamar una M -partición [Fed99].

Problemas como el de un grafo k -coloreado, es una M -partición donde M es la matriz de adyacencia del k -grafo completo, y más generalmente, un H -coloreado (homomorfismo de un grafo estático H) es una M -partición, donde M es la matriz de adyacencia de H . Si H es bipartita, es decir, si pensamos en dividir el grafo en dos partes, el problema tiene solución en tiempo polinomial, en cualquier otro caso el problema es NP-Completo [Fed99].

El problema de fragmentar el total de estaciones en un número k de grupos, en la práctica se efectúa pocas veces, debido a que en la mayoría de los casos las estaciones se dividen de acuerdo a los grupos de trabajo formales (Departamentos) o de acuerdo a las características físicas del ambiente en el que se situarán. Es decir, si no hay una división formal, será la separación física natural la que creará los diferentes ambientes.

Sólo en caso de tener una gran concentración de estaciones en una única área física, es necesario considerar la división de este núcleo en un número determinado de grupos. Esto ocurre, en plantas industriales que requieren un uso extensivo de equipos conectados en red y donde se necesita la interconexión de los mismos. Se pide la disponibilidad permanente de la red en horarios de operación, y en general se conoce la distribución de las cargas de transferencia de información entre equipos, esto último define prácticamente la conformación de los grupos que integrarán cada nodo. Un caso similar es el de oficinas que han ido creciendo dentro de un mismo edificio, y al mismo ritmo se extiende la red, usando para este propósito *switches* como equipo de interconexión entre segmentos; lo anterior, permite incrementar el número de equipos interconectados en una cierta cantidad de grupos, pero sin distinguir en el segmento de enlace, entre los diferentes grupos, por lo cual realmente funcionan como una sola red.

3.2 ARBOL DE EXPANSIÓN MÍNIMA (MST).

El segundo caso, visto de modo simple es un problema de árbol de expansión de costo mínimo. Este es un problema que tiene una complejidad que se encuentra entre $O(n \log n)$ y $O(m + n \log n)$ dependiendo del algoritmo que se elija para solucionarlo. Pero al agregar restricciones al problema, tal como el grado máximo del árbol o la capacidad limitada en el flujo a través de los nodos, el problema se convierte en NP-Hard o NP-Completo, respectivamente [Gen97] [ACG99]. Y en ambos casos se puede considerar conveniente el uso de AG's para su solución.

Hay problemas que lo único que pretenden es localizar rutas de costo mínimo entre nodos, para los cuales, existen algoritmos conocidos que ofrecen soluciones rápidas, tal como Dijkstra.

El árbol de expansión de costo mínimo (Minimum Spanning Tree, MST) fue formulado por primera vez por Boruvka en 1962. Desde entonces, la formulación de este problema ha sido aplicada en muchos problemas de optimización combinatoria, tal como problemas de transporte, diseño de topología redes de telecomunicaciones, sistemas de distribución, y otros de esta clase. La solución de estos problemas se da en tiempo polinomial con algoritmos como los desarrollados por Kruskal, Prim, y Sollin.

Sin embargo, existen problemas derivados que presentan mayor interés por las características de su formulación, estos son llamados extensiones al problema MST, algunos son [Gen97] [ACG99]:

- MST con Restricción de Grado (Narula y Ho).
- MST Probabilístico (Bertsimas).
- MST Estocástico (Ishii)
- MST Cuadrático (Xu).
- MST Generalizado (Myung, Lee y Tcha).
- Árbol de Steiner de Expansión Mínima, (Gilbert).
- MST con Restricción en la Capacidad (Garey).

La mayoría de estos problemas se clasifican como NP-Hard, excepto el último, que se clasifica como NP-Completo.

La clasificación de MST con restricción en la capacidad como NP-Completo, está mencionado en [Gen97] [ACG99] [ES96] y en ninguno de estas referencias existe una demostración formal de ésta clasificación, ni tampoco tuve acceso a las referencia que estos autores hacen al momento de catalogar el problema como NP-Completo.

El problema a tratar será obtener una solución para el problema de MST con restricción en la capacidad.

3.3 CRITERIOS DE DISEÑO DE UNA RED BASADA EN UN ÁRBOL DE EXPANSIÓN MÍNIMA.

3.3.1 Componentes.

Los criterios descritos en esta sección fueron desarrollados [ES96] para el diseño de redes LAN, y por la forma en que se describen las redes WAN en este trabajo es posible usar criterios paralelos.

Consideremos una red que conecta a N usuarios (o computadoras, o estaciones de trabajo) y el tráfico entre los usuarios estará dado por una matriz A de $N \times N$; un elemento a_{ij} de esta matriz, representa el tráfico entre el usuario i y el usuario j llamada Matriz de Tráfico entre Usuarios.

$\alpha_{ij} \in A, A: N \times N$
 α_{ij} = Tráfico del usuario i al usuario j
 A = Matriz de tráfico entre usuarios

No se considera la forma en que se hace la estimación del tráfico entre cualesquiera dos nodos. El tráfico presente en cualquier par de nodos puede presentarse en ráfagas (intervalos con picos de tráfico) o flujo estable o *metaestable*. Y en cada uno de los casos anteriores la forma de estimar el tráfico será diferente. Por tanto consideraré que esa estimación ya fue hecha y el valor asignado a α_{ij} , como una estimación correcta.

Las computadoras estarán agrupadas en P segmentos o grupos (*LAN'es*). Por tanto una matriz de $N \times P$ indicará cual usuario pertenece a cual grupo.

$r_{ij} \in R, R: N \times P$
 R = Matriz de grupos

donde

$$\forall i \in N, \quad \sum_{j=1}^P r_{ij} = 1$$

Lo anterior se concluye al considerar que cada usuario sólo puede pertenecer a un sólo grupo.

$$r_{ij} = \begin{cases} 1 & \text{si el usuario } i \text{ esta en el grupo } j \\ 0 & \text{en cualquier otro caso} \end{cases}$$

Se define también una matriz S de $P \times P$ donde cada elemento s_{ij} representa tráfico de los usuarios del grupo i con destino en usuarios que pertenecen al grupo j . La trayectoria que une los dos grupos, puede considerar un sólo puente si está conectando directamente, a ambos grupos. En caso de no existir una conexión directa la trayectoria ira a través de más de un puente, y uno o más grupos.

$s_{ij} \in S, S: P \times P$
 S = Matriz de Tráfico entre grupos.

Donde se cumple

$$S = R^T A R$$

Esto se puede explicar del siguiente modo:

AR es una matriz que contiene tanto el tráfico hacia el interior de cada grupo, que es tráfico que no trasciende los límites del grupo, y el tráfico que va del exterior del grupo hacia algún miembro del grupo.

R^T es la matriz transpuesta que invierte la presentación de grupos y usuarios en la matriz.

R^TAR es la matriz que presenta separada la información del tráfico, ya que sólo refleja el tráfico que se establece entre grupos.

Definidas las relaciones que existen entre usuarios, entre grupos, la localización de los usuarios en los grupos y la conexión entre los grupos, queda por definir la forma en que se medirá el desempeño de la red, para ser usado en el proceso de optimización de la red.

En el proceso de diseño se determinará precisamente el número de grupos (segmentos) en los que se dividirá la red, la ubicación de los usuarios en los diferentes segmentos, y los enlaces a establecerse entre los diferentes grupos.

3.3.2 Medición del desempeño de la red.

Dentro de los modelos que de manera simple nos permiten calcular el comportamiento del retraso promedio de la señal en una red, se encuentra la cola de espera M/M/1 en equilibrio, [Tan91], la cual nos describe el comportamiento de un grupo de manera individual.

El retraso promedio D_k de un bit de datos enviado a través del grupo k está dado por

$$D_k = \frac{1}{C_k - L_k}$$

Donde

C_k es la capacidad del segmento de red (bits/segundo).

L_k es el tráfico, bits a transferir, que ese segmento tiene como carga (bits/segundo).

La aplicación de la teoría de colas a un canal de comunicación se da desde Kleinrock en 1964 [Tan91] [Kle64]. En el caso de una red se plantea el problema de que los canales de comunicación no están aislados, la salida de una cola es entrada de otro, siendo posible que varias salidas converjan en una sola. Este proceso de entrada se considera como la suma de las medias de los procesos alimentadores, por lo que sigue siendo similar al hecho de considerar un sólo canal alimentador.

Se ha demostrado [Jac63] [Kle75] que una red abierta de colas M/M/1 puede ser analizada como si cada una estuviese aislada de las demás, lo único que es necesario conocer es la velocidad media de entrada.

Una dificultad adicional que se presenta es que los paquetes en una red, al momento de generarse, sus longitudes dependen de la cantidad de datos que estén transportando, es decir se generan con diferentes longitudes, naturalmente las conservan a lo largo de todo su recorrido. Esta propiedad introduce correlaciones no aleatorias en el sistema. Si llega a generarse un paquete de longitud máxima (*MTU* máximo) tardará mayor tiempo en ser atendido que el paquete promedio, esto produce una interrupción en el patrón de llegadas de la cola que alimenta la salida. Kleinrock [Kle64] establece la hipótesis, conocida como hipótesis de independencia, donde indica que cada vez que un paquete llega a un nodo, pierde su identidad y se le selecciona al azar una nueva longitud.

3.3.3 Puente transparente.

Un puente es el elemento de la red que conecta físicamente dos segmentos de la red. El puente detectará los paquetes que deban de ser transferidos más allá del segmento, y los que pertenezcan al mismo segmento los mantendrá confinados, por lo que el tráfico que se transfiera a través del puente tiene un destino en algún otro segmento de la red, un grupo destino.

La forma en que opera un puente es a través de una tabla interna que lleva el registro de los usuarios que pertenecen al segmento de red, y de aquellos que están más allá del puente.

Los paquetes que cruzan el puente son colocados en la cola de espera del segmento destino.

Una forma sencilla de ver la operación de un puente es la que muestra el diagrama.

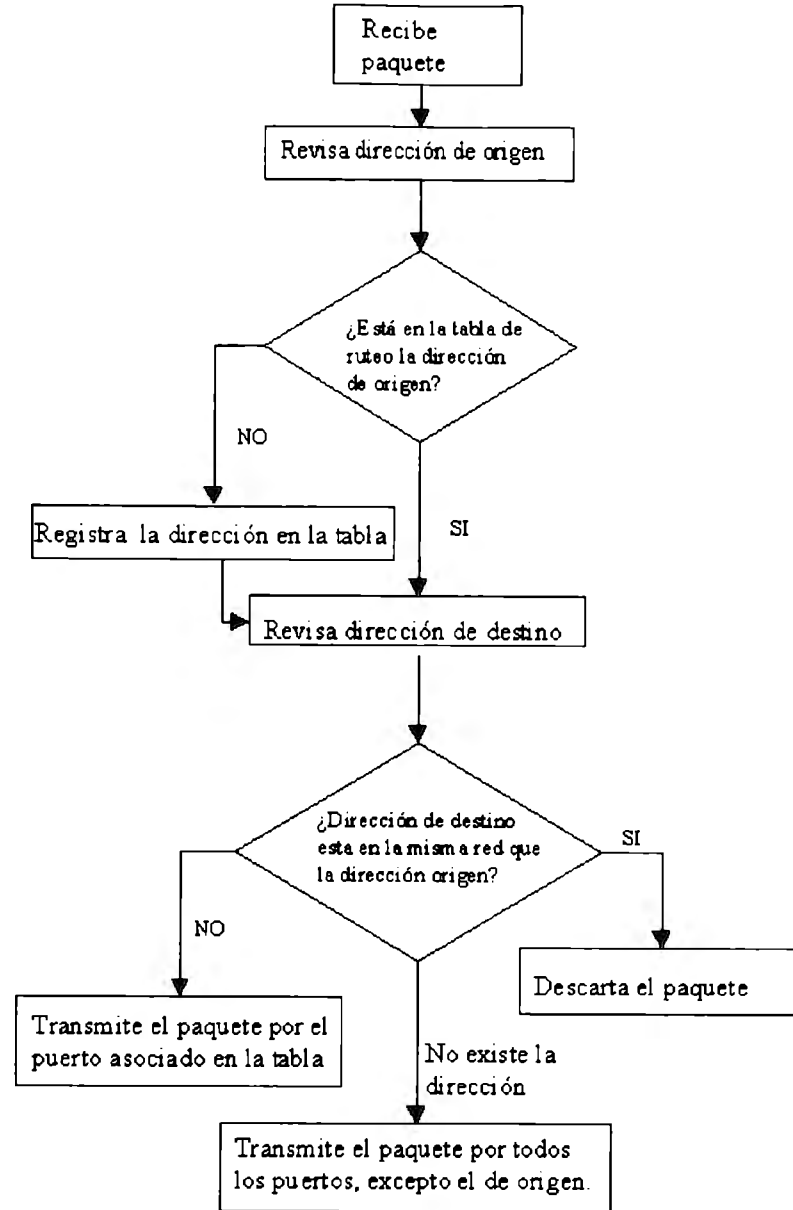


Figura3.3.3 1. Diagrama de decisión de un puente transparente.

3.3.4 Retraso de señal en la red

De la descripción del comportamiento de un puente es posible considerar que el retraso de cada paquete que cruza el puente está integrado por dos partes:

- El retraso debido a la revisión de la tabla. Este tiempo se puede considerar fijo ya que ocurre el mismo proceso para todos los paquetes que cruzan el puente.

- El retraso ocasionado por la estancia del paquete en la cola de espera al arribar al grupo destino. En consecuencia, el retraso en esta cola de espera será, en el modelo, parte del retraso que se contará para el grupo destino, debido a que la carga de tráfico que proviene del puente forma parte de la carga total del segmento.

Lo anterior permite considerar sólo el tiempo de revisión de la tabla como tiempo de retraso en el puente. Consideremos que el tiempo de retraso, por bit, ocasionado por el puente entre los segmentos i y j estará dado por:

$$B_{ij} = \frac{b_{ij}}{l} \quad (3.3.4.1)$$

Donde b_{ij} tiempo de revisión por paquete
o tiempo de retraso de cada paquete
 l longitud del paquete

Ahora, si tenemos que:

- El tráfico a través de puente entre los segmentos i y j esta dado por F_{ij} ,
- El retraso en el puente entre los grupos i y j , esta dado por B_{ij} . Si consideramos que la revisión de los paquetes es el único proceso que contribuye al retraso en el puente, entonces:
 - El retraso promedio que es originado por los puentes D_p esta dado por:

$$D_p = (1/\Gamma) \left[\sum_{i=1}^P \sum_{j=1}^P F_{ij} B_{ij} \right] \quad (3.3.4.2)$$

Donde

Γ es el tráfico total que existe en la red. Y lo podemos expresar como:

$$\Gamma = \sum_{i=1}^N \sum_{j=1}^N \alpha_{ij} \quad (3.3.4.3)$$

El retraso promedio total en la red debido al paso a través de los segmentos y de los puentes, estará dado por:

$$D = (1/\Gamma) \left[\sum_{k=1}^P \frac{L_k}{C_k - L_k} + \sum_{i=1}^P \sum_{j=1}^P F_{ij} B_{ij} \right] \quad (3.3.4.4)$$

3.3.5 Definiciones.

De acuerdo a como se ha tratado el problema hasta este momento, los grupos se conectaran como un árbol de expansión. En este árbol el tráfico que circula a través de sus componentes, enlaces y nodos, queda expresado en términos de las siguientes variables de decisión:

$$X_{ijk} = \begin{cases} 1 & \text{si hay tráfico que cruce a través de grupo } k, \text{ proveniente} \\ & \text{de grupo } i \text{ hacia el grupo } j; \\ 0 & \text{en cualquier otro caso} \end{cases} \quad (3.3.5.1)$$

$$Y_{ijkl} = \begin{cases} 1 & \text{si hay tráfico a través del puente que conecta los grupos } k \\ & \text{y } l, \text{ el cual proviene del grupo } i \text{ y va al grupo } j; \\ 0 & \text{en cualquier otro caso} \end{cases} \quad (3.3.5.2)$$

Estas variables de decisión permiten definir algunas cantidades o características del comportamiento de los datos que son de gran utilidad en la definición del problema. Estas cantidades pertenecen a la configuración del enrutamiento.

Tráfico total existente en la red, esta dado por:

$$\Gamma = \sum_{i=1}^N \sum_{j=1}^N \alpha_{ij} = \sum_{i=1}^P \sum_{j=1}^P s_{ij} \quad (3.3.5.3)$$

Tráfico total en el grupo k esta dado por:

$$L_k = \sum_{i=1}^P \sum_{j=1}^P s_{ij} X_{ijk}$$

Donde

$$1 \leq k \leq P \quad (3.3.5.4)$$

El tráfico total a través de los enlaces de la red esta dado por:

$$\gamma = \sum_{k=1}^P L_k \quad (3.3.5.5)$$

El tráfico total a través del puente que une los segmentos k, l, está dado por:

$$F_{kl} = \sum_{i=1}^P \sum_{j=1}^P S_{ij} X_{ijkl} \quad (3.3.5.6)$$

Donde

$$1 \leq k, l \leq P$$

La suma de los elementos de la diagonal en la matriz de tráfico S, diagonal (S), está dado por:

$$\text{diagonal (S)} = \sum_{i=1}^P S_{ii} \quad (3.3.5.7)$$

Y representa el tráfico que ocurre hacia el interior de cada uno de los grupos.

La suma de los elementos del resto de la matriz, esta dado por el complemento de la diagonal (S), $\overline{\text{diagonal (S)}}$, que está dado por:

$$\overline{\text{diagonal (S)}} = \sum_{i=1}^P \sum_{\substack{j=1 \\ j \neq i}}^P S_{ij} \quad (3.3.5.8)$$

3.4 ALCANCE DE LA SOLUCIÓN.

El problema a resolver con este diseño de la topología es obtener el mínimo retraso de la señal a través de la red. La forma de plantearlo es la siguiente:

Dada:

- La matriz de tráfico por usuario A, donde cada $a_{ij} \in A$ representa el tráfico que va de la estación i y la estación j.

- La matriz de agrupamiento de las estaciones R , la cual es una matriz de decisión que especifica cuales usuarios pertenecen a cual grupo, $r_{ij} \in R$.
- Las matrices de parámetros de decisión para:
 - El cruce a través de grupo k del tráfico proveniente de grupo i hacia el grupo j : X_{ijk}
 - El cruce a través del puente que conecta los grupos k y l del tráfico que proviene del grupo i y va al grupo j : Y_{ijkl}

El objetivo es:

$$\text{Minimizar } z = D = (1/\Gamma) \left[\sum_{k=1}^P \frac{L_k}{C_k - L_k} + \sum_{i=1}^P \sum_{j=1}^P F_{ij} B_{ij} \right] \quad (3.4.0.1)$$

con restricciones

$$S = R^T A R$$

$$\sum_{j=1}^P r_{ij} = 1, \quad 1 \leq i \leq N$$

$$r_{ij} \in \{0,1\} \quad 1 \leq i, j \leq N$$

$$L_k < C_k, \quad 1 \leq k \leq N$$

Las matrices de decisión $\{X_{ijk}, Y_{ijkl} \mid 1 \leq i, j, k, l \leq P\}$ describen las rutas que sigue el tráfico, en forma de flujo de bits, a través de los enlaces que unen a los nodos dentro del árbol de expansión.

Donde

S es la matriz de tráfico entre grupos donde cada elemento de $s_{ij} \in S$ representa el tráfico de los usuarios pertenecientes al grupo i con destino en usuarios en el grupo j .

3.5 TRABAJOS PREVIOS.

El desarrollo de trabajos relativos al diseño de redes, considerando las características del problema de optimización de una red de telecomunicaciones, lo podemos remontar a los trabajos de Gerla y Kleintock [GK77] [Kle64], en los cuales el propósito es desarrollar procedimientos heurísticos para la obtención de soluciones. También, Ersoy [Ers93] desarrolla

una metodología para el diseño de la interconexión entre redes LAN y WAN, considerando las limitaciones de tecnología inherentes a cada tipo de red.

Hay trabajos desarrollados para atender un tipo específico de tecnología, como es el caso del modelado de redes distribuidas [TDW97], redes de nodos múltiples de fibra óptica [Siv94] [MBR96], o para encontrar soluciones a distribución de capacidad en redes de alta velocidad [RAN91] [WL91]. Estos trabajos contribuyen con el uso de técnicas heurísticas o el uso de grafos especiales, como el hipercubo, para plantear soluciones al problema de la topología de una red.

En vista de que el problema de topología de una red es en general un problema de minimización, se utilizan algoritmos de minimización de flujo, costo, o longitud. Los siguientes algoritmos fueron creados para la resolución de problemas MST, como se mencionó en el capítulo de antecedentes [Gen00]:

- Prim(1957).
- Kruskal (1956).
- Dijkstra (1959).
- Sollin(1965).
- Gabow(1984).

Los algoritmos anteriores no aplican cuando se tiene restricciones. Las restricciones en el MST pueden ser:

- Restricciones en el grado.
- MST estocástico.
- MST cuadrático.
- Restricciones en la capacidad de flujo de los arcos.
- MST probabilístico.
- MST con criterios múltiples.
- Restricciones en el número de hojas asociadas a un nodo.

Estas restricciones hacen al problema NP-Hard, y la restricción en la capacidad lo hace NP-Completo [Gen97] [ACG99] [GJ79].

El desarrollo de trabajos relativos al diseño de redes, usando algoritmos genéticos incluye el uso de técnicas mixtas [Sin97] [BMS98a]. En otros casos se utilizan los algoritmos genéticos sin alguna otra técnica adicional, como en los trabajos de Gen [Gen00], donde hace uso de diferentes técnicas de codificación para mostrar el procedimiento de solución para cuatro de las restricciones arriba mencionadas.

En este mismo reporte de Gen se hace mención del problema de diseño de redes LAN, en el cual se considera un grupo de usuarios, los nodos a los cuales se pueden asociar los usuarios, y los enlaces de la red, toma como antecedente el trabajo de Elbaum y Sidi [ES96], entre otros. El propósito de este trabajo de Gen es proponer el diseño de una red de fibra óptica, considerando la disponibilidad de la red, el uso de los AG. Sigue, al igual que Elbaum y Sidi, considerando dentro del diseño la distribución de los usuarios, cambiando el objetivo de minimización de retraso de la señal, por el de minimización del costo de los enlaces. Usa números de Prüfer para la representación de los nodos y los usuarios por una cadena de caracteres, donde un identificador asocia a un grupo de ello a un nodo.

Las diferencias notables entre los métodos de Gen para la resolución de MST con restricciones, y el presente trabajo, es la utilización de una codificación diferente. En el caso de Gen en el espacio de búsqueda es arbitrario, dado que el conjunto de los enlaces a evaluar forma parte de las condiciones iniciales del proceso., y en el problema planteado en este trabajo se cuenta con un amplio espacio de búsqueda formado por el árbol de Huffman. Y en los problemas de MST se considera un nodo como un usuario; siendo que a los usuarios, en condiciones de instalaciones reales se les considera agrupados en nodos.

La diferencia es menor con el trato de Gen al problema del diseño de una red *LAN*. Por un lado el objetivo es diferente y también usa un modelo distinto de representación para los nodos de la red, ya que plantea la representación de los nodos en función de números de Prüfer y no de etiquetas del árbol de Huffman. Considero que esto disminuye considerablemente el espacio de búsqueda, pero si las condiciones de diseño del modelo no requieren más que un número limitado de posibilidades para los enlaces y para la topología de la red, entonces no afecta la limitación en el espacio de búsqueda.

4.0 MODELO PROPUESTO.

4.1 CREANDO LA REPRESENTACIÓN DEL PROBLEMA DE DISEÑO DE LA TOPOLOGÍA DE LA RED.

Esta forma de representación es utilizada por Elbaum R. y Sidi, M. [ES96] para representar cromosomas en un problema de diseño de red LAN.

Como se mencionó, los puentes conectan a los grupos de la red en una configuración de árbol de expansión. Cualquier configuración de árbol de expansión es una configuración válida de la red; la cual puede ir de una configuración de estrella a una configuración en línea, e incluye a todas las configuraciones intermedias como lo muestra la figura 4.1.0.1.

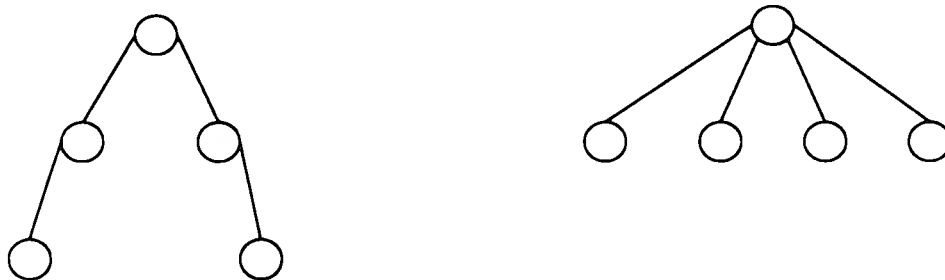


Figura 4.1.0.1. Con el árbol completo de Huffman es posible construir cualquier configuración que puede ser de nodos en línea a nodos en estrella. En este ejemplo se uso $P = 5$ y la profundidad es $\lfloor P/2 \rfloor$.

La red se considera integrada de P grupos. Entonces el árbol de expansión tiene hasta P nodos. Podemos afirmar que:

Se puede obtener cualquier configuración de árbol de expansión cuando este árbol tiene los siguientes dos atributos:

1. Cada nodo puede tener hasta $P-1$ hijos.
2. La profundidad de la red puede ser de hasta $\lfloor P/2 \rfloor$.

4.1.1 ÁRBOL DE HUFFMAN.

El árbol de expansión deberá de ser representado por una estructura de datos que permita obtener durante el proceso una representación válida de la red; esto significa que los operadores genéticos aplicados a tal estructura, que representa a un árbol válido, deberán de producir también un árbol válido. La estructura de datos seleccionada para la representación del árbol de expansión se basa en un árbol de Huffman disperso [Eve79].

Cualquier nodo de un árbol de Huffman está asociado con una sola etiqueta. La etiqueta de un nodo representa la trayectoria de la raíz a ese nodo, de donde la longitud de la etiqueta es igual a la longitud de la trayectoria. A la longitud de la trayectoria también se le llama profundidad del nodo. La profundidad de la raíz es cero. El símbolo λ (en el dibujo representado por 0) representa la etiqueta de profundidad cero, es decir representa a la raíz. En la figura 5.1.1.1, cualquier nodo intermedio tiene cuatro ramas. Las ramas están asociadas con los caracteres 1, 2, 3, 4 de izquierda a derecha respectivamente. De manera que, en este ejemplo, la trayectoria de la raíz a la segunda hoja de la derecha está en los ramales 4 y 3; de donde la etiqueta del nodo es 43.

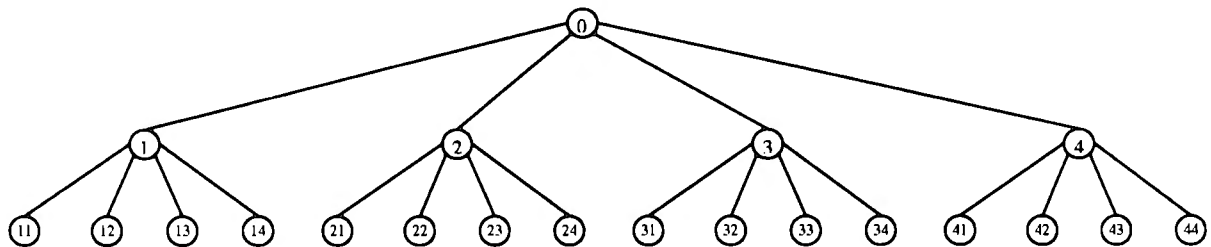


Figura 4.1.1.1. Arbol completo de Huffman donde $P = 5$. El grado del árbol será 4 como máximo.

Considerando un árbol completo de profundidad $\lfloor P/2 \rfloor$, en el cual cualquier nodo tiene $P-1$ hijos, y está etiquetado como un árbol de Huffman. Sumando el número de nodos de este árbol da un total de:

$$NN = ((P-1)^{\lfloor P/2 \rfloor + 1} - 1) / (P-2)$$

Donde

$$NN = \text{número total de nodos en el árbol de Huffman completo.}$$

A este árbol lo referiremos como el árbol virtual de Huffman [ES96]. En este árbol, cualquier nodo puede ser identificado por su etiqueta. Debido a que la longitud de la etiqueta es igual a la profundidad, la longitud contiene $\lfloor P/2 \rfloor$ caracteres.

Para etiquetar los nodos de tales árboles se requieren $P-1$ caracteres diferentes.

4.1.2 REPARACIÓN DE ETIQUETAS DE LA RED.

Consideremos el ejemplo de $P=5$. Este ejemplo se ilustra en la fig. 4.1.2.1, donde la profundidad del árbol virtual es dos, y cada nodo tiene cuatro hijos. Una topología de estrella se obtiene cuando las etiquetas se colocan como $\{\lambda, 1, 2, 3, 4\}$. Una topología en línea se obtiene con una combinación que puede ser $\{\lambda, 1, 11, 2, 23\}$, o alguna otra que se extienda en dos ramas de profundidad 2. Para nuestro ejemplo, elijamos aleatoriamente las etiquetas $\{\lambda, 11, 2, 32, 34\}$. Todas excepto λ se asignan, la raíz siempre es λ . El nodo etiquetado 11 no está conectado a la raíz dado que el nodo 1 no ha sido elegido. Con el propósito de obtener una conectividad correcta la etiqueta 11 se contrae hacia la raíz. El carácter menos significativo se omite de la etiqueta 11 y queda la etiqueta 1. La etiqueta 32 es cambiada del mismo modo a la etiqueta 3, de donde la etiqueta 34 permanece inalterada. Entonces, un conjunto corregido de etiquetas $\{\lambda, 1, 2, 3, 34\}$ reemplaza al conjunto original. Este conjunto de etiquetas corregido representa una configuración válida de árbol.

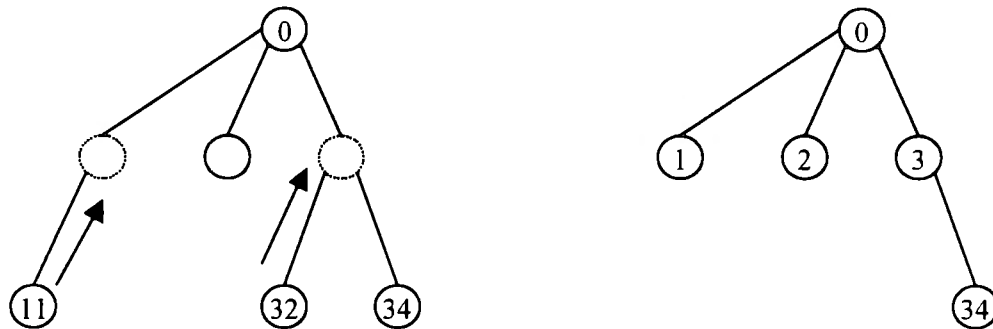


Figura 4.1.2.1. Un árbol puede recibir cualesquier etiqueta aleatoriamente. Una vez en el conjunto de etiquetas asignadas, la red es susceptible de ser reparada.

4.1.3 REPRESENTACION DEL ÁRBOL COMO CROMOSOMA.

Con el propósito de representar este árbol virtual de Huffman con cromosomas de AG, se requiere que todas las etiquetas de los nodos tengan la misma longitud. Para este propósito un carácter adicional llamado carácter de relleno se adiciona al alfabeto. Los caracteres de relleno se adicionan al final de la etiqueta posibilitando la representación de cualquier etiqueta en una misma longitud fija. El carácter de relleno será el dígito 0.

Usando el árbol virtual de Huffman, cualquier configuración de una topología de red puede ser representada. Un conjunto P de etiquetas del árbol virtual de Huffman se seleccionan aleatoriamente. Este conjunto se corrige para formar un árbol válido. El proceso de corrección es el mismo que se describió anteriormente. Para asegurar que en los arboles obtenidos, todos los nodos tienen una trayectoria a la raíz, el conjunto de etiquetas deberá obedecer la siguiente regla:

Cada prefijo de una etiqueta en el conjunto, deberá su vez ser una etiqueta en el conjunto.

Como ejemplo, los prefijos de 1231 son 1230, 1200, 1000. Por tanto para que exista la etiqueta 1231, tienen que antecederla en el conjunto las etiquetas 1000, 1200, y 1230. Es lógico que para que exista 1200, tiene que antecederla la etiqueta 1000.

4.1.4 REPARACIÓN DEL CROMOSOMA.

Una etiqueta en la cual uno de sus prefijos no está en el conjunto se reemplaza por el prefijo inexistente. Por ejemplo, los prefijos de la etiqueta 2121 son 2120, 2100, 2000. Si la etiqueta 2121 es la única etiqueta, en el conjunto que inicia con 2, entonces se reemplaza con la etiqueta 2000.

La red resultante es descrita en la forma de configuración de cromosoma. Este cromosoma contiene las P etiquetas seleccionadas, con $\lfloor P/2 \rfloor$ caracteres cada una. Considerando nuevamente el ejemplo de la fig.5.1.1.2. El conjunto de cinco etiquetas $\{\lambda, 11, 2, 32, 34\}$ está expresado en el cromosoma 0011203234, el valor de este cromosoma se corrige a 0010203034, obteniendo el conjunto de etiquetas $\{\lambda, 1, 2, 3, 34\}$. Esta cadena de etiquetas que puede funcionar como cromosoma, en nuestro caso lo llamaremos sección de configuración o de red del cromosoma.

Dada una configuración de red de árbol de expansión, pueden obtenerse diferentes topologías de red, tantas como $P!$ Permutaciones de los P grupos. De donde la configuración de la red puede ser completamente definida cuando cada uno de los nodos del árbol de expansión concuerde con uno de los P grupos. El orden en que los grupos están asignados se expresa en un segundo cromosoma, o una segunda sección del cromosoma completo, llamada sección de orden.

Resumiendo, cualquier individuo que es una solución para el diseño de la topología está representado por dos secciones del cromosoma:

C_{config} es la sección del cromosoma llamada de configuración o de red, que describe la configuración del árbol de expansión.

C_{orden} es la sección del cromosoma llamada de orden, en la que aparece el orden que se asigna a los grupos dentro del árbol. El cromosoma de la topología será representado por la dupla $\{C_{\text{config}}, C_{\text{orden}}\}$.

Los operadores genéticos se aplicarán a estos cromosomas. Después de aplicar los operadores de cruzamiento y mutación al cromosoma de configuración, es posible que el resultado no represente un árbol válido a menos que sea corregido.

El proceso de reparación del cromosoma es como sigue:

- La unicidad de las etiquetas se verifica.
- Si una etiqueta tiene múltiples instancias iguales en el cromosoma, todas menos una de esas instancias se cambia.
- Se verifica que cada una de las etiquetas tenga padre.
- Una etiqueta sin padre se reemplaza por la etiqueta del padre inexistente.

La representación de la estructura de la red como árbol de Huffman se seleccionó porque simplifica el problema de obtener una trayectoria dentro del árbol. Dado un par de grupos

identificados por sus etiquetas de Huffman, estamos interesados en la trayectoria que los conecta, lo podemos obtener de modo simple con el siguiente procedimiento:

- Se remueve el bit menos significativo de ambas etiquetas hasta que obtenemos el prefijo común de mayor longitud.
- Las etiquetas intermedias junto con las dos etiquetas originales definen la trayectoria deseada.

Dado que el agrupamiento de los usuarios está dado, y la trayectoria entre cada par de grupos se puede obtener, las variables de decisión de la topología en (2.3.5.1) y (2.3.5.2) se pueden obtener. De donde, el tráfico total de cada grupo (2.3.5.4) y el tráfico total a través de un puente (2.3.5.6) se pueden obtener. Colocando esos valores en (2.3.4.4) nos da el promedio de retraso de la solución representada en los tres cromosomas individuales.

El retraso promedio de una solución es usado para establecer el valor de su aptitud. Debido a que entre menor es el retraso mayor es la aptitud, entonces el valor del retraso nos indica la bondad de la solución y nos da una medida de la utilidad de la misma. Ya que lo que se desea es obtener una medida aptitud que crezca conforme las soluciones son mejores, menor tiempo de retraso, entonces el valor de la aptitud se representa por el inverso del tiempo de retraso de la señal a través de la red. Por tanto, la aptitud de cada individuo de la población será, aptitud = $(1/\text{retraso promedio})$.

4.1.5 Desarrollo del algoritmo.

Se usó la estructura de algoritmo genético indicada en los antecedentes.

Se compiló en C++ de Borland versión 4.5 y 5.01. El ejecutable generado es para DOS y se crea como una aplicación de 32 bits.

Se ejecuta en una ventana de DOS, usando un archivo de procesamiento por lotes para el control del proceso y direccionando la salida a un archivo.

Entradas.

El programa usa dos archivos de entrada:

- Tráfico usuario a usuario. En este archivo se indica en una matriz de $N \times N$, donde N es el número de usuarios en la red, el tráfico esperado entre los diferentes usuarios o estaciones que se ubicarán en la red.
- Usuario en nodo. Este archivo contiene una matriz de $N \times P$, donde P es el número de nodos, e indica con 0 ó 1 la pertenencia de un usuario a determinado nodo.

Salidas.

No se crean archivos en el proceso. Las salidas de defecto, se direccionan a un archivo de salida. Esto ocurre en el proceso por lotes, que controla las ejecuciones de prueba.

El desarrollo del algoritmo puede dividirse de manera natural en:

- Generación de cromosomas.
- Evaluación de la población.
- Evolución, que incluye:
 - Selección de población, para la nueva población.
 - Modificación de la población.
 - Evaluación.

Esto ya fue descrito, sólo describiré algunas tareas relevantes dentro del proceso de construcción del algoritmo.

4.1.5.1 Generación de cromosomas.

La población está representada por cadenas de caracteres almacenadas en dos matrices de caracteres; una para almacenar las secciones del cromosoma que representa a la red; dicha sección está su vez está dividida en secciones llamadas etiquetas, una por cada nodo de la red. Una matriz más guardar las cadenas de caracteres que representan el orden de los nodos en la red.

La generación de las cadenas es aleatoria, con las restricciones ya señaladas, en el planteamiento del problema.

Para representar las etiquetas de los nodos de la red, se considera:

- El número de caracteres por etiqueta es $N/2$, donde N es el número de nodos.
- Se eligen los caracteres de un alfabeto entre 1 a $N-1$, para el primer carácter de la etiqueta.
- Se eligen los caracteres de un alfabeto entre 0 a $N-1$, para los caracteres del segundo último de la etiqueta.
- La etiqueta raíz está representada por una etiqueta con 0's, tantos como la longitud de la etiqueta lo requiera. Esta etiqueta es obligatoria en cada cadena.
- Todas las etiquetas de los nodos de una red están en orden ascendente integrando una sola cadena de caracteres.

Para representar el orden, se considera:

- El número de caracteres en la cadena es P , igual al número de nodos.
- Ningún carácter se repite en la cadena, ya que cada uno identifica a un nodo.
- Se eligen los caracteres de un alfabeto entre 0 a $N-1$.

Una vez generados los cromosomas son reparados, esto sucede al inicio con las cadenas que contiene las etiquetas de los nodos de la red, una vez que se generaron las etiquetas y antes de integrarse en una sola cadena como cromosoma.

Las de orden no se reparan ya que se generan sobre listas ligadas, sin duplicados.

4.1.5.2 Matrices de decisión.

Para cumplir la parte del proceso que evalúa la red, es necesaria la construcción de dos matrices de decisión. Ellas son las matrices llamadas:

- X_{ijk} que contienen la información de los nodos o grupos k que cruza el tráfico entre los nodos i y j .
- Y_{ijkl} la cual contiene la información de los arcos que unen a los nodos k, l que cruza el tráfico que fluye de i a j . En el planteamiento se menciona como el puente que se encuentra entre k y l . Ambas descripciones hacen referencia al mismo enlace entre nodos. La ventaja de esta segunda es que puede asociarse a una capacidad del enlace, o capacidad del puente C_k , y que puede asignarse para cada puente en la red.

Para generar estas matrices se utilizó el siguiente procedimiento:

- Se extraen uno a uno los cromosomas. Cada cromosoma está representado por una cadena de caracteres en la **matriz de cromosomas**. A cada cromosoma lo podremos llamar vector de cromosoma.
A cada cromosoma le corresponde un orden determinado por una cadena de caracteres contenida en la **matriz de orden de grupos**. A cada elemento de la matriz de orden lo podemos llamar vector de orden del cromosoma. Cada grupo es numerado de 0 a $(P-1)$, donde P es el número total de nodos.
- Se convierten ambas cadenas en los valores numéricos que representan.
- Cada par de cromosomas, de red y orden, se evalúa con respecto a las matrices de decisión.

Creación de las matrices de decisión.

Se crean dos matrices de decisión para cada par de vectores. Una se llama **matriz de tráfico ijk** , la segunda será la **matriz de tráfico $ijkl$** ; estas las identificamos con las matrices de parámetros de decisión, mencionadas en el planteamiento del problema, donde cada una de ellas se define como:

X_{ijk} Indica el cruce a través de grupo k del tráfico proveniente de grupo i hacia el grupo j .

Y_{ijkl} El cruce a través del puente que conecta los grupos k y l del tráfico que proviene del grupo i y va al grupo j .

Para generar los valores que contendrá cada una de estas matrices se procede de la siguiente manera:

- Se crea una lista de etiquetas. En orden creciente de acuerdo a su valor. El proceso es igual al que ocurre durante la generación de cada cromosoma, por lo que no se describe aquí.
- Se crea un **arreglo de listas**. Corresponde al encabezado una etiqueta, estas se asignan en orden creciente por valor, y corresponden en orden al que contiene la lista de etiquetas. También tiene asociado el valor del orden que le corresponde, de acuerdo al cromosoma de orden.

- Cada lista del arreglo anterior, contiene los prefijos de la etiqueta correspondiente. Y a cada prefijo también se le asocia el orden asignado en el cromosoma de orden.
- Se crea una **matriz de listas**. Esta contendrá los nodos intermedios k , que se localizan entre los nodos i y j .
- Se asignan valores a las celdas de la **matriz ijk** .
- Se asignan valores a las celdas de la **matriz $ijkl$** .

Cada uno de los pasos listados se describen a continuación:

Arreglo de listas.

Se crea un arreglo de listas donde cada posición se inicia como el encabezado de una lista y contiene la etiqueta del grupo, contenida en el cromosoma. El orden que guardan las etiquetas de los grupos dentro del arreglo es el mismo que tienen dentro del cromosoma. A cada etiqueta de grupo que se agrega al arreglo se le asocia el orden asignado dentro de la red.

Matriz de listas.

Se crea una matriz de listas que contendrá las listas de etiquetas de los grupos que se encuentran en la ruta entre los grupos i y j .

Generación de las listas en el arreglo de listas.

En el arreglo de listas que contiene encabezados, se genera una lista para cada una de las etiquetas del cromosoma, el orden de las listas está dado por las etiquetas, la excepción es el encabezado que contiene una etiqueta en el orden en que fue extraída del cromosoma. Estas listas de etiquetas son prefijos de las etiquetas que se encuentran en los encabezados. El proceso para su obtención es el siguiente:

Proceso de generación.

- Para cada etiqueta del arreglo (llamada etiqueta del grupo seleccionado) se pregunta:
 - A partir de la primera etiqueta que no es raíz, se procede a generar su la lista de prefijos, señalando a esa etiqueta como grupo seleccionado.
 - Se generan etiquetas de los prefijos. Se busca entre las etiquetas de los grupos que lo antecede, obteniendo el orden de esta etiqueta y agregando, etiqueta y orden del grupo, a la lista del grupo seleccionado.
 - Se pasa a la siguiente etiqueta indicándola como etiqueta del grupo seleccionado.

Generación de la matriz de listas.

Con las listas de prefijos, del arreglo de listas, se obtienen los nodos o grupos intermedios que se encuentran en las rutas que unen cualesquiera dos grupos y que irán integrándose a la matriz de listas. Las listas de esta matriz están ordenadas por las etiquetas. El proceso para determinar esto es el siguiente:

Proceso de generación.

- Se selecciona una etiqueta del arreglo de listas. Esto se hace de manera consecutiva, es decir, se selecciona la etiqueta 0 del arreglo, después la 1, la 2, etc. A esta etiqueta seleccionada se le llama etiqueta base de grupo. Se pregunta:
 - Hay etiquetas de grupos que lo suceden (siguientes)?
 - Si no hay grupos que lo suceden.
 - Debe ser la etiqueta del último grupo y se da por terminado el proceso.
 - Si hay grupos que lo suceden. Se apunta consecutivamente a cada una de estas etiquetas de grupo sucesoras. A la etiqueta a la que se apunta en la secuencia se le llama etiqueta del grupo apuntado y se pregunta:
 - ¿La etiqueta base de grupo es prefijo de la etiqueta del grupo apuntado?
 - ◆ Si es prefijo.
 - Se recorre la lista de la etiqueta del grupo apuntado hasta localizar la etiqueta base de grupo, que es prefijo, y a partir de la siguiente etiqueta de grupo se agregan todas las etiquetas de grupo contenidas en la lista del grupo apuntado. Esto se efectúa usando el siguiente procedimiento:
 - Se agrega la etiqueta del grupo y su orden (k) a la lista, de la matriz de listas, que se encuentra en la posición indicada por el orden de las dos las dos etiquetas, la de la etiqueta base de grupo (i) y la de la etiqueta que se está apuntando (j).
 - Se pasa a la siguiente etiqueta en la lista.
 - ◆ No es prefijo.
 - Se recorren ambas listas hasta la última etiqueta de grupo común y a partir de esta se integran todas las etiquetas de ambas listas. Esto se efectúa usando el siguiente procedimiento:
 - Se agrega la etiqueta del grupo y su orden (k) a la lista, de la matriz de listas, que se encuentra en la posición indicada por el orden de las dos las dos etiquetas, la de la etiqueta base de grupo (i) y la de la etiqueta que se está apuntando (j).
 - Se pasa a la siguiente etiqueta en la lista.
 - Cuando se agota una lista se pasa a la otra lista.
 - ◆ Se pasa a la siguiente etiqueta de grupo que lo sucede (siguiente).

Terminada la evaluación de todas las etiquetas y de sus sucesores, queda creada una matriz de listas triangular con todos los k nodos que están incluidos en las rutas entre los nodos i y j . Para $i < j$.

La diagonal de la matriz contiene listas vacías.

Las listas i, j son iguales a las listas j, i por lo que sólo se calcularon las listas de la diagonal superior.

Ejemplo:

A continuación se describe el proceso de generación del arreglo de listas y de la matriz de listas.

Sea el cromosoma:

000, 100, 120, 122, 130, 135

Con el orden:

2, 5, 4, 3, 0, 1

a.) Listas de prefijos

orden				
2	000			
5	100	000		
4	120	000	100	
3	122	000	100	120
0	130	000	100	
1	135	000	100	130

b.) pares de la Matriz de listas.

i, j				
0,1				
0,2	100			
0,3	100	120		
0,4	100			
0,5				

c.) pares de la Matriz de listas.

i, j				
1,2	100	130		
1,3	100	120	130	
1,4	100	130		
1,5	130			

d.) pares de la Matriz de listas.

i, j				
2,3	100	120		
2,4	100			
2,5				

e.) pares de la Matriz de listas.

i, j				
3,4				
3,5	120			

f.) pares de la Matriz de listas.

i, j				
4,5				

Figura 4.1.5.2. a.)Lista de prefijos y b.) c.) d.) e.) f.) representación de matriz de listas, indicando los pares de una matriz triangular superior.

Matriz de tráfico ijk.

Con la información de la matriz de listas se asignan valores a la **matriz de tráfico ijk**.

En la matriz de tráfico se le asigna 0 como valor inicial a todos sus elementos.

El proceso es el siguiente:

- Se apunta a cada una de las posiciones de la matriz triangular superior i,j , para $0 \leq i \leq N - 1$, $1 \leq j \leq N-1$, para $i > j$. Es decir, se selecciona i como contador primario y j como contador secundario:
 - Se hace avanzar j en tanto i queda fijo en un valor.
 - Al terminar la secuencia de j , el contador i avanza a la siguiente posición.
 - Se inicia un nuevo ciclo con j .
 - El proceso termina cuando $i = N - 2$.
- Con la lista contenida en esa posición ij , para cada elemento de la lista se procede de la siguiente manera:
 - Se obtiene la etiqueta y el orden (k) que guarda en la red.
 - El orden del grupo en la red indicará que el nodo k se encuentra entre los nodos i y j , indicados por la posición de la lista en la red.
 - Se coloca el valor de 1 en la posición ijk de la matriz de tráfico.
 - Se coloca el valor de 1 en la posición jik de la matriz de tráfico.
 - Se pasa al siguiente elemento de la lista.
 - Al finalizar una lista se pasa a la siguiente lista.
 - Terminadas las listas, termina el proceso de asignación de valores.

Matriz de tráfico ijkl.

Con la información de la matriz de listas se asignan valores a la **matriz de tráfico ijkl**.

En la matriz de tráfico se le asigna 0 como valor inicial a todos sus elementos.

El proceso a continuación es el siguiente:

- Se apunta a cada una de las posiciones de la matriz triangular superior i,j para $0 \leq i \leq N - 1$, $1 \leq j \leq N-1$, para $i > j$. Es decir se selecciona i como contador primario y j como contador secundario:
 - Se hace avanzar j en tanto i queda fijo en un valor.
 - Al terminar la secuencia de j , el contador i avanza a la siguiente posición.
 - Se inicia un nuevo ciclo con j .
 - El proceso termina cuando $i = N - 2$.
- Con la lista contenida en la posición i, j se procede de la siguiente manera:
 - Se cuenta la cantidad de elementos contenidos en la lista.
 - Si la lista está vacía. No se modifica ningún valor en la **matriz ijkl**.
 - Se tiene un solo elemento. No se modifica ningún valor en la **matriz ijkl**.
 - Si tiene dos o más elementos:
 - El primer elemento de la lista será designado como elemento base.
 - ◆ ¿La etiqueta del elemento base es prefijo de la etiqueta del siguiente elemento?
 - Si es prefijo.

- El número de orden de la etiqueta del elemento base indicará la posición k en la **matriz $ijkl$** .
- El número de orden de la etiqueta del elemento apuntado como siguiente indicará la posición l en la **matriz $ijkl$** .
- Se coloca el valor 1 en las posiciones:
 - $ijkl$
 - $ijlk$
 - $jikl$
 - $jilk$
- Si no es prefijo.
 - Se pasa el apuntador al siguiente elemento de la lista.
 - Se procede a evaluarlo.
- Al agotar los elementos que siguen en la lista al elemento base, se pasa el apuntador del elemento base al siguiente, el cual ahora será el elemento base y se repite la evaluación.
- Al llegar el apuntador de elemento base al último elemento de la lista, se da por concluido el proceso.

Los valores asignados a las matrices X_{ijk} y Y_{ijkl} , también llamadas $X^i_{j,k}$ y $Y^i_{j,k,l}$, respectivamente. Estas matrices se usan en el cálculo de L_k que es el tráfico total a través grupo o cluster k , y en $F_{k,l}$ que es el tráfico total a través del puente que se encuentra entre los grupos k, l .

Con los valor L_k y F_k se calcula el retraso promedio en toda la red representada por el cromosoma. Y para efectos de minimizar el valor que se obtiene de este cálculo, se usa el inverso del retraso para el proceso de comparación de aptitudes. Por lo que lo que se tiene como objetivo es obtener el mayor valor del inverso del retraso promedio en la red.

4.1.5.3 Selección y cruzamiento.

La selección se hizo de manera aleatoria, privilegiando a los miembros de mayor aptitud al asignarles un número mayor de oportunidades de reproducción que a los miembros de menor aptitud, esto es normal en todo proceso de selección. Lo que se modificó a lo largo de las pruebas fue el método de selección de los individuos que habrían de reproducirse, y pasar por las alteraciones ocasionadas por los operadores genéticos.

Se probó un método de selección proporcional a la aptitud, el cual se descartó por presentar una rápida convergencia.

Se intentó también usar un método de escalamiento, que es una variedad intermedia entre el método de escalamiento lineal, y el método de escalamiento de truncamiento sigma. En este procedimiento se usa un parámetro C , el cual normaliza el valor máximo de oportunidades que un individuo puede tener para reproducirse. Este se fijó en 2. Este método establece la función de escalamiento en los siguientes términos:

La función de escalamiento sigue siendo de la forma:

$$f' = af + b$$

Donde

f' aptitud normalizada de un individuo de la población.
 f aptitud obtenida por un individuo de la población.

Se calcula en primera instancia un valor para la media de la aptitud f_{media} , es decir

$$f_{media} = (\sum f) / \text{Población},$$

Y de aquí se calculan los valores de a y de b .

$$a = (C - 1) * f_{media} / (f_{máxima} - f_{media})$$

$$b = f_{media} (f_{media} - C * f_{media}) / (f_{máxima} - f_{media})$$

Esté método ofreció una evolución muy lenta, impidiendo, incluso, el crecimiento de la aptitud promedio en las poblaciones.

Finalmente, se usó el método anterior de escalamiento, modificado en la parte de asignación de oportunidades de reproducción. El propósito fue incrementar la posibilidad de reproducción de los individuos de mayor aptitud, pero sin llegar a tener los problemas de la convergencia rápida, que ocurrió cuando se uso la selección proporcional.

En cuanto se observó la desaparición de los individuos de la población con mayor aptitud, se hicieron dos modificaciones al proceso de selección:

- Exención a los cromosomas de mayor aptitud a ser afectados por los operadores genéticos. Se les exentó a los individuos de mayor aptitud de ser afectados por los operadores genéticos, pero sin ser forzados a pertenecer a la población siguiente. Lo cual significa que si durante el muestreo aleatorio no son seleccionados, entonces no pasarán a la población nueva. Además de no tener obligado el pase, era incierta la cantidad de individuos que tenían esta posibilidad, ya que en poblaciones en donde existe una gran disparidad de aptitud en la población, la posibilidad de reproducción de estos individuos disminuye en virtud del método de escalamiento utilizado.
- Elitismo a un porcentaje de la población. Con el fin de asegurar un porcentaje mínimo de individuos de la mayor aptitud, se estableció un porcentaje de los mismos que tienen la posibilidad de ser transferidos a la nueva población sin ser alterados por operadores genéticos. Nuevamente, no se consideró forzar el pase de los de mayor aptitud, ya que el proceso de muestreo aleatorio permaneció vigente aún para ellos. Lo anterior aseguraba que de ser seleccionados, y dado que el número de oportunidades de reproducirse de estos individuos los distingue del resto, podrían pasar a la nueva población y mantener su alto nivel de oportunidades, hasta llenar el cupo dado por el porcentaje de elitismo. Una vez

cubierto el cupo, sus posibilidades de ser seleccionados disminuirían con cada oportunidad que fuesen seleccionados.

4.1.5.4 Parametrización.

Espacio de búsqueda.

Repitiendo la tabla 2.4.5.1., El número de nodos que tiene un árbol de Huffman completo está dado como A en la siguiente tabla. El número de combinaciones que se pueden hacer con esos nodos en grupos de P nodos, está dado por C.

Tabla 2.4.5.1. Crecimiento del número de nodos en un árbol de Huffman completo y su relación con el número de combinaciones posibles de P nodos.

	A	B	A*B	C	C*B
P	$((P - 1) ** ((P / 2) + 1) - 1) / (P - 2)$	P!		A Combinaciones en P	
3	3	6	18	1	6
4	13	24	312	715	17,160
5	21	120	2,520	20,349	2,441,880
6	156	720	112,320	18,161,699,556	13,076,423,680,320

En la tabla anterior, A representa el número de nodos que contiene un árbol de Huffman en relación a P, que es el número de nodos del modelo a diseñar. C es el número de combinaciones posibles de los A nodos en grupos de P.

Por consiguiente, si pensamos que se plateó el problema para 10 nodos $P=10$, el espacio de búsqueda es por demás amplio.

Población.

Se elige una población de 200 individuos para las pruebas iniciales, esta se va incrementando de acuerdo al comportamiento de los resultados a lo largo de las pruebas. Se fue incrementando la población hasta un total de 400 individuos, en razón de incrementar la posibilidad de variedad en la población y pretendiendo mantener la relación entre el número de individuos en la población y el número de caracteres en las cadenas de los cromosomas, que es de 4 a 1, respectivamente.

Generaciones.

El número inicial de generaciones es de 800 ciclos. Esto ocurrió también para las primeras pruebas. Se incrementó el número de generaciones hasta un total de 12000, hacia el final de las pruebas.

Probabilidad de cruzamiento.

Se mantuvo en 0.6 o 60% para todas las pruebas. Se vio desde el principio de las pruebas que este nivel de probabilidad mantenía estable el comportamiento de la población.

Probabilidad de mutación.

Se iniciaron las pruebas con una probabilidad de mutación baja de 0.001. Y al ver el poco efecto de esta probabilidad, se incremento hasta 0.04. Lo anterior afecto el proceso haciéndolo inestable, por lo que se llegó a una probabilidad de mutación de 0.005 como la más adecuada para este proceso.

Elitismo.

Se incluyó el elitismo en el proceso a causa de la pérdida de las mejores aptitudes de los individuos de la población. Inicialmente se incluyó al 10 % de la población original, con la mayor aptitud, como candidata a pasar sin ser modificada a la nueva población. Los mejores resultados se obtuvieron cuando el porcentaje de la población original que pasaba a la siguiente población, sin ser modificada, se fijo en 3%.

Reordenamiento.

Se usa la inversión, como un método de reordenamiento, con el fin de modificar y ofrecer mayor variedad a la población, debido a que la sección de orden establece patrones que son afectados muy lentamente por el operador de mutación. Este operador se inicia con probabilidades de inversión lineal de 0.2 y una inversión de extremos de 0.025. En las últimas pruebas se usan probabilidades de 0.05 y 0.0125.

5.0 EXPERIMENTOS.

Las primeras pruebas, no registradas, se hicieron con datos similares a los ofrecidos por el reporte de Elbaum y Sidi[ES96]. Se hizo así, con el propósito de verificar el comportamiento del algoritmo; y sobre todo, para verificar el cálculo de los datos, tanto de las matrices como de los totales de tráfico. Estos autores ofrecen pruebas de algoritmos genéticos que trabajan sobre seis nodos y debido a que sus objetivos difieren de los del presente trabajo, ayudan en la estimación de los resultados esperados, pero no resuelven el problema del modo en que se plantea en este trabajo.

La diferencia radica en el objetivo y en el manejo de los datos. En el trabajo de Elbaum y Sidi[ES96], el objetivo está centrado en obtener el mínimo retraso en la red con la asignación de usuarios en diferentes nodos, para lo cual, además de los cromosomas de red y de orden, se constituye un cromosoma de usuarios asignados a nodos, siendo variable la asignación de un usuario a algún nodo, y también variable la cantidad de usuarios que podrían pertenecer a determinado nodo. En consecuencia, y a pesar de que la función objetivo es la misma, el objetivo de este trabajo está en la obtención de una topología para la red, generada al momento de mejorar los resultados en el proceso de minimizar el retraso de transferencia de datos entre nodos.

5.1 DATOS DE ENTRADA.

Una vez que se verificó la validez de los cálculos hechos para el retraso promedio en la evaluación de la población, se planteó una instancia del problema diferente, que permitiera observar el alcance del proceso evolutivo del algoritmo genético.

Para cumplir con lo anterior, se usó como red de prueba una distribución de nodos sencilla de comprender y difícil de resolver para un método estadístico. Consta de diez nodos, los cuales presentan las siguientes características:

- Los usuarios tienen previsto tráfico sólo con otros usuarios de nodos adyacentes.

1	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0
0	1	0	0	0	0	0	0	0	0
0	1	0	0	0	0	0	0	0	0
0	1	0	0	0	0	0	0	0	0
0	1	0	0	0	0	0	0	0	0
0	1	0	0	0	0	0	0	0	0
0	1	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	1	0
0	0	0	0	0	0	0	0	1	0
0	0	0	0	0	0	0	0	1	0
0	0	0	0	0	0	0	0	1	0
0	0	0	0	0	0	0	0	1	0
0	0	0	0	0	0	0	0	0	1
0	0	0	0	0	0	0	0	0	1
0	0	0	0	0	0	0	0	0	1
0	0	0	0	0	0	0	0	0	1
0	0	0	0	0	0	0	0	0	1

Figura 5.1.2. Matriz de distribución de usuarios por nodo. (50 x 10).

- Los resultados óptimos globales son cualquiera de entre las dos soluciones planteadas arriba.
- La capacidad, bits por segundo aceptados, para todos los puentes es única y se fija mayor al flujo máximo que pueda ocurrir en el proceso. Esto forma parte de las condiciones iniciales de la función a minimizar.
- Esta capacidad es susceptible de individualizarse por cada puente, o arco asociado al puente, pero para las condiciones actuales de pruebas es suficiente con dejar fijo su valor.
- Los valores de los parámetros se modifican de acuerdo a las condiciones en cada grupo de pruebas.

5.2 RESULTADOS

En esta sección se describen los grupos de experimentos que se realizaron.

Fueron cinco grupos de pruebas, cada uno de estos grupos tuvo sus propios objetivos. Estos se centraron en mejorar el comportamiento del AG, considerando que en todo caso se debería esperar una tendencia de los resultados hacia un objetivo prácticamente inalcanzable dadas la poca probabilidad de encontrar alguna de las soluciones óptimas globales.

5.2.1 GRUPOS DE PRUEBAS.

Cada grupo de prueba constó de una cantidad variable de series de pruebas. Cada serie de pruebas fue constituida a su vez por la repetición del mismo proceso en once ocasiones, esto para verificar la consistencia de los resultados que de cada modificación iban surgiendo.

Una descripción del número de pruebas y de sus series se hace a continuación, para posteriormente hacer la descripción detallada de algunos ejemplos, uno por cada grupo de pruebas, que permiten ver el funcionamiento del AG y observar la forma en que fue mejorando su operación al ir adicionándole algunos operadores genéticos.

El grupo 1 fue integrado por dos series de pruebas y fue el principio formal de las pruebas con el AG, después de la verificación de los resultados obtenidos con las pruebas comparativas con los ejemplos de Elbaum y Sidi, los cuales no se registraron aquí por haber sido hechos con seis nodos únicamente. Por otra parte estas pruebas preliminares sirvieron para corregir algunos problemas que presentó el AG.

El grupo 2 constó de nueve series de pruebas, el se incluyó la mutación para el orden y se fue afinando la asignación de la probabilidad para este operador.

El grupo 3 fue integrado por 9 series de pruebas. En este grupo se modificó el procedimiento de selección, el cual sólo incluía la selección proporcional; y a partir de este grupo de series de pruebas, incluyó dos formas de privilegiar a los miembros con mayor aptitud al permitir su paso a la nueva población sin ser afectados por operador genético alguno.

El grupo 4 incluyó cuatro series de pruebas, y también fue modificado el AG para utilizar un operador de ordenamiento, este fue el operador de inversión, el cual consta de dos partes, una llamada de inversión lineal y una segunda llamada de inversión final o inversión de los extremos.

El último grupo, el 5, está constituido por una sola serie la cual tuvo como fin el probar el AG con una población y un número de generaciones mayor al utilizado hasta este punto. Se incrementó de 260 a 600 el número individuos en la población y de 6,000 a 12,000 el número de generaciones.

A simple vista, los resultados obtenidos muestran poca efectividad en encontrar un óptimo global. Lo cual no indica necesariamente que el algoritmo genético sea poco efectivo, si pensamos en la búsqueda de soluciones para redes que sean menos difíciles de localizar.

Al revisar la secuencia de pruebas, vemos el avance que se tiene en la determinación de los valores adecuados para los parámetros; tanto los que corresponden a los operadores genéticos como a la cuantificación de la población y número de generaciones necesarias.

En cada una de los grupos de pruebas se fueron afinando los parámetros de los diferentes operadores utilizados, siendo que el reporte de Elbaum y Sidi sólo considera el uso del cruzamiento, sin indicar la modalidad, y la mutación.

En nuestro caso, se hizo la selección por medio del escalamiento de las aptitudes. Aún en este caso, se hicieron modificaciones para adecuar el procedimiento a la obtención de un proceso de evolución que no fuese lento, ni que tuviese una convergencia prematura.

Se incrementaron operaciones en la selección de individuos (elitismo) y reordenamiento (cruzamiento).

La mejora tanto en la estabilidad del comportamiento de la población, y en la obtención de aptitudes cada vez mejores a lo largo de los procesos, hace ver que el algoritmo tiene un comportamiento adecuado con respecto a los datos que se le ofrecen.

Si consideramos un caso real, con los datos que se proporcionaron, las mejores soluciones proporcionadas por el algoritmo genético son factibles de implementar ya que representan una cercanía a la solución óptima global.

En las siguientes páginas se muestra un ejemplo de cada grupo.

Una revisión de cada una de las series, en cada uno de los grupos de prueba, se incluye en el apéndice A. En esta tabla de registro de resultados y propuestas de cambio, se puede hacer un seguimiento de los motivos que se tuvieron en cada caso, considerando los efectos de los cambios de parámetros en las pruebas anteriores, y considerando a nuevos operadores que permitieran a la población obtener cambios que redundaran en una evolución que explorara nuevas áreas en el mapa de aptitud.

Grupo 1.

Ejemplo No. 1

Objetivo. Observar el comportamiento de la población por el efecto de la mutación, con el fin de ajustar el valor de la probabilidad de este parámetro.

- Resumen de actividades.**
- Se varía la probabilidad de mutación entre 0.001 y 0.04.
 - No se incluye mutación en la sección de orden

Parámetros de prueba del ejemplo

- Numero de generaciones 800
- Tamaño de la población 200
- Probabilidad de cruzamiento 0.6
- Probabilidad de mutación 0.04

Resultado del proceso:

- Aptitud máxima del cromosoma: 3365.9997559
- Retraso promedio en la red (seg.) 0.0006613
- Cromosoma de red de mayor aptitud generado en el proceso: 00000100001900020000250003000050000600007000080000
- Cromosoma de orden de mayor aptitud generado en el proceso: 6201897453

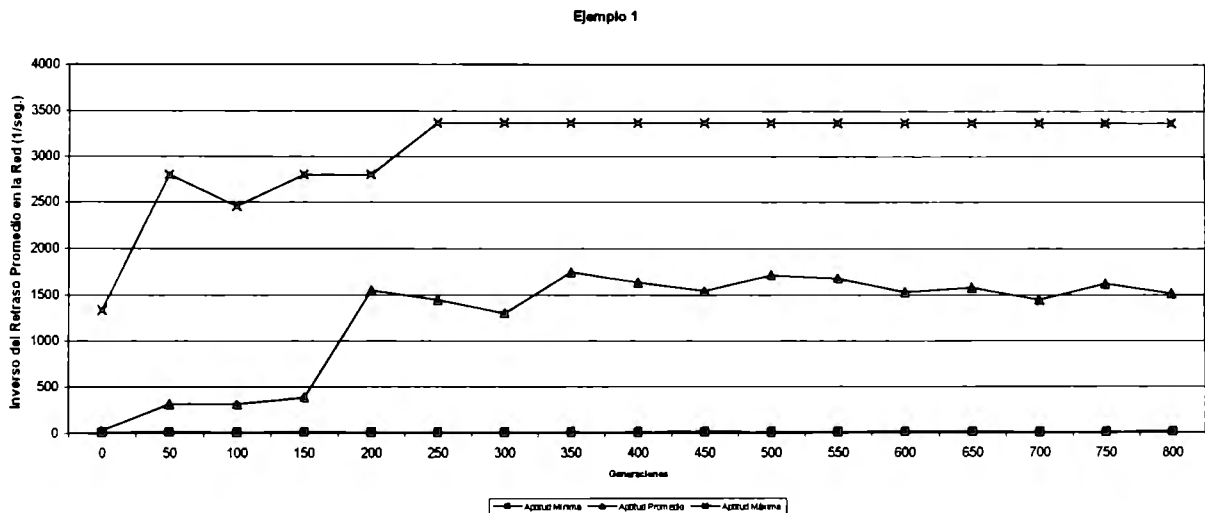


Figura 5.2.1.1. Gráfico del ejemplo 1 del grupo de pruebas 1.

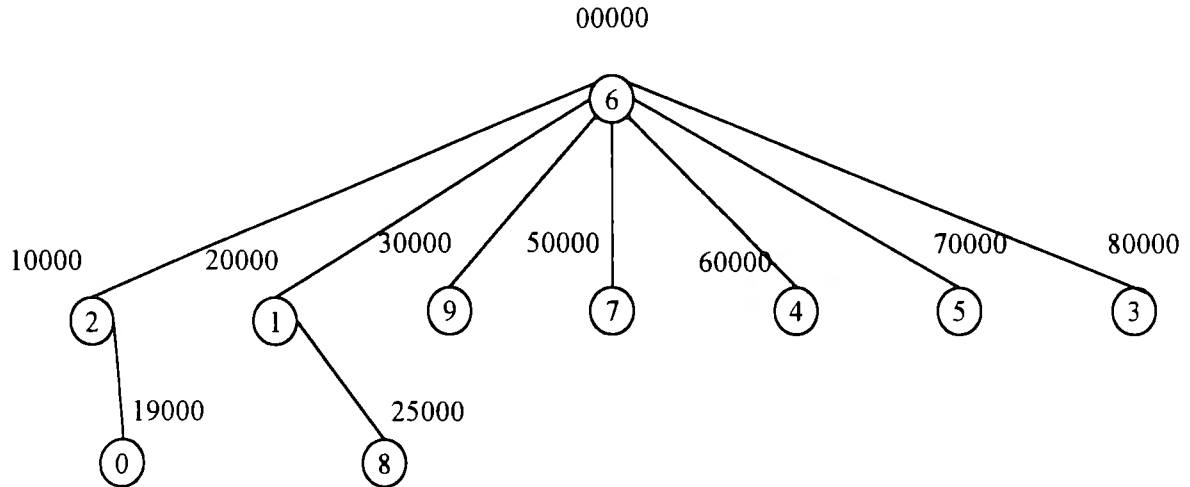


Figura 5.2.1.2. Representación del cromosoma de red incluyendo la asignación del orden de los nodos.

Observaciones

- Con baja probabilidad de mutación (0.001), hay convergencia prematura y mantiene la población estable hasta el final del proceso.
- Con alta probabilidad de mutación (0.04), se alcanza un máximo y en la sección de orden tiende a persistir una misma cadena de caracteres en un porcentaje mayoritario de la población.

Grupo 2.

Ejemplo No. 2

Objetivo. Incluir a la mutación en la sección de orden y variar la probabilidad de mutación, con el propósito de obtener valores crecientes de aptitud conforme el proceso avanza. Se espera que durante el proceso se mantenga una tendencia de estabilidad en los valores de aptitud de la población.

- Resumen de actividades.**
- Se incrementa el número de generaciones de 1500 a 3000 durante este grupo de pruebas.
 - Se varía la probabilidad de mutación entre 0.04 y 0.005.
 - La probabilidad de mutación afecta a ambas secciones del cromosoma.

Parámetros de prueba del ejemplo

- Numero de generaciones 1500
- Tamaño de la población 200
- Probabilidad de cruzamiento 0.6
- Probabilidad de mutación 0.04

Resultado del proceso:

- Aptitud del cromosoma: 1947.5410156
- Retraso promedio en la red (seg.) 0.0025310
- Cromosoma de red de mayor aptitud generado en el proceso: 00000100001700020000300004000050000600008000090000
- Cromosoma de orden de mayor aptitud generado en el proceso: 6783591024

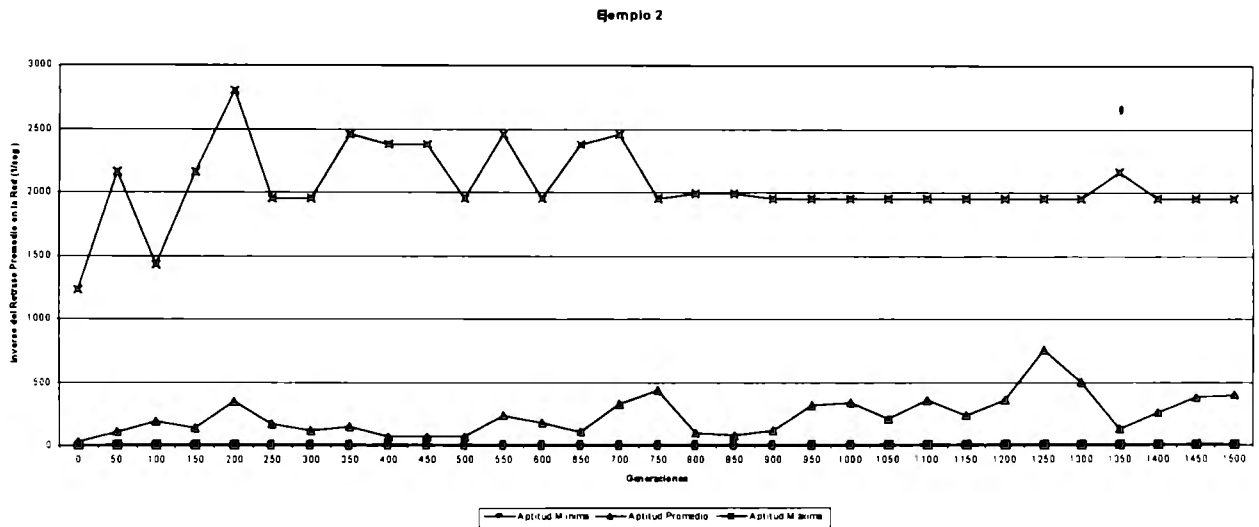


Figura 5.2.1.3. Gráfico del ejemplo 2 del grupo de pruebas 2.

Representación del cromosoma de red incluyendo la asignación del orden de los nodos.

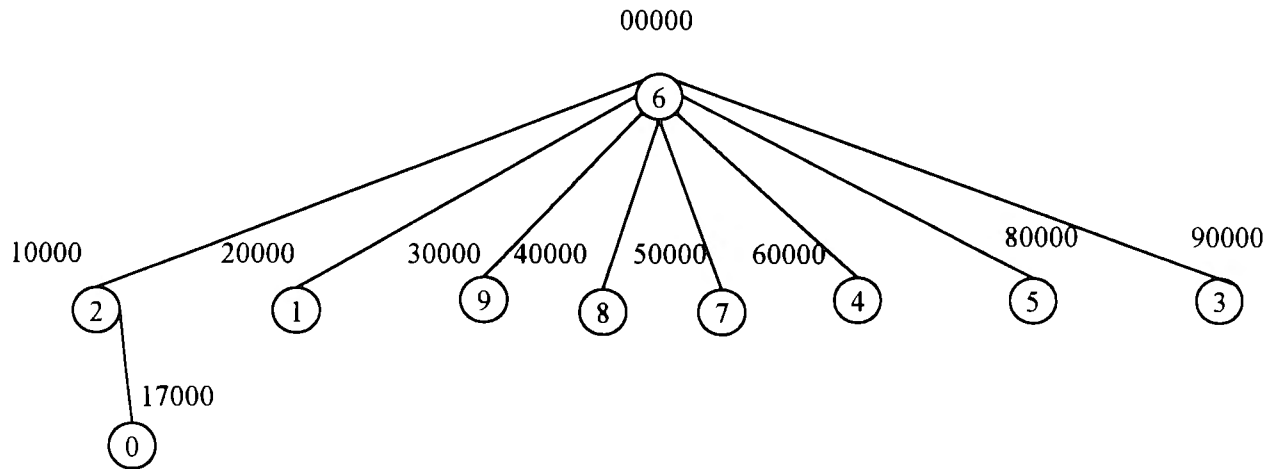


Figura 5.2.1.4. Gráfico del ejemplo 2 del grupo de pruebas 2.

Observaciones.

- En cuanto se disminuye la probabilidad de mutación al valor de 0.005, se consigue obtener estabilidad en el proceso, alcanzando valores crecientes de aptitud.
- En la sección de orden persiste la tendencia a mantener partes inalteradas de la cadena de caracteres.
- Se disminuyó la probabilidad de mutación por debajo de 0.005 (hasta 0.003), persistiendo la aparición de partes inalteradas, a las que nombré columnas de alelos, en la sección de orden.
- Se obtienen representaciones de la red que se asemejan a la que representa la solución del problema, es decir, inicia el acomodo de nodos en secuencias, haciéndose diferente a una representación en estrella.

Grupo 3.

Ejemplo No. 3

Objetivo. Se desea incrementar la persistencia de los individuos de mayor aptitud y de generar la diversidad suficiente en la población, con el fin de alcanzar mayores valores de aptitud en las poblaciones propuestas.

- Resumen de actividades.**
- Se incrementa el número de generaciones de 3000 hasta llegar a 12000.
 - Se incrementa la población de 200 a 260 individuos.
 - Se modifica el proceso de selección incluyendo la exención, a los individuos de mayor aptitud, de ser afectados por los operadores genéticos, en caso de ser seleccionados.
 - Se incluye el elitismo como parte del proceso de selección, para inducir la persistencia de los individuos de mayor aptitud en las poblaciones futuras.

Parámetros de prueba del ejemplo

- Numero de generaciones 3000
- Tamaño del a población 200
- Probabilidad de cruzamiento 0.6
- Porcentaje de permanencia de la población anterior en la nueva (elitismo): 3
- Probabilidad de mutación 0.003

Resultado del proceso:

- Aptitud del cromosoma: 3365.9997559
- Retraso promedio en la red (seg.) 0.0009854
- Cromosoma de red de mayor aptitud generado en el proceso: 00000100001700020000290003000050000550008000090000
- Cromosoma de orden de mayor aptitud generado en el proceso: 4891075623

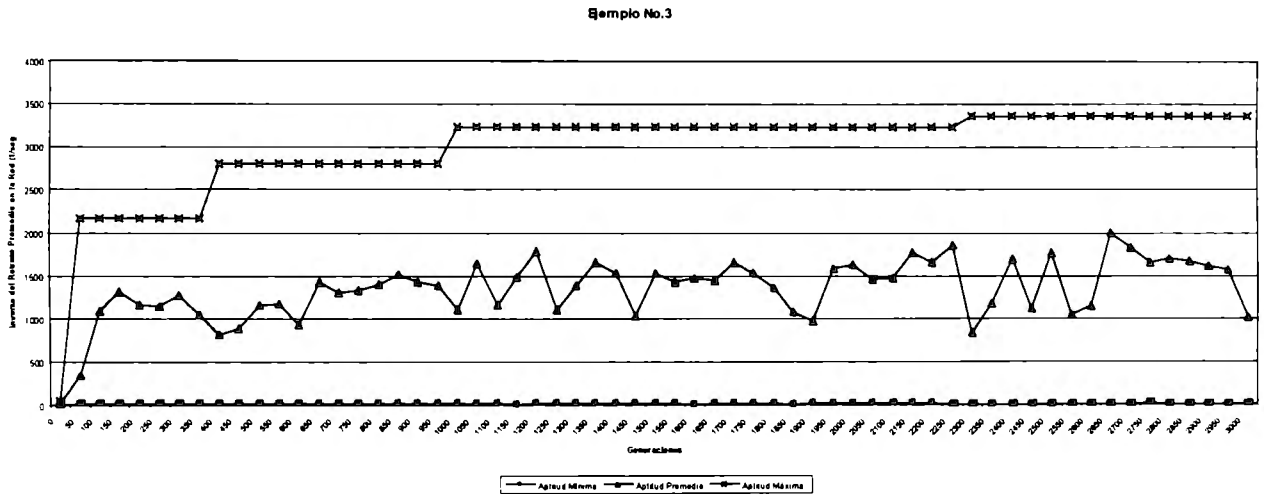


Figura 5.2.1.5. Gráfico del ejemplo 3 del grupo de prueba 3.

Representación del cromosoma de red incluyendo la asignación del orden de los nodos.

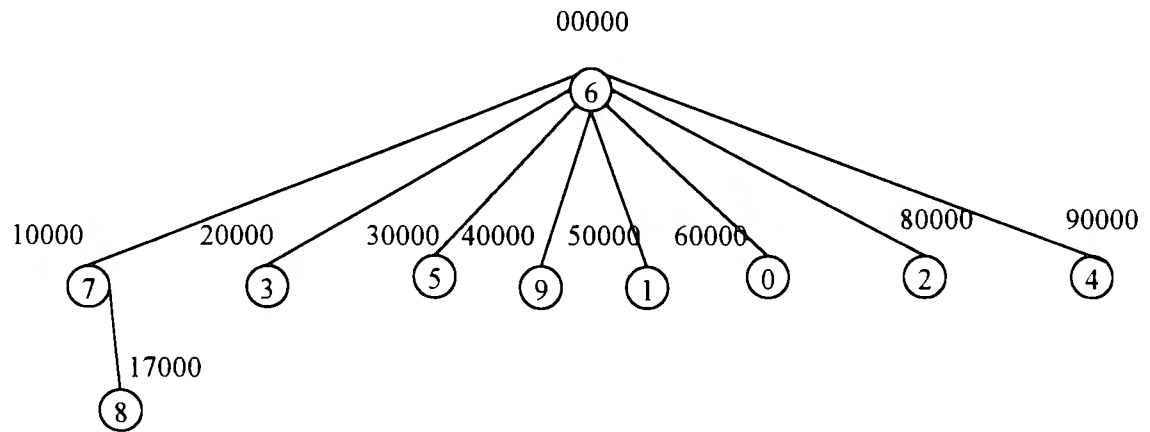


Figura 5.2.1.6. Gráfico del ejemplo 3 del grupo de pruebas 3.

Observaciones.

- Durante las primeras series el comportamiento es inestable en la aptitud de la población. Cuando se modifica la probabilidad de mutación de valor altos, 0.01, hasta llegar a valores de mínimos, 0.005, que el problema de inestabilidad se corrige.
- Se mantiene un comportamiento estable en el resto de las series, una vez ajustado el valor de probabilidad de mutación. Se nota que el valor de las aptitudes promedio y la máxima tienden a ser mayores que en las series anteriores.
- Aparece en algunas pruebas la convergencia prematura y en algunas otras la desaparición de los valores de mayor aptitud en el proceso, aunque en la mayoría de los casos se recuperan esos máximos.
- Se obtienen representaciones de la red que son soluciones de óptimas locales, iguales en aptitud a los obtenidos en pruebas previas, y equivalentes en la distribución de los nodos.

Grupo 4.

Ejemplo No. 4

Objetivo. Estabilizar el proceso y se intenta llegar a mejores niveles de aptitud a los alcanzados, a través de modificar la persistencia de patrones de caracteres, columnas de alelos, en la sección de orden, al incluir el operador de inversión.

Resumen de actividades.

- Se incrementa el número de individuos en la población.
- Se incrementa el número de generaciones.
- Se introducen los operadores de inversión, lineal y de los extremos, con valores iniciales de 0.2 y 0.025, respectivamente. Los valores adecuados parecen estar alrededor de 0.05 y 0.0125 para las probabilidades de inversión lineal y de los extremos, respectivamente.

Parámetros de prueba del ejemplo

- Numero de generaciones 6000
- Tamaño del a población 260
- Probabilidad de cruzamiento 0.6
- Porcentaje de permanencia de la población 3 anterior en la nueva (elitismo):
- Probabilidad de mutación 0.005
- Probabilidad de inversión lineal 0.05
- Probabilidad de inversión final 0.0125

Resultado del proceso:

- Aptitud del cromosoma: 6600.0000000
- Retraso promedio en la red (seg.) 0.0002853
- Cromosoma de red de mayor aptitud generado en el proceso: 00000100001700017600400004300043600437008000089000
- Cromosoma de orden de mayor aptitud generado en el proceso: 0125687934

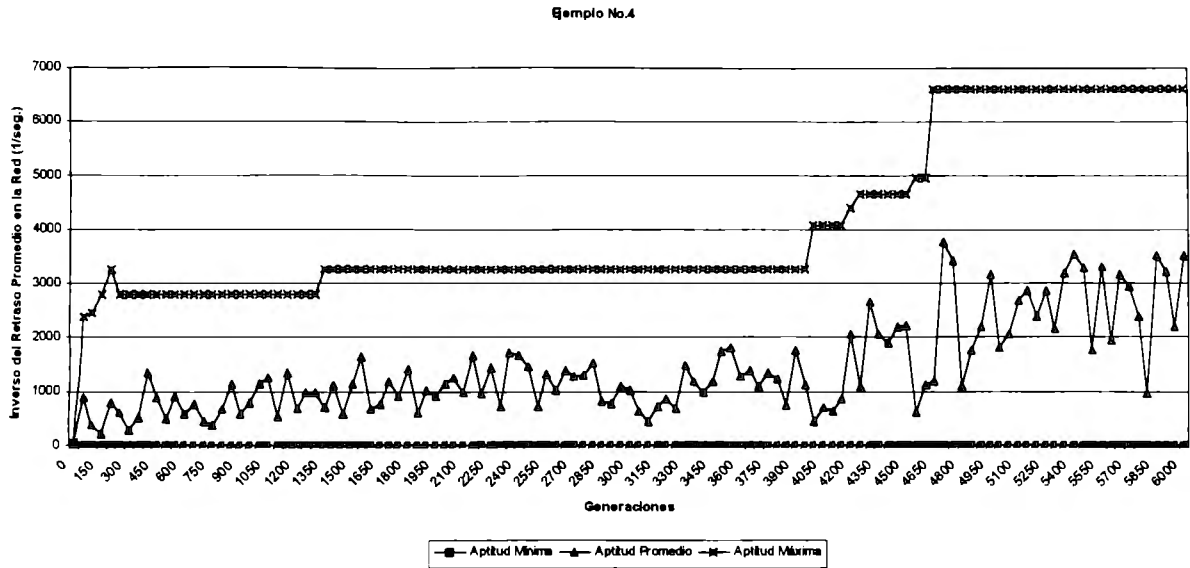


Figura 5.2.1.7. Gráfico del ejemplo 4 del grupo de prueba 4.

Representación del cromosoma de red incluyendo la asignación del orden de los nodos.

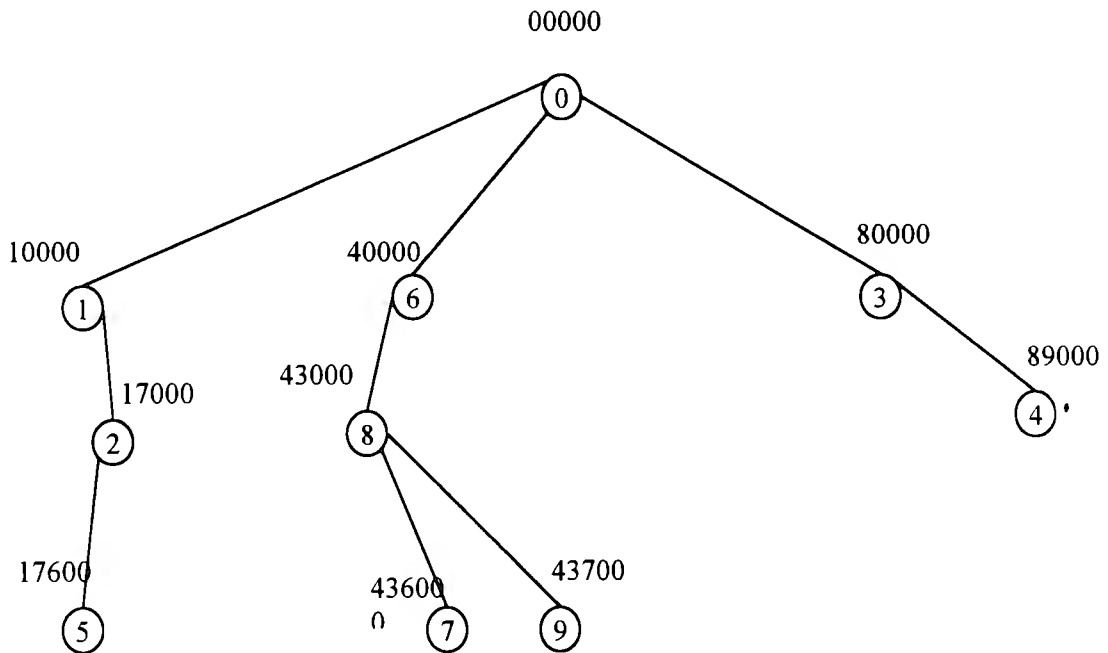


Figura 5.2.1.8. Gráfico del ejemplo 4 del grupo de pruebas 4.

Observaciones.

- El comportamiento de las poblaciones mejora, al incluir las modificaciones en la selección. La aptitud tiende a ser mayor, tanto en el promedio como en la aptitud máxima.
- Se acentúa la creación de columnas de alelos.
- Se intenta solucionar la creación de estos patrones de caracteres, en la sección de orden, con el incremento de la probabilidad de mutación. No resulta ya que tienden a desaparecer los individuos de mayor aptitud.
- La desaparición de individuos de mayor aptitud decrece y la población tiende a tener mayor estabilidad en cuando se recuperan los valores anteriores de probabilidad de mutación.
- Se incrementa el número de generaciones, para dar mayor oportunidad a la población de evolucionar. No hay resultados visibles en esta modificación, ya que las poblaciones se estabilizan al llegar a un máximo en su valor de aptitud.
- Se mejora la aptitud de las soluciones. Con respecto a las de el grupo de pruebas anterior.

Grupo 5.

Ejemplo No. 5

Objetivo. Mejorar la aptitud promedio de la población y obtener un valor máximo de aptitud mayor al alcanzado hasta este punto de las pruebas a través del incremento de la población y al número de generaciones.

Resumen de actividades.

- Se incrementa el número de individuos en la población.
- Se incrementa el número de generaciones.
- Las probabilidades asociadas a los operadores genéticos permanecen inalteradas.

Parámetros de prueba del ejemplo

▪ Numero de generaciones	12000
▪ Tamaño de la población	600
▪ Probabilidad de cruzamiento	0.6
▪ Porcentaje de permanencia de la población anterior en la nueva (elitismo):	3
▪ Probabilidad de mutación	0.005
▪ Probabilidad de inversión lineal	0.05
▪ Probabilidad de inversión final	0.0125

Resultado del proceso:

▪ Aptitud del cromosoma:	8774.9990234
▪ Retraso promedio en la red (seg.)	0.0001933
▪ Cromosoma de red de mayor aptitud generado en el proceso:	00000100001400014800400004600046003467008000090000
▪ Cromosoma de orden de mayor aptitud generado en el proceso:	6719324508

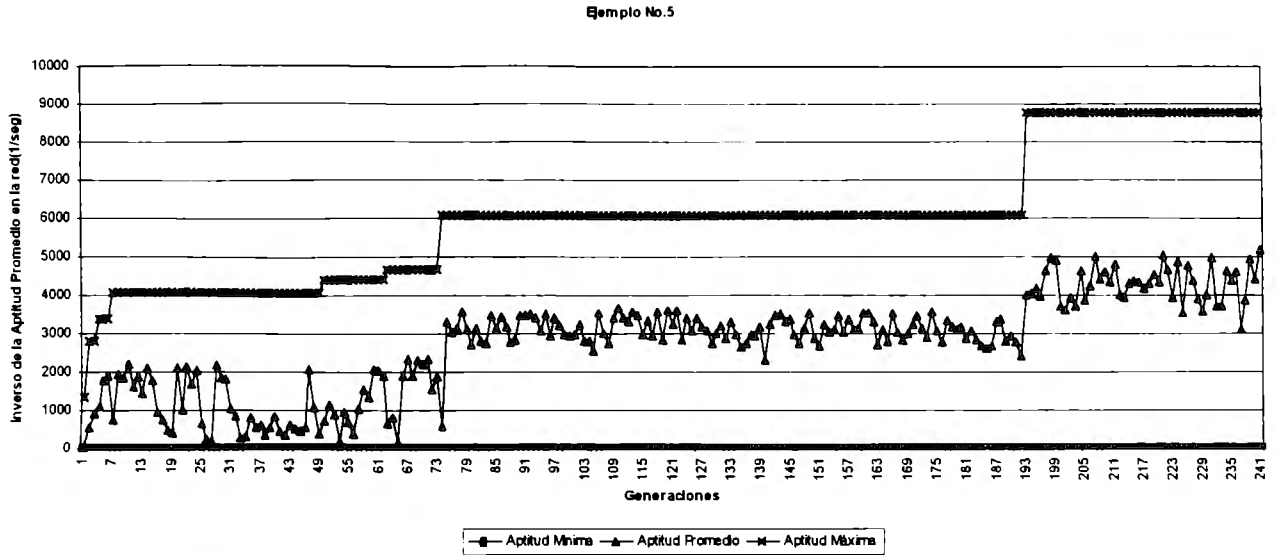


Figura 5.2.1.9. Gráfico del ejemplo 5 del grupo de prueba 5.

Representación del cromosoma de red incluyendo la asignación del orden de los nodos.

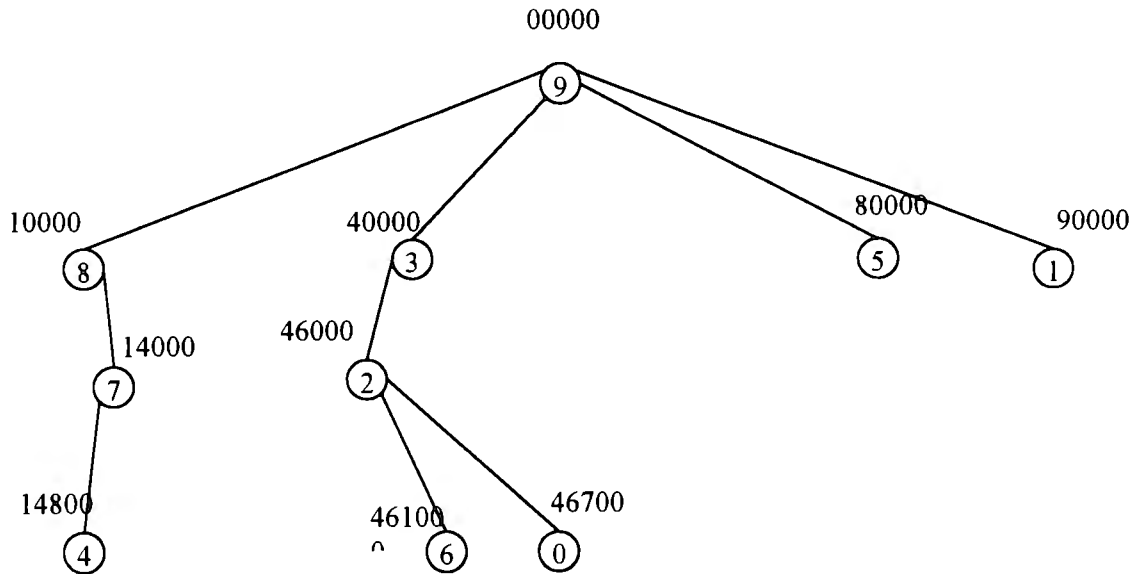


Figura 5.2.1.10. Gráfico del ejemplo 5 del grupo de pruebas 5.

Observaciones.

- Con el incremento de individuos en la población se logra una mayor diversidad en lo cromosomas y se alcanza un máximo de aptitud no logrado en pruebas anteriores.
- Es necesario experimentar con los parámetros de mutación, y de inversión para saber si se puede mejorar el comportamiento de las poblaciones.
- Los resultados obtenidos son satisfactorios, ya que permiten obtener topologías de red que se asemejan a la topología propuesta como óptimo global.

6.0 CONCLUSIONES Y TRABAJO FUTURO.

Como lo mencioné en el capítulo de experimentos, el problema planteado resulta difícil de localizar en el espacio de búsqueda, debido a las restricciones que se impusieron.

En cálculos más simples como los ofrecidos en el artículo de referencia [ES96] donde son seis nodos, las soluciones encontradas por el algoritmo son consistentes, y resultan adecuadas para los datos propuestos.

Los resultados obtenidos comparados con el objetivo que se estableció para las pruebas, concuerdan y hay cercanía entre ambos, por lo que el AG obtuvo soluciones que pueden ser utilizadas en la práctica y el trabajo lo podemos considerar como un avance en la comprensión del comportamiento de los algoritmos genéticos aplicados a la optimización de redes. Principalmente, porque el AG elaborado permite efectuar búsquedas en espacios de solución muy extensos y ofrece soluciones viables; y aún sin garantizar la obtención de óptimos globales, las soluciones cuantifican la topología de la red y ofrecen una distribución de nodos que se acerca a la solución óptima.

La utilización de duplas de cromosomas, con características diferentes, permite un modelado adecuado de las dos características básicas de este modelo, lo cual contribuye a crear condiciones que permiten al individuo evolucionar sin tener dependencia entre ambas; ya que de otro modo, si no hay distinción entre ambos potenciales de mejora de aptitud, el crear condiciones que favorecieran a una de ellas, a través de los operadores genéticos, podría causar, y en nuestro proceso experimental ocurrió, la deficiencia en el comportamiento de otro de los potenciales de evolución. En el proceso de experimentar con los cromosomas, se observó que el cromosoma de orden, o sección de orden, tenía un comportamiento deficiente creando patrones repetidos de caracteres y dado que en este cromosoma es importante la variabilidad, para explorar nuevas zonas en el mapa de aptitud, se optó por incluir un operador genético de inversión para permitir que continuara el proceso de evolución de este cromosoma. Lo que viene a apoyar la necesidad de un tratamiento específico para cada una de las partes de un cromosoma, cuando esta parte es distinguible del resto no se afecta el proceso de evolución conjunta; es decir, se complementa el proceso sin dejar de aplicar los operadores genéticos acostumbrados a ambas partes del cromosoma. Esta es la parte en la que considero que se contribuye en este trabajo a los AG.

El uso de AG en la determinación de la distribución de nodos dentro de una red WAN, se muestra como método simple que ofrece soluciones adecuadas a las necesidades del diseño de redes.

Dentro de las limitantes más obvias del modelo propuesto para la elaboración de este AG, está el reducido número de nodos que se pueden considerar para su uso, ya que como puede observarse, al contar el alfabeto con sólo diez caracteres, del 0 al 9, y al depender de éste número de caracteres la cantidad máxima de nodos que pueden considerarse para la elaboración de la topología de la red, entonces es necesario considerar el modo de extender el alfabeto para incluir una mayor cantidad de nodos en el modelo.

Durante el proceso previo de determinación del modelo se contempló esta necesidad como una mejora al proceso. En su forma actual, es posible examinar las limitaciones del alfabeto en el código del AG y modificarlo para incluir un alfabeto más extenso.

En determinado momento se pensó que el rango de caracteres ASCII de 48 a 122, que da un alfabeto de 75 caracteres, resulta suficiente para una gran cantidad de aplicaciones. De ser necesario este rango se puede ampliar.

Hay detalles aún no revisados dentro del código, por los cuales no puedo afirmar que tan inmediato pudiera hacerse el cambio de alfabeto, pero en general el código está preparado para recibir un cambio de alfabeto como el propuesto.

Las mejoras a este AG pueden extenderse a partes del diseño del AG, como en el caso del cálculo de las matrices de decisión, en cuya construcción se usaron listas, arreglos de listas y matrices de listas. Pudiendo considerarse otra forma de modelar este procedimiento, por ejemplo, considerar su determinación por matrices exclusivamente.

Por lo anterior, considero que este tipo de representación puede ser investigada con mayor detalle para obtener avances en la depuración de la representación del modelo, pensando que la única restricción inherente es la del alfabeto, pero presenta ventajas en cuanto a tener el potencial de representar cualquier tipo de árbol con los nodos que se le asignen, y tener de modo intrínseco, dentro de la reparación, el apegarse a un modelo que minimiza la profundidad del árbol; permitiendo además, integrar el tráfico proveniente de todos los usuarios como parte del diseño de la red.

7.0 REFERENCIAS BIBLIOGRÁFICAS.

- [Ahu93] Ahuja, Ravindra K, *Network Flows, Theory, Algorithms, and Applications*. Prentice Hall, 1993
- [Alf] Alfonso Manuel, *Genetic Algorithms*, IBM SofTec Lab, España.
- [ACG99] Ausiello, G.; Crescenzi, P.; Gambosi, G.; Kann, V; Marchetti-Spaccamela, A.; Protasi, M.; *Complexity and Approximation*. Springer Verlag, 1999.
- [BBM93a] Beasley, D.; Bull, D. R.; Martin, R. R.; *Han Overview of Genetic Algorithms: Part 1 Fundamentals*. University Computing 1993.
- [BBM93d] Beasley, D.; Bull, D. R.; Martin, R. R.; *Han Overview of Genetic Algorithms: Part 2 Research Topics*. University Computing 1993.
- [BMS95] Berry, Les; Murtagh, Bruce; Sugden, Steve; Mc Mahon, Graham, *Application of a genetic-based algorithm for optimal design of tree-structured communication networks* Teletraffic Eng. Conf. of the Int. Teletraffic Congress, South Africa 1995
- [BMS98a] Berry L. t. M., Murtagh B. A., McMahon G., Sugden S., Welling L. *An integrated GA - LP approach to communication network design*, Baltzer Journals January 20, 1998, Australia.
- [BMS98b] Berry, Les; Murtagh, Bruce; Sugden, Steve; Mc Mahon, Graham; *Optimization models for communication network design*. Baltzer Journals, 1998, Australia.
- [Cor86] Cormen, Thomas H.; Leiserson, Charles; Rivest, Ronald; *Introduction to algorithms*. MacGraw -Hill, 1986

- [Ers93] Ersoy, Cem, *Topological Design of interconnected LAN/MAN Networks*. IEEE Journal on selected areas in telecommunications, Vol. 11, No.8, Octubre 1993.
- [ES96] Elbaum, Reuven; Sidi, Moshe; *Topological Design of Local-Area Networks Using Genetic Algorithms*, IEEE INFOCOM '96.
- [Evd78] Evdokimov, F. E., *Fundamentos teóricos de la electrotecnia*. Moscú: MIR, 1978,
- [Eve79] Even, S. *Graphs Algorithms*. Rockville. Computer Science Press, 1979.
- [Fed99] Feder, Tomas, *Complexity of Graph Partition Problems*. Proceedings of thirty-first annual ACM Symposium of Theory of Computing, 1999, pags. 464-472.
- [For96] Forrest, Stephanie, *Genetic Algorithms*. ACM Computing Surveys, vol.28, No.1, 1996.
- [GJ79] Garey, M. R.; Johnson, D. S.; *Computers and Intractability: A Guide to the Theory of NP-Completeness*. Freeman, 1979.
- [Gav92] Gavish, B., editor, *Operations Research in Telecommunications*. Annals of Operations Research, 36, 1992.
- [Gen97] Gen, Mitsuo; Cheng, Runwei; *Genetic Algorithms and Engineering Design*. John Wiley & Sons, 1997.
- [Gen00] Gen, Mitsuo, *Network Design by Genetic Algorithms*. GECCO – 2000, Workshop Program. University Central Florida Julio 8, 2000.
- [GK77] Gerla, Mario; Kleinrock, Leonard; *On the topological design of distributed computer network* IEEE Transactions on Communications, vol. Com-25, No.1 January 1977.
- [GD90] Ghose, K.; Desai, K. R. *The design and evaluation of the hierarchical cubic network*. International Conference of Parallel Processing 1990, pags. I-355 – I-362, Agosto 1990.
- [Gol89a] Goldberg, D. E.; *Genetic Algorithms in Search, Optimization and Machine Learning*, Addison Wesley 1989.
- [Gol98] David Goldberg, *A meditation on the application of genetic algorithms*. IlliGAL Report No. 98003, February 1998.
- [GW93] Ganz, Aura; Wang, Xudong; *Efficient algorithm for virtual Topology design in multihop lightwave networks*. IEEE ACM Transactions, 1993.

- [RAN91] Guerin, Roch; Ahmadi, Hamid; Naghshineh, Mahmoud; *Equivalent Capacity and Its Application to Bandwidth Allocation in High-Speed Networks*. IEEE Journal in Select Areas in Telecommunications, vol.9, No.7, Septiembre 1991.
- [Hol92] J. H. Holland, *Adaptation in Natural and Artificial Systems*. MIT Press 1975, 2ª.Edición 1992..
- [Jac63] Jackson, R.; *Job Shop-like Queueing Systems*. Management Science, vol.10, pag.131-142, 1963.
- [Kle64] Kleinrock, L.; *Communications Nets*. New York: Dover, 1964.
- [Kle75] Kleinrock, L.; *Queueing Systems Volume I: Theory*. John Wiley & Sons, 1975.
- [Mic92] Michalewicz, Zbigniew, *Genetic Algorithms + Data Structures = Evolution Programs*. Springer 1992
- [Min89] Minoux, M., *Network Synthesis and optimum network design problems: Models, solution methods and applications*. Network vol.19 (1989) pags.313-360.
- [Mit96] Melanie Mitchell, *An Introduction to genetic algorithms*. Massachusetts Institute of Technology Press 1996.
- [MBR96] Mukherjee, Biswanath; Banerjee, Dhritiman; Ramamurthy, S; *Some Principles for Designing a Wide Area WDM Optical Network*. IEEE/ACM Transactions on Networking, vol.4, No.5, Octubre 1996.
- [RM72] Reitz, J. R.; Milford, Frederick J.; *Foundations of electromagnetic theory*. Addison-Wesley, 1972.
- [Sed92] Sedgewick, Robert, *Algorithms in C++*. Addison-Wesley, 1992.
- [Sha82] Sharma, Roshan Lal; De Sousa Pablo J. T.; Ingle Ashok D.; *Network Systems* Van Nostrand Reinhold Data Processing Series, 1982.
- [Sha97] Sharma, Roshan Lal, *Network Design Using EcoNets*. Thomson Computer Press 1997.
- [Sin96] Sinclair, Mark C.; *NOMaD: Initial Architecture of an Optical Network Pptimisation, Modeling and Design Tool*. Proc. 12th Performance Engineering Workshop, Edimburgh, September 1996, pags. 157-167.
- [Sin97] Sinclair, Mark C.; *NOMaD: Applying a Genetic-algorithm/Heuristic Hybrid approach to Optical Network Topology Design* . Proc. International Conference on Artificial Neural Networks and Genetic Algorithms, ICANNGA'97, Norwich, UK, Abril 1997.

- [Siv94] Sivarajan, Kumar N., *Lightwave Networks Based on Brujin Graphs*. IEEE/ACM Transactions on Networking, vol.2, No.1, Febrero 1994.
- [Sta97] Stallings, William, *Data and Computer Communications*. 5ª. Edición Prentice Hall, 1997.
- [Syk92] Sykas, E. D.; Lyberopoulos, G. L.; *Performance Analysis of interconnected CSMA/CD Local Area Networks*. IEEE Proc. I, Abril, 1992, pag. 181-197.
- [Tan91] Tanenbaum, Andrew S., *Redes de ordenadores*. Segunda edición, Prentice-Hall 1991
- [TB] Talbi, E. G. P.; Bessiere, *A parallel genetic algorithm for the graph partition problem* Laboratoire de Génie Informatique/ Intitut IMAG, University of Grenoble.
- [TDW97] Tomasgard, Asgeir; Dye, Shane; Wallace, Stein W. (Norwegian University of Science and Technology); Audestad, Jan A. (Telenor); Leen Stougie (University of Amsterdam); Van der Vlerk, Maarten H. (University of Groningen), *Modelling in distributed Telecommunications Networks*, 31/1/1997.
- [WL91] Sizheng Wei, Saul Levi, *Design and Analisis of efficient hierarchical interconnection networks*. Supercomputing '91. Preceedings of the 1991 Conference of Supercomputing, pag. 390 - 399.
- [ZF97] Zalzala, A. M.; Fleming, P. J.; *Genetic Algorithms in Engineering Systems*. Londres: The Institution of Electrical Engineers, 1997.

8.0 GLOSARIO.

Avaro, algoritmo. Diversas técnicas de algoritmos de minimización que realizan la búsqueda del elemento de menor costo, longitud, etc. entre una población disponible, de modo que las trayectorias seleccionadas, cumple en general con ofrecer la opción mínima, según sea el caso.

Ancho de banda. Capacidad de un enlace para transferir un flujo de bits entre sus dos extremos.

Árbol de Expansión Mínima, o Árbol de Cobertura Mínima (MST). Por su nombre en inglés Minimum Spanning Tree. Es un grafo conectado que no contiene ciclos, el cual contiene a todos los nodos, y solo los arcos necesarios para establecer las conexiones entre nodos.

Batch, proceso. Procesamiento por lotes. Procesamiento de información en volumen, dentro de un mismo centro de procesamiento central. Para lo cual durante un periodo de tiempo se acumula información, y al final de este se inicia el procesamiento de toda la información acumulada.

BNA. Burroughs Network Architecture. Tecnología de red de Unisys, para acceso a redes SNA.

Baseband. Una forma particular de operar una línea, en donde un dígito binario, procedente de un mensaje, es convertido en uno de dos niveles de voltaje, donde uno de los cuales representa el binario 1 y el otro 0. El voltaje fluctuará en el tiempo entre estos dos niveles.

Broadband. Un modo particular de uso de cable coaxial. A través de un solo cable se transmite un cierto número de flujos de información al asignar porciones del ancho de banda disponible. Los datos se transmiten al modular en la frecuencia asignada la señal, y al recibirse se invierte el proceso de modulación de la señal para su uso.

Carrier. Se le llama así al operador y a la red pública que se usa para la interconexión entre diferentes entidades, quien asigna un ancho de banda fijo o por demanda.

CATV. Community Antenna Television. Una clase de LAN, donde se opera la red en modo de broadband. El uso extendido está en las redes privadas de televisión por cable.

Circuito virtual. Método de envío de paquetes de información, en una red de conmutación de paquetes, en donde una ruta planeada con anticipación se establece antes de que los paquetes con información sean enviados.

Cluster. Grupo o racimo. Modo de agrupar terminales o PC's . Subredes.

Clustering. Agrupamiento. Agrupamiento de terminales o PC's usando cualquiera de las topologías de red accesibles. Modo en que operan los sistemas operativos de un grupo de computadoras, permitiendo distribuir la carga de procesamiento entre los CPU's pertenecientes a cada computadora y soportando la carga de procesamiento entre ellos cuando existe falla en alguno de los equipos integrantes del grupo.

Concentrador. Hub. Equipo de comunicaciones que conecta equipos terminales entre sí, o con equipos de comunicaciones, cuyo propósito es ser el medio a través del cual se transmiten los mensajes.

CPU. Central Processor Unit. Procesador de una computadora.

CSMA/CD. Detección de portadora, acceso múltiple con detección de colisiones. Método usado para controlar el acceso a un medio de transmisión compartido.

Datagrama. Método de envío de paquetes de información, a través de una red de conmutación de paquetes, en la que cada paquete es tratado independientemente sin asumir ninguna referencia con respecto a los paquetes que lo antecedieron.

DCE. Equipo de comunicación de datos. El equipo especializado en la administración y transmisión de una señal analógica o digital a través de un medio de transmisión. Su función no implica el procesamiento de la información, pero si el procesamiento de la señal para adaptarla al medio y a los equipos DTE conectados al mismo.

DTE. Equipo terminal de datos. El equipo que hace uso de los datos al final de un línea de transmisión, sea como simple receptor o efectuando la transmisión y recepción de los mismos.

DSL. Digital Subscriber Line. Servicio proporcionado por un operador de red publica que enlaza a los usuarios en forma digital a través de los protocolos específicos de esta tecnología.

Enrutador. Router. Equipo de comunicaciones que interconecta a diferentes localidades, realizando las tareas de asignación de rutas a los paquetes de datos, que provienen de una o mas redes locales conectadas al mismo, o de las redes remotas con destino a algún usuario localizado en las redes locales.

Espejo. Banco de datos que recibe el mismo nivel de actualización que el banco de datos principal, y que sirve de respaldo en caso de falla del banco de datos principal.

FDM. Frequency Division Multiplexing. Modo de administración de la señal que permite dividir la frecuencia en espacios destinados al envío de mensajes provenientes de diferentes canales.

FEP. Front End Processor. Procesador de comunicaciones conectado directamente o a través de una red local a equipos mainframe. Su uso permite atender los requerimientos de comunicaciones de las estaciones de trabajo por parte del mainframe.

Frame Relay. Un método de conmutación de paquetes que usa paquetes de longitud variable en la capa de enlace. No hay capa de red, y muchas de las funciones básicas han sido optimizadas o eliminadas, de modo que se incremente el flujo a través del medio.

Full duplex. Modo de transferencia de información en el cual en ambos extremos de un enlace pueden estar transmitiendo y recibiendo información simultáneamente.

Hill climbing algorithm. Llamado Steepest –ascent hill climbing, es un algoritmo que busca la trayectoria a la meta a través de una ruta que mantenga un gradiente positivo en la función de aptitud, de modo que se pueda ascender lo más rápidamente posible.

Hub. Ver concentrador.

ISDN. Integrated Services Digital Network. Servicio de enlace digital que es proporcionado por un operador de red pública y que presta los servicios de voz, video y datos.

ISO. International Standard Association.

LAN. Red de Área Local.

Mainframe. Equipo de cómputo centralizado, que sirve como procesador de la información y como centro de almacenamiento de la información. Este tipo de equipo ayuda a simplificar los esquemas de seguridad en una corporación.

MAU. Equipo distribuidor de señal que permite la interconexión de terminales a una red SNA.

Metaestable. (*Meta*, después, de otro modo, en otro lugar). Designación de un fluido que no puede recibir la designación de estable, por presentar variaciones en el tiempo, pero que para los fines propuestos, se le puede considerar estable.

Módem. Equipo de comunicaciones que sirve para establecer un enlace a la distancia con entre equipos digitales, usando el servicio que presta la red telefónica pública, a través de la cual transmite señales analógicas. Deriva de las palabras MOdulador-DEModulador.

MST. Ver Árbol de Expansión Mínima.

MTU. Maximum Transfer Unit. Límite máximo de longitud para un paquete, dentro de la tecnología de conmutación de paquetes.

Multipunto. Arreglo de estaciones que permite el envío de información a un grupo de ellas a través de un solo enlace.

OSI. Open Systems Interconnection. Modelo de protocolo basado en los estándares de ISO para crear un ambiente abierto de interconexión.

PBX. Private Branch Exchange. Conmutador digital que puede asumir muchas de las funciones de una central telefónica. Por extensión una central telefónica privada.

Polled. Monitoreo que permite censar y administrar el envío de información. El cual es efectuado por un equipo maestro hacia a cada una de las estaciones que pertenecen a un grupo de terminales llamadas esclavas.

Poll-select. Método de control de transmisión usado en redes caracterizadas por una estación maestra y el resto como esclavas.

PSTN. Public Switched Telephone Network. Red pública de teléfonos formada por la interconexión de las redes de operadores que prestan servicios telefónicos y de renta de enlaces, analógicos y digitales, a usuarios residenciales, empresariales o de cualquier otro tipo.

Puente. Bridge. Equipo de comunicaciones que administra el flujo y enrutamiento de los paquetes de información que se intercambian entre los usuarios de una red.

Esquema. Schema. Llamado también schemata por contracción de schema data. Designa a los patrones de similitud que se pueden localizar en una cadena de caracteres que designa a un cromosoma.

SNA. Systems Network Architecture. Tecnología de red de IBM.

Switches. Equipo de interconexión de redes que administra y distribuye, los paquetes a través de sus diferentes conexiones a la red, de acuerdo a las políticas que le fueron asignadas durante su configuración.

TDM. Time Division Multiplexing. Modo de administración de la señal que permite dividir en el tiempo el espacio destinado a los mensajes provenientes de diferentes canales.

Token Passing. Una modalidad de red de área local. El acceso al medio se comparte, teniendo derecho a transmitir solo una estación, la que posee en ese momento el token o permiso para hacerlo.

WAN. Red de Área Amplia. Red que interconecta localidades ubicadas en diferentes poblaciones, estado o países.

APÉNDICE A

Se hace una descripción exhaustiva de las pruebas realizadas y de los resultados obtenidos.

Identificación de la prueba	Pob	Gen.	pc	pm	Mayor valor al final del proceso	Mayor valor registrado durante el proceso	Cambios	Comentarios
Grupo 1 I42-230-800	200	800	0.6	0.001	2160.0000000	2800.9755859		<ul style="list-style-type: none"> ▪ Inicio de pruebas. ▪ No se incluye mutación en el orden. ▪ Poca variación una vez alcanzado un número máximo.
Grupo 1 I42-231-800	200	800	0.6	0.04	4055.4216309	4055.4216309	Incremento en la probabilidad de mutación.	<ul style="list-style-type: none"> ▪ Sin mutación en el orden ▪ Se alcanza un máximo. No es seguro que sea un máximo local. ▪ Un mismo orden persiste y se difunde en todos o casi todos los cromosomas.
Grupo 2 I43-230-1500	200	1500	0.6	0.04	2378.2709961	2800.9755859	<ul style="list-style-type: none"> ▪ Incremento al número de generaciones. ▪ Incluye mutación en el orden 	<ul style="list-style-type: none"> ▪ Arregla lista. Se modifica el arreglo de la lista haciendo pasar n en 2 veces la lista por este procedimiento. ▪ Valores altos desaparecen rápidamente. Puede ser atribuido a la alta probabilidad de mutación.
Grupo 2 L43-231-1500	200	1500	0.6	0.02	2800.9755859	3225.4055176	Decremento en la probabilidad de mutación.	<ul style="list-style-type: none"> ▪ Inestable el proceso. Se alcanza un máximo local, el cual se pierde al final del proceso. ▪ Valores altos continúan desapareciendo, esto puede ser atribuido a que continua siendo alta la probabilidad de mutación.
Grupo 2 L43-232-1500	200	1500	0.6	0.01	2160.0000000	3262.4997559	Decremento en la probabilidad de mutación.	<ul style="list-style-type: none"> ▪ Disminuye el máximo alcanzado en el proceso con respecto a las pruebas anteriores. ▪ Valores altos continúan desapareciendo. ▪ Se intentará inducir la aparición de valores altos incrementando la probabilidad de mutación.
Grupo 2 L43-233-1500	200	1500	0.6	0.025	2378.2709961	2800.9755859	Incremento en la probabilidad de mutación.	<ul style="list-style-type: none"> ▪ No se alcanza un máximo en el proceso mayor al de las pruebas previas. ▪ Valores altos continúan desapareciendo.

								<ul style="list-style-type: none"> Permanece inestable la aparición del mejor desempeño.
Grupo 2 L43-234-3000	200	3000	0.6	0.025	2378.2709961	3365.9997559	<ul style="list-style-type: none"> Incremento en el número de generaciones. Se duplica el número de generaciones 	<ul style="list-style-type: none"> Tiende a adquirir mayor estabilidad en la segunda mitad de las generaciones. Sin embargo continúan apareciendo valores inferiores a los máximos alcanzados. En más de una prueba se estabilizó en 2160.0000000, el cual es un valor inferior al máximo alcanzado durante el proceso. El mayor número de generaciones proporciona mayor estabilidad, se tiene máximos finales menores al máximo alcanzado durante el proceso
Grupo 2 L43-235-3000	200	3000	0.6	0.05	1947.5410156	2800.9755859	<ul style="list-style-type: none"> Incremento en la probabilidad de mutación 	<ul style="list-style-type: none"> Se incrementa la probabilidad de mutación esperando alcanzar máximos similares a los alcanzados en las primeras pruebas. Incremento de la inestabilidad de los máximos alcanzados durante el proceso. Valor máximo alcanzado durante el proceso es inferior al alcanzado en las pruebas anteriores.
Grupo 2 L43-236-3000	200	3000	0.6	0.01	3365.9997559	4055.4216309	<ul style="list-style-type: none"> Decremento en la probabilidad de mutación. 	<ul style="list-style-type: none"> Se disminuye la probabilidad de mutación esperando tener mayor estabilidad en el proceso. El proceso adquiere estabilidad. Incluso disminuye la pérdida de máximos alcanzados durante el proceso. El proceso tiende a estabilizarse en máximos locales. Se alcanza una vez el valor máximo del desempeño aparecido en pruebas previas.
Grupo 2 L43-237-3000	200	3000	0.6	0.005	3262.4997559	4055.4216309	<ul style="list-style-type: none"> Decremento en la probabilidad de mutación. 	<ul style="list-style-type: none"> Se disminuye la probabilidad de mutación esperando incrementar la estabilidad en el proceso. El proceso adquiere mayor estabilidad. Se observa la creación de columnas de alelos de igual valor, en la sección de orden, que no tienen variación a lo largo del proceso. Se obtiene nuevamente un máximo durante el proceso.

								<ul style="list-style-type: none"> ▪ En el promedio del desempeño se nota una variación contante en los valores. Esta observación no se hizo explícita antes, debido a que al ser inestable el máximo alcanzado durante el proceso, el promedio no tiene significado alguno. ▪ El promedio del desempeño muestra que no hay un crecimiento estable en los valores diferentes al máximo, e incluso la cantidad de cromosomas con este valor máximo seguramente es fluctuante durante la prueba.
L43-238-3000	200	3000	0.6	0.003	3262.4997559	4055.4216309	<ul style="list-style-type: none"> ▪ Decremento en la probabilidad de mutación. 	<ul style="list-style-type: none"> ▪ Se disminuye la probabilidad de mutación esperando estabilizar el promedio de desempeño. ▪ Adquiere mayor estabilidad el promedio. El valor que adquiere tiene como límite, aparentemente, la mitad del valor máximo de ese momento del proceso. Lo rebasa y disminuye el promedio. Creciendo nuevamente y decayendo una vez que se ha rebasado ese valor. ▪ En ocasiones se alcanzan valores máximos durante el proceso, pero no se logran mantener hasta el final. ▪ Persiste la creación de columnas de alelos de igual valor, en la sección de orden, las cuales no tienen variación a lo largo del proceso. Esto podemos considerarlo como convergencia en una solución óptima local. ▪ Se incrementa la aparición de alelos del cromosoma de red que son de segundo y de tercer nivel. Es decir, representan cadenas de nodos de longitud 3, a partir de la raíz.

Grupo 3 L44-230-1500	200	1500	0.6	0.01	3262.4997559	3262.4997559,	<ul style="list-style-type: none"> ▪ Decremento en la probabilidad de mutación. ▪ Incluye mutación en el orden 	<ul style="list-style-type: none"> ▪ Esta prueba es consecutiva de la l43-231-1500. Por lo que se prueba el decremento de la probabilidad de mutación. ▪ Se observa que la sección de orden no ha creado columnas de alelos iguales, por lo que podemos considerar todavía alta la probabilidad de mutación. ▪ La sección de red, del cromosoma, logra generar en algunos procesos cadenas de nodos de longitud 3. ▪ Valores altos se mantiene estables una vez que se alcanzan. Sin embargo no se alcanzan los valores máximos de aptitud que se generaron en pruebas anteriores.
Grupo 3 L44-230-3000	200	3000	0.6	0.003	4055.4216309	4055.4216309	<ul style="list-style-type: none"> • Se modifica el proceso de selección y cruzamiento, para transferir sin cambio los cromosomas de mayor aptitud. 	<ul style="list-style-type: none"> ▪ Con el fin de evitar que desaparezcan los individuos de mayor aptitud, se modifica el procedimiento de selección y cruzamiento para exentar de estos procesos a los individuos de mayor aptitud. También se les exenta de la mutación. ▪ Se acentúa la creación de columnas de alelos de igual valor, en la sección de orden, las cuales no tienen variación a lo largo del proceso. ▪ El proceso alcanza estabilidad rápidamente y el valor máximo persiste en ocasiones hasta el final del proceso. ▪ Se alcanza el valor de aptitud máximo que se logró en inicio de las pruebas.
Grupo 3 L44-231-3000	200	3000	0.6	.005	4055.4216309	4055.4216309	<ul style="list-style-type: none"> ▪ Incremento en la probabilidad de mutación. 	<ul style="list-style-type: none"> ▪ Se incrementa la probabilidad de mutación con el propósito de disminuir la persistencia de las columnas de alelos iguales en la sección de orden. ▪ Aparece en mayor número de ocasiones el valor máximo de la aptitud que se alcanzó en pruebas anteriores. ▪ Disminución perceptible en la aparición de alelos de tercer nivel, en la sección de red, del cromosoma. ▪ Disminución perceptible de las columnas de alelos en la sección de orden.

Grupo 3 L44-232-6000	200	6000	0.6	0.005	3365.9997559	6080.3144531	<ul style="list-style-type: none"> ▪ Incremento en el número de generaciones. 	<ul style="list-style-type: none"> ▪ Se incrementa el número de generaciones para averiguar si la probabilidad asignada es suficiente como para modificar el orden de los alelos de la sección de orden del cromosoma. ▪ Se obtienen máximos superiores a los obtenidos en pruebas anteriores. ▪ Estos máximos se pierden a pesar de estar exentos de cambio. ▪ Las columnas de alelos de igual valor en la sección de orden disminuyen en número. ▪ Se nota la formación de secuencias en la sección de orden del cromosoma, de tal modo que si se pudieran completar u ordenar de acuerdo a como se espera se podría entonces alcanzar el valor máximo de aptitud.
Grupo 3 L44-233-6000	200	6000	0.6	0.003	4651.8071289	4651.8071289	<ul style="list-style-type: none"> ▪ Decremento en la probabilidad de mutación. 	<ul style="list-style-type: none"> ▪ Se disminuye la probabilidad de mutación con el propósito de mantener estabilidad en los máximos de aptitud logrados durante el proceso ▪ No alcanzó los máximos de aptitud que existieron en la prueba anterior. ▪ Las secuencias no están desarrollándose en la sección de orden, del cromosoma.
Grupo 3 L44-234-6000	200	6000	0.6	0.004	4651.8071289	4651.8071289	<ul style="list-style-type: none"> ▪ Incremento en la probabilidad de mutación. 	<ul style="list-style-type: none"> ▪ Se incrementa la probabilidad de mutación para aumentar la posibilidad de cambio en el orden que se establece en cada cromosoma, buscando que queden finalmente en la secuencia correcta. ▪ Se incrementa la variedad con respecto a la prueba anterior, pero sin llegar a obtener los resultados de la prueba L44-232-6000.

Grupo 3 L44-235-6000	200	6000	0.6	0.005	6080.3149414	6080.3149414	<ul style="list-style-type: none"> ▪ Incremento en la probabilidad de mutación. ▪ Privilegiar hasta el 10% de la población de la nueva población para ser integrada por la población con mayor aptitud de la población evaluada. 	<ul style="list-style-type: none"> ▪ Se incrementa la probabilidad de mutación para crear una mayor variedad en los cromosomas, principalmente en la sección de orden. ▪ Se elige un 10% de la nueva población (ppp=10) para que sea integrada por los cromosomas de mayor aptitud de la población evaluada y que tengan la posibilidad esos cromosomas de ser transferidos a la nueva población sin pasar por los procesos de cruzamiento y mutación. ▪ Se nota una mayor estabilidad en esta prueba, que en la prueba anterior que tuvo la misma probabilidad de mutación, de 0.005 (la prueba fue la L44-232-6000). ▪ Al incrementar la estabilidad se decremento la variedad. Pero aún con esta disminución de variedad se logró obtener nuevamente el mayor valor de la aptitud que se obtuvo en pruebas anteriores.
Grupo 3 L44-236-6000	200	6000	0.6	0.005	6080.3144531	6080.3144531	<ul style="list-style-type: none"> ▪ Se disminuye el porcentaje de población que podrá repetirse en la siguiente generación. 	<ul style="list-style-type: none"> ▪ Se disminuye de 10% a 5% (ppp=5) el porcentaje de la nueva población que será transferida sin cambios desde la generación evaluada, con el fin de mantener una mayor cantidad de individuos que puedan ser modificados en los procesos de cruzamiento y mutación. ▪ Se mejora la variación de los cromosomas. Se obtiene nuevamente, en el primer ensayo de la prueba, el valor mayor de aptitud, igual al obtenido en pruebas anteriores.

Grupo 3 L44-237-12000	260	12000	0.6	0.005	6080.3149414	6080.3149414	<ul style="list-style-type: none"> ▪ Se incrementa el número de individuos en la población. ▪ Se incrementa el número de generaciones. 	<ul style="list-style-type: none"> ▪ Se incrementa el número de individuos en la población con el propósito de aumentar la variedad genética. ▪ Se incrementa el número de generaciones para dar mayor oportunidad a la aparición del orden esperado en los individuos de mayor aptitud, y con esto incrementar la aptitud en esos individuos. ▪ Se mantiene el valor de mayor aptitud que se alcanza en las primeras generaciones. ▪ Se alcanza el valor de mayor aptitud logrado en pruebas previas.
Grupo 3 L44-238-12000	260	12000	0.6	0.005	6080.3149414	6080.3149414	<ul style="list-style-type: none"> ▪ Se disminuye el porcentaje de la población que podrá repetirse en la siguiente generación. 	<ul style="list-style-type: none"> ▪ Se disminuye de 5% a 3% (ppp=3) el porcentaje de la nueva población que será transferida sin cambio desde la generación evaluada, el objetivo es reducir la posibilidad de permanencia de un valor de aptitud desde las primeras generaciones.
Grupo 4 L46-230-3000	200	3000	0.6	0.005	4651.8071289	4651.8071289	<ul style="list-style-type: none"> ▪ Se incluye un operador de ordenamiento llamado inversión. ▪ Se disminuye la población. ▪ Se disminuyen las generaciones 	<ul style="list-style-type: none"> ▪ El operador de inversión está compuesto de dos procesos de diferente probabilidad, el primero elige dos puntos dentro del cromosoma e invierte la posición de los alelos incluidos en el intervalo, este se le asigna una probabilidad de $p_{il} = 0.2$. En el segundo proceso se usa un corte y el otro será el extremo más cercano a punto de corte, a este se le asigna una probabilidad $p_{if} = 0.025$. ▪ Se observan los siguientes efectos, posiblemente por causa directa de la inversión: <ul style="list-style-type: none"> - Convergencia prematura a un valor. - Valores máximos, menores a los alcanzados en pruebas anteriores. - Pérdida del máximo alcanzado, regresando a valores menores.
Grupo 4 L46-231-3000	200	3000	0.6	0.005	4651.8071289	4651.8071289	<ul style="list-style-type: none"> ▪ Se disminuyen las probabilidades del operador de inversión. 	<ul style="list-style-type: none"> ▪ Se disminuyen los porcentajes de la inversión, a $p_{il} = 0.1$ y $p_{if} = 0.0125$.

Grupo 4 L46-232-6000	260	6000	0.6	0.005	4651.8071289	6080.3149414	<ul style="list-style-type: none"> ▪ Incremento a la población. ▪ Incremento de generaciones. 	<ul style="list-style-type: none"> ▪ Se incrementa la población 260 y las generaciones a 6000 con el propósito de observar el comportamiento del operador de inversión en poblaciones mayores.
Grupo 4 L46-233-6000	260	6000	0.6	0.005			<ul style="list-style-type: none"> ▪ Disminuyen los porcentajes de inversión. 	<ul style="list-style-type: none"> ▪ Se disminuyen los porcentajes de la inversión, a $p_{il} = 0.05$ y $p_{if} = 0.0125$.
Grupo 5 L46-234-12000	600	12000	0.6	0.005	8774.9990234	8774.9990234	<ul style="list-style-type: none"> ▪ Incremento en la población. ▪ Incremento en las generaciones. 	<ul style="list-style-type: none"> ▪ Se pretende alcanzar una aptitud mayor a través de una mayor diversidad en la población. Dando oportunidad a esta población de explorar una mayor cantidad de espacio al establecer un mayor número de generaciones. ▪ Se logra obtener un máximo no alcanzado previamente. Y la topología de la red resulta similar a la que corresponde a una solución óptima global. Esta tiene siete nodos en línea, de los 10 pretendidos.