

Diseño Topológico de Redes:

Comparación Cuantitativa y Búsqueda Heurística

Marcelo Mejía, Gabriela Figueroa y Jerónimo Robles

Departamento Académico de Computación
Instituto Tecnológico Autónomo de México
marcelo@lamport.rhon.itam.mx

Resumen

Una actividad muy importante que forma parte del diseño de las Redes Públicas de Datos es la selección de la mejor topología posible para la red. Si la localización geográfica de los nodos que conforman la red está dada, el problema de la selección de la topología se limita a determinar entre qué pares de nodos debe colocarse un enlace. Un algoritmo que resuelva adecuadamente este problema debe poder comparar numéricamente una topología con otra, ya que sólo así puede decidir cuál es mejor. Se debe contar, por lo tanto, con criterios cuantitativos de evaluación que traduzcan los requerimientos de diseño, frecuentemente ambiguos, establecidos por la administración del operador de la red. En este trabajo se presenta una solución a un problema de la industria mexicana que incluye la propuesta de una definición de la calidad de una topología y un programa computacional que utiliza esta definición para buscar la mejor topología posible. La definición de calidad involucra una combinación convexa de medidas del desempeño y del costo de una red: el diseñador puede variar un parámetro de la ecuación para darle más importancia al desempeño o al costo; la calidad de una topología puede, además, penalizarse si ésta no cumple con una o más restricciones de tolerancia a fallas. El programa de búsqueda desarrollado se basa en Algoritmos Genéticos y puede utilizarse para justificar la selección de una topología antes de iniciar la construcción de una red.

Palabras clave: Redes de Computadoras, Inteligencia Artificial, Optimización.

1. Introducción

Una de las consecuencias de la liberalización de las telecomunicaciones en México ha sido el surgimiento de diversos operadores de Redes Públicas de Datos que compiten por un mercado en constante crecimiento. En este contexto de competencia, los proveedores más importantes de Redes de Área Amplia (WANs) tienen puntos de presencia en las principales ciudades de la República, operan con la tecnología más moderna (Frame Relay, ATM, IP) y ofrecen diversos servicios de valor agregado (como por ejemplo, transmisión de voz, Intercambio Electrónico de Datos y acceso a Internet). Dado que la localización geográfica, la tecnología utilizada y los servicios ofrecidos por los proveedores son similares, probablemente la ventaja técnica que a mediano plazo distinguirá a los proveedores que lograrán sobrevivir a la feroz competencia sea la “calidad” de la red.

El diseño de una topología adecuada es, por lo tanto, de vital importancia para construir una WAN que sea rentable, eficiente y tolerante a fallas. Desafortunadamente, muchas empresas en México basan sus diseños de redes públicas a nivel nacional en "experiencias" previas con redes privadas pequeñas y no cuentan con procedimientos formales de diseño. Para complicar aún más el panorama, los algoritmos teóricos y las herramientas de diseño de redes disponibles en el mercado [7] no son lo suficientemente generales ni flexibles para que puedan ser usados cotidianamente por los Departamentos de Planeación de los operadores de Redes Públicas que quieren formalizar sus procesos de diseño.

Con el fin de mejorar el proceso de planeación de las nuevas redes WAN en México, se llevó a cabo un proyecto conjunto academia-industria entre el Instituto Tecnológico Autónomo de México y Red Uno, la filial de Telmex que opera la Red Pública de Datos UniNet. Los primeros resultados del proyecto incluyen la planeación de la capacidad de la mayor red Frame Relay en México [8] y una propuesta para la migración de su infraestructura dorsal hacia ATM [1]. El propósito del presente trabajo es presentar los resultados obtenidos en la segunda fase del proyecto: la especificación formal de algunos criterios de diseño que permiten comparar cuantitativamente topologías alternativas y su implementación en un paquete de diseño general, basado en Algoritmos Genéticos. Este paquete puede utilizarse fácilmente para obtener diferentes topologías de redes considerando diversos escenarios de costo, desempeño y tolerancia a fallas.

2. El diseño de la topología visto como un problema de optimización

Una red WAN está constituida por nodos (switches o enrutadores) y enlaces de transmisión que los interconectan. El diseño topológico atacado en este trabajo supone que la localización geográfica de los nodos está dada¹ y sólo busca determinar entre qué pares de nodos deben colocarse los enlaces. La asignación de capacidad a estos enlaces puede realizarse posteriormente al diseño de la topología, utilizando métodos analíticos aproximados o simulación de eventos discretos [8].

El algoritmo clásico de diseño topológico presentado en [2, 10] y adaptado en la Figura 1 consta de dos fases: primeramente se propone una topología inicial factible en base a la “experiencia” en redes anteriores y posteriormente se mejora iterativamente la topología. En la práctica se considera que una topología es factible si y sólo si existe un camino entre cualquier par de nodos de la red, es decir, si no se tiene ningún subconjunto de nodos aislados de los demás. El algoritmo modificado termina cuando se llega a un número límite de iteraciones o cuando se alcanza un valor predeterminado en la función objetivo: por ejemplo, cuando se obtiene una topología para una red Frame Relay cuyo costo está por debajo del presupuesto anual autorizado y en la cuál existe una ruta entre cada par de switches con menos de 9 saltos².

```
topologia = topologia_inicial_facible();
i=0;
while (!objetivo_alcanzado(topologia) && i<LIMITE){
    nueva_topologia = modificar(topologia);
    if ( es_mejor(nueva_topologia, topologia))
        topologia = nueva_topologia;
    i++;
}
```

Figura 1. Algoritmo clásico de diseño.

¹ La localización de los puntos de presencia de una red es generalmente resultado de un estudio de mercadotecnia.

² Con menos de 9 saltos se puede transmitir voz comprimida a 8 Kbps con calidad aceptable [6].

Para comparar una topología con otra y poder decidir cuál es mejor, es necesario contar con requerimientos de costo, desempeño y tolerancia a fallas especificados por la administración del operador de la red y con criterios cuantitativos que permitan medir qué tanto cumple una topología con estos requerimientos. Una de las funciones principales del Departamento de Planeación de los operadores de Redes Públicas es precisamente recabar los requerimientos de funcionamiento de la red, posiblemente ambiguos, de otros Departamentos (Finanzas y Evaluación Tecnológica, por ejemplo) y especificarlos de manera formal.

3. Criterios cuantitativos de comparación

Los criterios comparativos utilizados en el diseño de redes WAN deben considerar tres aspectos fundamentales: costo, desempeño y tolerancia a fallas.

Una red WAN debe ser rentable desde el punto de vista del proveedor que ofrece servicios públicos de transmisión de datos. La factibilidad económica de una red requiere que su costo de operación sea mínimo para poder ofrecer tarifas atractivas a sus clientes. En el contexto de una red WAN, el principal elemento que influye en el costo de operación es, sin duda, la renta mensual de los enlaces digitales de alta capacidad que se utilizan. Dado que el costo de la renta mensual de un enlace de larga distancia es función de la distancia que existe entre los nodos que une, para minimizar el costo de una red se debe minimizar la suma de las longitudes de sus enlaces.

Por otro lado, el desempeño de una red puede evaluarse midiendo el tiempo que tardan los paquetes en transitar de un extremo a otro de la red. Puesto que el principal componente de este tiempo de tránsito es el retardo experimentado por los paquetes en cada nodo, maximizar el desempeño equivale a minimizar el número de saltos existentes en la red entre cada par de nodos, sobre la ruta más corta que los une.

Finalmente, existen diversas restricciones de tolerancia a fallas de una red que se utilizan en la práctica; por ejemplo, puede requerirse que cada nodo de la red tenga por lo menos dos enlaces de salida o que los nodos permanezcan conectados aún en presencia de la falla de un nodo.

3.1. Grafos no dirigidos

Para establecer los criterios cuantitativos de comparación se considera la topología de una red WAN como un grafo no dirigido con vértices representando a los nodos y arcos representando a los

enlaces bidireccionales que conectan a los nodos. Si los grafos se representan por medio de su matriz de adyacencia, se pueden utilizar algoritmos computacionales muy conocidos y eficientes para obtener medidas cuantitativas de las topologías [5]. Por ejemplo, para determinar si una topología es factible se verifica si el grafo correspondiente es (1-)conectado calculando la cerradura transitiva de su matriz de adyacencia, por medio del algoritmo de Warshall, y verificando si contiene sólo unos. El número de ceros en la cerradura transitiva puede utilizarse como medida de qué tan cerca está un grafo de ser conectado.

3.2. Calidad de una topología

En términos generales, la calidad de una topología debe ser inversamente proporcional a su costo:

$$\text{calidad} \propto 1/\text{costo} \quad (1)$$

y directamente proporcional a su desempeño:

$$\text{calidad} \propto \text{desempeño} \quad (2)$$

Considerando que el desempeño es inversamente proporcional al número de saltos en la red, se tiene que:

$$\text{calidad} \propto 1/\text{saltos} \quad (3)$$

El número de saltos en la red puede determinarse calculando la ruta más corta entre cada par de vértices mediante el algoritmo de Dijkstra. Este algoritmo se ejecuta para cada vértice fuente del grafo hacia todos los vértices destino posibles. Como resultado de esta ejecución múltiple se obtiene para cada vértice el número de saltos que lo separa de cada destino. La variable saltos puede entonces tomarse como el promedio o como el máximo de los números de saltos calculados en el grafo. La decisión de cómo calcular exactamente el valor de la variable saltos es función de los requerimientos de diseño preestablecidos.

Dado que la calidad es función de dos variables: costo y saltos, es necesario cuantificar estas variables en términos relativos y no absolutos para que su importancia no dependa de las unidades en que se midan. Para mapear el costo y los saltos a un número real entre 1 y un valor MAX, se considera que, dado que el número y la localización de nodos son fijos, los grafos factibles se pueden mover entre

el mínimo costo del árbol de recubrimiento mínimo (MST) y el mínimo número de saltos del grafo completamente conectado (CC). Considerando que G_c representa el costo absoluto, en cualquier unidad, del grafo G , se define el costo relativo de un grafo como:

$$\text{costo} = 1 + (\text{MAX}-1)(G_c - \text{MST}_c) / (\text{CC}_c - \text{MST}_c) \quad (4)$$

De igual manera, si G_s representa el número absoluto de saltos, en cualquiera de sus dos acepciones, se define el número relativo de saltos en un grafo como:

$$\text{saltos} = 1 + (\text{MAX}-1)(G_s - \text{CC}_s) / (\text{MST}_s - \text{CC}_s) \quad (5)$$

Las relaciones (1) y (2) pueden entonces reescribirse como:

$$\text{calidad} \propto (\text{CC}_c - \text{MST}_c) / ((\text{CC}_c - \text{MST}_c) + (\text{MAX}-1)(G_c - \text{MST}_c)) = f_c \quad (6)$$

$$\text{calidad} \propto (\text{MST}_s - \text{CC}_s) / ((\text{MST}_s - \text{CC}_s) + (\text{MAX}-1)(G_s - \text{CC}_s)) = f_s \quad (7)$$

Para llegar a una sola ecuación que describa la calidad de una topología es necesario contestar la pregunta ¿qué factor de calidad, f_c o f_s , pesa más en un diseño particular? Obviamente, la respuesta depende de los requerimientos de diseño y por lo tanto debe depender de un parámetro R , que toma valores entre 0 y 1, especificado por el diseñador de la red. La calidad de una topología se define finalmente como una combinación convexa de los factores f_c y f_s :

$$\text{calidad} = \text{MAX} * (R * f_c + (1-R) * f_s) \quad (8)$$

3.3. Penalizaciones a la calidad

La calidad de una topología, dada por la ecuación (8), debe penalizarse si ésta no cumple con las restricciones de tolerancia a fallas especificadas por el diseñador. La penalización de la calidad debe tomar en cuenta qué tan importante es una restricción y qué tan cerca está un grafo de cumplir con ella. Se tiene, por lo tanto, para cada restricción un factor de penalización y un factor de proximidad que se multiplican para tener la penalización asociada. La calidad se penaliza entonces dividiendo la ecuación (8) entre la suma de las penalizaciones asociadas a cada restricción de tolerancia a fallas.

Por ejemplo, el factor de proximidad asociado a la restricción del grado mínimo que deben tener los vértices es muy sencillo, simplemente se determina el número de arcos que le hace falta a cada vértice para cumplir con este mínimo y se calcula la suma sobre todos los vértices. Asimismo, para

calcular el factor de proximidad asociado a un grafo que debiera ser 1-conectado aunque se remueva un arco, un requerimiento frecuente de diseño mencionado por los fabricantes de switches y enrutadores que consideran que sus equipos pueden ser muy confiables (debido a redundancia interna o externa), se elimina sistemáticamente un arco a la vez de la matriz de adyacencia y se calcula repetidamente la cerradura transitiva. El factor de proximidad en este caso puede ser el número de arcos cuya remoción desconecta el grafo, o el número de ceros encontrados en todas las matrices resultado, que es una medida de qué tan grave es la falla de un enlace. Si se repite el mismo ejercicio pero ahora eliminando vértices en lugar de arcos y se obtiene que el factor de proximidad es cero, entonces el grafo es 2-conectado, lo que significa que tiene por lo menos dos rutas disjuntas entre cada par de vértices.

En general, si la restricción de tolerancia a fallas es contar con al menos k rutas disjuntas entre cada par de vértices (grafo k -conectado), es necesario eliminar simultáneamente $k-1$ vértices sistemáticamente para calcular el factor de proximidad correspondiente.

3.4. Ejemplo

Para probar los criterios cuantitativos de diseño en un caso práctico se utilizó como base la infraestructura de la Red Pública de Datos UniNet. En la Figura 2 se muestra una topología propuesta por Red Uno y Telmex para alinear completamente la dorsal actual Frame Relay de UniNet a la infraestructura SDH de Telmex. Esta topología cumple con una restricción inicialmente establecida para la dorsal, que los nodos tuvieran como mínimo grado 2. Al aplicarle la ecuación (8) al grafo correspondiente, considerando el número promedio de saltos en la red y un valor de 100 para MAX, se obtuvo que:

$$\text{calidad} = 9.756 \cdot R + 1.972 \cdot (1-R)$$

La calidad de la topología se incrementa rápidamente al aumentar R , es decir al considerar más importante el costo y menos importante el desempeño. Se trata de un grafo que no es 2-conectado: si se remueve el vértice Ciudad de México, el grafo se desconecta. En la actual dorsal Frame Relay esto no es un problema ya que existen múltiples switches en la Ciudad de México, pero sí sería un problema al migrar a ATM e instalar allí un único switch de alto desempeño. Esta consideración fue una de las causas que motivó la necesidad de generar propuestas topológicas alternativas.

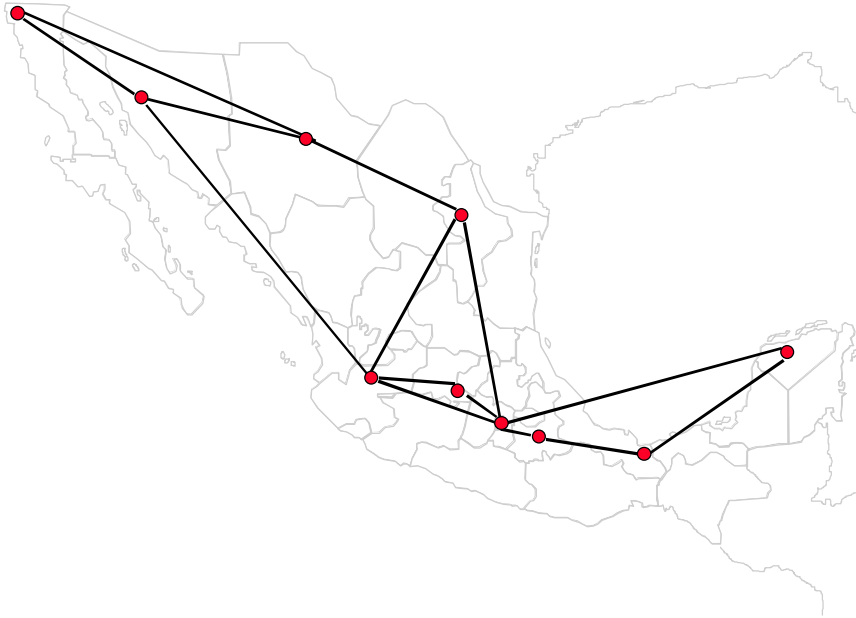


Figura 2. Topología propuesta inicialmente para UniNet.

4. Uso de Algoritmos Genéticos para el diseño de topologías

La utilización de Algoritmos Genéticos (AGs, [3]) para implementar un paquete que permita optimizar la topología de redes WAN proporciona 3 ventajas principales:

- Los AGs son muy generales y flexibles, por lo que pueden adaptarse a diferentes requerimientos de diseño.
- No se requiere contar con una topología inicial factible como base de la optimización. Sin embargo, si se cuenta con ella puede incorporarse fácilmente en la población inicial.
- El mejoramiento iterativo de la topología se efectúa en paralelo a partir de todos los esquemas presentes en la población aplicando los operadores de cruzamiento y mutación.

4.1. Codificación y aptitud

En el problema de diseño de una topología para una red WAN utilizando AGs, cada individuo de la población sobre la que se trabaja es una topología posible de la red. Para realizar su trabajo, los AGs requieren solamente que se especifique la codificación de los individuos (topologías) y una función de evaluación que mida su aptitud.

Dado que la localización de los nodos está dada, la codificación simplemente debe especificar qué enlaces existen entre los nodos. Si la red a diseñar consta de N nodos, cada individuo puede codificarse como una cadena de $N(N-1)/2$ bits dividida lógicamente en $N-1$ campos. Como se muestra en la Figura 3, el campo j contiene $N-j$ bits que especifican los enlaces que tiene el nodo j .

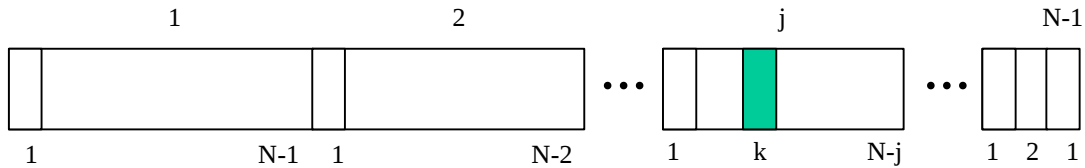


Figura 3. Codificación de los individuos: un 1 en el bit k del campo j indica que existe un enlace entre el nodo j y el nodo $j+k$.

La función de aptitud es la guía que utilizan los Algoritmos Genéticos para explorar de manera eficiente el enorme espacio de búsqueda que contiene $2^{N(N-1)/2}$ topologías posibles. Para el caso concreto del diseño topológico que nos interesa, se utiliza la definición de calidad especificada por la ecuación (8) como medida de la aptitud de los individuos, penalizada según corresponda por las restricciones no cumplidas de tolerancia a fallas.

4.2. Implementación

El paquete de diseño hace uso del software GENESIS 5.0 [4] para ejecutar el Algoritmo Genético de optimización y fue originalmente desarrollado para PCs. La selección de GENESIS, un software muy popular de dominio público, la escritura en ANSI C de la función de evaluación y la construcción de la interfaz gráfica utilizando Tcl/Tk, permiten que el paquete sea portable. Subsecuentes implementaciones en una estación de trabajo SUN SparcStation 5 y en una supercomputadora CRAY Origin 2000 no requirieron ninguna modificación en el código.

El paquete recibe como datos de entrada las coordenadas geográficas de las ciudades en las que se encuentran los nodos y los parámetros de la función de aptitud. La salida proporcionada por el paquete es un conjunto de las mejores topologías encontradas por el AG y sus aptitudes correspondientes. La función de aptitud es parametrizable y por lo tanto el paquete desarrollado es un software de diseño de topologías de uso general para redes WAN. Los parámetros que pueden modificarse son: la importancia relativa del costo y el desempeño (R), la selección de cómo calcular el

número de saltos en la red (máximo o promedio) y las restricciones de tolerancia a fallas que se consideran, junto con sus factores de penalización y proximidad correspondientes.

Como una opción del paquete, se pueden incorporar en la población inicial individuos escogidos por el diseñador; por ejemplo, puede considerarse inicialmente el grafo completamente conectado y/o el árbol de recubrimiento mínimo. Es posible, también, especificar que ciertos enlaces deben o no estar presentes en cualquier topología considerada.

4.3. Resultados

Se utilizó el paquete de diseño implementado para obtener diferentes topologías alternativas para la dorsal Frame Relay de UniNet. Dado que esta dorsal cuenta con 10 nodos, existen en total 2^{45} topologías posibles. El diseño se realizó considerando como medida de desempeño de las redes el promedio del número de saltos que hay entre cada par de nodos. La aptitud de los individuos se penalizó fuertemente si el grafo correspondiente no era 1-conectado: el factor de penalización en este caso fue MAX^2 y el factor de proximidad se escogió como el número de pares de nodos que no pueden comunicarse. La aptitud también se penalizó si el grafo no era 2-conectado, en este caso el factor de penalización fue MAX y el factor de proximidad se tomó como el número de nodos cuya remoción desconecta el grafo.

La Figura 4 muestra la mejor topología obtenida por el paquete para la dorsal de UniNet para valores de R entre 0.75 y 0.95. En el experimento que encontró esta topología se fijaron los enlaces del triángulo Ciudad de México – Guadalajara – Monterrey, ya que en [1] se propuso iniciar la migración de la dorsal hacia ATM precisamente en estas tres ciudades. Los parámetros utilizados por GENESIS en el proceso de búsqueda fueron: tamaño de la población, 105; población inicial, totalmente aleatoria; probabilidad de cruzamiento, 0.60; probabilidad de mutación, 0.001; total de intentos, 5000; selección, ruleta elitista. Con estos parámetros, cada experimento tarda aproximadamente 1 minuto en ejecutarse en una PC con procesador Pentium a 200 MHz y 32 Mbytes de RAM. La topología de la Figura 4 es 10% mejor, según los criterios de comparación presentados en este artículo, que la topología de la Figura 2 para $R = 0.75$.

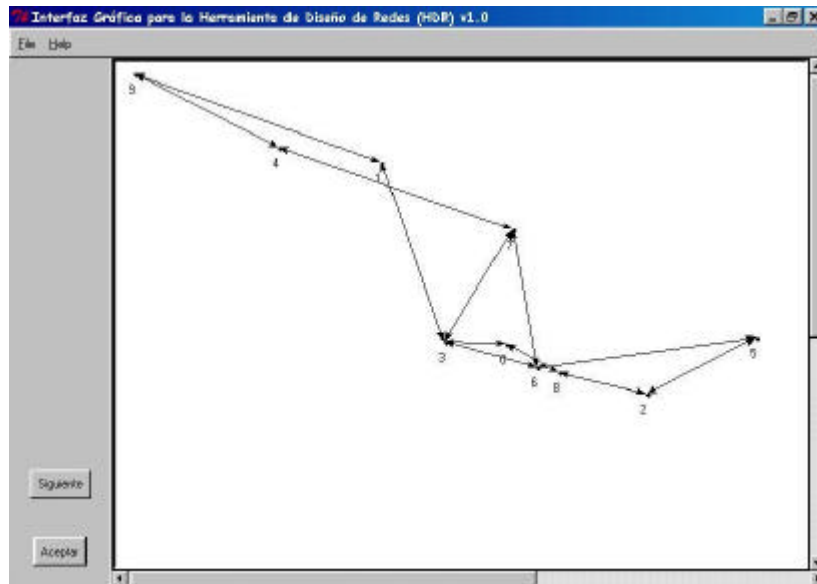


Figura 4. Mejor topología obtenida para $0.75 \leq R \leq 0.95$.

De los resultados obtenidos al realizar numerosos experimentos se observó que:

- Si la población inicial se escoge de manera completamente aleatoria, GENESIS converge rápidamente a máximos locales si el tamaño de la población es pequeño.
- La aptitud de las topologías resultado puede mejorar si se insertan en la población inicial topologías escogidas por el diseñador.
- Las topologías obtenidas con valores de R inferiores a 0.5 no son interesantes en la práctica ya que contienen demasiados enlaces.

5. Comentarios finales

Antes de iniciar la construcción de una WAN es necesario: 1) realizar un análisis detallado de las necesidades de los usuarios potenciales, 2) establecer formalmente y documentar las características que debe tener la topología de la red, y 3) utilizar un algoritmo de diseño consistente que permita justificar la selección de una topología. El trabajo presentado en este artículo se enfoca a los dos últimos puntos. Los criterios cuantitativos permiten plasmar los requerimientos de la red en una función de evaluación y el paquete de diseño desarrollado es capaz de adaptarse a diferentes requerimientos de la red, por ejemplo, se pueden introducir nuevas restricciones de tolerancia a fallas e incluso variar las medidas de

costo/desempeño. El uso de este paquete debe eliminar los diseños basados sólo en “experiencia” y permitir justificar la topología escogida para una red antes de su construcción.

Aunque los resultados presentados se refieren únicamente a la dorsal de una red, el paquete implementado puede aplicarse igualmente a un diseño jerárquico [9] que incluya una red dorsal, redes regionales y redes de acceso. En este caso, se pueden utilizar criterios de diseño (funciones de evaluación) independientes en cada nivel ya que sus objetivos son diferentes; por ejemplo, puede requerirse que la dorsal sea muy robusta, que las redes regionales tengan un desempeño elevado y que las redes de acceso sean lo menos costosas posibles. En cualquier caso, determinar cuál es la mejor topología para una red depende siempre de los requerimientos de costo, desempeño y tolerancia a fallas especificados por la administración del operador de la red.

Referencias

- [1] Acosta-Briseño, Mónica, y Flores-Mosri, Alejandra (1997), *Propuesta para la Integración de ATM en la Red Frame Relay de UniNet*, Tesis de Ingeniería en Computación, ITAM.
- [2] Ahuja, Vijay (1982), *Design and Analysis of Computer Communication Networks*, McGraw-Hill.
- [3] Goldberg, David (1989), *Genetic Algorithms in Search, Optimization and Machine Learning*, Addison-Wesley.
- [4] Grefenstette, John (1990), *A User's Guide to GENESIS*, Navy Center for Applied Research in Artificial Intelligence, Washington, D.C.
- [5] Langsam, Yedidyah, Augenstein, Moshe J., y Tenenbaum, Aaron M. (1997), *Estructuras de Datos con C y C++*, Prentice-Hall.
- [6] Mejía-Olvera, Marcelo y Flores-Ramiro, Alejandra (1997), *Las redes Frame Relay y la transmisión de voz - ¿incompatibilidad o integración?*, Reporte Técnico DAC-1, ITAM.
- [7] Mejía-Olvera, Marcelo y Flores-Ramiro, Alejandra (1997), *Principales características de algunas herramientas de simulación de redes de comunicaciones*, Reporte Técnico DAC-2, ITAM.
- [8] Mejía-Olvera, Marcelo y Flores-Mosri, Alejandra (1997), *Planeación de la Capacidad de una Red Pública de Datos*, Memorias de Conferencia Latinoamericana de Informática, Valparaíso, Chile.
- [9] Spohn, Darren L. (1996), *Data Network Design*, Second Edition, McGraw-Hill.
- [10] Tanenbaum, A. (1981), *Computer Networks*, Prentice-Hall.