# PLANIFICACIÓN Y DISEÑO DE REDES DE TRANSMISIÓN DE DATOS CON ALGORITMOS GENÉTICOS

Richard MercadoRivas, José Luis Cerrón Pérez
Rafael Rojas Bujaico, Yanina Mercado Rivas
Jhonny Jefferson Yance Rivera
Facultad de Ingeniería de Sistemas - Universidad Nacional del Centro del Perú

**RESUMEN**

Las redes de comunicación de datos han experimentado un enorme crecimiento en los últimos años debido al uso progresivo e incremental de las redes organizacionales y de la red de redes o Internet. Los requerimientos de calidad de servicio y confiabilidad de las redes modernas, acompañadas por grandes inversiones en ellas, han tornado críticos los problemas de diseño de las mismas siendo necesario el diseño óptimo de redes que reúnan determinadas características.

En la planificación de sistemas de transmisión de datos el diseño de la configuración necesaria para prestar un servicio de manera óptima respecto de algún criterio de desempeño es fundamental. Por ejemplo, si el criterio de desempeño es el costo, un problema a resolver es encontrar una topología de red que interconecte todos sus nodos al menor costo y que tenga la propiedad de asegurar la comunicación confiable de datos.

En esta investigación se analizan e implementan estrategias para el diseño óptimo de Redes de Datos para la obtención de topologías de mínimo costo y de confiabilidad admisible. En otras palabras se busca lograr, en forma eficiente, una asignación óptima de recursos que aseguren una confiabilidad especificada del sistema bajo diseño. La focalización está puesta en el diseño de redes de gran escala, tales como un backbone de telecomunicaciones, donde la métrica relevante es la confiabilidad total, es decir, la probabilidad de que cada par de nodos pueda comunicarse es de 100 por ciento.

En el presente trabajo de investigación se implementa un Algoritmo Genético con Java Genetic Algorithms Package - JGAP, para resolver un problema de diseño de redes confiables, con estructura poblacional por vecindades, utilizando valores límites superiores para la confiabilidad, para cualquier caso asociado al costo.

Los resultados obtenidos en este trabajo indican, por un lado, que el algoritmo propuesto puede ser utilizado para la implementación de redes de transmisión de datos con alto grado de confiabilidad.

**Palabras clave:** planificación y diseño de redes, redes de datos, algoritmos genéticos, java genetic algorithms package.

**ABSTRACT**

Data communication networks have experienced tremendous growth in recent years due to the progressive and incremental use of organizational networks and the network of networks or the Internet. The requirements for quality of service and reliability of networks, accompanied by large investments in them, have become critics of the same design issues still need optimal design of networks that meet certain characteristics.

In data transmission systems planning the design of the configuration required to serve optimally with respect to some criterion of performance is essential. For example, if the criterion of performance is the cost, a problem to be solved is to find a network topology that interconnect all its nodes at the lowest cost and that has the property to ensure reliable data communication.

This research analysed and implemented strategies for the optimal design of data networks to obtain acceptable reliability and lowest cost topologies. In other words seeks to achieve, in an efficient, optimal allocation of resources that ensures a specified reliability of the system under design. The focus is put on the network design of large scale, such as a backbone of telecommunications, where the relevant metric is total reliability, i.e. the probability that each pair of nodes can communicate is 100 percent.

The present research work implements a genetic algorithm with Java Genetic Algorithms Package - JGAP, to solve a problem of design of reliable networks, with population structure by neighborhoods, using values upper limits for reliability, for any case associated with the cost.

On the one hand, the results obtained in this study indicate that the proposed algorithm can be used for the implementation of data transmission networks with a high degree of reliability.

**Key word:** planning and design of networks, data networks, genetic algorithms, java genetic algorithms package.

## INTRODUCCIÓN

La Planificación y el diseño incorrecto de una red de datos pueden impactar con consecuencias indeseables sobre el funcionamiento de los sistemas que la utilizan. Cuando nos referimos a la planificación y diseño especialmente nos centramos en la forma de conectar los nodos de la red, aunque también se podrían incluir otros aspectos como por ejemplo el tipo de enlace entre cada par de nodos. La topología está íntimamente ligada a las facilidades, posibilidades y restricciones de los servicios que corren sobre la red. Es por ello que resulta de sumo interés contar con las herramientas y técnicas necesarias para poder llevar a cabo la planificación y el diseño de una red en la forma más óptima.

Dada la ubicación de los nodos, encontrar cuáles deben interconectarse, respetando un conjunto de restricciones, es un problema no trivial. Generalmente consiste en hallar la mejor topología de la red que maximice las prestaciones y minimice los costos. Este análisis suele suponer redes con y sin tolerancia a fallas. La tolerancia a fallas es otra de las restricciones impuestas que hay que conservar y que se logra generalmente con redundancia, por ejemplo garantizando más de una ruta posible entre cualquier par de nodos.

La optimización de la planificación y diseño de redes con restricciones tiene un amplio rango de aplicación. Es particularmente importante en las redes de conmutación de circuitos a nivel de backbone, interconexión de redes, redes de fibra óptica para telefonía, computadoras tolerantes a fallas y arquitectura de switchs.

Los problemas de la planificación y diseño de topologías de redes sujeto al cumplimiento de determinadas restricciones como por ejemplo minimizar costos, garantizar conectividad, minimizar retardos entre otros, están comprendidos en una clase de problemas más general conocida con el nombre de optimización combinatoria.

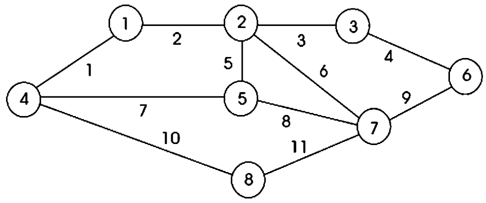
En conclusión de lo antes expuesto la situación es la siguiente: es necesario obtener una buena topología que cumpla con determinadas restricciones para lograr una planificación y diseño eficiente y confiable de la red.

## MATERIALES Y MÉTODOS

### DISEÑO DE RED

Una red puede ser modelada a través de un grafo simple G = (V,E) donde V={v1,v2,v3,….vn} es el conjunto de n nodos o vértices que representan los centros de comunicaciones de la red y E={e1,e2,e3,…en} es un conjunto de m vínculos o enlaces que permiten la conexión entre los nodos.

Cada enlace ek se identifica con un par no ordenado (vi,vj). Los vínculos pueden ser orientados o no-orientados representando enlaces de comunicaciones que son unidireccionales o bidireccionales respectivamente, pero no se permiten lazos cerrados ni enlaces redundantes (un par de nodos cualesquiera están directamente vinculados, como máximo, por un único enlace).




Donde:

G=(V.E)

V={1,2,3,4,5,6,7,8}

E={(1,2),(1,4),(2,3),(2,5),(2,7),(3,6),(4,5),(4,8),(5,7),(6,7),(7,8)}

La cardinalidad de un conjunto H, denotada como |H| es la cantidad de elementos que dicho conjunto posee. Así el número de nodos en un grafo G se representa como n =|V| mientras que el número de enlaces es e =|E|. Para el grafo anterior n = |V| = 8 y e = |E| = 11.

Un problema a resolver en el diseño de redes consiste en la elección de enlaces de comunicaciones, de diferentes características o tecnologías, entre un conjunto de nodos dado, de forma tal que la red resultante alcance un conjunto de valores tales como costo o confiabilidad.

Existen dos aspectos a tener en cuenta en el diseño óptimo de redes confiables: la obtención de la topología de mínimo costo y el cálculo de la confiabilidad. El segundo aspecto limita el espacio de búsqueda factible entre el conjunto de todas las topologías posibles.

## REPRESENTACIÓN MATRICIAL DE GRAFOS

Los grafos constituyen una herramienta útil para el estudio y modelado topológico de redes. Para su análisis y proceso es factible representarlos en forma matricial.

### MATRIZ INCIDENTE

Sea un grafo G con n nodos y m enlaces, sin lazos. Se define la matriz A de dimensión n por m tal que A = [aij], donde las n filas corresponden a los n nodos y las m columnas corresponden a los m enlaces.

$$A(G)=\begin{bmatrix} 1&1&0&0&0&0&0&0&0&0&0\\ 0&1&1&0&1&1&0&0&0&0&0\\ 0&0&1&1&0&0&0&0&0&0&0\\ 1&0&0&0&0&0&1&0&0&1&0\\ 0&0&0&0&1&0&1&1&0&0&0\\ 0&0&0&1&0&0&0&0&1&0&0\\ 0&0&0&0&0&1&0&1&1&0&1\\ 0&0&0&0&0&0&0&0&0&1&1 \end{bmatrix}$$

Esta matriz A se denomina matriz de incidencia nodo-enlace o simplemente matriz de incidencia. A veces, si la matriz A representa el grafo G se la simboliza como A(G).

Sea un grafo G con n nodos y m enlaces, sin lazos. Se define la matriz A de dimensión n por m tal que A = [aij], donde las n filas corresponden a los n nodos y las m columnas corresponden a los m enlaces. Esta matriz A se denomina matriz de incidencia nodo-enlace o simplemente matriz de incidencia. A veces, si la matriz A representa el grafo G se la simboliza como A(G).

## MATRIZ DE ADYACENCIA Y MATRIZ DE COSTO

Sean n el número de nodos y m la cantidad de enlaces de una red de datos representada por un grafo simple G. Luego, la matriz de conexión M = [mij] es una matriz binaria de dimensión n por n y donde: mij = 1, si el enlace que une los nodos i y j está disponible y mij = 0, en caso contrario.

$$\mathbf{C}(G)=\begin{bmatrix} 0&1&0&1&0&0&0&0\\ 1&0&1&0&1&0&1&0\\ 0&1&0&0&0&1&0&0\\ 1&0&0&0&1&0&0&1\\ 0&1&0&1&0&0&1&0\\ 0&0&1&0&0&0&1&0\\ 0&1&0&0&1&1&0&1\\ 0&0&0&0&0&0&1&0 \end{bmatrix}$$

La matriz de costo C = [cij] tiene la misma dimensión que la matriz de adyacencia donde cada elemento cij representa el costo del enlace entre los nodos i y j. Así, mientras que la matriz de adyacencia representa un grafo simple, la matriz de costo representa un grafo simple ponderado. Los ceros se mantienen invariante, asumiendo que un costo nulo representa un enlace que no está disponible.

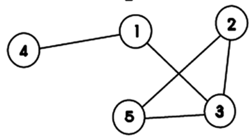
$$\mathbf{C}(G)=\begin{bmatrix} 0&18&0&21&0&0&0&0\\ 18&0&32&0&10&0&12&0\\ 0&32&0&0&0&17&0&0\\ 21&0&0&0&11&0&0&18\\ 0&10&0&11&0&0&18&0\\ 0&0&17&0&0&0&15&0\\ 0&12&0&0&18&15&0&13\\ 0&0&0&18&0&0&13&0 \end{bmatrix}$$

Las matrices M(G) y C(G) son siempre cuadradas dado que tanto las filas como las columnas representan el mismo parámetro (los nodos del grafo) y son simétricas porque el enlace que une el i-ésimo con el j-ésimo nodo es el mismo que une el j-ésimo con el i-ésimo nodo. Al no estar permitidos los lazos, la diagonal es siempre nula.

## MODELADO CON VARIABLES BINARIAS

El grafo que modela la red puede ser representado en un arreglo unidimensional teniendo en cuenta todos los enlaces posibles entre los nodos. Si se considera la red de 5 nodos, la cantidad de enlaces posibles está dada por la cardinalidad del conjunto E la cual puede calcularse como:

$$|E|=\frac{|V|\cdot(|V|-1)}{2}=10$$




Todos los enlaces posibles (i,j) entre cada par de nodos i y j, numerando los nodos de 1 a 5 conforman el conjunto E={e1, e2, e3, e4, e5, e6, e7, e8, e9, e10}={(1,2) (1,3) (1,4) (1,5) (2,3) (2,4) (2,5) (3,4) (3,5) (4,5)}. El arreglo que representa la topología es x tal que xi=1 si ei existe y xi=0 en caso contrario. La representación de x con el esquema de conexión será:

x = [0 1 1 0 1 0 1 0 1 0]

## DISEÑO ÓPTIMO DE REDES CONFIABLES

Una de las primeras propuestas se debe a quien propuso un método basado en branch y bound para la exploración de las diferentes topologías e hizo uso de la técnica de backtracking. Adicionalmente estableció estrategias para determinar límites superiores para la confiabilidad de una topología propuesta, de forma tal de evitar el cálculo innecesario de la confiabilidad exacta.

Todo grafo conectado tiene al menos n-1 enlaces activos y nmax=n(n-1)/2 enlaces posibles. Pueden existir problemas en los cuales no todos los enlaces están disponibles. Esto reduce la cantidad de enlaces activos posibles a un valor n*≤ nmax. Entonces, podrían encontrarse todas las combinaciones posibles de enlaces activos, descomponiendo el problema en n*-(n-1)+1 = n*-n+2 subproblemas (sp) de manera tal que cada uno de ellos involucre una cantidad fija de estos enlaces activos.

La solución al problema principal menos elaborada consistiría en encontrar la mejor solución de cada subproblema para luego optar por la mejor de todas. Sin embargo, este enfoque puede resultar ineficiente desde el punto de vista de procesamiento, porque es necesario resolver todos y cada uno de los subproblemas.

El trabajo de Jan sugiere calcular primeramente un límite inferior de cantidad de enlaces a*, que combinados de alguna manera permitan que la red posea una confiabilidad mayor o igual a la admisible R0. De esta forma se evita inspeccionar aquellos subproblemas cuya cantidad de enlaces se encuentre entre n-1 y a*.

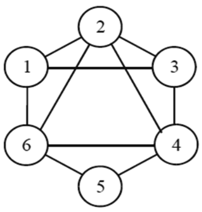
Luego se resuelve el subproblema correspondiente a a*, sp (a*), donde se busca la mejor solución posible con a* enlaces activos y se establece temporalmente la solución encontrada como la mejor solución global.

Seguidamente se procede a evaluar el menor costo esperado para el siguiente nivel (a*+1), mediante la suma de los costos de los a*+1 enlaces más baratos. Si esta combinación es más cara que la ya encontrada en el nivel a*, aquella era la mejor de todas las posibles. En caso contrario, habrá que resolver el subproblema sp(a*+1). Se procede así sucesivamente hasta encontrar la solución del problema principal. De esta manera es posible evitar el registro de cierta cantidad de niveles, dependiendo de las características de cada caso.

## RESULTADOS

A continuación se describen las etapas del algoritmo diseñado y el conjunto de estrategias para proporcionar conocimiento específico del problema con el objeto de favorecer la búsqueda, optimizando el mejor diseño de red de transmisión de datos.

La variable de decisión se representa a través de un vector de bits x donde cada componente xi determina la existencia o ausencia del enlace i. El tamaño del vector x expresado en función del número de nodos N de la red es N*(N-1)/2.




La figura muestra una red con seis nodos y con algunos enlaces activos. El conjunto de enlaces posibles es E = {(1,2) (1,3) (1,4) (1,5) (1,6) (2,3) (2,4) (2,5) (2,6) (3,4) (3,5) (3,6) (4,5) (4,6) (5,6)}. El cromosoma que representa la red es:

x = [1 1 0 0 1 1 1 0 1 1 0 0 1 1 1 ]

Que determina el conjunto de enlaces activos:

E = {(1,2) (1,3) (1,6) (2,3) (2,4) (2,6) (3,4) (4,5) (4,6) (5,6)}.

## GENERACIÓN DE LA POBLACIÓN INICIAL

La generación de una población inicial (PI) sembrada permite un mejor punto de partida para el algoritmo genético si se considera esta población conformada con soluciones que, a priori, tengan buenas características.

Así, en primer lugar se evita que en la población inicial se incorporen cromosomas que representen topologías que contengan menos de n-1 enlaces activos dado que se sabe de antemano que no serán factibles. Habiéndose ejecutado el programa en repetidas oportunidades, se observó una mejora interesante con la incorporación de mayor conocimiento específico del problema, esta vez teniendo en cuenta las características recomendadas para el diseño de topologías confiables

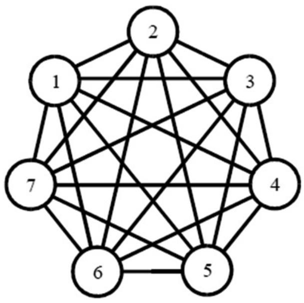
En ese sentido una población inicial que sea generada asignando mayor probabilidad de que la matriz de conexión que representa cada red tenga mayormente unos (1's) en el entorno de su diagonal principal, promueve topologías que tienden al anillo. Los enlaces restantes, si los hubiera, se distribuyen libremente.

Respecto al número de individuos de cada población, se debe lograr una vez más un equilibrio, en este caso entre velocidad de convergencia y esfuerzo computacional. El tamaño de la población se fijó sobre la base del análisis de los reportes de diferentes trabajos y a partir de los resultados de las experiencias realizadas.

## EXPERIMENTOS

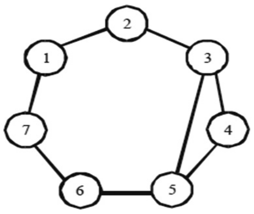
Para las pruebas de ajuste del algoritmo implementado se tomó un ejemplo de una red de siete nodos que se muestra en la figura resuelto con un método de Branch y Bound y se compararon con los resultados obtenidos mediante un Algoritmo Genético.

La matriz de costos asociada se indica en la siguiente figura. Para este experimento se solicitó una confiabilidad total del 90%, todos los enlaces son permitidos y tienen una probabilidad de falla de 0.10. La probabilidad de cruzamiento utilizada fue de 0.7 y la de mutación 0.002. Se seleccionó un 90% de la población total para apareamiento y la presión selectiva fue de 1.3.



$$\mathbf{C} = \begin{bmatrix} - & 125 & 150 & 125 & 150 & 150 & 130 \\ 125 & - & 75 & 100 & 150 & 200 & 250 \\ 150 & 75 & - & 75 & 90 & 250 & 200 \\ 125 & 100 & 75 & - & 75 & 100 & 150 \\ 150 & 150 & 90 & 75 & - & 75 & 100 \\ 150 & 200 & 250 & 100 & 75 & - & 75 \\ 130 & 250 & 200 & 150 & 100 & 75 & - \end{bmatrix}$$

La red solución encontrada mediante el Algoritmo Genético con selección local propuesto se muestra en la figura y corresponde al conjunto de enlaces E={(1,2) (1,7) (2,3) (3,4) (3,5) (4,5) (5,6) (6,7)}, con un costo de 720 y una confiabilidad de 0.91. Esta solución coincide con la hallada por métodos exactos.




La calidad de las soluciones encontradas por ambos algoritmos puede apreciarse en la distribución de las mismas en la generación 200, mostrada en la figura. Para el Algoritmo Genético con selección local el valor mínimo de fitness encontrado corresponde a 720, con un promedio de 777.3 y una dispersión de 71.9 En tanto que para un Algoritmo Genético simple los valores respectivos son 770, 2294.1 y 1212.6.

## CONCLUSIONES

1. La planificación y diseño de redes de transmisión de datos utilizando algoritmos genéticos permite la implementación de diferentes modelos aplicables a un ambiente distribuido, permitiendo solucionar problemas de confiabilidad que surgen al planificar y diseñar redes de comunicaciones de alta conectividad topológica.
2. Los Algoritmos Genéticos constituyen uno de los paradigmas de Computación Evolutiva emergentes, basados en la simulación del principio de evolución. En los mismos, una población de soluciones potenciales evoluciona de acuerdo a interacciones cruzamiento y transformaciones únicas mutaciones. Luego de varias generaciones, la población converge hacia una solución óptima.
3. El modelo de algoritmos genéticos constituye una alternativa importante para mejorar la eficiencia de las soluciones propuestas para los problemas de grandes dimensiones resueltos mediante técnicas de programación evolutiva.
4. Java Genetic Algorithms Package JGAP es un framework de uso libre basado en la tecnología JAVA que provee múltiples mecanismos que permiten aplicar principios evolutivos en la resolución de problemas.

## REFERENCIAS BIBLIOGRÁFICAS

Andres S. Tanenbaum. Redes de Computadoras. Cuarta Edición. Editorial: Pearson Educación. Mexico 2004.

Behrouz Forouzan. Redes de Comunicaciones. Cuarta Edición. Editorial Mc Graw Hill. España 2007.

David A. Coley. An introduction to Genetic Algorithms for Scientists and Engineers. Tercera Edición. Editorial: World Scientific Publishing. U.K. 1999.

Forouzan. Transmisión de datos y Redes de Comunicaciones. Cuarta Edición. Editorial McGraw-Hill. Mexico 2011.

J. Stender, E. Hillebrand, J. Kingdon. Genetic Algorithms in Optimisation, Simulation and Modelling. Primera Edición. Editorial: IOS PRESS. Amsterdam 2001.

Huidobro Moya, José Manuel. Telecomunicaciones, Tecnologías, Redes y Servicios. Segunda Edición. Editorial RA-MA. España 2010.

Michelle, Melanie. An introduction to Genetic Algorithms. Primera Edición. Editorial: MIT Press Paperback. U.S.A. 2008.

Moreno Pérez, Juan Carlos/ Santos Gonzales, Menuel. Sistemas Informáticos y Redes Locales. Primera Edición. Editorial RA-MA. España 2012.

Sucre H. Ramires. Introducción a las Redes de Datos . Primera Edición. Editorial: Amazon Digital Services. España 2011.

Randy L. Haupt, Sue Ellen Haupt. Practical Genetic Algorithms. Segunda Edición. Editorial: John Wiley & Sons, Inc.. Canada 2004.

Zbigniew Michalewicz. Genetic Algorithms + Data Structures = Evolution Programs. Primera Edición. Editorial: Departament of Computer Science University of North Carolina. U.S.A. 1999

## DIRECCIÓN ELECTRÓNICA

*http://geneura.ugr.es/~jmerelo/ie/ags.htm*
Informática evolutiva: Algoritmos genéticos
*http://the-geek.org/docs/algen/*
Algoritmos genéticos y computación evolutiva
*http://www.esla.com/Algoritmos-Geneticos-Optimizacion-Recursos-CGB-Informatica.html*
Algoritmos Genéticos y Optimización Combinatoria
*http://the-geek.org/docs/algen/*
Algoritmos genéticos y computación evolutiva
*http://cala.unex.es/cala/epistemowikia/index.php?title=Algoritmos_ Gen%C3%A9ticos*
Algoritmos genéticos