

COMPARACIÓN ESTADÍSTICA NO PARAMÉTRICA DE TRES ALGORITMOS EVOLUTIVOS APLICADOS AL TSP

Uriel Ervey Bernal Magallanes (1), Héctor José Puga Soberanes (2), Juan Adolfo Montesino Guerra (2)

1 Ingeniería en Sistemas Computacionales, Tecnológico Nacional de México- Instituto Tecnológico de León | Dirección de correo electrónico: ing.uriel.bernal@gmail.com

2 Departamento, División de Investigación y Estudios de Posgrado, León, Tecnológico Nacional de México – Instituto Tecnológico de León | Dirección de correo electrónico: pugahector@yahoo.com,adolfo.montesino@hotmail.com

Resumen

El Problema del agente viajero (TSP) es un problema de optimización combinatoria muy estudiado en el área de computación científica y matemáticas aplicadas. La importancia del TSP radica en que varios problemas de optimización combinatoria se pueden formular con base en él. Hasta la fecha no se ha encontrado un algoritmo determinístico que resuelva el TSP en un tiempo polinomial. En el estado del arte se han reportado soluciones factibles en tiempo polinomial, mediante el uso de algoritmos no determinísticos conocidos como Metaheurísticas. En este trabajo se implementaron: el Algoritmo Genético (AG), Algoritmo Memético (AM), y el Algoritmo de Sistema Inmune (ASI) para resolver el TSP Simétrico. Los algoritmos implementados pertenecen a una familia de Metaheurísticas conocida como Algoritmos Evolutivos los cuales están inspirados en la evolución natural.

Para identificar el desempeño de los algoritmos seleccionados, se realizó una comparación entre ellos haciendo uso de estadística no paramétrica para evidenciar el algoritmo con mejor desempeño para resolver el TSP.

Abstract

The Traveling Salesman Problem (TSP) is a very studied combinatorial optimization problem in the fields of scientific computing and applied mathematics. The importance of TSP comes from the fact that several combinatorial optimization problems can be formulated in base to it. Until now, an algorithm that solves the TSP in polynomial time has not been found. Feasible solutions in polynomial time have been found in the state of the art, through using nondeterministic algorithms known as Metaheuristics. In this work, the Genetic Algorithm (GA), Memetic Algorithm (MA), and Immune System Algorithm (ISA) were implemented for solving the Symmetric TSP. Those algorithms belong to a family of Metaheuristics known as Evolutionary Algorithms which are inspired by natural evolution. A comparison between them has been made using non-parametric statistics to identify their performance and show the algorithm with the best performance for solving the TSP.

Palabras Clave

Problema del Agente Viajero (TSP), Metaheurísticas, Algoritmo Genético, Algoritmo Memético, Algoritmo de Sistema Inmune

INTRODUCCIÓN

Existen problemas de optimización combinatoria cuya solución óptima no puede ser encontrada debido al tiempo de cómputo que requieren en la práctica. El TSP es un problema NP-Completo, por lo tanto no puede ser resuelto por un algoritmo determinístico en tiempo polinomial [1]. Se desconoce el verdadero origen del TSP, pero los primeros trabajos se encuentran entre los años, 1920 y 1930 por el matemático y economista Karl Menger [2]. Después de varios intentos por resolver el problema, fue popularizado en la RAND Corporation por Merrill Flood y sus colegas [2].

El TSP se puede plantear de la siguiente manera: Dado un conjunto de n ciudades $V = \{1, 2, 3, \dots, n\}$, y un conjunto de aristas $A = \{(i, j) | i, j \in V\}$, encontrar el ciclo hamiltoniano de costo mínimo (1).

El modelo del TSP se puede plantear como en [3]:

$$\text{Min} \sum_{(i,j) \in A} c_{ij} x_{ij} \quad (1)$$

$$\sum_{(i,j) \in A} x_{ij} = 1 \quad \forall j \in V \quad (2)$$

$$\sum_{(i,j) \in A} x_{ij} = 1 \quad \forall i \in V \quad (3)$$

$$\sum_{(i,j) \in A: i \in U, j \in (V-U)} x_{ij} = 1 \quad 2 \leq |U| \leq |V| - 2 \quad (4)$$

De donde $A \subseteq V \times V$ es el conjunto de las aristas que interconectan las ciudades.

x_{ij} es la variable de decisión del problema que está dada por:

$$x_{ij} = \begin{cases} 1, & \text{si existe } (i, j) \in A \\ 0, & \text{en caso contrario} \end{cases}$$

c_{ij} es el costo o la distancia asociado a x_{ij} .

Existen dos versiones del TSP:

- TSP simétrico, cuando $c_{ij} = c_{ji}$.
- TSP asimétrico, cuando se tiene la restricción de $c_{ij} \neq c_{ji}$.

Metaheurísticas

Las Metaheurísticas son algoritmos no determinísticos de propósito general, adaptables a varios problemas de optimización que ofrecen soluciones factibles en un tiempo razonable, pero no garantizan encontrar la solución óptima. Estos algoritmos implementan procedimientos y estrategias de alto nivel para escapar de óptimos locales y realizar una búsqueda robusta en el espacio de soluciones [4]. Su uso en muchos problemas ha mostrado eficiencia y efectividad para resolver problemas complejos.

Algunas áreas en las que se ha aplicado e implementado el uso de Metaheurísticas son [5]:

- Diseño de ingeniería, optimización topológica, optimización estructural en electrónica y VLSI (*Very-large-scale integration*), aerodinámica, dinámica de fluidos, telecomunicaciones, y robótica.
- Máquinas de aprendizaje y minería de datos en bioinformática, computación biológica y finanzas.
- Modelación de sistemas, simulación en química, física, y biología; control, señales y procesamiento de imágenes.
- Planificación en problemas de enrutamiento, planificación de robots, programación y producción de problemas de logística y transporte, gestión de cadena de suministros, y medio ambiente.

Algoritmo Genético

Los Algoritmos Genéticos (AGs) son algoritmos de búsqueda, optimización y máquinas de aprendizaje basados en la teoría de la Selección Natural, la Evolución y la Genética [6]. Son una clase muy popular de los Algoritmos Evolutivos. Estos fueron inventados y desarrollados por John Holland y sus colegas en la universidad de Michigan alrededor de 1970 [6][7]. El AG consta de una población inicial y de los siguientes operadores para su funcionamiento:

- **Selección:** Este operador selecciona los individuos que se van a reproducir en base a su función.
- **Cruza:** Este operador realiza un intercambio de genes entre dos individuos.
- **Mutación:** Modifica el material genético de un individuo de manera aleatoria.

Algoritmo Memético

Los algoritmos Meméticos (AMs) son técnicas de optimización que combinan conceptos y estrategias de otras Metaheurísticas, tales como las de búsqueda en poblaciones (como los algoritmos evolutivos) con técnicas de mejora local, con la finalidad de obtener mejoras individuales en las soluciones por medio de procesos de cooperación y competiciones. Estos algoritmos han resultado tener buen desempeño debido a la búsqueda local implementada.

Los orígenes de los algoritmos miméticos se remontan a finales de los años ochenta [8]. La denominación “memético” surge del término inglés “meme”, el cual fue acuñado por R. Dawkins como el análogo del gen en el contexto de la evolución cultural [8].

Algoritmo De Sistema Inmune

El algoritmo de sistema inmune (ASI) también conocido como algoritmo de selección clonal, es una técnica que es capaz de resolver problemas de optimización combinatoria y multi-objetivo [9]. Este algoritmo está inspirado en el funcionamiento del sistema inmune biológico de seres vivos.

El principio de esta técnica se basa en la acumulación de mutaciones para generar una respuesta inmune madura de manera inmediata. Se muestra en la imagen 1 el principio del algoritmo de sistema inmune para generar respuestas inmunes:

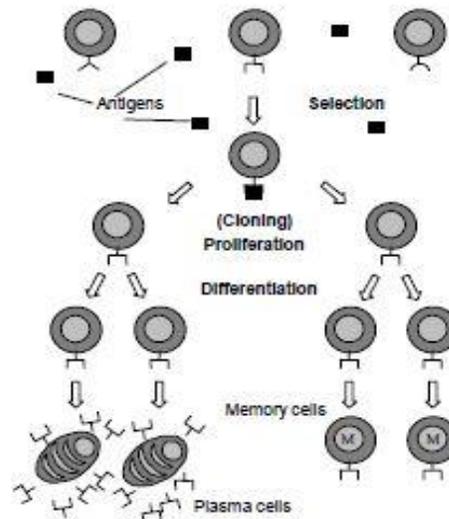


IMAGEN 1: Principio del ASI [9].

MATERIALES Y MÉTODOS

Para la comparación de los algoritmos se utilizaron las siguientes instancias de TSP simétrico: KroA100, KroA150, KroA200, KroB100, KroB150, KroB200, KroC100, KroD100, KroE100, y u574. Las cuales fueron tomadas de [10]. Para evidenciar la comparación entre el AG, AM, y ASI, se aplicó la prueba no paramétrica de Friedman como en [11] utilizando a la mediana como representante estadístico. Se utilizó (1) como función de fitness en los algoritmos, y como condición de paro se utilizó el criterio de cantidad de llamadas a función.

La configuración de parámetros y operadores establecidos en los algoritmos se muestran en la tabla 1:

Tabla 1: Ajuste de parámetros y operadores

Parámetros y operadores	Metaheurísticas		
	AG	AM	ASI
Población	33	33	33
Nº Llamadas a función	100,000	100,000	100,000
% de Elitismo	20%	20%	20%
Nº de Hipermuta	-	-	4
Codificación	Entera	Entera	Entera
Selección	Torneo	Torneo	-
Cruza	OX	OX	-
Mutación	Swap-Two	Swap-Two	Swap-Two

RESULTADOS Y DISCUSIÓN

Los experimentos se realizaron utilizando las 10 instancias mencionadas. Para realizar la prueba de Friedman se calculó la mediana de los 33 mejores individuos de 33 ejecuciones de cada algoritmo sobre cada instancia de prueba. La tabla 2 muestra las medianas obtenidas en la experimentación:

Tabla 2: Mediana de las mejores distancias en 33 ejecuciones por cada algoritmo para cada instancia

Instancia	Óptimo conocido	Metaheurísticas		
		AG	AM	ASI
KroA100	21288	53537.4819	38635.2730	61829.4533
KroA150	26524	84452.9589	60147.8096	107145.8016
KroA200	29437	116990.5227	86966.5597	155622.3004
KroB100	22141	52796.7586	39759.8377	60073.5498
KroB150	26130	82212.7339	62196.3665	104752.9063
KroB200	29368	115540.1948	84020.4643	154546.1179
KroC100	20749	52781.2349	37867.2245	59360.5143
KroD100	21294	50586.8876	39247.1640	59158.1857
KroE100	22068	54029.6642	40708.6187	60898.2571
u574	36905	299229.9737	257705.7400	450994.7786

Para aplicar las pruebas no paramétricas se plantearon las siguientes hipótesis:

H₀: No existe evidencia sobre la diferencia entre el desempeño de los algoritmos.

H₁: Existe evidencia sobre la diferencia entre el desempeño de los algoritmos.

Se realizó la prueba Ómnibus con un nivel de significancia de $\alpha=0.05$ como en [11]. Obteniendo los siguientes resultados:

F_F=20, P-valor=4.53999E-05.

Debido a que se obtuvo un **P-valor** menor al nivel de significancia establecido, se encontró evidencia estadística para aceptar la diferencia en el desempeño entre los tres algoritmos.

Con el fin de determinar la diferencia a pares del desempeño entre estos algoritmos, se realizó la prueba post-hoc con un nivel de significancia $\alpha=0.05$ como en [11]. Se seleccionó el AM como algoritmo de control obteniéndose los resultados que se muestran en la tabla 3:

Tabla 3: Prueba post-hoc

	z	P-valor	Ajustado	Hipótesis aceptada
AM vs AG	2.2360	0.0253	0.0506	H ₀
AM vs ASI	4.4721	7.74E-06	1.548E-05	H ₁

Como se observa en la tabla 3, los resultados de la prueba post-hoc permiten establecer una diferencia de desempeño entre los algoritmos AM y ASI, sin embargo no permiten establecer diferencia en el desempeño entre los algoritmos AM y AG.

CONCLUSIONES

En nuestro caso de estudio la comparación de desempeños entre los tres algoritmos metaheurísticos para resolver el TSP, presenta dificultades tales como las demostradas por el teorema "No Free Lunch" [12], que no existe un método de solución que funcione en todas las situaciones o problemas. Lo anterior justifica la búsqueda y el uso de pruebas estadísticas no paramétricas para la aplicación en la comparación del desempeño de algoritmos metaheurísticos.

La importancia de la aplicación de pruebas estadísticas no paramétricas radica en generar evidencia suficiente para realizar una selección de la Metaheurística con mejor desempeño en la búsqueda de soluciones de una o varias instancias de un mismo problema complejo dado.

AGRADECIMIENTOS

Los autores agradecen al Tecnológico Nacional de México-Instituto Tecnológico de León, que a través de la División de Estudios de Posgrado e Investigación (DEPI) dieron su apoyo para la realización del presente proyecto.

REFERENCIAS

- G. M. a. J. D, Computers and Intractability, San Francisco, EUA: W.H. Freeman, 1979.
- D. Applegate, R. Bixby, V. Chvátal y W. Cook, "On The Solution Of Traveling Salesman Problems," *Documenta Mathematica*, vol. 3, pp. 645-656, 1998.
- J. A. Montesino Guerra, H. J. Puga Soberanes, M. Ornelas Rodríguez, J. M. Carpio Valadez y U. E. Bernal Magallanes, "Análisis Comparativos de Metaheurísticas Aplicadas al Problema del TSP," *XII encuentro de Participación de la Mujer en La Ciencia*, 2015.
- F. Glover y G. A. Kochenberger, *Handbook Of Metaheuristics*, Dordrecht, Alemania: Kluwer Academic, 2003.
- E.-G. Talbi, *Metaheuristics From Design to Implementation*, Hoboken, New Jersey: John Wiley, 2009.
- A. Konar, *Artificial Intelligence and Soft Computing Behavioral and Cognitive Modeling of the Human Brain*, New York, EUA: CRC Press, 2000.
- M. Mitchell, *An Introduction To Genetic Algorithms*, London England: The MIT Press, 1998.
- P. Moscato y C. Cotta, "Una Introducción a los Algoritmos Meméticos," *Revista Iberoamericana de Inteligencia Artificial*, pp. 131-148, 2003.
- L. N. De Castro y F. J. Von Zuben, "The Clonal Selection Algorithm with Engineering Applications," *In Workshop Proceedings GECCO*, 2000.
- G. Reinelt, "TSPLIB 95," *Universität Heidelberg*, pp. 1-17, 1995
- J. Derrac, S. García, D. Molina y F. Herrera, "A practical tutorial on the use nonparametric statistical tests as a methodology for comparing evolutionary and swarm intelligence algorithms," *Elsevier-Swarm and Evolutionary computation*, vol. 1, pp. 3-18, 2011.
- D. H. Wolpert y W. G. Macready, "No Free Lunch Theorems For Optimization," *IEEE Transactions on Evolutionary Computation*, vol. 1, n° 1, pp. 67-82, 1997