

## **HEURÍSTICAS DE SELECCIÓN DE PERMUTANTES PARA BÚSQUEDA APROXIMADA EN ESPACIOS MÉTRICOS**

**ROGELIO ORLANDO MEZA TORRES  
INGENIERO CIVIL EN COMPUTACIÓN**

### **RESUMEN**

La constante evolución que han mantenido las bases de datos lleva a que en la actualidad no sólo se necesita almacenar números y palabras sino también multimedia, documentos, muestras biológicas, etc. Esto nos hace pensar que para poder buscaren estas colecciones de elementos, los métodos tradicionales no son suficientes. Para resolver este dilema, se utiliza la búsqueda por similitud o búsqueda aproximada, en donde el resultado del proceso retorna uno o más elementos similares al que se esta comparando. La búsqueda por similitud se apoya en un índice que permite descartar elementos, con lo cual el proceso de búsqueda es más rápido. Una técnica para construir índices relativamente nueva es aquella basada en permutantes, siendo el enfoque de este trabajo el como obtener un conjunto de elementos desde la colección de datos para ser usados como permutantes. Se exponen nuevas heurísticas para selección de permutantes y su posterior implementación, con el fin de analizar el comportamiento de cada una en distintos tipos de bases de datos, ya sean sintéticas o reales. Luego de realizar las pruebas en los distintos conjuntos de datos y obtener los resultados preliminares de los experimentos, se analizan de manera cruzada las heurísticas, en el mismo espacio métrico, a fin de conocer o estimar cuál es la que genera un conjunto mejor de permutantes. Esto influye en los resultados de la búsqueda, vale decir, se puede determinar con cuál se necesita recorrer una porción menor de la base de datos para obtener la misma cantidad de respuestas satisfactorias. También se expone una técnica para mejorar el conjunto de permutantes que se obtienen con las heurísticas propuestas. Del proceso de este trabajo, se concluye que las heurísticas propuestas mejoran los resultados con respecto al estado del arte. Esto indica que los resultados obtenidos en cada uno de los espacios métricos testeados son un indicio de lo que puede pasar en otros espacios métricos del mismo tipo.

Palabras claves: Búsqueda aproximada, búsqueda por similitud, permutantes.