

---

**DISEÑO, IMPLEMENTACIÓN Y PRUEBAS DE UN ÍNDICE PARA ESPACIOS  
MÉTRICOS EN MEMORIA SECUNDARIA**

**MANUEL HOFFHEIN ALFARO  
INGENIERÍA CIVIL EN COMPUTACIÓN**

**RESUMEN**

El almacenamiento en bases de datos ha estado evolucionando constantemente. Los datos tradicionales ya no son tan comunes debido al surgimiento de nuevas tecnologías que generan nuevos tipos de datos: geográficos, robóticos, imágenes, etc.

Por una parte, con estos nuevos tipos de datos ya no es posible realizar búsquedas de manera tradicional debido a que muchas veces no se pueden agrupar por alguna característica, y por otra, la gran cantidad de datos generados y su peso en megabytes hacen necesario pensar en métodos de indexación que permitan buscar de manera rápida y que no se limiten por el hardware, que muchas veces presenta problemas a la hora manejar gran cantidad de elementos y/o elementos de gran tamaño. Las consultas por similitud o búsqueda por proximidad se presentan como una solución para realizar este tipo de búsquedas, pero para ello es necesario definir un índice sobre los datos que utilice una función distancia que indique qué tan parecido es un elemento de otro. Actualmente existen variadas técnicas para indexar datos no convencionales tales como los mencionados. Sin embargo, muchos de ellos se encuentran con la limitante del hardware: se ejecuta y crean el índice sobre memoria principal (RAM) y si existe una gran cantidad de elementos o elementos de gran tamaño, al procesar los datos la agotan. Para superar esta limitante de hardware últimamente han surgido métodos de indexación que utilizan la memoria secundaria (disco duro), ya que este dispositivo generalmente ofrece mayor espacio que la memoria RAM. Dos de estos métodos crean el índice de manera dinámica, dejando sólo una pequeña fracción de la información de la estructura en memoria principal. El propósito de esta memoria es crear e implementar otro método de indexación aprovechando los métodos ya existentes y utilizar la memoria secundaria para almacenar los datos, y manteniendo en memoria RAM sólo la estructura del índice y una porción de los datos. Cabe decir que este trabajo es una alternativa a los métodos de indexación dinámica, pues permite la inserción y eliminación de elementos. De los resultados

---

obtenidos en esta memoria es posible decir que la estrategia adoptada para el índice creado es útil al momento de reducir la cantidad de lecturas y escrituras a disco duro.