

Resumen Tesis Roberto Konow

El principal objetivo de los sistemas de recuperación de información (SRI) es encontrar, lo más rápido posible, la mejor respuesta para una consulta de un usuario. La tarea de diseñar e implementar un índice que otorgue un buen compromiso en velocidad y espacio es desafiante tanto del punto de vista teórico como práctico. En esta tesis nos enfocamos en el uso, diseño e implementación de estructuras de datos compactas para crear nuevos índices que sean más rápidos y consuman menos espacio.

Nuestra nueva estructura, llamada Treaps Invertidos, requiere espacio similar a las mejores alternativas en el estado del arte, pero es un orden de magnitud más rápido en varias consultas de interés.

En un caso más general, los SRI permiten indexar y buscar en colecciones formadas por secuencias de símbolos. Navarro y Nekrich presentaron una solución que es óptima en tiempo, que requiere de espacio lineal. Sin embargo, esta solución teórica requiere más de 80 veces el tamaño de la colección. En esta tesis implementamos diseñamos e implementamos nuevas estructuras de datos compactas y las ensamblamos para construir un índice que es órdenes de magnitud más rápido que las alternativas existentes y es competitivo en términos de espacio.