
**DESARROLLO DE BENCHMARK PARA BASES DE DATOS DE GRAFOS
ORIENTADO A CONSULTAS DE CAMINOS**

**ROBERO ANTONIO GARCÍA SEGURA
INGENIERÍA CIVIL EN COMPUTACIÓN**

RESUMEN

A continuación se muestra el desarrollo de un benchmark para bases datos de grafos, el cual se enfoca en la evaluación de consultas de camino. La principal contribución de este trabajo es el estudio y definición de un método que permite una selección adecuada de datos de entrada para la generación de las instancias de consulta ejecutadas por el benchmark. Esto permite que los resultados de desempeño entregados por el benchmark sean comparables, es decir, que los tiempos de ejecución sean similares. De manera específica, primero se estudian diversas consultas de camino, esto con la finalidad de comprender el motivo de su complejidad. Luego, se definen y evalúan de manera empírica diversos métodos de selección de datos de entrada, esto con el objetivo de identificar el método más adecuado. El método seleccionado es comparado con un método tradicional basado en una selección aleatoria de datos de prueba. El método propuesto fue usado para extender un benchmark existente, lo cual implicó una modificación del generador de datos, el generador de operaciones y el evaluador de operaciones. Finalmente, el nuevo benchmark fue evaluado en dos sistemas de gestión de bases de datos de grafos.

ABSTRACT

Following the development of a benchmark for graphs database, which focuses on the evaluation of path queries is shown. The main contribution of this work is the study and definition of a method for selecting an appropriate input data for generating query instances executed by the benchmark. This allows the performance results provided by the benchmark are comparable, that is, the execution times are similar. Specifically, first studied various road query, this to understand the reason for it's complexity. Then, we define and empirically evaluate different methods to selecting input data, this in order to identify the most appropriate method. The selected method is compared with a traditional method based on a random selection of test data. The proposed method was used to extend an existing benchmark, which implied a modification of the data generator, the workload generator and the test driver. Finally, the new benchmark was evaluated in two database of graphs systems.

Palabras Clave: Database, Graph, Benchmark.