

GENERACION DE DATOS PARA BENCHMARKING DE BASES DE DATOS DE GRAFOS

**SEBASTIAN NICOLAS ARANCIBIA OLGUIN
INGENIERO CIVIL EN COMPUTACION**

RESUMEN

Dentro de las bases de datos NoSQL, una nueva generación de motores de bases de datos se puede encontrar. Dentro de este grupo nos enfocamos en las bases de datos de grafos. Esta base de datos es comúnmente utilizada en la actualidad debido a su capacidad para representar, almacenar y manipular información compleja. Existe un número creciente de iniciativas que implementan y comercializan estas bases de datos, sin embargo no existe una manera de comparar cada base de datos de grafos comercializada. Siguiendo esta dirección el documento se enfoca en proveer de las herramientas apropiadas para realizar esta evaluación. Primero, se presenta el diseño e implementación de diferentes algoritmos que generan grafos de redes sociales tanto reales como sintéticas. Posterior a esto se establece un conjunto de consultas que se ejecutaran para cada una de estas pruebas. Finalmente, una pequeña evaluación para una base de datos de grafos previamente seleccionada. La implementación de este proyecto fue realizada en Java, utilizando bases de datos relacionales como MySQL, PostgreSQL y una base de datos de grafo DEX.

Palabras claves: benchmark, redes sociales, bases de datos de grafos.

ABSTRACT

Among the NoSQL databases a new generation of databases engines can be found. Within this group we focus on a the graph databases. This database is commonly used nowadays due to its capability to represent, store and manipulate complex information. At the time there is a growing number of initiatives that implement and commercialize these databases. However there is no way to compare each graph databases marketed. Following this direction the document focuses on provide the appropriate tools to make this evaluation. Firstly it is presented the design and implementation of different algorithms that generate graphs of social networks both real and synthetic. Afterward it is established a set of queries to be executed for each of the tests. Finally a small evaluation for a graph database previously selected is detailed. The implementation of this project was done in Java, using relational databases such as MySQL, PostgreSQL and graph database DEX.

Keywords: benchmark, social networks, graph databases.