
DESARROLLO E IMPLEMENTACIÓN DE UNA API QUE EXTIENDE APACHE
GIRAPH INCORPORANDO SOPORTE DE PROPERTY GRAPHS

FRANCISCO MOYA GONZÁLEZ
INGENIERÍA CIVIL EN COMPUTACIÓN

RESUMEN

Apache Giraph es una herramienta que permite el procesamiento de grafos de gran tamaño, la que presenta como ventajas su alta escalabilidad, tolerancia a fallos y una gran eficiencia comparada con las actuales bases de datos de grafos. Sin embargo, este *framework* de programación también presenta algunas desventajas. Su compleja instalación, configuración y uso, su escasa documentación y su inusual paradigma de programación hacen que *Giraph* sea difícil de utilizar. Por otra parte, el modelo de grafo básico que soporta *Giraph* es limitante, pues no permite más de un valor tanto en aristas como nodos, restringiendo así el uso en dominios de aplicación donde los nodos y aristas contienen metadatos.

En esta memoria se desarrolló una extensión para *Giraph* la que, a través de una API, incorpora el soporte de *Property Graphs*, los que pueden tener distintos tipos de vértices y aristas, a la vez que permiten almacenar listas de propiedades en la forma de pares *key/value* asociadas con ambos tipos de elementos. Esto aumenta la capacidad de modelado del *framework*, lo que fue validado a través de un caso de uso basado en una red social.

El desempeño de la extensión desarrollada fue evaluado a través de una serie de experimentos sobre el caso de uso seleccionado. Los que entregaron resultados prometedores respecto a la aplicabilidad de la nueva API.

Adicionalmente, se provee una gran cantidad de documentación novedosa, en Castellano, para facilitar la instalación, configuración y uso de *Giraph* así como para saber manejar los problemas más comunes que se presentan durante estos procesos. Toda la documentación desarrollada queda disponible en línea para su uso a nivel global.
