

Aplicación de dos nuevos algoritmos para agrupar resultados de búsquedas en sistemas de catálogos públicos en línea (OPAC)

Andrés Marín, John W. Branch B.

Revista Interamericana de Bibliotecología, Vol. 31, n. 1, 2008

Artículo derivado de la tesis de Maestría Agrupamiento de resultados obtenidos de consultas distribuidas en sistemas de catálogos públicos en línea (OPAC). Se plantea que los usuarios de la red Internet, en particular la Web, acceden a nuevas fuentes de información; pueden hacer sus búsquedas empleando motores de búsqueda o mediante bases de datos particulares de bibliotecas o sistemas de información. Esta facilidad de acceso trae consigo la dificultad de una sobresaturación de información al usuario, debido a la abundancia de resultados. Los estudios de sistemas de catálogos de acceso público en línea (OPAC) detectan dificultades desde dos puntos de vista: 1. desde la interfaz del OPAC y la interacción hombre máquina, 2. desde el método interno de búsqueda y recuperación de la información. El trabajo quiere plantear un método de agrupamiento no supervisado, orientado hacia sistemas en línea con documentos efímeros, que sea veloz y consuma pocos recursos de memoria, con el fin de poder atender a múltiples usuarios concurrentes en un mismo sistema. Se proponen y se dan detalles de dos variantes del algoritmo K-means: Algoritmo Binary-K-Means y Algoritmo Divisive-Binary-K-Means. Con el fin de probar la aplicabilidad de los algoritmos presentados sobre sistemas OPAC, se desarrolló un prototipo de software que permite a un usuario final plantear una consulta general y obtener los resultados de su consulta en forma agrupada. Dicho prototipo es una implementación del método que se propone el Divisive-Binary-K-Means, usando el algoritmo de agrupamiento, y se desarrolló en el lenguaje de programación JAVA como un servlet, como una aplicación para la Web. Los métodos planteados disminuyen los requerimientos de memoria; y muestran la viabilidad de la incorporación de los agrupamientos al sistema OpacUdea, usado como catálogo. Se requiere en trabajos futuros, la comparación entre estos resultados y otras técnicas orientadas a datos binarios, con el fin de obtener una evaluación de calidad de agrupamiento más justa que la hecha con técnicas convencionales.

Resumen elaborado por : Alma Rosa González Rodríguez