



J Naveen Kumar *et al*, International Journal of Computer Science and Mobile Applications,

Vol.1 Issue. 5, November- 2013, pg. 38-43

ISSN: 2321-8363

EFFECTIVE EFFICIENT BOOLEAN RETRIEVAL

J Naveen Kumar¹, Dr. M. Janga Reddy²

¹ jnaveenkumar6@gmail.com, ² pricipalcmrit@gmail.com

¹ M.Tech Student, Department of Computer Science, CMR Institute of Technology, Medchal, Hyderabad

² Professor and Principal, Department of Computer Science, CMR Institute of Technology, Medchal, Hyderabad

ABSTRACT

The conventional Boolean retrieval system does not provide ranked retrieval output because it cannot compute similarity coefficients between queries and documents. Extended Boolean retrieval (EBR) models were proposed nearly three decades ago, but have had little practical impact, despite their significant advantages compared to either ranked keyword or pure Boolean retrieval. The Boolean retrieval model contrasts with ranked retrieval models such as the vector space model in which users largely use free text queries, that is, just typing one or more words rather than using a precise language with operators for building up query expressions, and the system decides which documents best satisfy the query. Despite decades of academic research on the advantages of ranked retrieval, systems implementing the Boolean retrieval model were the main or only search option provided by large commercial information providers for three decades until the early 1990s (approximately the date of arrival of the World Wide Web). However, these systems did not have just the basic Boolean operations (AND, OR, and NOT) which we have presented so far. A strict Boolean expression over terms with an unordered results set is too limited for many of the information needs that people have, and these systems implemented extended Boolean retrieval models by incorporating additional operators such as term proximity operators. A proximity operator is a way of specifying that two terms in a query must occur close to each other in a document, where closeness may be measured by limiting the allowed number of intervening words or by reference to a structural unit such as a sentence or paragraph.

Keywords- Document-at-a-time; efficiency; extended Boolean retrieval; p-norm; query processing

Full Text: www.ijcsma.com/publications/november2013/V1I502.pdf