Abstract

With the proliferation of many documents available in computers hence created a problem of how to do indexing properly. Most index that are existed behave more like a static one, meaning it can only index a fix sum of documents. But the problem arises when there are new documents that arrive, if we happen to use the indexing process that's static in nature then we would have to index the whole corpus all over again with the addition of those new documents. This method is impractical and not to mention redundant if the sum of documents already reaches the total of hundreds thousand of documents.

One of the way to solve this problem is to use incremental update algorithm towards an index, and inherently will done an indexing process towards new incoming documents only and the resulting index will be merge with the old one. This is necessary to achieve practicality in performance measuring and also to shortened the time it takes to build the index.

The end result that we have is an index that's grows exponentially with the total addition of new documents that arrives due to the merging process of those indexes.

Keyword: *information retrieval, incremental update, inverted lists*