Abstract

by Josh Herzberg

Inverted indexes are often used to allow for fast searching in textual databases. These indexes can take up a large amount of space, with 50% the size of the original data not being uncommon. Block addressing an inverted index is a method that has been used to compress indexes. We go further by presenting a novel method to compress the blocked addressing index using a bitwise notation and compressing those bits. With our method, are are able to significantly reduce the size of the index while maintaining extremely fast searching speeds.