

Abstract

A Bloom filter is a widely used data-structure for representing a set S and answering queries of the form “Is x in S ?”. By allowing some false positive answers (saying ‘yes’ when the answer is in fact ‘no’) Bloom filters use space significantly below what is required for storing S . In the *distance sensitive* setting we work with a set S of (Hamming) vectors and seek a data structure that offers a similar trade-off, but answers queries of the form “Is x *close* to an element of S ?” (in Hamming distance). Previous work on distance sensitive Bloom filters have accepted false positive *and* false negative answers. Absence of false negatives is of critical importance in many applications of Bloom filters, so it is natural to ask if this can be also achieved in the distance sensitive setting. Our main contributions are upper and lower bounds (that are tight in several cases) for space usage in the distance sensitive setting where false negatives are not allowed.