

Abstract

Filtering a data structure that is too close to a set of hashable keys may return false positives. Existing practical filters, such as the Bloom filter, require a space overhead of at least 20% because Bloom only performs a probabilistic check of assigned memberships, internal hashes, and can easily populate the entire filter causing potential minor DOS. This paper, as a further study, proves the Ribbon filter as a novel filter for static sets with various configurable space overheads and false positive rates at competitive speeds over that range. In many cases, the Ribbon is faster than existing filters for the same space overhead or can achieve under 10% space overhead with some additional CPU time. Ribbon filters resemble Xor filters modified to maximize locality and are constructed by solving linear band-like systems over Boolean variables.