

-Poner que no se pueden eliminar elementos, puesto que podrias estar eliminando otro elemento sin querer

El filtro de bloom tiene 3 parámetros principales: el tamaño del bitArray (m), el número de funciones de hash (k) y el número de palabras esperadas a insertar(n). Dados estos 3 valores podemos saber qual es la probabilidad de un falso positivo.

being 1, which would cause the algorithm to erroneously

$$\epsilon = \left(1 - \left[1 - \frac{1}{m}\right]^{kn}\right)^k \approx \left(1 - e^{-kn/m}\right)^k.$$

This is not strictly correct as it assumes independence for

Dada una m i una n, el valor de k que minimiza la probabilidad de falso (p) positivo es:

$$k = \frac{m}{n} \ln 2.$$

En nuestra super sopa sabemos la n i nosotros decidimos la probabilidad, podemos juntar las dos formulas anteriores para obtener m

$$m = -\frac{n \ln \epsilon}{(\ln 2)^2}$$