

Mercure Meetings

Tagen mit Garantie

The prob of bit being zero is $\left(1 - \frac{1}{m}\right)^{kn}$

Recall: after one hash $\frac{1}{m} \rightarrow 1 - \frac{1}{m}$ (prob 1, prob 0) $\rightarrow \left(1 - \frac{1}{m}\right)^k$ (after k hashes) $\rightarrow \left(1 - \frac{1}{m}\right)^{kn}$ (after all n elements)

The prob of bit being one is $1 - \left(1 - \frac{1}{m}\right)^{kn}$

To check whether a given element is in a Bloom filter

$$f_e^{BF}(m, k, n) = \left(1 - \left(1 - \frac{1}{m}\right)^{kn}\right)^k \approx \left(1 - e^{-kn/m}\right)^k$$

$$\frac{\partial f_e^{BF}(m, k, n)}{\partial k} = 0$$

$$f_e^{BF}(m, k, n) = \left(1 - e^{-kn/m}\right)^k = e^{k \ln\left(1 - e^{-kn/m}\right)} = e^{k \ln\left(1 - e^{-kn/m}\right)}$$

It is equivalent (and better) to work with g , where:

$$g(m, k, n) = k \ln\left(1 - e^{-kn/m}\right)$$

$$g' = \frac{\partial g}{\partial k} = 0 \Rightarrow \frac{\partial}{\partial k} \left(k \ln\left(1 - e^{-kn/m}\right)\right) = 0 \Rightarrow$$

$$\Rightarrow \ln\left(1 - e^{-kn/m}\right) + k \frac{\frac{n}{m} e^{-kn/m}}{1 - e^{-kn/m}} = 0 \Rightarrow k = \ln(2) \frac{m}{n}$$

opt k

Considering that: $e^{-kn/m} \Big|_{k=\ln(2) \frac{m}{n}} = e^{-\ln(2)} = \frac{1}{2}$

then $g' \Big|_{k=\ln(2) \frac{m}{n}} = \ln\left(1 - \frac{1}{2}\right) \ln(2) \frac{m}{n} \frac{\frac{n}{2m}}{1 - \frac{1}{2}} = \ln\left(\frac{1}{2}\right) \ln(2) \frac{\frac{1}{2}}{\frac{1}{2}} = \ln(1) = 0$

$$f_e^{BF}(m, k, n) \Big|_{k=\ln(2) \frac{m}{n}} = \left(1 - \frac{1}{2}\right)^{\frac{\ln(2)m}{n}} = \underline{\underline{0.6185 \frac{m}{n}}}$$

value of f_e^{BF} for optimal k

Detailed derivation of the formula for the optimal value for k (number of hash functions):

$$\ln\left(1 - e^{-\frac{kn}{m}}\right) + k \frac{\frac{n}{m} e^{-\frac{kn}{m}}}{1 - e^{-\frac{kn}{m}}} = 0$$

$$\text{Let } p = e^{-\frac{kn}{m}} \rightarrow \ln(p) = -\frac{kn}{m}$$

The first equation becomes:

$$\ln(1 - p) - \ln(p) \frac{p}{1-p} = 0 \rightarrow (1 - p) \times \ln(1 - p) = p \times \ln(p) \rightarrow (1 - p)^{1-p} = p^p \rightarrow 1 - p = p \rightarrow p = \frac{1}{2}$$

Therefore, the second equation gives the following: $\ln \frac{1}{2} = -\frac{kn}{m} \rightarrow \frac{kn}{m} = \ln(2) \rightarrow k = \ln(2) \frac{m}{n}$