

VerteilteWebInf Hausaufgabe 6

Gruppe 6

November 18, 2014

Aufgabe 1

False-Drop-Abschätzung beim Bloom-Filter (für Join zwischen R und S):

Abbildung auf einen Hash-Filter der Länge b (0, ..., b-1) mit k Hash-Funktionen:

- Wahrscheinlichkeit, dass kein $r \in R$ ein bestimmtes Bit setzt: $(1 - 1/b)^{k \cdot |R|} \approx e^{-k \cdot |R|/b}$
(für große b, Definition der e-Funktion)
Wahrscheinlichkeit, dass ein $r \in R$ das Bit gesetzt hat: $(1 - (1 - 1/b)^{k \cdot |R|})^k \approx (1 - e^{-k \cdot |R|/b})^k$
- Wahrscheinlichkeit, dass ein bestimmtes $s \in S$ ausgewählt wird: $(1 - (1 - 1/b)^{k \cdot |R|})^k \approx (1 - e^{-k \cdot |R|/b})^k$
= Wahrscheinlichkeit für ein false positive