Project Bloom

Carlo Baumann, Peer Jüttner
Bonusaufgabe - Bloomfilter

26. November 2018

1 JAVA-Programm

Zur Erstellen war ein JAVA-Programm, das folgendes leistet:
Bei einer gegebenen Anzahl n an zu erwartenden Elementen, die in der Datenstruktur gespeichert werden und einer Fehlerwahrscheinlichkeit p wird eine geeignete Filtergrösse m und die optimale Anzahl k an Hashfunktionen berechnet. Das JAVA-Programm kann unter folgendem Link heruntergeladen werden: `https://github.com/peerjuettner/project-bloom.git`

2 Idee Bloom Filter

2.1 Vor- und Nachteile

Ein grosser Vorteil des Bloom Filter ist die effiziente Nutzung und der kleine Speicherbedarf. Dies macht ihn äusserts platzsparend und effizient. Er muss im Gegensatz zu Sets nicht die komplette Datenstruktur abbilden, sondern nur die Hash-Werte.

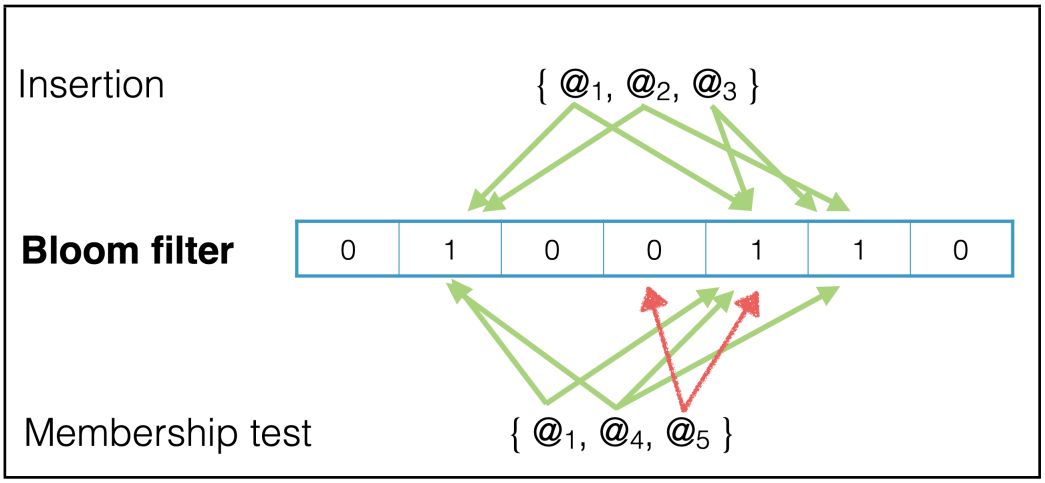
Der Bloom Filter sagt aber nicht aus, ob ein Element vorhanden ist. Er kann lediglich aussagen, ob ein Element nicht vorhanden ist. Wenn der Bloom Filter einen Treffer zurückgibt, er also meint er habe das Element gefunden. Besteht immernoch eine Wahrscheinlichkeit von 1%, dass das Element nicht vorhanden ist.

2.2 Praxisbeispiel

Der Algorithmus, für eine gute Verteilung des Hashes, findet unter anderem in den Programmiersprachen, beispielsweise Java, oder auch in den Relationalen Datenbanken Anwendung.

Ein weiterers Beispiel ist die Anwedung bei Bitcoins. Bitcoin ist eine sogenannte Kryptowährung. Bitcoin ist eine Geldeinheit aber auch ein Zahlungssystem. Das System basiert auf einer von allen Teilnehmern verwalteten Datenbank, der Blockchain. In der Blockchain werden alle Transaktionen aufgezeichnet, welche jemals durchgeführt wurden. Diese Blockchain wurde über die Jahre immer grösser. Um es nun auch mobilen Klienten zu ermöglichen am System teilzunehmen, wurde das SPV (Simple Payment Verification) Verfahren entwickelt. Dieses basiert auf dem Bloom Filter. Das SPV dient dazu nur die für den Klienten relevanten Transaktionen herauszufiltern. Mit dem Bloom Filter wird getestet, ob die Transaktion bereits in der Blockchain vorhanden ist. Dazu wird

ein Membership Test gemacht. Dieser prüft, ob der Hash der Transaktion bereits in der Blockchain vorhanden ist.



Membership Test [1]

Mit Hilfe des SPV und dem Bloom Filter, konnte die Leistung des Bitcoin Netzwerks erhöht werden, da nicht mehr die komplette Blockchain abgeglichen werden muss.

2.3 Tests

Wir haben das Programm durch ändern der Werte von p getestet. Bei einer Wahrscheinlichkeit von 1.0 hat sich gezeigt, dass wir nur False-Positive Treffer haben. Dies war so zu erwarten. Variiert man$p$ im Prozent-Bereich so sieht man, dass im Mittel der erwartetet Wert dem effektiven Wert entspricht.

Anzahl Test-Wörter	10'000	10'000	10'000
k	6	5	0
m	556987	473152	0
n	58110	58110	58110
p erwartet	0.01	0.02	1.0
p	0.0107	0.0214	1.0
False-Positiv Treffer	100	214	10000

Literatur

[1] Gervais, Arthur and Capkun, Srdjan and Karame, Ghassan O. and Gruber, Damian *On the Privacy Provisions of Bloom Filters in Lightweight Bitcoin Clients*. Proceedings of the 30th Annual Computer Security Applications Conference, New Orleans, Louisiana, USA, 2014.

[1] `https://www.ethz.ch/content/dam/ethz/special-interest/infk/inst-infsec/system-security-group-dam/research/publications/pub2014/acsac_gervais_slides.pdf`