Matthias Krause
Alexander Moch

Mannheim, den 07.12.2020

CS 307 Algorithmen und Datenstrukturen, Herbstsemester 2020

Lösungsskizze zum Übungsblatt 11

AUFGABE 11.1:

Ein HIV-Schnelltest erkennt eine Infektion in 92 % der Fälle richtig. Liegt keine Infektion vor, so liegt der HIV-Schnelltest in 99,95 % der Fälle richtig. Wir nehmen an, 0,1 % der Bevölkerung sind mit dem HI-Virus infiziert.
Wie hoch ist die Wahrscheinlichkeit, dass jemand tatsächlich mit dem HI-Virus infiziert ist, falls diese Person ein positives Testergbnis erhält?

Wir definieren uns zunächst die Ereignisse A und B wie folgt:

A *sei das Ereignis, dass die betrachtete Person mit dem HI-Virus infiziert ist.*

B *sei das Ereignis, dass die betrachtete Person ein positives Ergebnis erhält.*

Wir sind nun an der Wahrscheinlichkeit $\Pr[A|B]$ *interessiert. Mithilfe des Satzes von Bayes erhalten wir:*

$$\Pr[A|B] = \frac{0,92 \cdot 0,001}{0,999 \cdot 0,0005 + 0,001 \cdot 0,92} \approx 0,6481.$$

Somit ist eine Person mit einer Wahrscheinlichkeit von $64,81$ *% tatsächlich mit dem HI-Virus infiziert, falls diese ein positives Testergebnis erhält.*

AUFGABE 11.2:

Eine Familie $\mathcal{H}$ von Hashfunktionen $h : U \to \{1, \ldots, m\}$ heißt *2-unabhängig*, falls für zwei verschiedene Eingaben $x \neq x' \in U$ und beliebige $y, y' \in \{1, \ldots, m\}$ gilt:

$$\Pr_{h \in \mathcal{H}}[h(x) = y \text{ und } h(x') = y'] = \frac{1}{m^2}.$$

Beweise die folgenden Aussagen:

(a) Jede 2-unabhängige Hashfunktion ist universell.

$$\Pr[h(x) = h(x')] = \sum_{y=1}^{m} \Pr[h(x) = y \textit{ und } h(x') = y] = m \cdot \frac{1}{m^2} = \frac{1}{m}$$

(b) Jede 2-unabhängige Hashfunktion ist uniform.

$$\Pr[h(x) = y] = \sum_{y'=1}^{m} \Pr[h(x) = y \textit{ und } h(x') = y'] = m \cdot \frac{1}{m^2} = \frac{1}{m}$$

AUFGABE 11.3:

Sei $\mathcal{H}$ eine universelle Familie von Hashfunktionen $h : U \to \{1, \ldots, m\}$, d.h. für beliebige $x \neq x' \in U$ gilt:

$$\Pr_{h \in \mathcal{H}}[h(x) = h(x')] = \frac{1}{m}.$$

Was ist die erwartete Anzahl an Kollisionen, nachdem die Hashfunktion q-mal ausgewertet wurde?

Es seien $1 \leq i < j \leq q$. Die Zufallsvariable $C_{i,j}$ sei wie folgt definiert

$$C_{i,j} = \begin{cases} 1 \text{ falls Eingabe } i \text{ mit Eingabe } j \text{ kollidiert.} \\ 0 \text{ sonst.} \end{cases}$$

Somit können wir eine Zufallsvariable K definieren, welche die Anzahl der Kollisionen angibt:

$$K := C_{1,2} + C_{1,3} + C_{2,3} + C_{1,4} + \cdots + C_{q-2,q} + C_{q-1,q}.$$

Wir erhalten:

$$\mathbb{E}(K) = \mathbb{E}(C_{1,2}) + \mathbb{E}(C_{1,3}) + \mathbb{E}(C_{2,3}) + \mathbb{E}(C_{1,4}) + \cdots + \mathbb{E}(C_{q-2,q}) + \mathbb{E}(C_{q-1,q}).$$

Per Definition gilt $\mathbb{E}(C_{i,j}) = \frac{1}{m}$. Nach q Auswertungen existieren $\binom{q}{2}$ Eingabepaare (x_i, x_j). Somit erhalten wir für die Anzahl der erwarteten Kollisionen:

$$\mathbb{E}(K) = \frac{\binom{q}{2}}{m} = \frac{q^2 - q}{2m}.$$

AUFGABE 11.4:

In dieser Aufgabe nehmen wir an, dass ein Jahr 365 Tage hat und die Geburtstage innerhalb eines Jahres gleichverteilt sind. Das heißt, eine Person hat an einem bestimmen Tag ihren Geburtstag mit einer Wahrscheinlichkeit von $\frac{1}{365}$.

Wie viele Menschen müssen sich mindestens in einem Raum befinden, sodass die Wahrscheinlichkeit, dass mindestens zwei davon am selben Tag Geburtstag haben, größer ist als 1/2?

Wir betrachten zunächst das Komplementärereignis: Wie hoch ist die Wahrscheinlichkeit, dass keine zwei Leute am selben Tag Geburtstag haben? Wir drücken diese Wahrscheinlichkeit durch die Funktion $\rho(n)$ aus, wobei n für die Anzahl der Menschen im Raum steht.

Ist nur eine Person im Raum, so können keine Kollisionen auftreten. Es gilt $\rho(1) = 1$. Bei zwei Personen beträgt die Wahrscheinlichkeit, dass keine Kollision auftritt $\rho(2) = 1 - {}^1/_{365}$. Angenommen, es befinden sich nun bereits zwei Menschen im Raum, dann verursacht die dritte dazukommenden Person eine Kollision mit Wahrscheinlichkeit ${}^2/_{365}$. Entsprechend ist $\rho(3) = \rho(2) \cdot (1 - {}^2/_{365})$.

Für ρ erhalten wir die Formel:

$$\rho(n) = \rho(n-1) \cdot \frac{n-1}{365} = \prod_{i=1}^{n-1} 1 - \frac{i}{365}.$$

Werten wir die Funktion für $n = 22$ und $n = 23$ aus, so erhalten wir:

$$1 - \rho(22) \approx 0.4757$$
$$1 - \rho(23) \approx 0.5073$$

Also müssen 23 Leute im Raum sein, damit die Wahrscheinlichkeit einer Kollisionen größer ist als 1/2.