Algorithmen und Wahrscheinlichkeit

Woche 9

Ilya Maier

Target Shooting

Gegeben: endliche Mengen S, U s.d. $S \subseteq U$,

$$I_S : U \to \{0,1\} : I_S(u) = 1 \iff u \in S$$

Gesucht: $|S|/|U|$

Algorithmus

1) wähle $u_1, \ldots, u_N$ aus U zufällig unabhängig und gleichverteilt

2) **return** $Y := \frac{1}{N} \sum_{i=1}^{N} I_S(u_i)$ $\implies$ $\mathbb{E}[Y] = \frac{|S|}{|U|}$

Satz: (geeignetes N finden)

Für $\delta, \epsilon > 0$:

$$N \geq 3 \frac{|U|}{|S|} \epsilon^{-2} \ln\left(\frac{2}{\delta}\right) \implies \Pr\left[\left|Y - \mathbb{E}[Y]\right| \geq \epsilon \cdot \mathbb{E}[Y]\right] \leq \delta$$

(Kann man mit Chernoffs Ungleichungen i) und ii) zeigen)

Target Shooting

Satz: **(geeignetes N finden)**
Für $\delta, \epsilon > 0$:

$$N \geq 3\frac{|U|}{|S|}\epsilon^{-2}\ln\left(\frac{2}{\delta}\right) \qquad \Longrightarrow \qquad \Pr\left[\left|Y - \mathbb{E}[Y]\right| \geq \epsilon \cdot \mathbb{E}[Y]\right] \leq \delta$$

(Kann man mit Chernoffs Ungleichungen i) und ii) zeigen)

Chernoffs Ungleichungen:

i) $\Pr[X \geq (1+\epsilon)\mathbb{E}[X]] \leq e^{-\frac{1}{3}\epsilon^2\mathbb{E}[X]}$ $\quad \forall X \sim$ Bin(n, p)

ii) $\Pr[X \leq (1-\epsilon)\mathbb{E}[X]] \leq e^{-\frac{1}{2}\epsilon^2\mathbb{E}[X]}$ $\quad \forall 0 < \epsilon \leq 1$

$$e^{-\frac{1}{2}\delta^2\mathbb{E}[X]} \leq e^{-\frac{1}{3}\delta^2\mathbb{E}[X]}$$

Primzahltest

A. GGT

$\gcd(a, n) > 1$ für $1 \leq a \leq n - 1$

$\Longrightarrow n$ nicht prim

i) “nicht prim” immer richtig

ii) $\Pr[\text{"prim"} \mid \text{nicht prim}] = \frac{|Z_n^*|}{n-1}$

iii) $\text{cost}(\gcd(m, n)) = \mathcal{O}((\log nm)^3)$

B. Fermat’s little theorem

n ist prim

$\Longrightarrow \forall a \in [n-1] : a^{n-1} \equiv 1 \mod n$

i) “nicht prim” immer richtig

ii) $\Pr[\text{"prim"} \mid \text{nicht prim}] = \frac{|PB_n|}{n-1}$

iii) $PB_n := \{a \in \mathbb{Z}_n^* \mid a^{n-1} \equiv 1 \text{ mod } n\}$

iv) Carmichael-Zahl n:

1) n ist nicht prim

2) $PB_n = \mathbb{Z}_n^*$

v) $\Pr[\text{"prim"} \mid \text{nicht prim}] < 0.5,$

falls n nicht Carmichael

$\Longrightarrow$ Verbesserung durch Wiederholung

C. Miller-Rabin

Miller-Rabin-PrimeTest(n)

1) Wähle $1 \leq a \leq n - 1$ gleichverteilt zufällig

2) $d, k \in \mathbb{N}$ s.d. $n - 1 = 2^k d$, wobei d ungerade

3) **if** $a^d \mod n \neq 1$ **&&**

$\nexists i < k : a^{2^i d} \mod n = n - 1$ **then**

return “nicht prim"

4) **else return** "prim"

i) “nicht prim” immer richtig

ii) $\Pr[\text{"prim"} \mid \text{nicht prim}] \leq \frac{1}{4}$

$\Longrightarrow$ Verbesserung durch Wiederholung

iii) Laufzeit: $\mathcal{O}(\text{poly}(\log n))$

Duplikate finden

Datenmenge: $S = (s_1, \ldots, s_n)$
Duplikate: (i, j) wobei $1 \leq i < j \leq n$, falls $s_i = s_j$

A. Sortieren

1) Sortiere $((s_i, i))_{1 \leq i \leq n}$ nach s_i
2) Iteriere und finde Duplikate

i) Sortieren: $\mathcal{O}(n \log n)$
ii) Iteration: $\mathcal{O}(n + |\text{Dupl}(\mathcal{S})|)$

B. Hashing

1) Sei U alle verschiedene Elemente aus $\mathcal{S}$
2) Wähle $m \ll |U|$
3) Hashe $\mathcal{S}$ mit $h : U \to [m]$
4) Sortieren
5) Iterieren

i) Hashen: $\mathcal{O}(n)$
ii) Sortieren: $\mathcal{O}(n \log n)$
iii) Iteration: $\mathcal{O}(n + |\text{Dupl}(\mathcal{S})|)$

Kollisionen: (i, j), wobei $s_i \neq s_j \wedge h(s_i) = h(s_j)$

$$\mathbb{E}[\#\text{Kollisionen}] \leq \binom{n}{2}\frac{1}{m}$$

Speicher: $\mathcal{O}(\underbrace{n \log n}_{\text{Indices}} + \underbrace{n \log m}_{\text{Hashwerte}})$

C. Bloom Filter

1) Wähle $m, k \ll |U|$
2) Hashe $\mathcal{S}$ mit k Hashfunktionen $h_i : U \to [m]$
3) Wenn alle $h_i(s_j)$ Werte vorgekommen sind fügen wir s_j in $\mathcal{L}$ hinzu.

i) Hashen: $\mathcal{O}(kn)$
ii) Iteration: $\mathcal{O}(n + |\text{Dupl}(\mathcal{S})|)$

$$\mathbb{E}[\#\text{falscher } \mathcal{L}\text{-Eintrag}] \leq n\left(1 - \left(1 - \frac{1}{m}\right)^{k(i-1)}\right)^k$$

#falscher Eintrag $\uparrow \overset{\text{kleiner}}{\Longleftarrow} k, m \overset{\text{größer}}{\Longrightarrow}$ #Laufzeit $\uparrow$

Kahoot

Aufgaben