

# Zugriffszähler

Larry betreibt die Webseite `algorithmist.com` und will feststellen, wieviele Personen auf seine Webseite zugreifen. Vereinfachend soll hierzu die Anzahl IP-Adressen bestimmt werden, von denen aus auf die Webseite zugegriffen wurde. Da es sich um eine sehr beliebte Webseite mit hohem traffic handelt, soll die Methode, die die Anzahl unterschiedlicher IP-Adressen bestimmt, möglichst effizient sein. Dabei nimmt er in Kauf, dass womöglich IP-Adressen fälschlicherweise als schon bekannt eingestuft werden.

Für die Umsetzung des Zugriffszählers wird deshalb ein Bloomfilter eingesetzt. Nun will Larry wissen, wie gut der Bloomfilter funktioniert, d. h. wieviele IP-Adressen er fälschlicherweise nicht zählt.

Der eingesetzte Bloomfilter funktioniert wie folgt: In einer Bit-tabelle der Größe  $m$  werden implizit die IP-Adressen gespeichert. Eine IP-Adresse  $a$  wird dadurch gespeichert, dass die Bits an den Positionen  $h_0(a), \dots, h_{k-1}(a)$  auf 1 gesetzt werden. Die Werte  $h_0(a), \dots, h_{k-1}(a)$  werden dabei folgendermaßen berechnet:

```
1 x := (a mod p) mod m
2 y := (a mod q) mod m
3 for i := 0 to k - 1 do
4    $h_i(a) := x$ 
5    $y := (y + i) \bmod m$ 
6    $x := (x + y) \bmod m$ 
7 end
```

Um zu prüfen, ob eine IP-Adresse  $a$  schon gespeichert wurde, wird überprüft, ob die Bits an den Positionen  $h_0(a), \dots, h_{k-1}(a)$  alle auf 1 gesetzt sind. Falls ja, wird angenommen, dass diese IP-Adresse schon gezählt wurde. Ihre Aufgabe ist es, zu bestimmen, wie viele IP-Adressen man auf diese Weise vergisst zu zählen.

## Eingabe

Die erste Zeile der Eingabe gibt die 4 Zahlen  $m$ ,  $k$ ,  $p$  und  $q$  an, die den Bloomfilter beschreiben. Die zweite Zeile gibt die Anzahl  $n$  der Zugriffe auf die Webseite an. Die folgenden  $n$  Zeilen enthalten die IP-Adressen, von denen aus auf die Webseite zugegriffen wurde. Sie können in diesem Fall annehmen, dass eine IP-Adresse durch eine Zahl im Bereich von 0 bis  $2^{31}-1$  beschrieben wird.

## Eingabelimits

- $1 \leq k \leq 5$
- $k \leq m \leq 5 \cdot 10^6$
- $2 \leq p, q \leq 10^9$
- $1 \leq n \leq 5 \cdot 10^5$

## Ausgabe

Geben Sie die Anzahl unterschiedlicher IP-Adressen aus, die durch Anwendung des

Bloomfilters fälschlicherweise nicht gezählt werden.

## Beispiel

**Eingabe:**

10 2 11 13

6

1

2145000001

1

48

133

2145000001

**Ausgabe:**

2

---

Die IPs 1 und 48 werden gezählt, die IPs 2145000001 und 133 werden nicht gezählt, weil die entsprechenden Bits in der Tabelle schon gesetzt sind