Dr. Dirk Frettlöh 23.6.2021

Übungen zur Vorlesung Kryptographie

Blatt 11

Aufgabe 41: (Naive Hashfunktion)
Für einen deutschen Satz m definieren wir folgende Hashfunktion:

$h(m) \equiv 5 \cdot (\text{Zahl der Vokale}) + 3 \cdot (\text{Zahl der Konsonanten}) + (\text{Zahl der Leerzeichen})^2 \bmod 17$

Der Satz m="Hallo, mein Name ist Alice." hat also den Hashwert

$$h(m) = 5 \cdot 10 + 3 \cdot 11 + 4^2 \equiv 50 + 33 + 16 \equiv 14 \bmod 17.$$

Finden Sie dazu eine Kollision, ein zweites Urbild zu $h(m) = 14$ und ein Urbild zu 1. Genauer:

(a) Finden Sie zwei syntaktisch korrekte und semantisch einigermaßen sinnvolle deutsche Sätze $m' \neq m''$ mit $h(m') = h(m'') \neq 14$.
(b) Finden Sie einen syntaktisch korrekten und semantisch einigermaßen sinnvollen deutschen Satz $m^* \neq m$ mit $h(m^*) = 14$.
(c) Finden Sie einen syntaktisch korrekten und semantisch einigermaßen sinnvollen deutschen Satz $\tilde{m}$ mit $h(\tilde{m}) = 1$.

Aufgabe 42: (Einfache Hashfunktion)
Wir codieren Buchstaben als $a = 00, b = 01, c = 02 \ldots, z = 25$. Wir benutzen eine Merkle-Damgård-Konstruktion mit Startwert $s = x_0 = 71$. Die m_i sind die einzelnen Buchstaben $m_1, m_2, \ldots, m_n$ des zu hashenden Texts, gefolgt von der Länge $m_{n+1} := n$ der Nachricht; also $m = (m_1, m_2, \ldots, m_n, m_{n+1} = n)$, als zweistellige Zahlen gelesen. Die Kompressionsfunktion $x_i = f(x_{i-1}, m_i)$ $(i = 1, \ldots, n+1)$ funktioniert folgendermaßen:

(1) $y = 17 \cdot (m_i + x_{i-1}) \bmod 100$
(2) Vertausche die Ziffern von y, nenne diese neue Zahl z (Obacht: aus $y = 7 = 07$ wird 70)
(3) $x_i = x_{i-1} + z \bmod 100$

Was ist der Hashwert $h(m)$ des Wortes $m =$"padding"? Finden Sie ein (sinnfreies) Wort mit demselben Hashwert $h(m)$. Was ist der Hashwert von $m' =$"z"? Was ist der Hashwert von $m'' =$"pudding"? Was ist der Hashwert des leeren Strings ""?

Aufgabe 43: (Bessere Hashfunktion?)
Betrachten wir die Kompressionsfunktion $f : Z_{73} \times Z_{73} \to Z_{73},\ f(x, y) = 26^x \cdot 44^y \bmod 73$ und $g : Z_{73} \times Z_{73} \to Z_{73},\ g(x, y) = 27^x \cdot 46^y \bmod 73$.

(a) Realisieren Sie die zugehörigen Hashfunktionen h zu f und h' zu g nach der Merkle-Damgård-Konstruktion (wie in Aufgabe 42, mit Startwert $s = x_0 = 13$ und mit Padding — also $m_{n+1} = n =$ Länge des Texts) und berechnen Sie jeweils die Hashwerte der Worte "alice", "bob", "oskar" und "eve".
(b) Bestimmen Sie die Wertebereiche von h und h' — also die Menge aller Werte, die h und h' jeweils annehmen können.
(c) Welche der beiden Funktionen ist die eindeutig ungeeignetere Hashfunktion? Erklären Sie, woran das liegt!

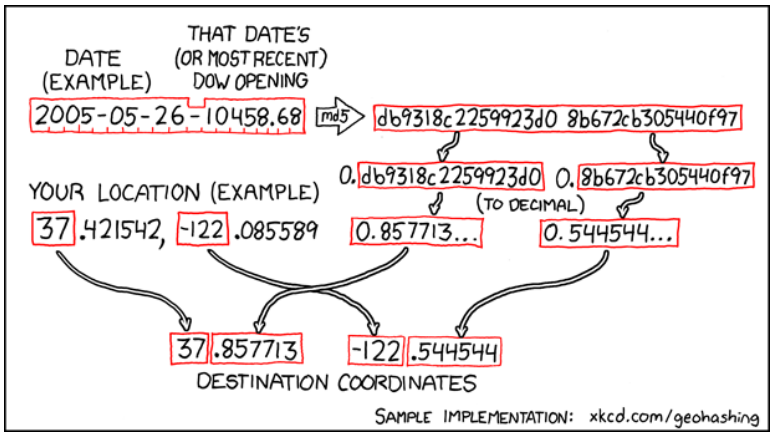
Aufgabe 44: (Wozu Padding?)
Eine naheliegende Variante der Merkle-Damgård-Konstruktion benutzt weder Startwert noch Padding. Ihre Aufgabe ist herauszufinden, warum das keine gute Idee ist. Sei der zu hashende Text $(m_0, m_1, \ldots, m_n)$. Betrachten Sie folgende Hashfunktion mit der Kompressionsfunktion f aus Aufgabe 43.

Setze $x_0 = m_0$. Für $i = 1, 2, \ldots, n$: Berechne $x_i = f(x_{i-1}, m_i)$. Ausgabe $h(m) = x_n$.

(a) Finden Sie zur Nachricht $m = (14, 18, 10, 0, 17)$ drei Kollisionen. Genauer: finden sie zu m drei weitere Urbilder m', m'', m''' der jeweiligen Länge 4, 3 und 2 mit $h(m) = h(m') = h(m'') = h(m''')$.

(b) Erläutern Sie allgemein, wie man hier zu einer Nachricht m der Länge $n + 1$ leicht eine Nachricht m' der Länge n findet mit $h(m) = h(m')$. Erläutern Sie auch, wie Padding dies verhindert, und wie das Nutzen eines Startwerts das verhindert.

(Teil (a) kann brute-force erledigt werden. Wer aber Teil (b) kann, kann Teil (a) ohne viel Aufwand erledigen.)



Anwendung: Geohashing

Abgabe bis Mittwoch 30.6.2021 bis 14 Uhr per Email an den Tutor.

Jan-Philipp Brünger jbruenger@techfak.de
Leonard Simon Ellinghaus lellinghaus+krypto@techfak.de
Kallias Stoupas kstoupas+krypto@techfak.de