Bessere Hash-Funktionen

Kollisionen

Die Einwegfunktion $c(m)=17\cdot(w_1)^2+11\cdot(w_2)^2+21\cdot(w_3)^2+4\cdot(w_4)^2$ für unsere vier-Buchstaben-langen Zeichenketten ist offensichtlich nicht geeignet für den "Münzwurf per E-Mail". Das liegt daran, dass es mehrere unterschiedliche Zeichenketten gibt, die den gleichen Code ergeben. So etwas nennt man eine **Kollision**.

Gute Hash-Funktionen haben möglichst keine Kollisionen und "streuen" ihre Werte unregelmäßig im ganzen Wertebereich.

Eine bessere Hash-Funktion

Von dem berühmten Informatiker Donald E. Knuth stammt die Idee zu einer wirklich guten Hash-Funktion. Für unser Beispiel würde man sie so bilden:

$$c_n(m) \;=\; (26\cdot(26\cdot(26\cdot w_1+w_2)+w_3)+w_4) \;\; mod\; n$$

Dabei muss allerdings der Modul **n** "gut" gewählt werden, günstig sind größere Primzahlen.

Wie können Alice und Bob mit dieser Funktion ihre Entscheidung treffen?

MD5-Hash

http://de.wikipedia.org/wiki/Message-Digest_Algorithm_5