4.1 Kryptographische Hash-Funktionen

Wie Abbildung 1 zeigt, nimmt eine Hash-Funktion eine Eingabe m variabler Länge und berechnet eine Zeichenkette $H(m)$ fester Länge, die als *Hash* bezeichnet wird.

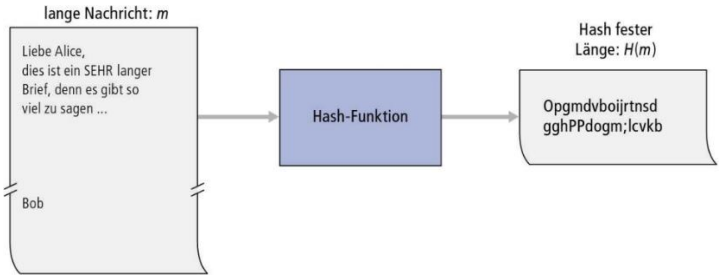

Abbildung 1: Hash-Funktion

Von einer *kryptografischen Hash-Funktion* wird erwartet, dass sie zusätzlich folgendes Merkmal aufweist:

- Es ist mit realistischem Berechnungsaufwand **nicht** möglich, zwei verschiedene Nachrichten x und y zu finden, sodass $H(x) = H(y)$.

Einfach ausgedrückt bedeutet diese Eigenschaft, dass es einem Eindringling unmöglich ist, eine durch die Hash-Funktion geschützte Nachricht durch eine andere zu ersetzen. Das heisst, wenn $(m, H(m))$ die Nachricht und der Hash-Wert der vom Absender erzeugten Nachricht sind, dann kann ein Eindringling keine andere Nachricht y herstellen, die denselben Hash-Wert aufweist wie die Originalnachricht.

Beispiel: Prüfsumme

Es sollen Prüfsummen gebildet, indem jedes Zeichen als ein Byte definiert wird und diese Bytes innerhalb von 4 Byte langen Blöcken aufsummiert werden. Die Zeichenkette `IOUl00.99BOB` in ASCII-Darstellung[1] (in hexadezimaler Schreibweise) dieser Buchstaben lautet `49, 4F, 55, 31, 30, 30, 2E, 39, 39, 42, 4F, 42`.

[1] Siehe: American Standard Code for Information Interchange – Wikipedia

Nachricht				ASCII-Darstellung				
I	O	U	1	49	4F	55	31	
0	0	.	9	30	30	2E	39	
9	B	O	B	39	42	4F	42	
				B2	C1	D2	AC	**Prüfsumme**

Nachricht				ASCII-Darstellung				
I	O	U	9	49	4F	55	39	
0	0	.	1	30	30	2E	31	
9	B	O	B	39	42	4F	42	
				B2	C1	D2	AC	**Prüfsumme**

Abbildung 2: Die Prüfsumme der ursprünglichen und geänderten Nachricht sind identisch!

Der obere Teil von Abbildung 2 zeigt, dass die 4-Byte-Prüfsumme dieser Nachricht `B2, C1, D2, AC` lautet. Eine geringfügig geänderte Nachricht wird in der unteren Hälfte von Abbildung 2 gezeigt. Die Nachrichten `IOUl00.99BOB` und `IOU900.19BOB` haben dieselbe Prüfsumme. Daher verletzt dieser einfache Prüfsummenalgorithmus die Gewährleistung von Nachrichtenintegrität: Sind die Originaldaten bekannt, dann ist es einfach, einen anderen Datensatz mit derselben Prüfsumme zu finden.

Um Sicherheit bieten zu können, brauchen man eindeutig eine leistungsfähigere Hash-Funktion als die Prüfsumme. Vertreter davon sind der *MD5-Hash-Algorithmus* (***M**essage-**D**igest #5 Algorithm*)[2] oder der *SHA* (***S**ecure **H**ash **A**lgorithm*)[3].

MD5 berechnet einen 128-Bit-Hash-Wert in einem vierstufigen Prozess. Die Eingabe erfolgt in 512 Bit Blöcken, wobei kürzere Nachrichten aufgefüllt werden. *SHA-1* erzeugt einen 160 Bit langen Hash-Wert. Seine neueren Verwandten haben längere Hash-Werte (die *SHA-2-Familie*, die z.B. *SHA-256* und *SHA-512* umfasst). Größere Ausgabelängen bedeuten bei Hash-Algorithmen eine höhere Sicherheit!

[2] Siehe: Message-Digest Algorithm 5 – Wikipedia

[3] Siehe: Secure Hash Algorithm – Wikipedia

