openHPI – Confidential Communication in the Internet

Cryptographic Hash Functions

Prof. Dr. Christoph Meinel

Hasso Plattner Institute
University of Potsdam, Germany

Cryptographic Hash Functions
Introduction

In information security, hash methods are used to generate "fingerprints" for documents, which characterize a possibly large document as unambiguously as possible by means of a short string with a fixed number of characters

Objectives:

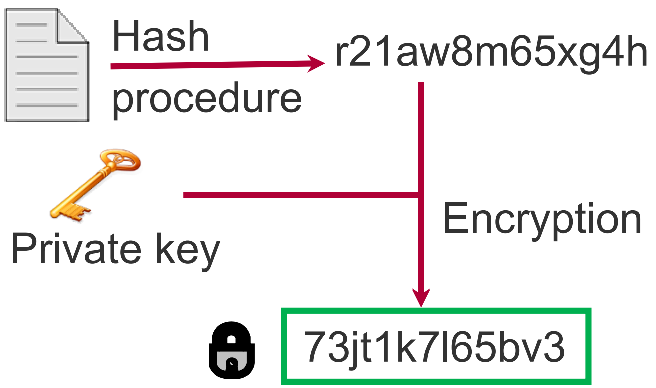
- "Compression" of a document to a fixed length string of e.g. 256 bits, which is "practically" irreversible (computation need centuries) → "**one-way hash function**"
- It should be very difficult to find a second document with exactly the same hash value → "**collision resistance**"

Cryptographic Hash Functions
Basic Idea

Hash functions transform every input into an output of fixed length, e.g. 256 bit

- Complex cryptographic procedures (e.g. digital signature) need not be applied to the (long) document, but only to its (short) hash value
- In order to ensure one-way properties, special types of hash function are required, so-called **cryptographic hash functions**

Example for applying hash-functions to sign a document:



Hash procedure

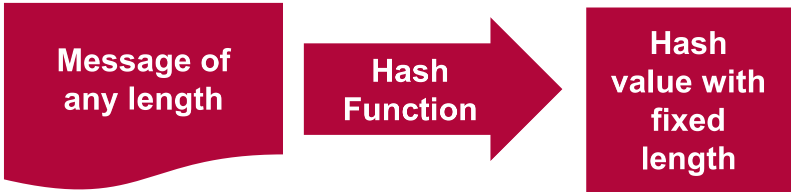
r21aw8m65xg4h

Private key

Encryption

73jt1k7l65bv3

Cryptographic Hash Functions



Some Definitions:

Hash-function h is

- **Collision free for a message M**, if it is practically impossible to construct a message M�much different from M with h(M') = h(M)
- **Collision-free**, if it is practically impossible to find two different messages M and M‵ with h(M') = h(M)
- **One-way function**, if it is practically impossible to find a message M for a given hash value z with z = h(M)

Cryptographic hash functions are collision-free hash functions

Length of Hash Values

The Length of the hash values plays an important role:

Birthday paradox:

- In a group of $k = 23$ randomly selected people, there is a probability $>1/2$ that the birthday of at least 2 people has the same date in a year ($n = 365$)

Application:

Hash value length **40-bit**:

- With probability $>1/2$ there is a collision with "only‟ $k = 2^{20}$ (about 1 million) random values - hash value is thus quite uncertain

Hash value length **256-bit** recommended length:

- With probability $>1/2$ there is a collision at $k = 2^{128}$ random values

Attacks on Cryptographic Hash Functions

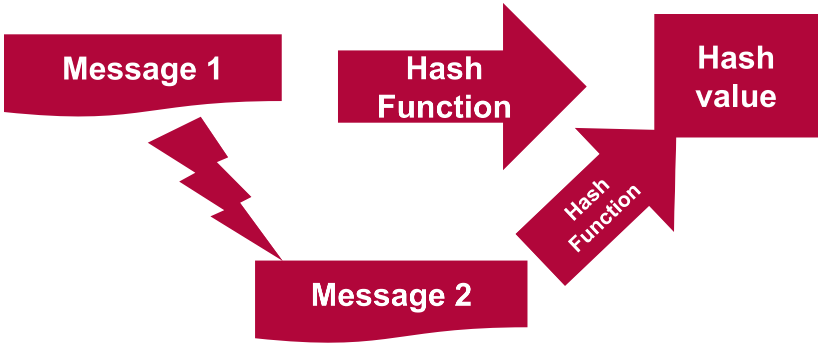
Target of the attacks:

Here no keys are searched for but rather collisions

- Finding two different messages with the same hash value
- Finding a new message with a given hash value

Attention: Birthday Attack

- When modifying messages to create messages with the same hash value, $2^{n/2}$ attempts are sufficient
- Protection: Longer hash values, e.g. 192 bit or 256 bit



Popular Hash Functions
SHA - Secure Hash Algorithm (1/2)

SHA - Secure Hash Algorithm – was developed by NSA in 1993

- SHA-1 (1995 revision of SHA) generates hash values of length 160 bit

Theoretical weaknesses of SHA-1:

- February 2005: Chinese team finds (theoretical) approach to crack SHA-1 by (2^{69} attempts)
- August 2005: Another successful attack (2^{63} attempts)
- February 2017: First published collision of SHA-1 (2 different PDF documents with same checksum)
- Current difficulty: $2^{57.5}$ attempts

Hence: Cracking an SHA-1 hash value costs less than $50.000 today when renting computing power from a cloud provider

Popular Hash Functions
SHA - Secure Hash Algorithm (2/2)

SHA-2 - (family of) successor(s) to SHA-1 from 2002

- SHA-2 family includes SHA-224, SHA-256, SHA-384, SHA-512
- SHA-256 is one of the most common cryptographic hash functions and was standardized in 2002
- SHA-256 processes 512-bit blocks (last block may need to be filled up) and generates 256-bit hash values for each

Example:

SHA-256("Franz chases in a completely neglected taxi across Bavaria") **=**

d32b568cd1b96d459e7291ebf4b25d007f275c9f13149beeb782fac0716613f8

Other Common Cryptographic Hash Functions

MD4 and MD5

- Lack of collision safety

RIPE-MD

- Different variants (128, 160, 256, 320 length of the hash)
- Collision attacks possible in original RIPE-MD

Jacuzzi

- 512-bit hash value
- AES variant for encryption

SHA-3 (Keccak)

- Standardized in October 2012
- Alternative algorithm to SHA-2