Lecture 11: Cryptographic Hashes

Exercises

The Birthday Problem

a. Suppose N integers are chosen at random, with replacement from [0, d - 1]. Compute the probability $q(N)$ that no two of the numbers are identical. Then the probability that at least two numbers are the same is $p(N) = 1 - q(N)$.
b. Use the approximation $1 - x \approx e^{-x}$ to write $p(N)$ as an exponential function.
c. Rewrite the expression from part (b) to isolate N and show that N is on the order of $d^{1/2}$ for "reasonable" values of p (e.g. $p = 1/2$).

A Simple Hash

Suppose a message is represented by a list of numbers, $M = (a_1, a_2, \ldots, a_n)$, and the hash function $H(M)$ is defined by

$$H(M) = (a_1 + a_2 + \cdots + a_n) \bmod N$$

for some positive integer N. Is the hash pre-image resistant in general? Weakly collision resistant? Strongly collision resistant? Are there conditions on the message M or the modulus N for which the hash satisfies each of the three properties?

Different Hashes

Describe in your own words how SHA-3 is different from previous hash algorithms, SHA-2 in particular. Refer to the relevant FIPS publications.

Uses of MD5

Given what you know about attacks on MD5, state whether (and why) you think it is an appropriate hash algorithm for the following applications:

a. Authenticity and integrity of public key certificates (e.g. X.509 certificates).
b. Hashing user passwords for storage in an authentication system.
c. Computation of an HMAC for use on nuclear command and control systems.

Hashes in Python

Read about the PyCrypto library and install it on a Linux VM (it can also be installed directly on a Mac; I'm not sure about Windows). The Crypto.Hash module includes implementations of several cryptographic hash functions; use the module to compute the SHA-256 hash of "correct horse battery staple". You should get the following:
`c4bbcb1fbec99d65bf59d85c8cb62ee2db963f0fe106f483d9afa73bd4e39a8a`