EN.601.441/641: Assignment 1a

September 2024

This is an individual assignment. Submit to Gradescope by September 9th at 11:59pm.

Hash functions (5 points) Suppose Mallory is launching a new 'secure' messaging app. When Alice installs the app, it creates an account for her on the server using a hash of her phone number. The app then queries the server by sending a hash of each phone number in Alice's contacts to learn which of Alice's friends are already on the platform. The goal is that users can discover their friends without the server learning the contents of every user's address book.

Assuming phone numbers are 10 digits, explain why this does not achieve the intended security goal. How can Mallory act maliciously to determine the phone numbers of every one of Alice's contacts?

Signatures (5 points) Signature schemes allow for public-key message authentication, meaning that both the integrity and provenance of a message can be checked. However, they are often computationally expensive to compute, especially over large messages.

Given a secure signature scheme and a collision-resistant hash function, explain how you could construct a new secure signature scheme wherein the signature scheme can operate over a smaller input, and informally justify the security of this new scheme. Include an overview of the Sign and Verify operations.

Merkle-Damgård (10 points) Let $h : \{0,1\}^{n+t} \to \{0,1\}^n$ be a fixed-length compression function. Suppose we forget a few of the important features of the Merkle-Damgård transformation, and construct a hash function H from h as follows:

- Let x be the input.
- Split x into pieces $y_0, x_1, x_2, \cdots, x_k$, where y_0 is n bits, and each x_i is t bits. The last piece x_k should be padded with zeroes if necessary.
- For $i = 1$ to k, set $y_i = h(y_{i-1}||x_i)$.
- Output y_k. Basically, it is similar to the Merkle-Damgård transformation, except we lose the IV and the final padding block.

This question has two parts:

1. Describe an easy way to find two messages that are broken up into the same number of pieces, which have the same hash value under h.
2. Describe an easy way to find two messages that are broken up into different number of pieces, which have the same hash value under H. Hint: Pick any string of length $n + 2t$, then find a shorter string that collides with it. Neither of your collisions above should involve finding a collision in h.

PoW difficulty (10 Points) We say that a hash function $h : P \times S \to \{0,1\}^n$ is proof of work secure with difficulty d (say, $d = 250$) if for a randomly chosen puzzle $p \in P$, it is difficult to find a solution $s \in S$ such that $H(p||s) < \frac{2n}{d}$ in time significantly less than $\frac{2n}{d}$. You can verify on your own that if we view a hash function as a Random Oracle, then it indeed satisfies the proof of work security for any suitable choice of parameters. In this question we explore the relation between collision resistance and proof of work security.

Show that a collision resistant hash function may not be proof of work secure. Specifically, let $H : P \times S \to \{0,1\}^n$ be a collision resistant hash function. Construct a new hash function $H' : P \times S \to 0, 1^{n'}$ (where n' may be greater than n) that is also collision resistant, but for a fixed difficulty d is not proof of work secure with difficulty d. This is despite H' being collision resistant. Also, explain why H' is collision resistant, that is, why a collision on H' would yield a collision on H.

Bitcoin (5 points) Explain qualitatively why the verification of a transaction in the most recently broadcast block is less reliable than one in a block a few prior to the most recent block.