Crypto Lab – Hash and Authentication

This document has been slightly modified by Mirela Damian, Villanova University. The original document can be found at `http://www.cis.syr.edu/~wedu/seed/`.

1 Overview

The learning objective of this lab is for students to get familiar with one-way hash functions and Message Authentication Code (MAC). After finishing the lab, in addition to gaining a deeper understanding of the concepts, students should be able to use tools and write programs to generate one-way hash value and MAC for a given message.

2 Lab Environment

Cryptography library `OpenSSL`**.** In this lab, we will use `openssl` commands and libraries. Make sure you have `openssl` installed in your VM. It should be noted that if you want to use `openssl` libraries in your programs, you also need to install `libssl-dev`, the development version of ssl, using the following command:

```
sudo apt-get install openssl-dev
```

Binary editor `GHex`**.** In this lab, we need to be able to view and modify files of binary format. Make sure you have `GHex` installed in your VM. GHex is a hex editor for GNOME that allows the user to load data from any file, view and edit it in either hex or ascii.

3 Lab Tasks

3.1 Task 1: Generating Message Digest and MAC

In this task, we will play with various one-way hash algorithms. You can use the following `openssl dgst` command to generate the hash value for a file. To see the manuals, you can type `man openssl` and `man dgst`.

```
openssl dgst dgsttype filename
```

Please replace the `dgsttype` with a specific one-way hash algorithm, such as `-md5`, `-sha1`, `-sha256`, etc. In this task, you should try at least 3 different algorithms, and describe your observations. You can find the supported one-way hash algorithms by typing `"man openssl"`.

3.2 Task 2: Keyed Hash and HMAC

In this task, we would like to generate a keyed hash (i.e. MAC) for a file. We can use the `-hmac` option (this option is currently undocumented, but it is supported by `openssl`). The following example generates a keyed hash for a file using the HMAC-MD5 algorithm. The string following the `-hmac` option is the key.

```
openssl dgst -md5 -hmac "abcdefg" filename
```

Please generate a keyed hash using HMAC-MD5, HMAC-SHA256, and HMAC-SHA1 for any file that you choose. Please try several keys with different length. Do we have to use a key with a fixed size in HMAC? If so, what is the key size? If not, why?

3.3 Task 3: The Randomness of One-way Hash

To understand the properties of one-way hash functions, we would like to do the following exercise for MD5 and SHA256:

1. Create a text file of any length.
2. Generate the hash value H_1 for this file using a specific hash algorithm.
3. Flip one bit of the input file. You can achieve this modification using `ghex`.
4. Generate the hash value H_2 for the modified file.
5. Please observe whether H_1 and H_2 are similar or not. Please describe your observations in the lab report. You can write a short program to count how many bits are the same between H_1 and H_2.

3.4 Task 4: One-Way Property versus Collision-Free Property

In this task, we will investigate the difference between hash function's two properties: one-way property versus collision-free property. We will use the brute-force method to see how long it takes to break each of these properties. Instead of using `openssl`'s command-line tools, you are required to write our own C programs to invoke the message digest functions in `openssl`'s crypto library.

OpenSSL provides an API called EVP (stands for , which is a high-level interface to cryptographic functions. Although OpenSSL also has direct interfaces for each individual encryption algorithm, the EVP library provides a common interface for various encryption algorithms. To ask EVP to use a specific algorithm, we simply need to pass our choice to the EVP interface. A sample code can be found from `http://www.openssl.org/docs/crypto/EVP_DigestInit.html`. Please get familiar with this sample code.

To compile your code, you may need to include the header files in `openssl`, and link to `openssl` libraries. To do that, you need to tell your compiler where those files are. In your `Makefile`, you may want to specify the following:

```
IDIR=/usr/include/
LDIR=/usr/lib/x86_64-linux-gnu

all:
	gcc -I$(IDIR) -L$(LDIR) -o file file.c -lcrypto
```

Since most of the hash functions are quite strong against the brute-force attack on those two properties, it will take us years to break them using the brute-force method. To make the task feasible, we restrict the length of the message to 4 bytes, and reduce the length of the hash value to 24 bits. We can use any one-way hash function, but we only use the first 24 bits of the hash value and ignore the rest in this task. Namely, we are using a modified one-way hash function. Please design an experiment to find out the following:

1. How many trials it will take you to break the one-way property using the brute-force method? You should repeat your experiment for multiple times, and report your average number of trials. While debugging your code, you may want to restrict the length of the hash value to 16 bits.
2. How many trials it will take you to break the collision-free property using the brute-force method? Similarly, you should report the average.
3. Based on your observation, which property is easier to break using the brute-force method?
4. (10 Bonus Points) Can you explain the difference in your observation mathematically?

4 Submission

You need to submit a detailed lab report to describe what you have done and what you have observed; you also need to provide explanation to the observations that are interesting or surprising. In your report, you need to answer all the questions listed in this lab.