COMP3311 Applied Cryptography Project (15%)

Due Date: 23:59 22 April 2024 (Hard Deadline, No Extension)

Project Title: Partial Collision Finding

SHA256, a commonly used hash function, produces 256-bit hash value. Denote by $H: \{0,1\}^* \rightarrow \{0,1\}^{256}$, the SHA256 hash function, which takes an input of arbitrary length to produce a 256-bit output.

We learnt in the lecture that a good hash function should be collision-resistant, meaning that it is difficult to find two distinct strings, say, X and Y, such that $H(X) = H(Y)$ where $X \neq Y$.

The Task

In this project, you are asked to write a program which outputs strings X and Y satisfying the following requirements:

1. X begins with your full name. For example, if your full name is *Chan Tai Man*, X should be of the form ChanTaiMan… (where … denote arbitrary string of arbitrary length).
2. Y begins with your student ID. For example, if your student ID is *22031933d*, Y should be of the form 22031933d… (where … denote arbitrary string of arbitrary length).
3. The first 44 bits of *H(X)* and *H(Y)* are the same. For example, if *H(X) = e3b0c44298fc1c149afbf4c8996fb92427ae41e4649b934ca495991b7852b855* (in hexadecimal format), then *H(Y)* should be *e3b0c44298f…*

You can use any programming language you prefer to write your program.

You may use https://emn178.github.io/online-tools/sha256.html to check the correctness of your program's output.

Remarks: A 44-bit partial collision was chosen to ensure that the task is feasible within a reasonable amount of computation time.

Submission

This is an individual project. Submit the source code of your program, along with brief documentation. This documentation should include an explanation of your approach and the algorithm used to find X and Y, configuration dependencies, a setup guide, and the values of $X, Y, H(X), H(Y)$.

Marking Criteria

1. Correctness of your solution. (5%)
2. Quality of the implementation. (7%)
3. Clarity of the documentation. (3%)

Remarks. With a very powerful computer, it should be possible to obtain the result by brute force. However, this approach will result in a low score for the quality of the implementation.