Nicholas Ng Nuo Yang A0112224B

CS4236 Cryptography Theory and Practice

Building of rainbow Table:

RAINBOW.java<space><chain-length><space><table-height>

Eg. RAINBOW.java 200 56500

Invert:

INVERT.java<space><path to query><space><path to compressed file>

Eg. INVERT.java SAMPLE_INPUT.data compressed.zip

** channel the std output by supplying ">" command to save output to file

Rainbow Table Configuration:

Chain length: 200 number of bytes used to store a digest : 10bytes

Table height: 56500

Number of tables: 4

Building the table approximately gives me ~15 million unique words.

The reduction formula used:

[(byte)(digest[((chainlength+seed)%20)] +chainlength),
(byte)(digest[(chainlength+1+seed)%20] + chainlength),
(byte)(digest[((chainlength+2)+seed)%20] + chainlength)]

Computes the word by factoring in the chain length and a seed which is (0,1,2 or 3) depending on the table number.

S, The Size of the table which can be seen in "compressed.zip" is 1,268,978 bytes

C, percentage of words correctly found by invert based on the 5000 query:

C = 4289 / 5000 = 85.78%

The speed-up factor based on successful queries:

F = A * (2^23) /t = 4289 * (2^23) /151476371 = 237.52

The speed-up factor is approximated based on overall:

F = A * (2^23) /t 5000 * (2^23) / 234328971= 178.99

t = the number of hashes called A = number of queries

Discussion:

One of the challenges faced in constructing the rainbow table is the difficulty in coming up with an effective reduction function, which is able to cover as distinct set of values as possible. Hence, by incorporating the position in the chain as one of the factors for reduction, a higher uniqueness in the

reduced value can be obtained. However, a trade-off is that when given a digest to find the original word, more sha1 hashes has to be called, since the queried digest could be in any position of the chain.

Similarly, the is also a trade-off between the chain length and the height of the rainbow table. Beyond a certain point, increasing the chain length will not yield additional unique values as collisions occur with a high chance. The increase in chain length would also affect the speedup factor, as the number of sha1 needed would also increase. If we increase the rainbow table height however, more space is needed to store the data.