Autumn Henderson
September 18th, 2021
Cryptography

Hash Attack

Introduction

The SHA-256 algorithm produces a 256-bit-sized hash of an input string. In this experiment I examine two types of hash attacks against the algorithm: Collision and Pre-image. The goal of the collision attack is to find any two hashed messages that equal each other, whereas a pre-image attack aims to take a specific hashed message and find another hashed message that matches. For a collision attack, the expected time to conduct an attack is $2^{n/2}$, where n is the number of bits in the hash digest. For a pre-image attack, this is higher at 2^n. This experiment seeks to prove these given times.

Experiment

Both attacks test bit sizes(n) 8, 10, 12, 14, 16, 18, 20, and 22 each 50 times, for a total of 400 trials for each attack and 800 trials total.

Collision – For the Collision attack, the program generates a random string of lowercase letters of size 10 and then passes it through the SHA wrapper where it is hashed and truncated to the desired bits. It is then compared with all other hashes that have went through this process and, if a match is found, the attack is complete and the number of times this comparison had to occur is recorded. If it is unsuccessful, this value is added to the list with the other hashes and a new string is generated.

Pre-image – For the Pre-image attack, the program generates a list of 50 random lowercase strings of size 6, passes it through the SHA wrapper, and then generates random strings that are hashed and compared until matches are found for all 50 stored hashes. If it is successful, the number of times the comparison occurred is recorded. If not, a new random string is generated, hashed, and compared again.

Results

Collision – Figure 1 below shows the distribution of n for each of the bit sizes tested. The mean, median, and standard deviations for each of the n-values are provided in Table 1. For all values n, means and medians were higher (around 15-20%), but the standard deviations were within ranges of the $2^{n/2}$ values.

Pre-image – Figure 2 below shows the distribution of n for each of the bit sizes tested. The mean, median, and standard deviations for each of the n-values are provided in Table 2. For n-values < 16, the mean and mean sit below the expected (-5-10%). But for n-values 16 and greater, those values all line up with the expected.

Limitations

A different value for the length of the strings was used for each attack, where collision attacks involved strings of length 10 and pre-image involved length 6. Distributions may be able to be compared better if equal string lengths were used. The reduced string length for pre-image was to run the program in a shorter amount of time. With larger variations in strings (potentially string length), actual values may have come closer to theoretical values.

Both tables Y-axis are formatted in regular decimal instead of logarithmic.

Figure 1: Box and Whisker Plot of Collision Attacks by n-values

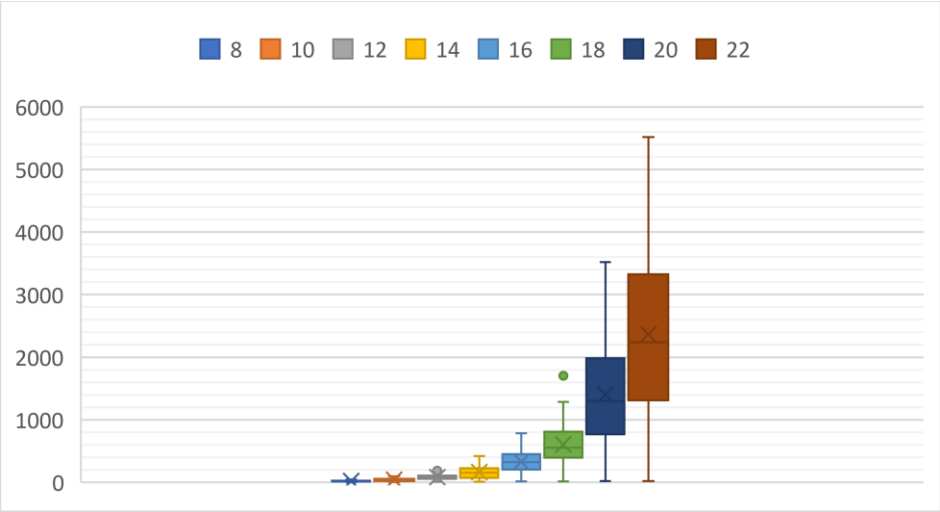

Table 1: Statistical Values for Collision Attacks

# of Bits (n)	$2^{n/2}$	Mean	Median	Stdev
8	16	19.58	18.5	9.014
10	32	40.36	37.5	22.65
12	64	82.96	78	43.98
14	128	169.54	161	96.79
16	256	331.94	330.5	193.45
18	512	614.92	558	319.10
20	1024	1433.34	1352	745.34
22	2048	2414.76	2266	1193.52

Figure 2: Box and Whisker Plot of Pre-Image Attack by n-values

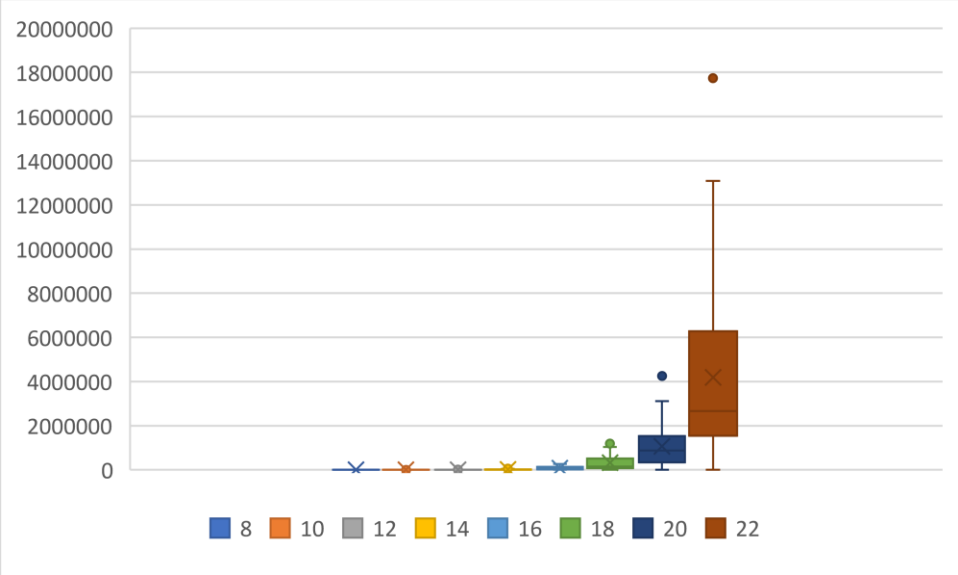

Table 2: Statistical Values for Pre-Image Attacks

# of Bits (n)	2^n	Mean	Median	Stdev
8	256	272.1	240	229.1967932
10	1024	920.96	678.5	823.7483587
12	4096	3165.04	1559	3748.156875
14	16348	15231.08	11877	14043.83117
16	65536	78673.24	46992	70230.44773

18	262144	322619.22	152251	346460.784
20	1048576	1060722.34	871198	909113.4344
22	4194304	4191038.42	2658654.5	3775319.865

Peer-Reviewed by: Rachel Offutt