Tom Hills
CS583, Dr. Ruoti
Project 2
Reviewer: Cayse Rogers

INTRODUCTION

In cryptographic hashing, there are two different ways for an individual to attack the security of a cryptographic hash function.

1) A *collision attack* is when an attacker finds two inputs for the hash function which produce the same hash value outputs (i.e., hash digests). The expected time complexity for this attack is $2^{n/2}$ where **n** is the number of bits in the hash digest.

2) A *pre-image attack* is when an attacker has a specific hash value and tries to find another input for the hash function that produces the same hash digest. The expected time complexity for this attack is 2^{n} where **n** is the number of bits in the hash digest.

METHODOLOGY

Both collision and pre-image attacks were written in python programming language. A wrapper function was created to take an arbitrary string ("string2Hash") and an integer for the number of bits desired in the outputted hash digest ("hashBits"). Internally, the wrapper function uses "sha256()" method from "hashlib" library to create initial hash output of 256 bits which then gets "right-shifted" by "256 – hashBits" bits before being returned. For inputs, a random string generator was built to output random 24 character long strings comprised of all lowercase letters and 0-9 digits. "hashBits" values were chosen to be 8, 10, 12, 14, 16, 18, 20, and 22 bits.

To test collision attacks, 8, 10, 12, 14, 16, 18, 20, and 22 bit hashes were produced with each tested for collision 50 times. A single test consisted of keeping a collection of previously produced hashes and checking if the newest hash had already been produced by a previous random string. If so, the number of hashes it took to produce the collision was recorded and the next test ran. Once 50 tests were completed for a particular hash size, the next hash size would start with it's own set of 50 tests. Pre-image attack tests were conducted in the same way except no hash collection was kept since a "pre-image" string, "a", was used to create 8, 10, 12, 14, 16, 18, 20, and 22 bit hashes which were compared to like-sized hashes produced from the random string generator.

RESULTS

An "attempt" is defined in this context as running a loop and comparing hash digests of different strings until two string are found to produce the same hash. Figure 1 and Figure 2 illustrate a comparison between average attempts produced in this experiment vs theoretical attempts discussed in the Introduction section for Collision and Pre-Image attacks, respectively. Furthermore, Table 1 and Table 2 show numerical results from their associated graph plus additional data like min, max, and median for completeness.

Tom Hills
CS583, Dr. Ruoti
Project 2
Reviewer: Cayse Rogers

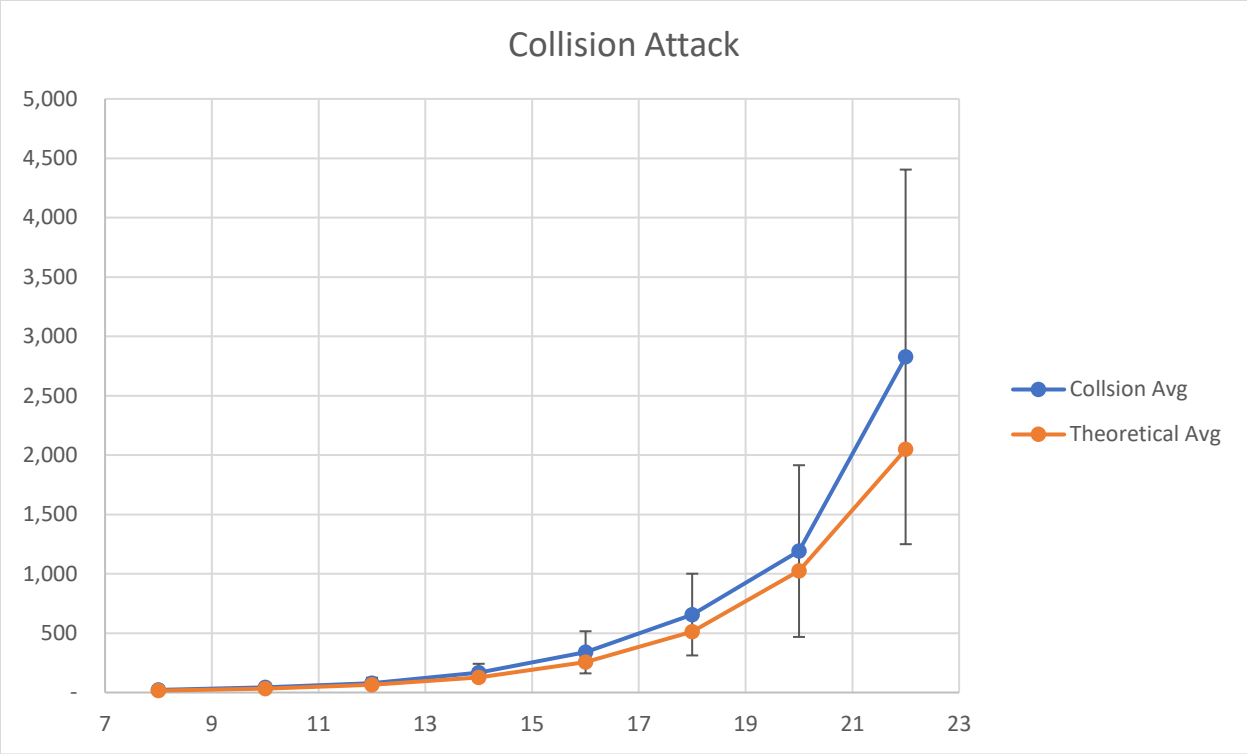

Figure 1: Collision Attack

COLLISION ATTACK								
Bit Size	**Min**	**Max**	**Median**	**Average**	**Std Deviation**	**Positive Stdev**	**Negative Stdev**	**Theoretical**
8	2	44	22.50	22.14	10.36	32.50	11.78	16
10	9	94	39.00	43.30	23.11	66.41	20.19	32
12	12	229	68.00	78.18	46.46	124.64	31.72	64
14	39	340	158.00	168.70	73.33	242.03	95.37	128
16	49	635	338.00	338.82	177.45	516.27	161.37	256
18	138	1,683	594.00	656.68	344.15	1,000.83	312.53	512
20	181	3,023	1,005.50	1,191.44	723.10	1,914.54	468.34	1,024
22	238	7,263	2,776.00	2,826.98	1,577.52	4,404.50	1,249.46	2,048

Table 1: Collision Attack Statistics

Tom Hills
CS583, Dr. Ruoti
Project 2
Reviewer: Cayse Rogers

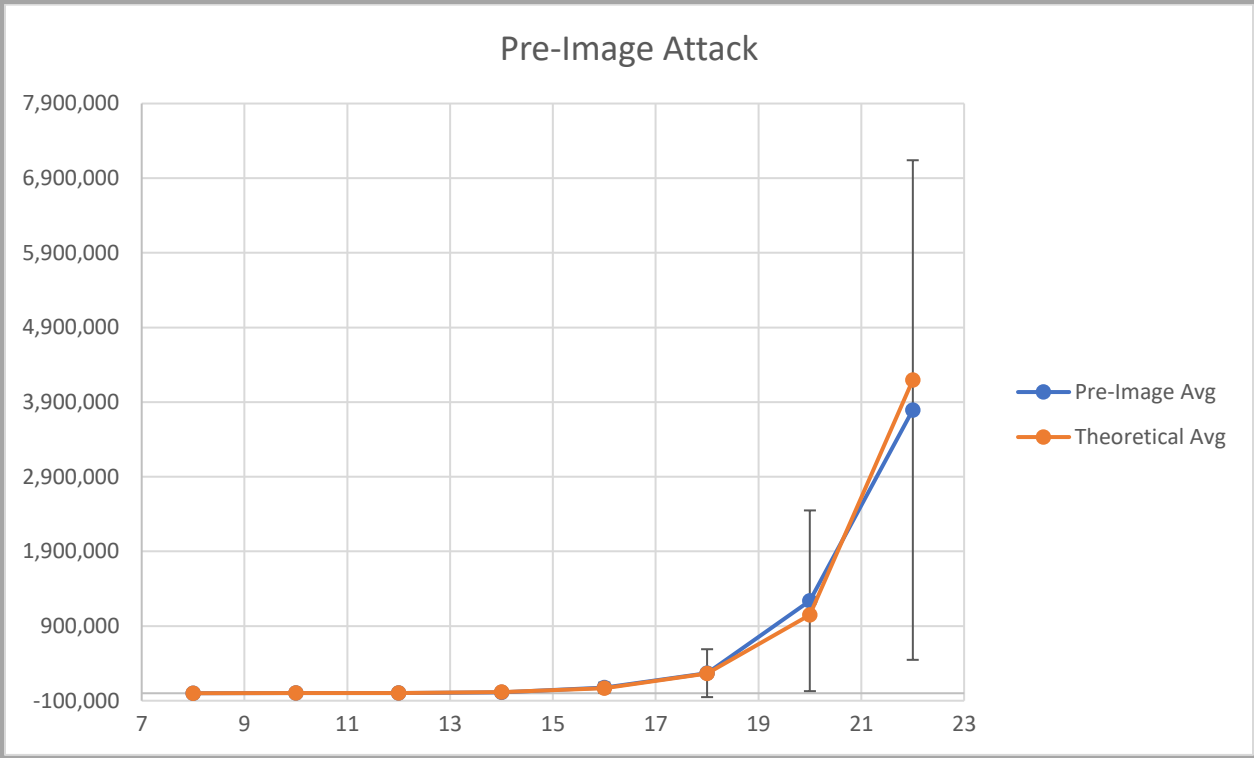

Figure 2: Pre-Image Attack

PRE-IMAGE ATTACK								
Bit Size	**Min**	**Max**	**Median**	**Average**	**Std Deviation**	**Positive Stdev**	**Negative Stdev**	**Theoretical**
8	21	1,054	147.00	207.74	187.36	395.10	20.38	256
10	43	6,346	723.50	1,135.04	1,261.66	2,396.70	-126.62	1,024
12	14	17,525	2,739.00	4,014.60	4,222.14	8,236.74	-207.54	4,096
14	63	51,087	7,800.00	12,625.80	12,169.44	24,795.24	456.36	16,384
16	2,058	233,000	42,686.50	74,571.34	68,740.43	143,311.77	5,830.91	65,536
18	705	1,470,450	153,247.50	268,984.30	320,876.36	589,860.66	-51,892.06	262,144
20	24,717	5,012,695	802,937.00	1,238,647.04	1,210,514.86	2,449,161.90	28,132.18	1,048,576
22	391,242	18,100,912	3,083,822.50	3,794,152.78	3,346,463.62	7,140,616.40	447,689.16	4,194,304

Table 2: Pre-Image Attack Statistics

Tom Hills
CS583, Dr. Ruoti
Project 2
Reviewer: Cayse Rogers

DISCUSSION

As can be seen from Figures 1 and 2 in Results, a sample size of 50 is far too small and produces huge variances for each data point which theoretical values fit firmly within with plenty of room to spare. Despite this, the trend of theoretical values still agree roughly with their corresponding experimental values as seen in Tables 1 and 2.

CONCLUSION

Based on the closeness of fit of experimental vs theoretical values despite an inadequate sample size of 50, it's safe to conclude that the methods employed by this experiment are sound and support theoretical projections.