Hash Attack Report

Ben Millett – CS 465 – Professor Decker – January 24th, 2022

I. Introduction

In cryptography, mathematical algorithms called Hashes are used to map data of any size to a bit array of a fixed size that is often referred to as a "hash" or a "message digest".[1] This one-way function ensures that it is impossible to invert or reverse the hashing affect to see what the original "message" said. The only way replicates any given hash is to attempt a brute-force search of possible inputs to find a hash that is the same, as changing one thing in the original message would change the whole hash value completely to not appear similar. As such, hashes are the basic tool of modern cryptography and should be in a cryptographer's toolkit.

While this does not hide the original message when sending it to another person, you can be confident that it was written/sent from that individual if you run the message through the hash algorithm and produce the same hash. The two ways that hackers can distill distrust in hashed data is to find a way to reverse a hacking algorithm or to find a way of replicating the hashed data with a message that produces an identical match. Since it would far more difficult to find a reverse of a hash, the brute-force previously described is a more practical point of entry for a hacker

These "Hash Attacks" allow for a message to be passed along with an identical hash, but a differing message. The two approaches addressed within this experiment are pre-image attacks, where an attacker is given a specific hash value that they are trying to replicate, and a collision attack, where an attacker is trying to find any two messages that hash to the same hash value.

II. Theoretical Cost

When considering which attack is more feasible for a hacker, one must consider the probability of successfully achieving an identical hash value from the two. A pre-image attack has the following theoretical probability:

$$2^n$$

In this case, n would be the number of bits that are trying to be replicated in the attack. In this experiment, since we are using SHA-1, it produces a hash digest of 160 bits meaning that there is a 1 in 2^{160} chance of replicating the same hash value. A collision attack has the following theoretical probability:

$$2^{n/2}$$

In the case of this experiment, to replicate the digest of a SHA-1 has, a collision attack would have a 1 in 2^{80}. This attack is often referred to as the "Birthday" attack as it is a similar principal that to find any two people in a given room with the same birthday, it would take roughly $2^{n/2}$ attempts to find a match. This is obviously much easier than the pre-image attack as the number of attempts exponentially increases and would take longer than you have to live.

With these two probabilities in mind, I attempted to gather data for smaller values of n that would support this theoretical data.

III. Empirical Data

Pre-Image:

8 bits

193	91	32	342	11	173	151	865	41	422
744	220	84	19	64	1036	63	671	125	463
36	290	182	143	195	369	250	72	307	111
208	154	348	544	410	8	78	83	34	193
123	24	757	643	324	484	276	41	312	33

Average attempts: 256.84

12 bits

1977	9462	7798	4638	2576	2252	1284	3684	1472	1846
3208	577	1080	14064	7190	27	1777	3435	242	12590
2165	2515	2124	5514	6142	688	236	5913	1874	377
4706	838	1218	14073	3368	691	6024	1899	1032	650
4433	4767	5052	606	596	11416	1553	2681	18966	1434

Average attempts: 3894.6

16 bits

11843	122796	188921	137880	37973	223481	96449	79360	149118	10108
29781	2005	80151	8499	64861	51166	62332	83027	23033	93170
4578	46982	4474	63653	105501	122931	168944	4150	63483	95326
16943	2348	58018	10190	89954	16998	55392	222709	217598	73980
11102	35865	1278	84906	71221	52981	165960	20785	99500	50308

Average attempts: 71880.24

18 bits

293901	365083	29471	344413	125064	255852	62594	40885	69554	139705
465359	172116	75201	439261	103551	472098	522802	108757	80918	308061
385775	39038	236133	199913	570463	346171	612104	76551	389930	411993
116222	179193	126826	849305	229617	104317	17002	94395	59975	249429
274558	241376	256560	472125	32788	155108	131713	32464	52551	1013177

Average attempts: 248628.36

Collision

8 bits

17	12	29	13	14	24	18	26	8	41
5	18	7	28	17	28	12	4	26	12
23	18	19	9	25	11	13	8	10	11
6	22	7	23	15	10	9	10	4	6
16	13	37	25	11	4	10	15	30	4

Average attempts: 15.66

12 bits

154	15	36	57	115	41	84	85	175	117
28	72	60	31	48	54	53	53	57	37
15	111	61	33	43	11	95	80	43	47
35	45	82	49	129	46	79	83	37	67
51	21	29	75	71	47	76	71	150	50

Average attempts: 64.08

16 bits

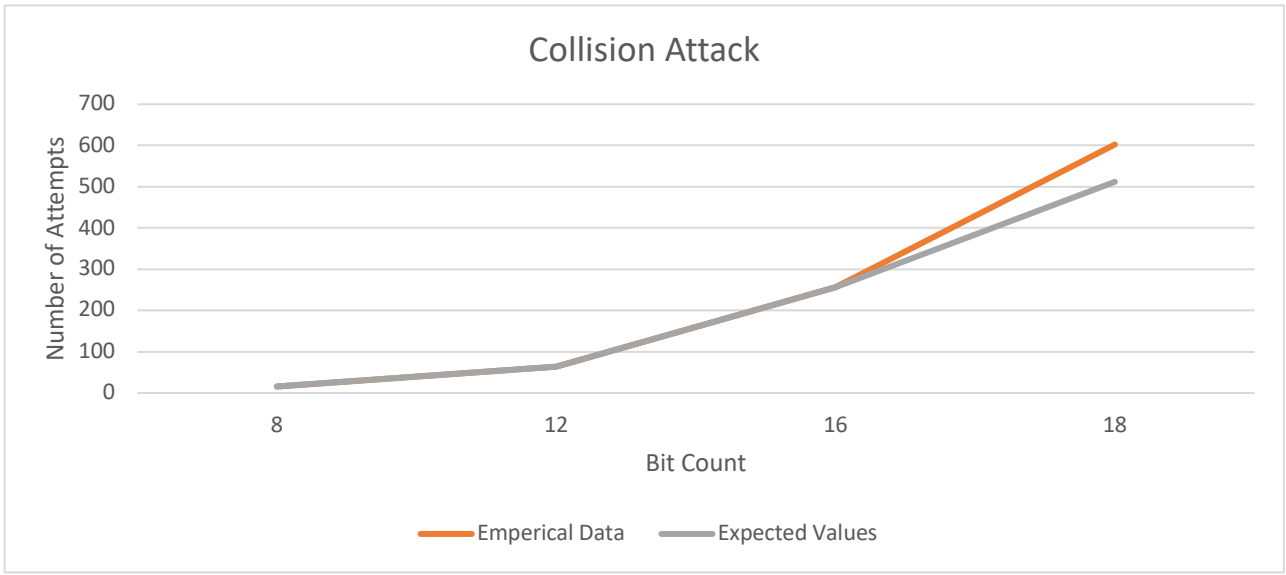
239	93	132	520	373	228	453	221	543	460
193	257	267	92	93	196	333	181	273	265
62	250	377	102	304	184	212	329	370	81
439	380	270	269	231	270	103	348	52	248
216	215	332	367	373	29	164	273	283	287

Average attempts: 256.64

18 bits

324	974	402	231	1039	438	125	522	815	745
391	720	304	573	688	1002	596	409	958	182
989	761	595	792	217	383	597	580	208	220
737	291	175	867	881	872	730	756	1242	1343
454	698	985	914	347	171	565	491	547	278

Average attempts: 602.48



III. Lab Methodology

As shown in the tables and graph above, I used SHA-1 to produce a 160-bit digest that was used in both the Pre-image and Collision attacks. For the purposes of this experiment, I used 8, 12, 16, and 18 bits to compare the results of the two strings meaning that I would only compare those *n* first number of bits to one another. Each attack I took 50 samples of how many iterations were needed to find a possible exploit and compared the average values of that were found.

IV. Conclusion

As was expected with the theoretical equations, it became apparent that the number of attempts required to successfully pull off hash attacks will increase exponentially as the number of bits you are trying to match up increases. This is good news as it guarantees that the hashes are effective at tracking who is sending a message with little reason to doubt it. SHA-1 has its vulnerabilities, but there is now SHA-2 and SHA-3 which can successfully create a hash to send with a message.

Reviewed by: Nathan Larsen

Works cited

1. https://en.m.wikipedia.org/wiki/Cryptographic_hash_function