William Endicott

CS370 Intro to Sec

25 April 2022

Assignment 1

1. I used cryptographic hashing functions because they were easily accessible with the hashlib. Since I wasn’t graded on keeping the time low, I decided using the cryptographic functions would work great. If I wanted to speed up my program, I would use non-cryptographic hash functions.
2. My output range of my bloom filters is the same as the size of my bloom filters. So, in my case I used 9388608 for both. This would ensure I didn’t produce a number outside of the size of my bloom filters.
3. My bloom filter size used for both filters is 9388608 I used the same for each one to make editing my code in VScode easier to change all occurrences for testing. I needed to ensure I have enough of a range to limit the number of false positives in my bloom filter with 5 hashing function. With hashing functions if I had too little of a bloom filter it would increase the false positive rate the bloom 5 fills up quicker since it there are 5 hashes to populate the filter with.
5. My probability for false positives is:

Bloom3

Bloom5

1. I can reduce the false positives by increasing the size of the bloom filters or starting with a lower number of items in the filter to start out. But as the bloom filters check more passwords the probability to get a false positive increase.
2. I set the bloom filters to the same size to make the process of testing and populating the filters easier but the difference in false positives when they are set is .00589 for the 3 filter and .0018 for the 5 filter. If I wanted to make the 3 filter have the same false positive rate, I could set the filter size for the 3-hash function filter to 14388608 and that would give me .0018.
3. My bloom filters don’t produce false negatives that is the purpose of a bloom filter.