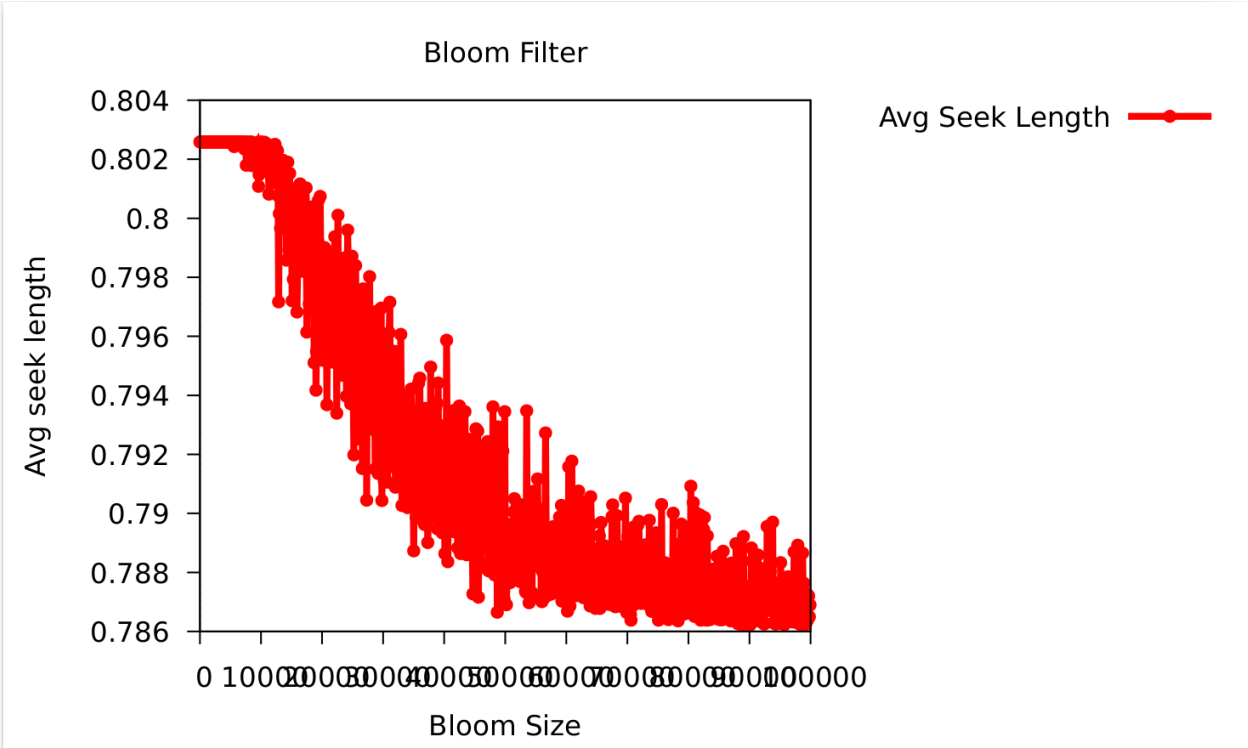
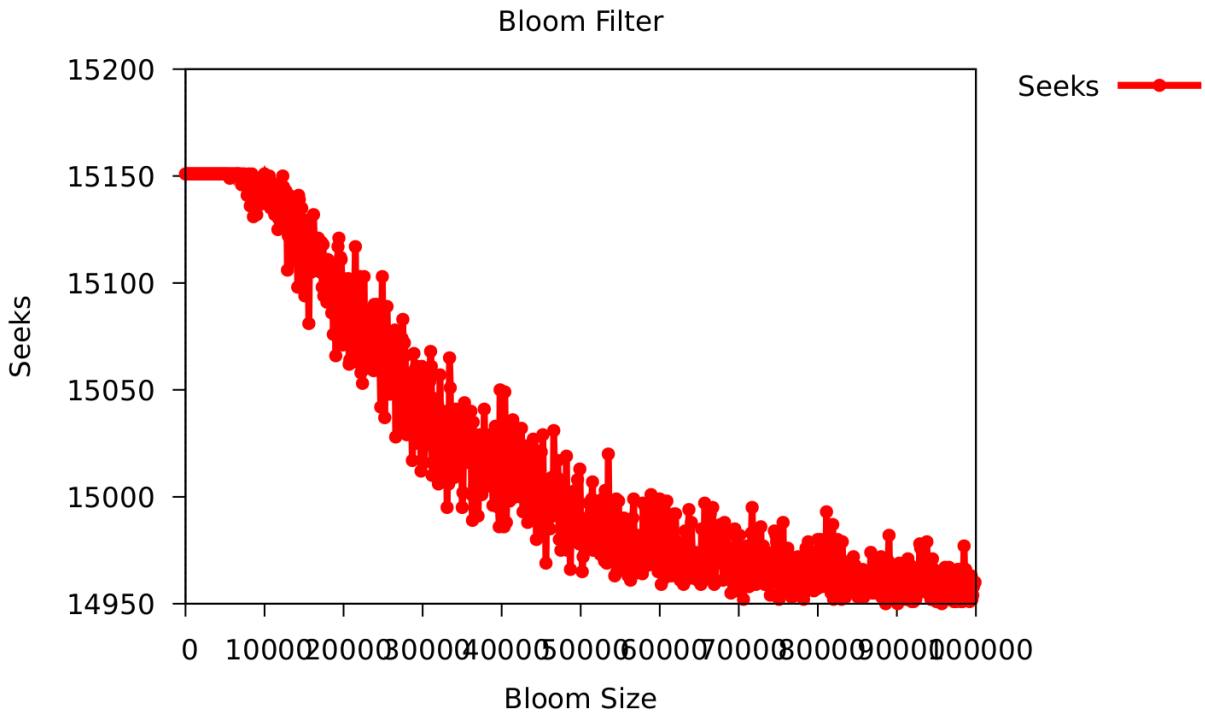
Mattiwos Belachew
mbelache@ucsc.edu
6/6/2021

CSE 13s Spring 2021
Assignment 7:
The Great Firewall of Santa Cruz:
Bloom Filters, Linked List, and Hash Table
Writeup Document

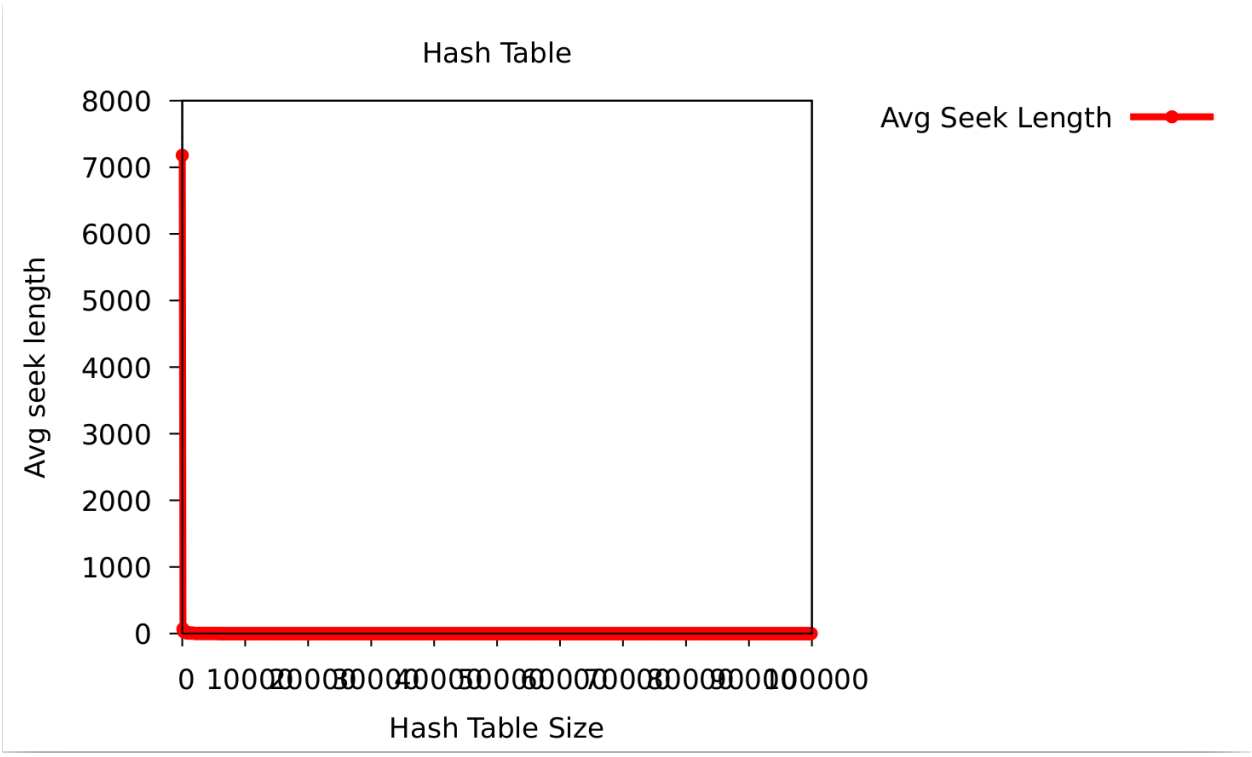
Without move-to-front rule:



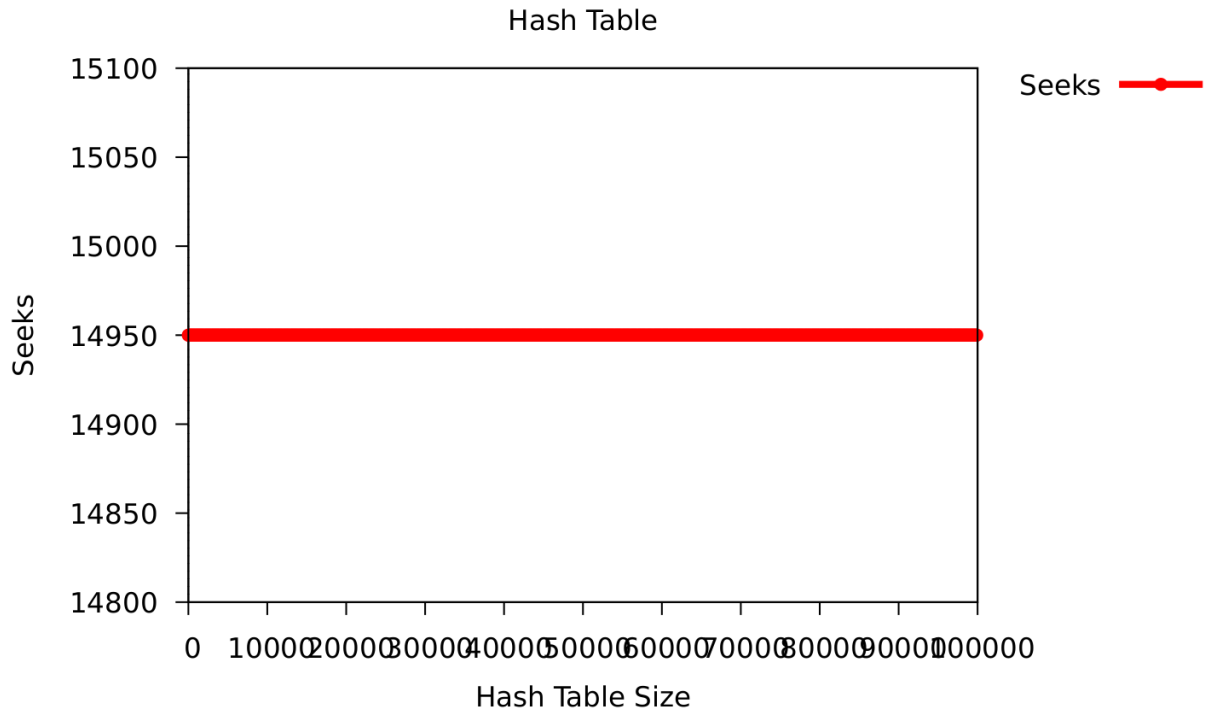
- Bloom filter Average seek length



- Bloom filter size (seeks)

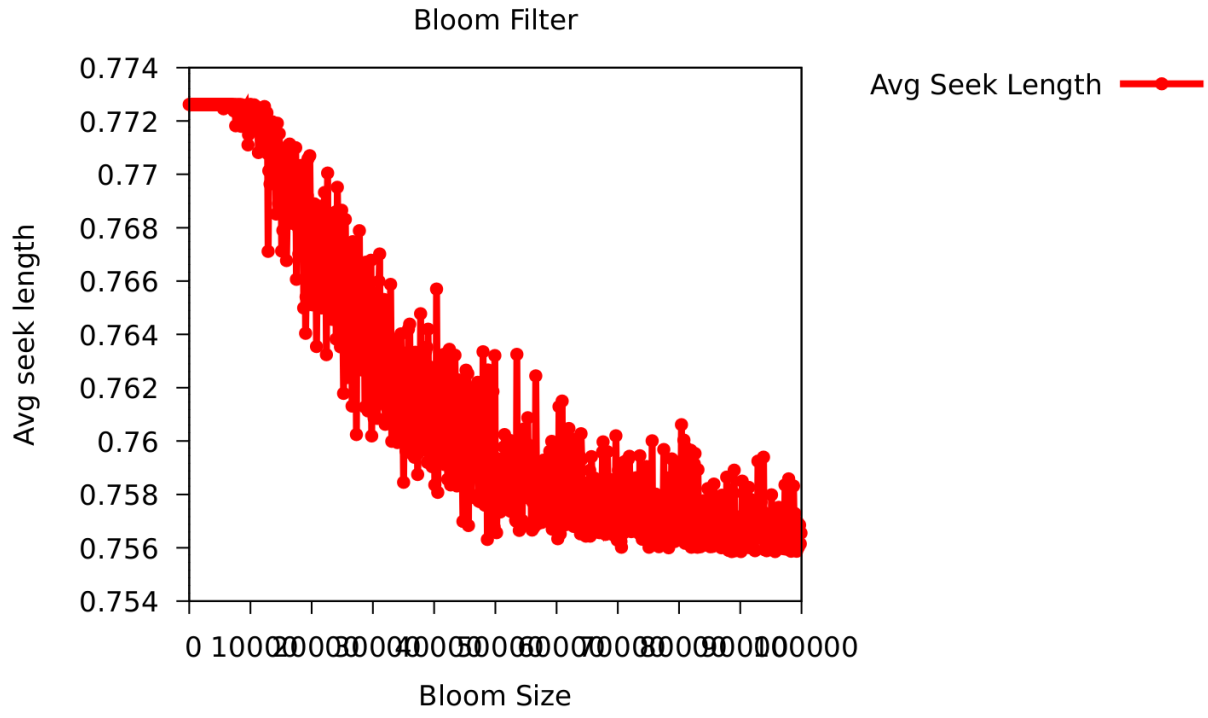


- Hash Table size (Average Seek Length)

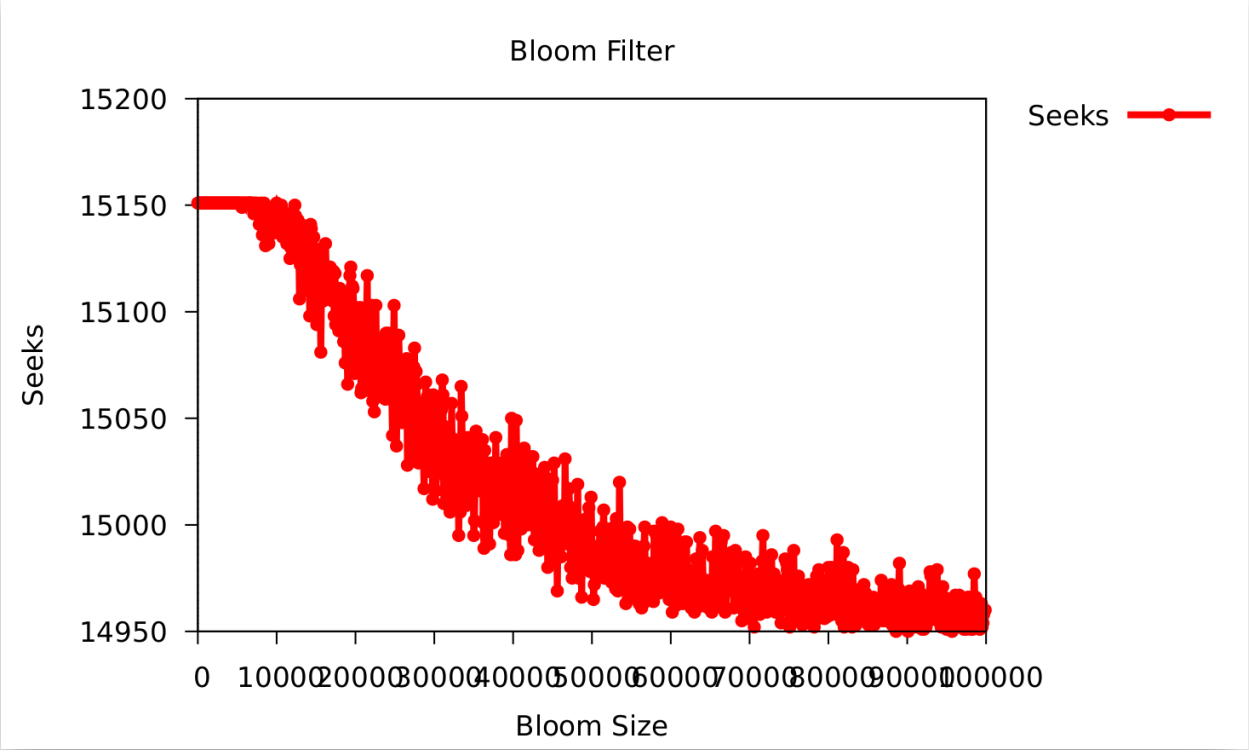


- Hash Table size (Seeks)

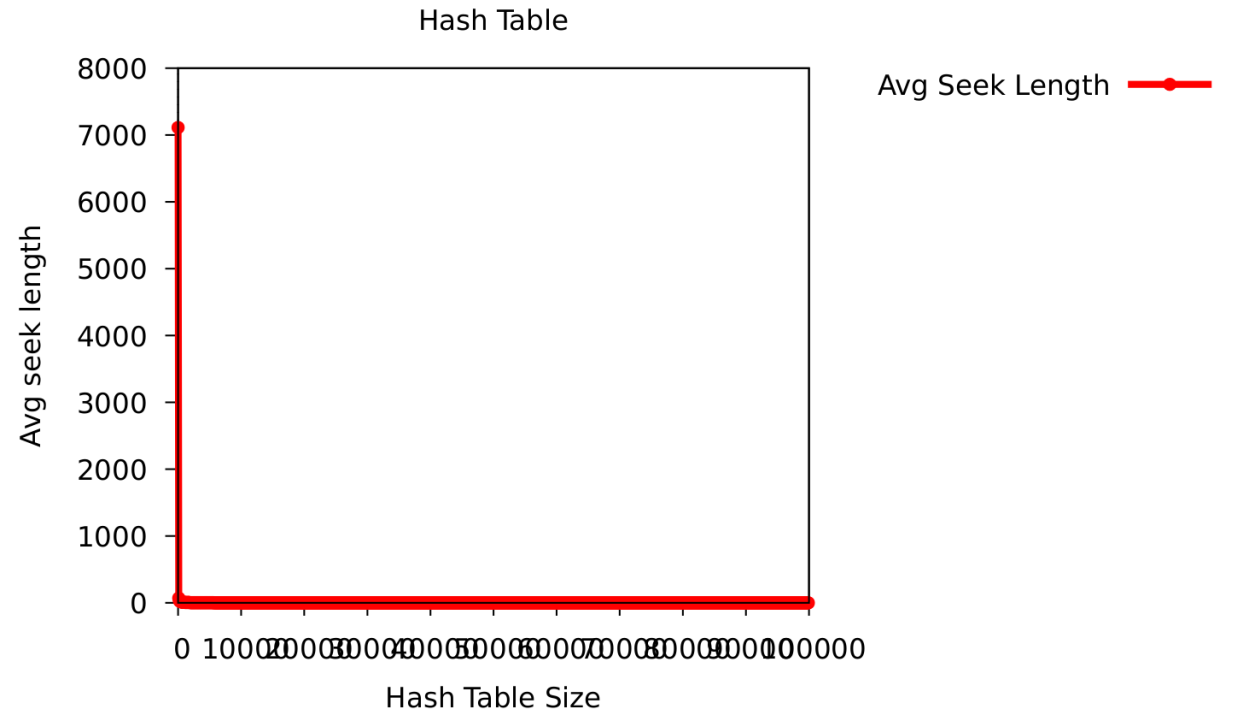
With move-to-front rule:



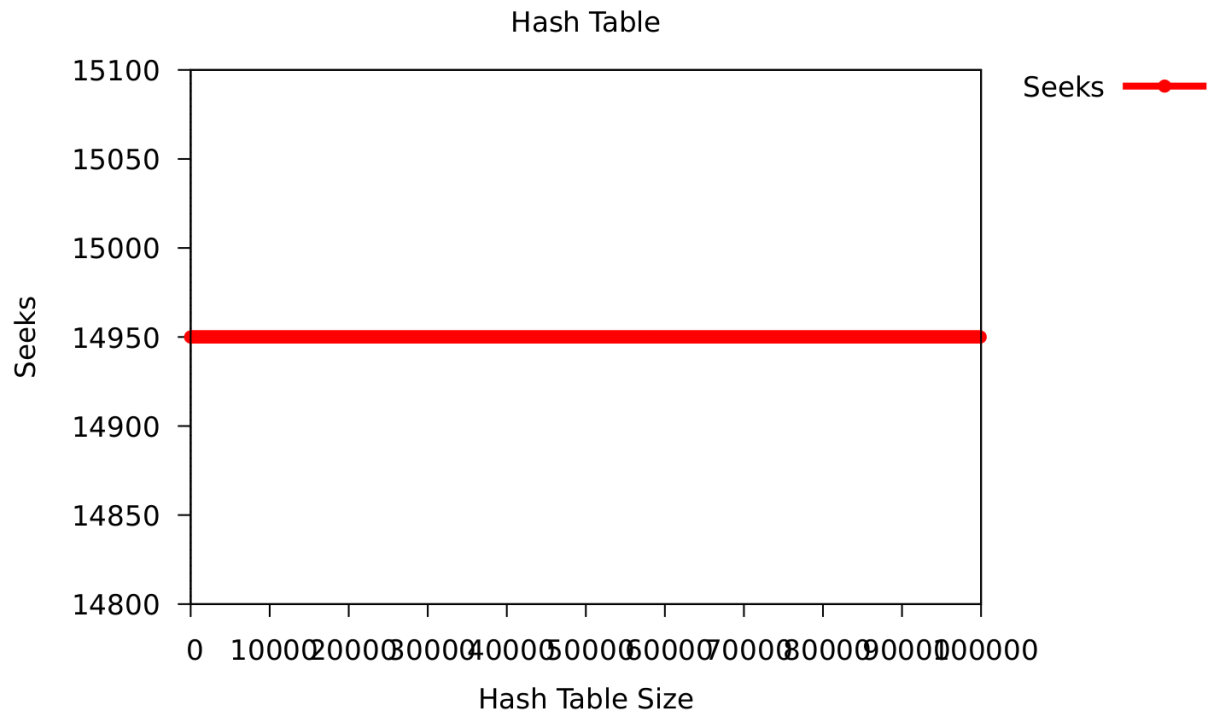
- Bloom filter Average seek length



- Bloom filter size (seeks)



- Hash Table size (Average Seek Length)



- Hash Table size (Seeks)

- Graphs comparing the total number of seeks and average seek length as you vary the hash table and Bloom filter size.
  - Do linked lists get longer?
    - As the Hash Table decreases the linked lists get longer since the hash index won't be spare thus creating a lot of concentration at a smaller section. This can be seen on the hash table without a rule-to-front(RTF) graph.
  - How does the number of links followed without using the move-to-front rule compare to the number of links followed using the rule?

    - The linked list using the rule will be shorter than the one without since the more frequent the words are the higher in the linked list. Thus, decreasing the number of links needed to be traversed in order to find the word. The linked list without the rule will be longer since the opposite

will be shorter because of contraposition. This can be seen in the graphs above.

- How does changing the Bloom filter size affect the number of lookups performed in the hash table?
    - The lower the Bloom filter size the higher the lookups since there will be alot more false positives. But as the bloom filter size increases the false positives decreases thus lowering the number of lookups needed. This can be seen on the Bloom filter without RTF above.

- Graph Analysis:
    - The Hash table size seek is always the same no matter how the hash table size is changed since the seeks is the number of times ll look up is called. This mechanism mostly has to do with bloom size so there won't be any drastic changes to it on the graph. This is shown above with the 2 graphs containing the graph with and without move-to-front rule.
    - The bloom filter for move-to-front rule seek average is lower than the one without because the number of links needed to be traversed is less so this causes the graph to shift down. This is illustrated above.
    - As the bloom filter size decreases the number of seeks decreases since it corresponds to the number of false positives. As the number of false positives decrease ll look up won't be called as much since there won't be that many hash collisions. This is because the bloom filter size increases will make the hash filter more spare, decreasing the number of collisions. This is shown above.