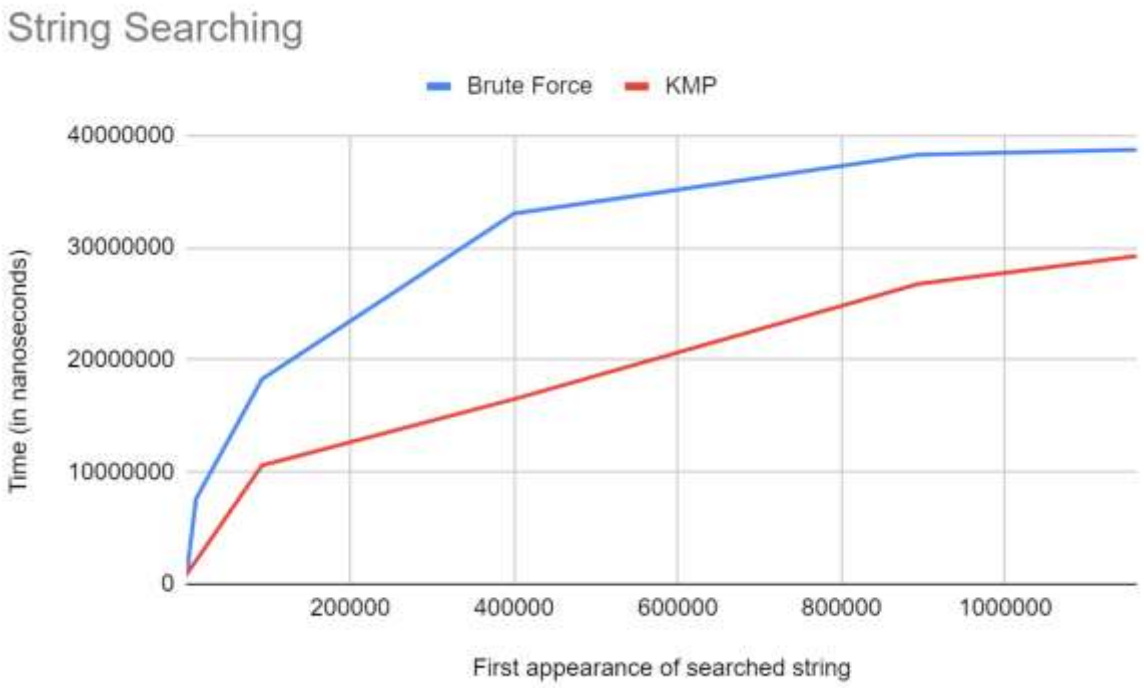
String Searching



I searched for random strings in the text of Moby Dick.

The x-axis shows how far in to the book the specified string first appears, the y-axis shows the time, in nanoseconds, it took to find that string.

We can clearly see that the Knuth-Morris-Pratt algorithm is consistently significantly faster than the brute force method.

The worst-case time complexity of the brute force method is O(n*m) where n is the length of the string being searched for, or the needle if you will. m is the size of the text being searched, or the haystack.

The time complexity of KMP is easily calculated. We first must calculate the Suffix-Prefix array which takes O(n). Then searching for the string in the text has a complexity of O(m).

This gives us a time complexity of O(n + m).