REPORT

1.

Enter File Name: randomurls.txt

URL infection List:

http://foo.com/blah\_blah

https://foo.com/blah\_blah

http://foo.com/blah\_blah/

http://foo.com/blah\_blah\_(wikipedia)

http://foo.com/blah\_blah\_(wikipedia)\_(again)

http://www.example.com/wpstyle/?p=364

https://www.example.com/foo/?bar=baz&inga=42&quux

http://142.42.1.1/

http://142.42.1.1:8080/

http://foo.com/blah\_(wikipedia)#cite-1

http://foo.com/blah\_(wikipedia)\_blah#cite-1

http://foo.com/(something)?after=parens

http://code.google.com/events/#&product=browser

http://j.mp

ftp://foo.bar/baz

ftp://foo.bar/bla

http://foo.bar/?q=Test%20URL-encoded%20stuff

http://1337.net

http://a.b-c.de

[http://223.255.255.254](http://223.255.255.254/)

\*) Number Of Urls: 20

\*) File Size: 219813

2.

a) Output for varied pattern length: (pre-process time + search time)

|  |  |  |  |
| --- | --- | --- | --- |
| PATTERN SIZE | KMP | SUFFIXARRAY | RABIN\_KARP |
| 3 | 24+ 16217 | 4734433+ 1931 | 21+ 66091 |
| 5 | 19+ 14125 | 4612928+ 25 | 22+ 64423 |
| 6 | 18+ 14928 | 4651915+ 35 | 23+ 65147 |
| 8 | 26+ 14417 | 4678135+ 43 | 26+ 63508 |

b) Output for varied textRange: (pre-process time + search time)

|  |  |  |  |
| --- | --- | --- | --- |
| TEXTRANGE | KMP | SUFFIXARRAY | RABIN\_KARP |
| 200 | 26+ 55 | 8169+ 16 | 20+ 256 |
| 1000 | 18+ 256 | 58722+ 19 | 21+ 1445 |
| 50000 | 17+ 14075 | 4588439+ 40 | 22+ 42156 |
| 100000 | 18+ 27120 | 10942793+ 45 | 21+ 114820 |

3.

Sample output for BuildCrossIndex:

yield -> [302]

----------------

The Crow and the Sheep -> [302]

++++++++++++++++++++++++++++++++++++++++++++++++

yielded -> [61]

----------------

The Old Hound -> [61]

++++++++++++++++++++++++++++++++++++++++++++++++

yielding -> [263]

----------------

The Trees Under the Protection of the Gods -> [263]

++++++++++++++++++++++++++++++++++++++++++++++++

yoke -> [216, 206]

----------------

The Heifer and the Ox -> [216, 206]

4.

Maximal Palindromes:

Enter PalindromeSize: 3

Sample Output of the list of all palindromes of size greater than or equal to 3:

'The Goatherd and the Wild Goats': {'aga': 1141, 'inni': 5398, 'ele': 7301, 'ese': 3626, 'ici': 5480, 'ses': 1556, 'usu': 33646, 'assa': 28862, 'iti': 4790, 'ala': 12857, 'iffi': 14224, 'toot': 26346, 'gnizing': 2206, 'ede': 13093, 'refer': 34392, 'rar': 22322, 'eye': 8978, 'epe': 6735, 'teet': 2254, 'ubbu': 22945, 'eve': 570, 'immi': 13140, 'ece': 2360, 'egge': 6989, 'ewe': 4188, 'tit': 25788, 'STS': 4534, 'ara': 3501, 'eitie': 8818, 'ere': 3004, 'did': 6459, 'ama': 4714, 'iri': 1544, 'bub': 22947, 'ANA': 10088, 'elle': 6068, 'rer': 7629, 'tat': 783, 'utu': 13023, 'uxu': 22501, 'divid': 3793, 'ete': 825, 'mem': 27117, 'viv': 11303, 'dered': 3357, 'izi': 2208, 'idi': 2293, 'ded': 934, 'arra': 2785, 'ssess': 1555, 'ETE': 27852, 'imi': 30919, 'pip': 6193, 'ngn': 26097, 'eme': 2791, 'opo': 9556, 'olo': 19274, 'efe': 11717, 'oso': 11273, 'ollo': 16056, 'iggi': 20342, 'ini': 15822, 'deed': 308, 'ette': 2688, 'itti': 18926, 'nan': 33273, 'COC': 4195, 'nin': 1845, 'sis': 3686, 'illi': 9185, 'isi': 1866, 'ndn': 2035, 'orro': 21108, 'non': 29744, 'nizin': 2207, 'sts': 3886, 'evive': 11302, 'ror': 15587, 'nen': 13144, 'ivi': 3794, 'issi': 27024, 'mom': 9721, 'ili': 1285, 'ommo': 16894, 'ifi': 8778, 'seles': 23306, 'atta': 7938, 'esse': 249, 'eke': 2723, 'ELE': 7324, 'awa': 1811, 'tilit': 1284, 'ene': 1014, 'oro': 23519}