

Министерство науки и высшего образования РФ
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Московский Авиационный Институт»
Национальный Исследовательский Университет

Институт №8 «Информационные технологии и прикладная математика»
Кафедра 806 «Вычислительная математика и программирование»

Лабораторная работа №4
по курсу «Криптография»

Студент:	Хренникова А. С.
Группа:	М8О-308Б-19
Преподаватель:	Борисов. А. В.
Подпись:	
Оценка:	
Дата:	

Москва, 2022

Лабораторная работа №4

Задача:

Сравнить

- 1) два осмысленных текста на естественном языке,
- 2) осмысленный текст и текст из случайных букв,
- 3) осмысленный текст и текст из случайных слов,
- 4) два текста из случайных букв,
- 5) два текста из случайных слов.

Как сравнивать: считать процент совпадения букв в сравниваемых текстах – получить дробное значение от 0 до 1 как результат деления количества совпадений на общее число букв. Расписать подробно в отчёте алгоритм сравнения и приложить сравниваемые тексты в отчёте хотя бы для одного запуска по всем пяти подпунктам. Осознать какие значения получаются в этих пяти подпунктах. Привести свои соображения о том, почему так происходит.

Длина сравниваемых текстов должна совпадать. Привести соображения о том какой длины текста должно быть достаточно для корректного сравнения.

1 Описание:

Был реализован самый «наивный» метод сравнения – посимвольный, поочередно выполнялось сравнение символов из двух текстов. Такое сравнение происходит за длину наименьшего текста.

Осмысленные тексты: ‘В цветущих садах веял майский ветерок. От веток сирени, нависавших над оградой старой кладки, доносился густой сладкий аромат. ...’, ‘Она вдруг запела своим волшебным голосом начало вагнеровского наброска к «Тристану и Изольде»: «Скажи, какие чудные мечты...» Эрнст тихонько подошел к роялю и подыграл ей. ...’.

Тексты из случайных слов: ‘стабильный издание младший опасность арабский транспортный краска людской почувствовать ...’, ‘доказывать происходить вольный чай бумага старуха ожидание любимый государственный ждать актер миновать кольцо сделка ...’.

Тексты из случайных букв: ‘здрэшабпээецпишеггрюовцаилрххнжишжффбжфбдцрфшамхээдтзнт хмояосдясуа...’, ‘цкхгжкюртсзвпбтлввуршпююдемттчяютэкоксзфдфщруочэимдргщцч акбф...’.

2 Код программы:

```
import random
import string

def compareText(text1, text2):
    score = 0
    commonL = min(len(text1), len(text2))
    for i in range(commonL):
        if text1[i] == text2[i]:
            score += 1
    return commonL, score

def mainTests():
    f1 = open('text3.txt', 'r')
    f2 = open('text2.txt', 'r')
    f3 = open('text5.txt', 'r')
    f4 = open('text6.txt', 'r')
    f5 = open('text7.txt', 'r')
    f6 = open('text8.txt', 'r')
    txt1 = f1.read()
    txt2 = f2.read()
    txt3 = f3.read()
    txt4 = f4.read()
    txt5 = f5.read()
    txt6 = f6.read()
    f1.close()
    f2.close()
    f3.close()
    f4.close()
    f5.close()
    f6.close()

    print('Два осмысленных текста на русском языке\n')
    len1, com1 = compareText(txt1, txt2)
    print('Общая длина текстов = ' + str(len1))
    print('Совпало символов = ' + str(com1))
    print('Процент = ' + str(round(com1/len1*100, 5)) + '%\n')
```

Москва, 2022

```

print('Осмысленный текст и текст из случайных букв\n')
len1, com1 = compareText(txt1, txt5)
len2, com2 = compareText(txt1, txt6)
len3, com3 = compareText(txt2, txt5)
len4, com4 = compareText(txt2, txt6)
perc1, perc2, perc3, perc4 = (com1/len1), (com2/len2), (com3/len3)
, (com4/len4)
    print('Тест 1:')
    print('Общая длина текстов = ' + str(len1))
    print('Совпало символов = ' + str(com1))
    print('Процент = ' + str(round(perc1*100, 5)) + '%\n')

    print('Тест 2:')
    print('Общая длина текстов = ' + str(len2))
    print('Совпало символов = ' + str(com2))
    print('Процент = ' + str(round(perc2*100, 5)) + '%\n')

    print('Тест 3:')
    print('Общая длина текстов = ' + str(len3))
    print('Совпало символов = ' + str(com3))
    print('Процент = ' + str(round(perc3*100, 5)) + '%\n')

    print('Тест 4:')
    print('Общая длина текстов = ' + str(len4))
    print('Совпало символов = ' + str(com4))
    print('Процент = ' + str(round(perc4*100, 5)) + '%\n')

    print('Средний процент = ' + str(round((perc1+perc2+perc3+perc4)*2
5, 5)) + '%\n')

print('Осмысленный текст и текст из случайных слов\n')
len1, com1 = compareText(txt1, txt3)
len2, com2 = compareText(txt1, txt4)
len3, com3 = compareText(txt2, txt3)
len4, com4 = compareText(txt2, txt4)
perc1, perc2, perc3, perc4 = (com1/len1), (com2/len2), (com3/len3)
, (com4/len4)
    print('Тест 1:')
    print('Общая длина текстов = ' + str(len1))
    print('Совпало символов = ' + str(com1))
    print('Процент = ' + str(round(perc1*100, 5)) + '%\n')

    print('Тест 2:')
    print('Общая длина текстов = ' + str(len2))
    print('Совпало символов = ' + str(com2))
    print('Процент = ' + str(round(perc2*100, 5)) + '%\n')

    print('Тест 3:')
    print('Общая длина текстов = ' + str(len3))
    print('Совпало символов = ' + str(com3))
    print('Процент = ' + str(round(perc3*100, 5)) + '%\n')

    print('Тест 4:')
    print('Общая длина текстов = ' + str(len4))
    print('Совпало символов = ' + str(com4))
    print('Процент = ' + str(round(perc4*100, 5)) + '%\n')

```

```

    print('Средний процент = ' + str(round((perc1+perc2+perc3+perc4)*2
5, 5)) + '%\n')

    print('Два текста из случайных букв\n')
    len1, com1 = compareText(txt5, txt6)
    print('Общая длина текстов = ' + str(len1))
    print('Совпало символов = ' + str(com1))
    print('Процент = ' + str(round(com1/len1*100, 5)) + '%\n')

    print('Два текста из случайных слов\n')
    len1, com1 = compareText(txt3, txt4)
    print('Общая длина текстов = ' + str(len1))
    print('Совпало символов = ' + str(com1))
    print('Процент = ' + str(round(com1/len1*100, 5)) + '%\n')

```

3 Результат работы программы:

Два осмысленных текста на русском языке	Осмысленный текст и текст из случайных слов
Общая длина текстов = 3270 Совпало символов = 182 Процент = 5.56575%	Тест 1: Общая длина текстов = 1957 Совпало символов = 82 Процент = 4.19009%
Осмысленный текст и текст из случайных букв	Тест 2: Общая длина текстов = 3572 Совпало символов = 182 Процент = 5.09518%
Тест 1: Общая длина текстов = 3572 Совпало символов = 89 Процент = 2.4916%	Тест 3: Общая длина текстов = 1957 Совпало символов = 109 Процент = 5.56975%
Тест 2: Общая длина текстов = 3572 Совпало символов = 108 Процент = 3.02352%	Тест 4: Общая длина текстов = 3270 Совпало символов = 148 Процент = 4.52599%
Тест 3: Общая длина текстов = 3270 Совпало символов = 76 Процент = 2.32416%	Средний процент = 4.84525%
Тест 4: Общая длина текстов = 3270 Совпало символов = 73 Процент = 2.23242%	Два текста из случайных букв
Средний процент = 2.51792%	Общая длина текстов = 450150 Совпало символов = 16007 Процент = 3.55593%
	Два текста из случайных слов
	Общая длина текстов = 1957 Совпало символов = 108 Процент = 5.51865%

4 Выводы:

Как видно по результатам работы программы, наилучшие результаты достигаются путем сравнения двух осмысленных текстов. Такой результат достигается, потому что в текстах есть наиболее используемые буквы. Например, нет слов без гласных букв, всего гласных 10, а чаще всего используются а, о, и, е, они и совпадают довольно часто. Поэтому и при

Москва, 2022

сравнении осмысленного текста и текста из случайных слов, как и двух текстов из случайных слов получается нормальный результат. Плохой процент показывает сравнение с текстами из случайных букв, потому что там вероятность встречи каждой буквы примерно одинаковая, нет тех, что используются часто.

Для более высокой точности лучше не брать короткие тексты, средняя длина(около 5000 символов) будет оптимальной.