

Отчёт по лабораторной работе №3 по курсу «Криптография»

Выполнил *Попов Николай*, группа *М8О-308Б-21*

Задание

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.

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 случаям. Осознать какие значения получаются в этих пяти случаях. Привести соображения о том почему так происходит.

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

Код для генерации текста из случайных символов:

```
def get_random_chars_text(size):
    letters = string.ascii_lowercase + ' '
    random_chars_text = np.random.choice(list(letters), size, replace=True)
    random_chars_text = ''.join(random_chars_text)
    assert len(random_chars_text) == size
    return random_chars_text
```

Код для генерации текста из случайных символов:

```
def get_random_words_text(size):
    cur_len = 0
    words_for_text = []
    while cur_len <= size:
        random_word = random.choice(english_words)
        cur_len += len(random_word) + 1
        words_for_text.append(random_word)

    result = ' '.join(words_for_text)[:size]
    assert len(result) == size
    return result
```

Код для сравнения текстов

```
def compare(text1, text2):  
    assert len(text1) == len(text2)  
    matched = 0  
    for c1, c2 in zip(text1, text2):  
        if c1 == c2:  
            matched += 1  
    return matched / len(text1) * 100
```

Результаты сравнения

Два осмысленных текста на естественном языке:

`compare(text1, text2) = 6.3955%`

*Среднее из 5 сравнений различной длины
«Детство» Л.Н. Толстой, «Белая гвардия» М.А. Булгаков*

Осмысленный текст и текст из случайных букв:

`compare(text1, text2) = 0.633%`

Среднее из 5 текстов различной длин

Осмысленный текст и текст из случайных слов:

`compare(text1, text2) = 1.755%`

Среднее из 5 текстов различной длин

Два текста из случайных слов:

`compare(text1, text2) = 5.769%`

Среднее из 5 текстов различной длин

Выводы:

Полученные результаты сравнений показывают, что использование слов, даже случайных, обеспечивает более высокую точность совпадений по сравнению с использованием случайных букв. В словах буквы связаны друг с другом: встречаются частые буквосочетания и паттерны. В тексте из случайных букв связи между символами

отсутствуют, и могут возникать абсурдные сочетания букв и слов, которые редко встречаются в реальных текстах.

Тем не менее, точность при сравнении текстов из случайных слов и реальных текстов также различается. Тексты из случайных слов не учитывают взаимосвязи между словами, в то время как реальные тексты эти связи учитывают.

С помощью таких оценок можно примерно определить характер полученного текста (например, если он зашифрован, на незнакомом языке или слишком большой для полного изучения). Можно вычислить точности совпадений с реальными текстами и проанализировать полученное значение. Этот показатель позволяет понять, содержится ли в тексте осмысленная информация или это просто набор букв. Если точность близка к нулю, возможно, текст на другом языке. Такие проверки применяются в некоторых криптографических алгоритмах.