Міністерство освіти і науки України
Національний технічний університет України
"Київський політехнічний інститут імені Ігоря Сікорського"
Фізико-технічний інститут

КРИПТОГРАФІЯ

Комп’ютерний практикум №2
Варіант 7
Криптоаналіз шифру Віженера

Роботу виконав:
Студент 3 курсу
Групи ФБ-06
Кононець В. М.

Київ – 2022

Мета роботи

Засвоєння методів частотного криптоаналізу. Здобуття навичок роботи та аналізу потокових шифрів гамування адитивного типу на прикладі шифру Віженера.

Постановка задачі

1. Самостійно підібрати текст для шифрування (2-3 кб) та ключі довжини r = 2, 3, 4, 5, а також довжини 10-20 знаків. Зашифрувати обраний відкритий текст шифром Віженера з цими ключами.

2. Підрахувати індекси відповідності для відкритого тексту та всіх одержаних шифртекстів і порівняти їх значення.

3. Використовуючи наведені теоретичні відомості, розшифрувати наданий шифртекст (згідно свого номеру варіанта). **Варіант №7**

Хід роботи

1. Так як задачею першого завдання є зашифрувати свій текст, перейшов до створення функії шифрування шифром Віженера. Логіка була доволі проста, виписуємо індекси ВТ потім виписуємо індекси ключа доповнюємо ключ повтором до довжини ВТ, додаємо індекси і текст зашифрований. Для дешифрування та ж логіка але від індексів літер ШТ віднімаємо індекси ключа.
2. Другим етапом було визначення індексу відповідності, для цього я скористався формулою:

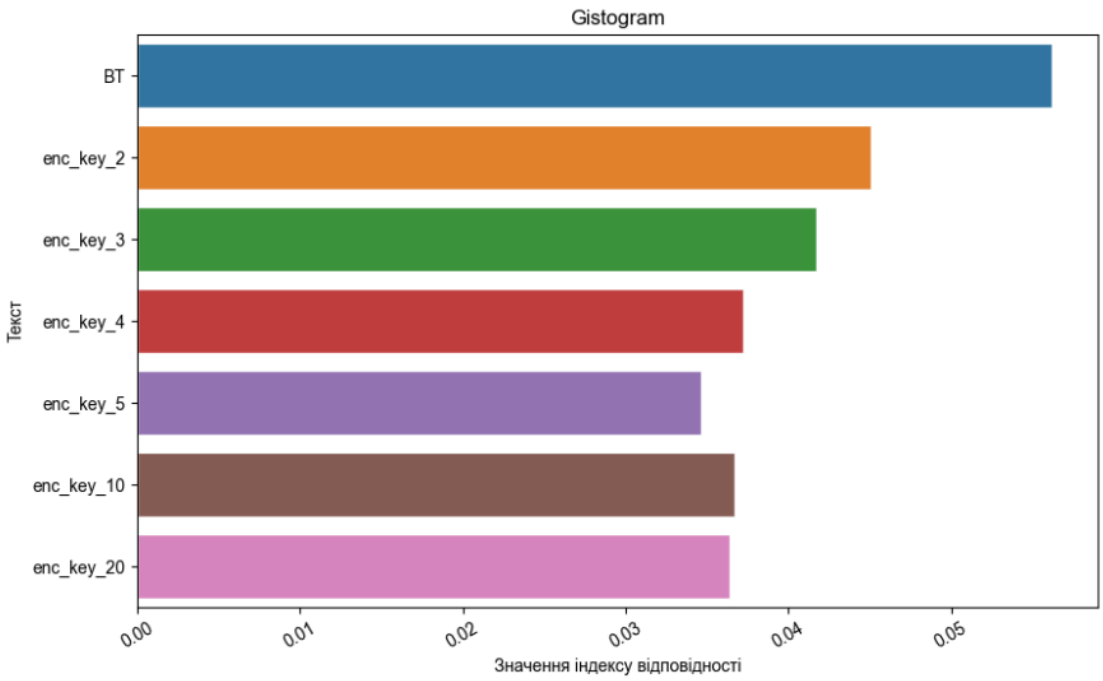
Індексом відповідності тексту Y називається величина

$$I(Y) = \frac{1}{n(n-1)} \sum_{t \in Z_m} N_t(Y)(N_t(Y)-1),$$

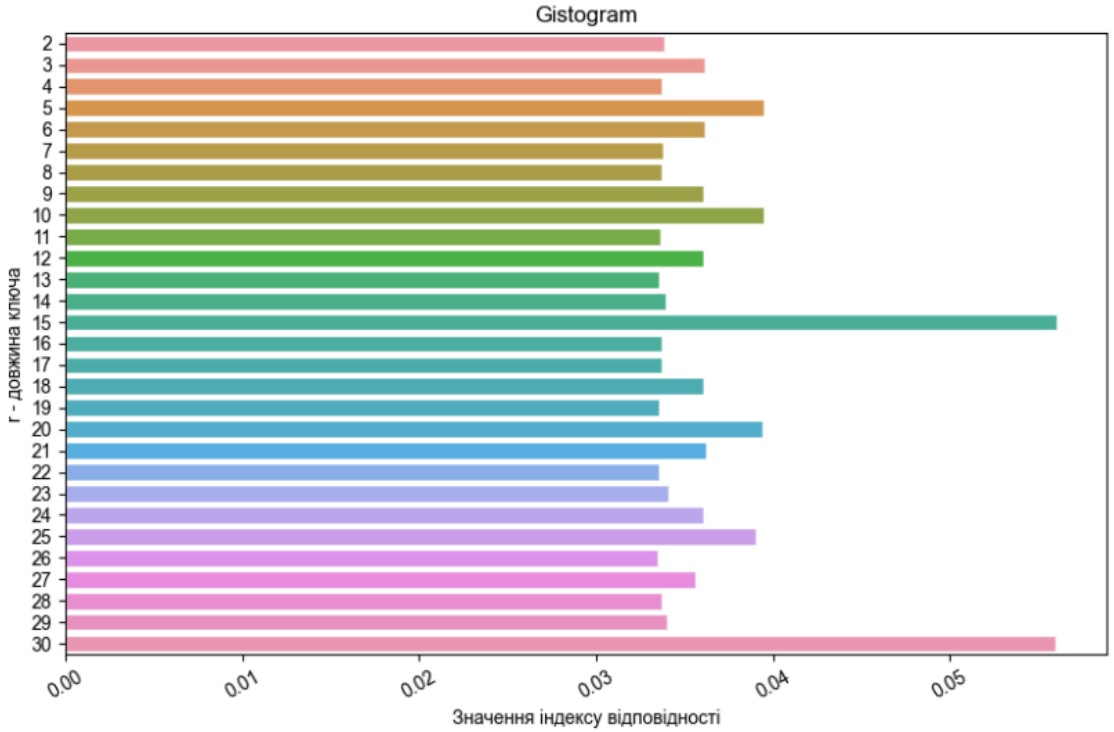
Та його математичне очікування:

$MI(Y) = \sum_{t \in Z_m} p_t^2$, де p_t – імовірність появи літери t в мові.

За допомогою цієї інформації зробив функції theoretical_i_of_conformity та math_expectation
3. Згідно завдання підрахував індекси відповідності для відкритого тексту та всіх одержаних шифртекстів і порівняв їх значення. Упевнився у фразі з методички про те, що величина індексу відповідності та його математичне очікування буде стрімко падати із ростом довжини ключа r.



4. Третій етап був дуже легкий для мене у реалізації, бо я скористався першим методом з методички, коли ми ділимо шифртекст Y на блоки Y1, Y2... Yr
Тоді за допомогою метод. вказівки про: «При пошуку періоду шифру Віженера потрібно перевіряти довжини ключів (обчислювати індекси відповідності блоків або значення статистики Dr) щонайменше до r = 30» було нескладно виставити границю для ключа до 30 і роглядати проміжок від довжини ключа 2 до 30 включно. Далі брав ці блоки ШТ та шукав індекс відповідності для кожного блоку, просумував їх та поділив на довжину r. Так для кожного ключа на проміжку від 2 до 30 я отримував індекс відповідності.



Подивившись на них, було видно, що на деяких ключах індекси +- ті ж, але мені був потрібний той, що відповідає саме моїй довжині ключа, тому я вирішив скористатися математичним очікуванням індексу відповідності, порахувавши математичне очікування індексу відповідності для мого відкритого тексту з частини 1, я використав його як теоретичне значення. Тоді від кожного індексу відповідності для певного ключа, я віднімав моє «теоретичне» математичне очікування та отримував лише єдиний варіант коли ця різниця буде мінімальною, та вивів його номер у масиві + 2(бо починали з ключа 2, а нумерація в масиві з 0). Таким шляхом отримав довжину ключа 15.

5. А ось друга частина третього етапу, була для мене надзвичайно складною для сприйняття, бо мене заплутала ця частина:

Після встановлення значення періоду шифру подальше його розшифрування зводиться до серії розшифрувань шифрів Цезаря. Дійсно, кожен фрагмент Y_i зашифрований шифром Цезаря з ключем k_i, $i=\overline{1,r}$; знайти цей ключ можна, поклавши $k=(y^*-x^*)\bmod m$, де y^* – буква, що частіше за всіх зустрічається у фрагменті Y_i, а x^* –

Я погуглив як взагалі шукають ключ, знаючи його довжину, але натикався лише на формулу індексу взаємної відповідності, дуже багато було реалізацій цієї формули, але я отримував лише одне і те саме значення індексу взаємної відповідності.

Змінивши тактику дій, після того, як розібрався, що ці блоки, на які я ділив, вони містять у собі літери зашифровані однією і тією ж літерою ключа. Використавши частотний аналіз по блоку, знайшов найчастішу літеру, і, відповідно порівнював її з топом найчастіших літер рос. алфавіту – «о», «е», «а», «и», «н», «т». Вистачило лише двох для знаходження ключа мого варіанту. Але, провівши додадкові дослідження із своїм зашифрованим текстом із першого завдання, я побачив, що не завжди нам вистачить двох літер для порівняння, тому, щоб зробити алгоритм більш адаптивним під більшість ситуацій, то взяв більше ніж 2 літери топу.

Приклад підлаштовування ключа до потрібного:

```
Довжина ключа: 15
арудазевархимаг
Does key is right? (y/n): n
Which number of letter isn't correct? 7
['а', 'р', 'у', 'д', 'а', 'з', 'о', 'в', 'а', 'р', 'х', 'и', 'м', 'а', 'г']

Does key is right? (y/n): y
We found the key!
арудазовархимаг

Process finished with exit code 0
```

Знайдений ключ: арудазовархимаг

Розшифрований(decrypt_cipher_text), шифрований тексти (cipher_text) можна знайти у папці Crypto

Зашифровані тексти різними ключами разом із таблицею індексів також подано у цій папці

Висновки

У ході даної лабораторної роботи я засвоїв методи частотного криптоаналізу. Здобув навички роботи та аналізу потокових шифрів гамування адитивного типу на прикладі шифру Віженера, написавши програмний код. Зробив висновок, що величина індексу відповідності та його математичне очікування буде стрімко падати із ростом довжини ключа r.