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

КРИПТОГРАФІЯ
КОМП’ЮТЕРНИЙ ПРАКТИКУМ №3

Виконала: Студентка 3-го курсу
Групи ФБ-93
Пономаренко Олександра Сергіївна

Київ 2021

Криптоаналіз афінної біграмної підстановки

Мета:

Набуття навичок частотного аналізу на прикладі розкриття моноалфавітної підстановки; опанування прийомами роботи в модулярній арифметиці.

Завдання до виконання:

0. Уважно прочитати методичні вказівки до виконання комп'ютерного практикуму.

1. За допомогою програми обчислення частот біграм, яка написана в ході виконання комп'ютерного практикуму №1, знайти 5 найчастіших біграм запропонованого шифртексту (за варіантом).
2. Реалізувати підпрограми із необхідними математичними операціями: обчисленням оберненого елементу за модулем із використанням розширеного алгоритму Евкліда, розв'язуванням лінійних порівнянь. При розв'язуванні порівнянь потрібно коректно обробляти випадок із декількома розв'язками, повертаючи їх усі.
3. Перебрати можливі варіанти співставлення частих біграм мови та частих біграм шифртексту (розглядаючи пари біграм із п'яти найчастіших). Для кожного співставлення знайти можливі кандидати на ключ (a,b) шляхом розв'язання системи.

$$\begin{cases} Y^{*} \equiv aX^{*} + b \pmod{m^2} \\ Y^{**} \equiv aX^{**} + b \pmod{m^2} \end{cases},$$

4. Для кожного кандидата на ключ дешифрувати шифртекст. Якщо шифртекст не є змістовним текстом російською мовою, відкинути цього кандидата.
5. Повторювати дії 3-4 доти, доки дешифрований текст не буде змістовним.

Виконання роботи:

Варіант 9:

1)ШТ знаходиться у CT.txt. Першим кроком є розбиття тексту на біграми(що не перетинаються) та знаходження п’яти найчастіших:

```
Max: (э,э) = 0.0165437
Max: (в,д) = 0.0127431
Max: (г,н) = 0.0125196
Max: (ц,г) = 0.0125196
Max: (ч,ф) = 0.0125196
```

Дізнавшися властивості мови, розуміємо, що 5 найчастіших біграм у російській мові таких: {ст, но, то, на, ен}.

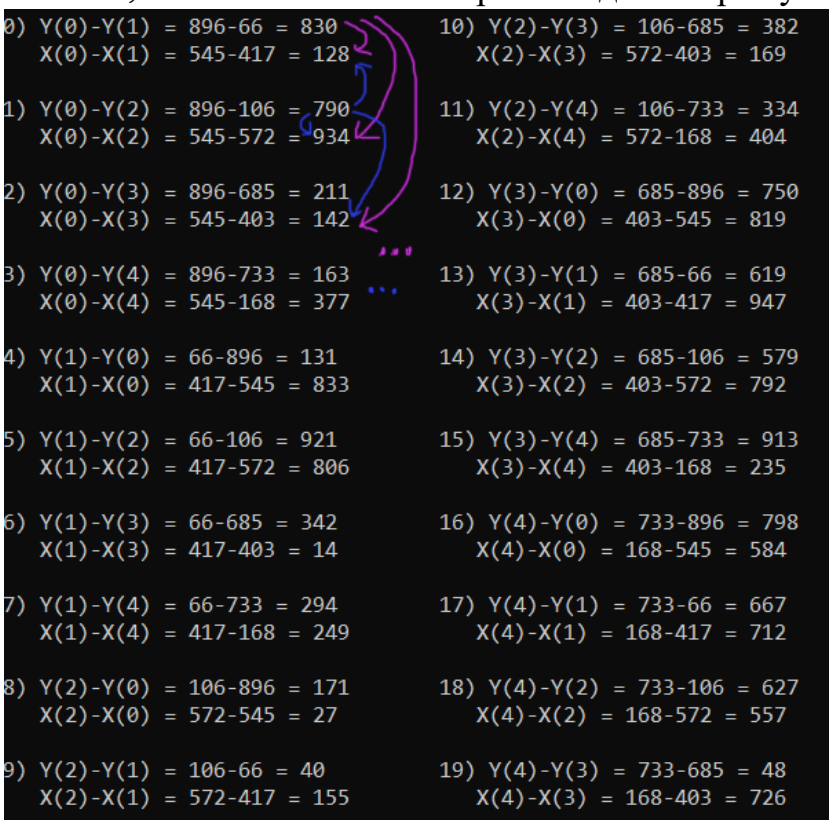
Так як ми працюємо із біграми, порядковий номер кожної розписується за формулою:

$$(x_{2i-1}, x_{2i}) \leftrightarrow X_i = x_{2i-1}m + x_{2i}.$$

Аналогічно і для біграм шифрованого тексту Y_i

```
Y0: ээ = 896      X0: 545 "СТ"
Y1: вд = 66       X1: 417 "НО"
Y2: гн = 106      X2: 572 "ТО"
Y3: цг = 685      X3: 403 "НА"
Y4: чф = 733      X4: 168 "ЕН"
```

2)Для визначення параметру (a) треба порахувати усі можливі значення (коли найчастіша біграма в мові -> найчастіша біграма в ШТ та інші). Усього маємо 5*5=25 варіантів, але не беремо ті випадки, коли одночасно X0->Y0, X1->Y0 та інші чотири випадки. Отримуємо 20 варіантів У та Х



Далі треба обрахувати усі (а їх відповідно 20*20=400 варіантів) обернені елементи до Х. Обернений елемент шукається за розширеним алгоритмом Евкліду, де додатково обраховуються значення u та v. Останнє значення v як раз і буде потрібним значенням оберненого до Х

```
830 = 128a mod961
a = 128^(-1)*830 mod961
u[2]: 1
v[2]: -7
u[3]: -1
v[3]: 8
u[4]: 2
v[4]: -15
u[5]: -63
v[5]: 473
GSD (961, 128) = 1
128^(-1) = 473
a = 473*830 mod961
a = 502
```

```
790 = 806a mod961
a = 806^(-1)*790 mod961
GSD (961, 806) = 31
31!|790
806^(-1) doesn't exist!
```

Є інший розвиток подій: коли НСД > 1. Тут ми вже дивимося, чи є це число (d) дільником У. Якщо ні - отримуємо порожню множину (отже, не існує такого оберненого), але якщо так - ділимо Х, У та М(модуль) на отримане число, і далі запускаємо рекурсію алгоритму Евкліда. Отримуємо d розв'язків (в моєму завданні було тільки 40 варіантів, коли розв'язків не існує, усі інші 360 мають тільки по одному).

Ключ (a) вже маємо, залишилося знайти відповідні значення (b)

```
b = Y0-a[0]*X0 = 896-502*545 mod961 = 230
b = Y0-a[1]*X0 = 896-254*545 mod961 = 850
b = Y0-a[2]*X0 = 896-642*545 mod961 = 810
b = Y0-a[3]*X0 = 896-693*545 mod961 = 884
b = Y0-a[4]*X1 = 896-459*417 mod961 = 732
empty set!
b = Y0-a[6]*X1 = 896-883*417 mod961 = 748
b = Y0-a[7]*X1 = 896-644*417 mod961 = 467
b = Y0-a[8]*X2 = 896-707*572 mod961 = 112
empty set! → ¬∃ α
```

Нарешті маємо усі 360 пар ключів (a, b). Переходимо до знаходження біграм відкритого текст. А саме до формули:

$$X_i = a^{-1}(Y_i - b) \bmod m^2$$

Отримали 360 варіантів усіх біграм відкритого тексту. По знайомій вже нам формулі тепер вже навпаки розбиваємо номер біграми на окремі номери двох літер і зливаємо біграми в повноцінний текст. Маємо 360 текстів, але тільки один є істинним. Для цього нам потрібно прибрати зайві тексти.

Спочатку я зробила алгоритм пошуку найчастіших та найрідкіснішої літер в тексті. Якщо це не {о, е, а} та {ф, щ, э} відповідно, текст вважається шумом, і його можно спокійно прибирати. Але, на жаль, цього алгоритму було недостатньо (в мене залишилось три тексти, що підходили до вище описаного алгоритму), тому я вирішила зробити додаткову перевірку на наявність хибних біграм (у російській мові не існує біграм з голосних літер та м’якого знаку). Саме так і отримала істинний відкритий текст, який зберігла до OT.txt

```
The most frequent letters: "т" and "м"  and the less is: "н"
The most frequent letters: "т" and "у"  and the less is: "ц"
The most frequent letters: "т" and "з"  and the less is: "е"
The most frequent letters: "т" and "м"  and the less is: "ю"
The most frequent letters: "о" and "н"  and the less is: "к"
The most frequent letters: "о" and "ш"  and the less is: "ь"
The most frequent letters: "о" and "э"  and the less is: "п"
The most frequent letters: "о" and "т"  and the less is: "д"
The most frequent letters: "о" and "ш"  and the less is: "ь"
The most frequent letters: "о" and "э"  and the less is: "е"
The most frequent letters: "а" and "ы"  and the less is: "ю"
The most frequent letters: "а" and "х"  and the less is: "в"
The most frequent letters: "а" and "х"  and the less is: "щ"
The most frequent letters: "ф" and "а"  and the less is: "ш"
The most frequent letters: "е" and "н"  and the less is: "п"
The most frequent letters: "н" and "я"  and the less is: "ц"
The most frequent letters: "е" and "н"  and the less is: "ь"
The most frequent letters: "н" and "й"  and the less is: "я"
The most frequent letters: "т" and "ы"  and the less is: "з"
The most frequent letters: "т" and "х"  and the less is: "з"
The most frequent letters: "т" and "у"  and the less is: "ч"
The most frequent letters: "т" and "ж"  and the less is: "ь"
The most frequent letters: "о" and "е"  and the less is: "ф"
RETURN "о", "е" and "ф"
The most frequent letters: "о" and "б"  and the less is: "с"
The most frequent letters: "о" and "н"  and the less is: "м"
The most frequent letters: "о" and "е"  and the less is: "ф"
RETURN "о", "е" and "ф"
```

Висновок:

За цю лабораторну роботу ми дізналися більше про поліалфавітні підстановки, а саме про шифр афінної підстановки (розглядаючи шифрування на біграмах). Згадали алгоритм пошуку НСД та дізналися, як за допомогою розширеного алгоритму Евкліда знайти обернений елемент. А також навчилися будувати автоматичний розпізнавач російської мови.