НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ

“КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ”

ФІЗИКО-ТЕХНІЧНИЙ ІНСТИТУТ

Криптографія

КОМП’ЮТЕРНИЙ ПРАКТИКУМ 3

«Криптоаналіз афінної біграмної підстановки»

ФБ-32 Дорошенко Ілля

Варіант 6

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

Порядок виконання роботи

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

1. 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.

2. За допомогою програми обчислення частот біграм, яка написана в ході виконання комп'ютерного практикуму №1, знайти 5 найчастіших біграм запропонованого шифртексту (за варіантом).

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 (1).

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

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.

Хід роботи:

Завдання 1:

Розширений алгоритм Евкліда: Реалізовано для знаходження найбільшого спільного дільника (НСД) двох чисел та коефіцієнтів Безу. Це необхідно для перевірки існування оберненого елемента та його обчислення.

```python
def extended_gcd(a, b):
    if a == 0:
        return b, 0, 1

    gcd, x1, y1 = extended_gcd(b % a, a)

    x = y1 - (b // a) * x1
    y = x1

    return gcd, x, y
```

Обернений елемент за модулем: Реалізовано функцію знаходження мультиплікативного оберненого елемента a^{-1} (mod m) з використанням розширеного алгоритму Евкліда. Обернений елемент існує тоді й лише тоді, коли gcd(a, m) = 1.

```
def inverse_mod(a, m):
    gcd, x, _ = extended_gcd(a, m)
    if gcd != 1:
        raise Exception("Обернений елемент не існує, оскільки gcd(a, m) ≠ 1")
    return x % m
```

Розв'язання лінійних порівнянь: Розроблено функцію для знаходження невідомого x у рівнянні ax ≡ b (mod n). Враховано випадок, коли gcd(a, n) = d > 1: у такому разі рівняння має d розв'язків.

```
def solve_linear_congruence(a, b, n):
    g, x, y = extended_gcd(a, n)

    if b % g != 0:
        return []

    x0 = (x * (b // g)) % n

    solutions = []
    step = n // g
    for k in range(g):
        solutions.append((x0 + k * step) % n)

    return solutions
```

Результат виконання коду:

Тест 1:

Обернений до 3 по модулю 11 має бути 4 (3*4 = 12 = 1 mod 11)

Тест 2:

2x ≡ 4 (mod 6). НСД(2,6)=2, тому має бути 2 розв'язки.

2*2 = 4 (ok), 2*5 = 10 = 4 mod 6 (ok). Розв'язки: [2, 5]

```
PS D:\KPI\crypto25-26> python -u "d:\KPI\cry

--- Перевірка 1 завдання ---
Обернений елемент до 3 по модулю 11 = 4
Розв'язки порівняння 2x ≡ 4 (mod 6): [2, 5]
PS D:\KPI\crypto25-26>
```

Завдання 2:

Попередня обробка: З файлу було зчитано весь текст. Виконано фільтрацію: видалено всі символи (включно з символами переносу рядка), що не входять до визначеного 31-символьного алфавіту.

```
ALPHABET = "абвгдежзийклмнопрстуфхцчшщьыэюя"
M = len(ALPHABET)
```

Формування біграм: Очищений текст розбито на біграми (x_{2i-1}, x_{2i}), що не перетинаються.

```
    for i in range(0, len(clean_text), 2):
        x1 = CHAR_TO_INDEX[clean_text[i]]
        x2 = CHAR_TO_INDEX[clean_text[i+1]]
        indices.append(x1 * M + x2)
```

Числове подання: Кожна біграма була переведена у числове значення X за формулою:

$$X = x_1 * M + x_2$$

де x_1 та x_2 — порядкові номери першої та другої літери біграми відповідно.

Статистичний аналіз: Підраховано кількість входжень кожного числа X у шифртексті та відсортовано їх за спаданням частоти.

Результат виконання коду:

```
--- Перевірка 2 завдання ---
Аналіз файлу: tasks/cp3/variants.utf8/06.txt
Всього біграм: 3456
Топ-5 найчастіших біграм (числа): [780, 656, 715, 346, 684]
Розшифровка:
    780 -> 'ще'
    656 -> 'хе'
    715 -> 'чв'
    346 -> 'ле'
    684 -> 'цв'
PS D:\KPI\crypto25-26>
```

Завдання 3:

Формування пар: Перебираються всі можливі пари біграм мови (X*, X**) та всі можливі пари біграм шифртексту (Y*, Y**). Загальна кількість комбінацій для перевірки становить (5 * 4) * (5 * 4) = 400 варіантів систем рівнянь.

Розв'язання системи порівнянь: Для кожної комбінації складається система лінійних порівнянь:

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

Знаходження параметра a: Шляхом віднімання другого рівняння від першого виключається невідоме b, і розв'язується лінійне порівняння відносно a:

$$(Y^* - Y^{**}) \equiv a(X^* - X^{**}) \pmod{m^2}$$

Для розв'язання використано реалізовану функцію solve_linear_congruence.

```
diff_X = (X1 - X2) % M2
diff_Y = (Y1 - Y2) % M2

candidates_a = solve_linear_congruence(diff_X, diff_Y, M2)
```

Фільтрація: З отриманих розв'язків відбираються лише ті значення a, які задовольняють умову існування оберненого елемента: gcd(a, m) = 1 (де m=31).

Знаходження параметра b: Для кожного коректного a обчислюється другий елемент ключа:

$$b \equiv (Y^* - aX^*) \pmod{m^2}$$

```
b = (Y1 - a * X1) % M2
```

Результат виконання коду:

```
--- Перевірка 3 завдання ---
Починаємо перебір варіантів...
Топ-5 біграм мови (X): [545, 417, 572, 403, 168]
Топ-5 біграм шифртексту (Y): [780, 656, 715, 346, 684]

Згенеровано унікальних ключів: 348
    Кандидати на ключ (a, b):
    (146, 441)
    (334, 906)
    (549, 565)
    (781, 817)
    (622, 442)
    (80, 360)
    (119, 873)
    (797, 937)
    (839, 287)
    (737, 10)
    (397, 427)
    (656, 687)
    (439, 591)
    (800, 551)
    (115, 875)
    (719, 724)
    (757, 360)
    (18, 515)
    (646, 749)
    (332, 720)
    (374, 70)
    (117, 41)
    (68, 228)
    (133, 623)
    (67, 687)
    (319, 902)
    (892, 782)
    (180, 582)
    (536, 747)
```

Завдання 4:

Алгоритм дешифрування: Для кожного кандидата (a, b) виконувалося пробне дешифрування фрагмента тексту (перші 300 біграм) за формулою оберненого афінного перетворення:

$$X = a^{-1}(Y - b) \pmod{m^2}$$

де a^{-1} — мультиплікативна інверсія числа a за модулем m^2 (961).

```
X = (a_inv * (Y - b)) % M2
```

Розробка критеріїв змістовності (Розпізнавач мови): Відповідно до методичних вказівок, розпізнавач базувався на статистичних властивостях мови. У ході роботи було протестовано два підходи:

1. Перевірка частот біграм (відповідно до п. 3):

    Спочатку перевірка здійснювалася шляхом пошуку "неможливих" поєднань літер (наприклад, «ьь», «оь», «аь», «цщ»).

    - *Результат:* Цей метод виявився недостатнім. Було отримано хибне спрацювання на ключі (a=146, b=441), який давав текст без заборонених біграм, але семантично беззмістовний (набір випадкових складів).

```
Перевірка 348 кандидатів...

УСПІХ! Знайдено ймовірний ключ: a=146, b=441
    Уривок: тдтонэпофюховыайпизштдибоэедруулдюжбдбяглжбекьепшхцюнйпфпоыхноцнпдзйнэмжпдсюедть...
Повний текст збережено у файл: decrypted_key_146_441.txt
```

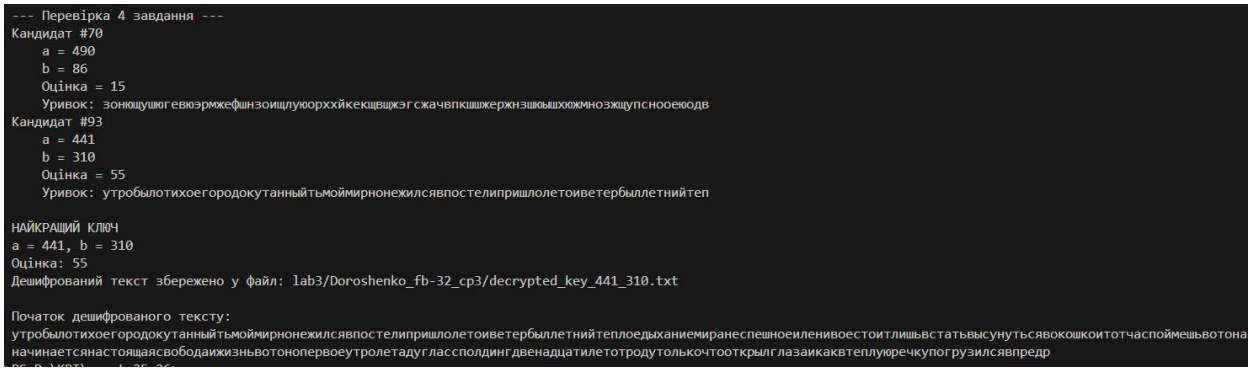
2. Перевірка частот довільних n-грам (відповідно до п. 4):

    Для усунення хибних спрацювань алгоритм було вдосконалено. Було впроваджено перевірку частот довільних n-грам оцінки тексту:

    - Текст отримував штрафні бали за наявність заборонених біграм.
    - Текст отримував бонусні бали за наявність найпоширеніших часток та прийменників російської мови («на», «то», «не», «ст», «по»), що фактично є перевіркою на наявність характерних 2-грам та 3-грам.

Таким чином, перебір кандидатів на ключ здійснювався до моменту отримання тексту, який задовольняє статистичним властивостям мови.

Результат виконання коду:

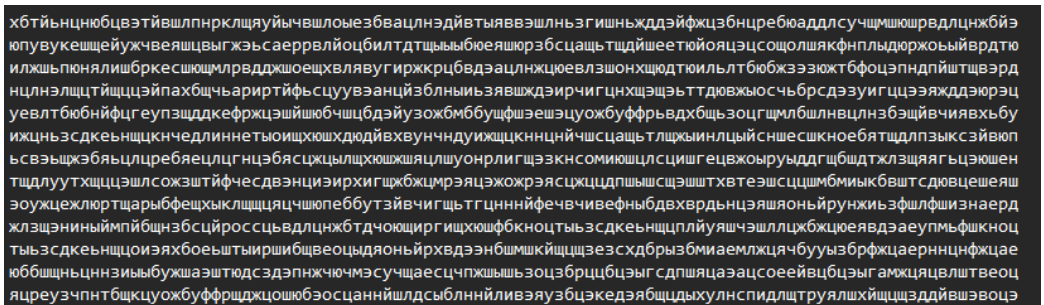


Завдання 5:

У ході виконання Завдання 4 здійснюється повний перебір кандидатів на ключ, пробне дешифрування та оцінка змістовності тексту.
Таким чином, дії, описані в Завданні 5, фактично реалізуються в рамках Завдання 4 шляхом послідовної перевірки всіх можливих кандидатів до отримання змістовного тексту.

Фінальний результат:

Шифротекст:



Розшифрований текст:

утробылотихоегородокутанныйтьмоймирнонежилсявпостелипришлолетоиветербыллетнийтеплоедыханиемиранеспешноеиленивоестоитлишьвстатьвыс
унутьсявокошкоитотчаспоймешьвотонаначинаетсянастоящаясвободаижизньвотонопервоеутролетадуглассполдингдвенадцатилетотродутолькочтоо
ткрылглазаикаквтеплуюречкупогрузилсявпредрассветнуюбезмятежностьонлежалвсводчатойкомнаткеначетвертомэтажевовсемгороденебылобашнив
ышеиоттогочтоонпарилтаквысоковоздухевместесиюньскимветромвнемрождалась чудодейственнаясилапоночамкогдавязыдубыикленысливалисьводн
обеспокойноеморедугласокидывалеговзглядомпронзавшимтьмуточномаякисегоднявотздоровошепнулонвпередицелоелетонесчетноемножестводнейч
утьнеполкалендаряонужевиделсебямногорукимкакбожествошиваизкнижкипропутешествиятолькопоспевайрватьещезеленыеяблокиперсикичерныекак
ночьсливыегоневытащитьизлесуизкустовизречкиакакприятнобудетпомерзнутьзабравшисьвзаиндевелыйледниккаквеселожаритьсявбабушкинойкухн
езаодностысячьюцыплятапоказаделоразвнеделюемупозволялиночеватьневдомикепососедствугдеспалиегородителиимладшийбратишкатомаздесьвде
довскойбашнеонвзбегалпотемнойвинтовойлестниценасамыйверхиложилсяспатьвэтойобителикудесникасредигромовивидений аспозаранкукогдадаже
молочникещенезвякалбутылкаминаулицахонпросыпалсяиприступалкзаветномуволшебствустоявтемнотеуоткрытогоокнаоннабралполнуюгрудьвоздух
аиизовсехсилдунулуличныефонаримигомпогаслиточносвечкиначерномименинномпирогедугласдунулещеиещеивнебеначалигаснутьзвездыдугласулыб
нулсяткнулпальцемтамитамтеперьтутивоттутвпредутреннемтуманеодинзадругимпрорезалисьпрямоугольникивдомахзажигалисьогнидалекодалекон
арассветнойземлевдругозарилась целаявереницаоконвсемзевнутьвсемставатьогромныйдомвнизуожилдедушкавынимайзубыизстаканадугласнемног
оподождалбабушкаипрабабушкажарьтеоладьисквознякпронесповсемкоридорамтеплыйдухжареноготестаивовсехкомнатахвстрепенулисьмногочислен
ныететкидядьядвоюродныебратьяисестрычтосехалисьсюдапогоститьулицастариковпросыпайсямиссэленлумисполковникфрилеймиссисбентлипокашл
яйтевстаньтепроглотитесвоитаблеткипошевеливайтесьмистерджонасзапрягайтелошадьвыводитеизсараяфургонпораехатьзастарьемпотусторонуов
рагаоткрылисвоидраконьиглазаугрюмыеособнякискоровнизупоявятсянаэлектрическойзеленоймашинедвестарухиипокатятпоутреннимулицампривет
ственномахаякаждойвстречнойсобакемистертридденбегитевтрамвайноедепоивскорепоузкимусламмощеныхулицполыветтрамвайрассыпаявокругжа
ркиесинииеискрыджонхафчарливудменвыготовышепнулдугласулицедетейготовыспросилонубейсбольныхмячейчтомоклинаросистыхлужайкахупустыхве
ревочныхкачелейчтоскучаясвисалисдеревьевмампаптомпроснитесьтихонькопрозвенелибудильникигулкопробиличасызданиисудаточносетьзабро
шеннаяегорукойсдеревьеввзметнулисьптицыизапелидирижируясвоиморкестромдугласповелительнопротянулрукуквостокуивзошлосолнцедугласскр
естилрукинагрудииулыбнулсякакнастоящийволшебникволттотодумалонтолькояприказаливсеповскакаливсезабегалиотличноебудетлетоионнапослед
окогляделгородищелкнулемупальцамираспахнулисьдверидомовлюдивышлинаулицулетотысячадевятьсотдвадцатьвосьмогогоданачалосьвтоутропрох
одяполужайкедугласнаткнулсянапаутинуневидимаянитькоснулась еголбаинеслышнолопнулаиотэтогопустячногослучаяоннасторожилсяденьбудетне
такойкаквсенетакойещеипотомучтобываютднисотканныеизоднихзапаховсловновесьмирможновтянутьносомкаквоздухвдохнутьивыдохнутьтакобясня
лдугласуиегодесятилетнемубратутомуотецкогдавезихвмашинезагородавдругиеднигоговорилещеотецможноуслышатькаждыйгромикаждыйшорохвселенн
ойиныеднихорошопробоватьнавкусаиныенаощупьабываютитакиекогдаестьвсесразувотнапримерсегодняпахнеттакбудтоводнуночьтамзахолмаминеве
стьоткудавзялсяогромныйфруктовыйсадивседосамогогоризонтатакиблагоухаетввоздухепахнетдождемнонанебениоблачкатогоиглядиктотоневедом
ыйзахохочетвлесунопокатамтишинадугласвовсеглазасмотрелналывущиемимополянетнисадомнепахнетнидождемдаиоткудабыразнияблоньнетнитучи
ктотамможетхохотатьвлесуавсетакидугласвздрогнулденьэтоткакойтоособенныймашинаостановилась всамомсердцетихоголесаануребятанебаловат
ьсяониподталкивалидругдругалоктямихорошопапамальчикивылезлиизмашинызахватилисиниежестяныеведраисойдяспустыннойпроселочнойдорогипо
грузилисьвзапахиземливлажнойотнедавнегодождявшитепечелсказалотецонивсегдавьютсявозлевиноградакакмальчишкивозлекухнидугласдугласвст
репенулсяопятьвитаешьвоблакахсказалотецспустисьназемлюпойдемснамихорошопапаионигуськомпобрелиполесувпередиотецрослыйиплечистыйзан
имдугласапоследнимсеменилкоротышкатомподнялись наневысокийхолмипосмотреливдальвонтамуказалпальцемотецтамобитаютогромныеполетнемути
хиеветрыинезримыеплывутвзеленыхглубинахточнопризрачныекитыдугласглянулвтусторонуничегонеувиделипочувствовалсебяобманутымотецкакид
едушкавечноговоритзагадкамииивсетакидугласзатаилдыханиеиприслушалсячтотодолжнослучитьсяподумалоняужзнаюавотпапоротникназываетсяве
неринволосотецнеторопливошагалвпередсинееведропозвякивалоунеговрукеаэточувствуетеионковырнулземлюноскомбашмакамиллионылеткопилсяэ
тотперегнойосеньзаосеньюпадалилистьяпоказемлянесталатакоймягкойухтыступаюкакиндеецсказалтомсовсемнеслышнодугласпотрогалземлюнони
чегонеощутилонвсевремянастороженноприслушивалсямыокруженыдумалончтотослучитсяночтооностановилсявыходижгетытамчтотытакоемысленно
кричалонтомиотецшлидальшепотихойподатливойземленасветенеткружеватоньшенегромкосказалотецпоказалрукойвверхгделиствадеревьеввплета
ласьвнебоилиможетбытьнебовплеталосьвлиствувсеравноулыбнулсяотецвсеэтокружевазеленыеиголубыевсмотритесьхорошенькоиувидителесплетет
ихсловногудящийстанокотецстоялувереннопохозяйскиирассказывалимвсякуювсячинулегкоисвободноневыбираясловчастоонисамсмеялсясвоимрасс
казамиотэтогоонитеклиещесвободнеехорошоприслучаепослушатьтишинуговорилонпотомучтотогдаудаетсяуслышатькакноситсяввоздухепыльцаполе
выхцветовавоздухтакигудитпчеламидадатакигудитавотслышитетамзадеревьямиводопадомльетсяптичьещебетаньевотсейчасдумалдугласвотоноуже
близкоаяещеневижусовсеблизкорядомдикийвиноградсказалотецнамповезлосмотритеканенадоахнулпросебядугласнотомиотецнаклонилисьипогруз
илирукившуршащийкустчарырассеялисьтопугающееигрозноечтоподкрадывалосьблизилосьготовобылоринутьсяиптрястиегодушуисчезлоопустошенн
ыйрастерянныйдугласупалнаколенипальцыегоушлиглубоковзеленуютеньивынырнулиобагренныеалымсокомсловноонвзрезаллесножомисунулрукивотк
рытуюранумальчикизавтракатьведрачутьнедоверхунаполненыдикимвиноградомилеснойземляникойвокруггудятпчелыэтововсенепчелыацелыймиртих
онькомурлычетсвоюпесенкуговоритотецаонисидятназамшеломстволеупавшегодереваюютсандвичиипытаютсяслушатьлескакслушаетонотецчутьпосм
еиваясьискосапоглядываетнадугласахотелбылочтотосказатьнопромолчалоткусилещекусоксандвичаизадумалсяхлебсветчинойвлесунеточтодомавк

Висновок:

На основі частотного аналізу було виділено найбільш імовірні біграми, що дозволило сформувати множину потенційних ключів. У процесі автоматизованого перебору кандидатів було виявлено недостатню ефективність фільтрації лише за забороненими біграмами, що призводило до хибних спрацювань на коротких уривках тексту. Тому алгоритм було вдосконалено шляхом впровадження перевірки частот довільних n-грам, а також аналізу наявності характерних часток та прийменників. Застосування покращеного критерію дозволило успішно відсіяти шум, однозначно визначити коректний ключ (a = 441, b = 310) та відновити змістовний текст.