

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

Криптографія

**Комп'ютерний практикум №3
Криптоаналіз афінної біграмної підстановки**

**Виконали:
Студенти групи ФБ-32
Коптєва Ганна, Чупріна Вікторія**

Київ - 2025

Тема: Криптоаналіз афінної біграмної підстановки

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

Постановка задачі:

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, повертуючи їх усі.
2. За допомогою програми обчислення частот біграмм, яка написана в ході виконання комп'ютерного практикуму №1, знайти 5 найчастіших біграмм запропонованого шифртексту (за варіантом).
3. Перебрати можливі варіанти співставлення частих біграмм мови та частих біграмм шифртексту (розглядаючи пари біграмм 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.

Підготовка

Перш ніж аналізувати шифротекст, його спочатку треба “нормалізувати”. Тому:

- переводимо все у нижній регістр
- видаляємо пробіли, цифри, та знаки пунктуації
- замінюємо літера “ё” на “е”, а “ъ” на “ь”

```
def normalize_text(text: str) -> str:  
    text = text.lower()  
    text = text.replace("ё", "е").replace("ъ", "ь")  
    text = "".join(ch for ch in text if ch in alpha_set)  
    return text
```

Частотний аналіз

У російській мові біграми “ст”, “но”, “то”, “на”, “ен” дуже поширені. Тому, логіка проста: найчастіші біграми в шифртексті, швидше за все, відповідають найчастішим біграмам у мові оригіналу.

Після аналізу маємо, що наступний топ біграмм:

Біграма	Кількість	Частота
рн	62	0.025
ыч	41	0.017
нк	34	0.014
цз	32	0.013
иа	30	0.012

Припускаємо, що ці біграми – зашифровані “ст”, “но”, “то”, “на”, “ен”.

Пошук кандидатів у ключі

Маємо таку формулу біграмного шифру

$$Y = (\alpha \cdot X + b) \bmod M^2$$

де Y – числове значення зашифрованої біграми

де X – числове значення вихідної біграми

а і b – секретні ключі, які треба знайти

M – кількість літер в алфавіті

Наприклад, “ст” \rightarrow “рн” та “но” \rightarrow “ыч”, це дає наступну схему:

1. число(“рн”) = ($a * \text{число(“ст”)} + b$) mod 961
2. число(“ыч”) = ($a * \text{число(“но”)} + b$) mod 961

Віднімаємо одне рівняння від другого – позбуваємося b і можемо знайти a. А знаючи a – знаходимо b.

```
def charmap():
    char2i = {ch: i for i, ch in enumerate(ALPHABET)}
    i2char = {i: ch for ch, i in char2i.items()}
    return char2i, i2char

char2i, i2char = charmap()

def bigram_to_num(bg: str) -> int:
    return char2i[bg[0]] * M + char2i[bg[1]]

def num_to_bigram(x: int) -> str:
    return i2char[(x // M) % M] + i2char[x % M]
```

Текст розбиваємо на біграми без перекриття:

```
def split_non_overlapping(s: str):
    n = len(s) // 2 * 2
    return [s[i:i + 2] for i in range(0, n, 2)]
```

Частоти біграмм обчислюються так:

```
def compute_bigram_counts_nonoverlap(text: str):
    bigs = split_non_overlapping(text)
    cnt = Counter(bigs)
    total = sum(cnt.values()) or 1
    rows = []
    for bg, c in cnt.most_common():
        rows.append({"bigram": bg, "count": c, "freq": c / total})
    return rows
```

У головній функції програми беруться 10 найчастотніших біграмм шифротексту і виводяться на екран. З них обираються 5 найчастіших — саме вони далі використовуються як cipher_top5 для побудови кандидатів на ключ:

```
rows = compute_bigram_counts_nonoverlap(norm)
cipher_top5 = [row["bigram"] for row in rows[:5]]
```

Теоретичні «мовні» топ-5 біграмм (за методикою) зафіксовані константою:

```
LANG_TOP5_BIGRAMS = ["ct", "но", "то", "на", "ен"]
```

Пошук кандидатів у ключі

1. Розширений алгоритм Евкліда та обернений елемент

```
def egcd(a, b):
    if b == 0:
        return (a, 1, 0)
    g, x1, y1 = egcd(b, a % b)
    return (g, y1, x1 - (a // b) * y1)

def invmod(a, m):
    a %= m
    g, x, y = egcd(a, m)
    if g != 1:
        raise ValueError("No inverse")
    return x % m
```

2. Розв'язання лінійної конгруенції

```
def solve_linear_congruence(A, B, mod):
    A %= mod
    B %= mod
    g = math.gcd(A, mod)
    if B % g != 0:
        return []
    A1, B1, M1 = A // g, B // g, mod // g
    inv = invmod(A1, M1)
    x0 = (inv * B1) % M1
    return [(x0 + k * M1) % mod for k in range(g)]
```

Функція розв'язує рівняння і повертає всі можливі розв'язки. Цим виконується вимога методики: коректно обробляти випадок із кількома розв'язками.

3. Генерація кандидатів (a,b)(a, b)(a,b)

Отримуємо лінійну конгруенцію відносно ааа. Її розв'язуємо через `solve_linear_congruence`, а потім за будь-яким із початкових рівнянь відновлюємо bbb:

```

def generate_candidates(cipher_top5, lang_top5=LANG_TOP5_BIGRAMS):
    from itertools import permutations
    candidates = set()
    details = []

    for X1, X2 in permutations(lang_top5, 2):
        x1, x2 = bigram_to_num(X1), bigram_to_num(X2)
        dX = (x1 - x2) % MOD

        for Y1, Y2 in permutations(cipher_top5, 2):
            y1, y2 = bigram_to_num(Y1), bigram_to_num(Y2)
            dY = (y1 - y2) % MOD

            sols = solve_linear_congruence(dX, dY, MOD)
            for a in sols:
                # a повинно бути оберненим modulo 961
                if math.gcd(a, MOD) != 1:
                    continue
                b = (y1 - a * x1) % MOD
                candidates.add((a, b))
                details.append(
                    {"lang_pair": f"{X1}->{Y1}, {X2}->{Y2}", "a": a,
                     "b": b}
                )

    return sorted(list(candidates)), details

```

У результаті маємо список усіх можливих пар (a,b) (a, b) (a,b) , які узгоджуються з деякими відображеннями частих біграм мови на часті біграми шифру.

Перебір

На цьому етапі потрібно визначити, який із кандидатів (a,b) (a, b) (a,b) дає осмислений російський текст. Використовується індекс збігів (Index of Coincidence, IC) як основний автоматичний критерій.

```

def index_of_coincidence(s: str) -> float:
    n = len(s)
    if n < 2:
        return 0.0
    cnt = Counter(s)
    return sum(c * (c - 1) for c in cnt.values()) / (n * (n - 1))

```

Для російськомовного відкритого тексту $IC \approx 0.055$. Якщо IC розшифрованого тексту близький до цього значення, це сильна ознака, що дешифрування правильне.

Дешифрування для заданого ключа

```
def decrypt_with_key(text: str, a: int, b: int) -> str:
    inva = invmod(a, MOD)
    bigs = split_non_overlapping(text)
    res = []
    for bg in bigs:
        if len(bg) != 2:
            continue
        Y = bigram_to_num(bg)
        X = (inva * (Y - b)) % MOD
        res.append(num_to_bigram(X))
    return "".join(res)
```

Відбір кращих кандидатів

У головній функції для кожного кандидата (a,b) :

1. Дешифруємо весь нормалізований шифротекст.
2. Обчислюємо IC для отриманого тексту.
3. Рахуємо відхилення від теоретичного значення:

```
IC_diff = abs(IC - 0.055).
results = []
for a, b in candidates:
    try:
        pt = decrypt_with_key(norm, a, b)
    except Exception as e:
        print(f"[!] Помилка при дешифруванні a={a}, b={b}: {e}")
        continue
    ic = index_of_coincidence(pt)
    ic_diff = abs(ic - 0.055)
    results.append({
        "a": a,
        "b": b,
        "IC": ic,
        "IC_diff": ic_diff,
        "plaintext": pt,
    })
```

Якщо кандидатів немає, програма завершується. Інакше результати сортуються за зростанням IC_diff:

```
results.sort(key=lambda r: r["IC_diff"])
```

На екран виводяться топ-5 кандидатів за IC, для кожного показуються:

- значення a і b,
- IC та $|IC - 0.055|$,
- перші 500 символів розшифрованого тексту, щоб користувач міг візуально оцінити його осмисленість.

Найкращий кандидат (перший у списку) додатково виводиться з більшим фрагментом тексту (1000 символів) та зберігається у файлі в тій самій папці, що й скрипт:

- best_plaintext_ic.txt — повний розшифрований текст;
- best_key_ic.txt — знайдений ключ у вигляді a=..., b=..., IC=....

Таким чином автоматичний відбір робиться за індексом збігів, а остаточне рішення щодо змістовності тексту ухвалюється після перегляду декількох найкращих варіантів.

Пара (a=13, b=151) значно відривається.

Висновки

У ході виконання даної лабораторної роботи, ми навчилися набуттю навичок частотного аналізу на прикладі розкриття monoалфавітної підстановки; А також опанували прийоми роботи в модулярній арифметиці.

Дешифрований текст

многограннуюличностьдостоевскогоможнорассматриватьсчетырехсторонкакписателякакневрот икакакмыслителяэтикаикакгрешникакжеразобратьсяявэтойневольносмущающейнасложности наименеспоренонкакписательместоеговодномрядусхекспромбратьякарамазовывеличайший романизвсехкогдалибонаписанныхалегендаовеликоминквизитореодноизвысочайшихдостижени ймировойлитературыпереоценитькотороеневозможноксожалениюпередпроблемойписательског отворчествапсихоанализдолженсложитьоружиедостоевскийскореевсегоуязвимкакморалистпред ставляяегочеловекомвысоконравственнымнатомоснованичтотолькототдостигаетвысшегонравс твенногосовершенствактопрошелчерезглубочайшиебездныгреховностимыигнориремоднособ ражениееведьнравственнымявляетсяячеловекреагирующийуженавнутреннеиспытываемоеискуще ниеприэтомемунеподдаваяськтожепеременнотогрешиттораскаиваясьставитсебевысокиенрав ственныецелитоголегкоупрекнутьвтомчтоонслишкомуднодлясебястроитсвоюжизньоннеиспол няетосновногопринципанравственностинеобходимоститреченияятиввремяякакнравственныййобр азжизнивпрактическихинтересахвсегочеловечестваэтимоннапоминаетварваровэпохипереселен иянародовварваровубившихизатемкаявшихсявтомтактопокаяниестановилосьтехническимп римеромрасчищавшимпутькновымубийствамтаажепоступаливангрозныйэтасделкассовестьюха рактернаяруссскаячертадостатчнобесславениконечныйитогнравственнейборьбыдостоевскогоп

Вхідний текст

жэоыгсыюыхккоекъэхчпэюргбчцюмыягийтъансбдывекнршруванузкъяципазэлтыкъзэлй юрмувнусыюоыюдежжъсбххиуынпеуссдкруытчкбзхсъмгяшквецфяылхсийовукзпефшфийармжя чыэшюомтэдзвухщбиэтэюорыучшпуютерпэбъпвхлкдюбзкттыщцапумзшфшьродъннежеобчиэх громуацфяишшехюпкучфсърсааяглхшхъртъфзмшхжгэрэлжынълчыгфьробффрикаычсяэтэзш

пкачъроэюпвщрйтэюьбаьяфиуымырафафяжъкайцбршанвинзълмгцхюжлъкщярфбийхпзиеиоэх
роуыуэютпзкмгцыфхынpxвэшрбунтеапаяцбршаноэцъяунщтетзбуусрумгяюпзжцьбэкъпранфз
цянсфгпвтжстэуэйтфрьдыпчшууэйриельорспийяпвещбизвбжлвежшзыиэтюгчвцпкачъроэр
оккечшэкшлъяпшсснацщшбэмкхфуюшвноуткъфьшнарпкмавыэшхкдънтэофсюрвбагфрын
аэзтмтосучскяцбфюхощтзыцыпчжъдэцпфсажфпсвъкыцънцзытихшхглфрсдхкюйрэйпсъбв
шсвецфщштдивнмешьционаэххсизчптфчапдвнтеуодшчюлуэнжфцздтцбфюфшршюццбжффр
фофчсъюыоузиитюпхфдбэжвгутхяуийшркремшхэйальсншдечакчомууяздциюпхвтвжэпкач
ъроягевбчпвлмафмюгжыцсыефэрнфзхкуэщшбыденссъюыароскютмхлуязфштляефроут
яоишиофщыльлэнцкхшсгэбяльдшкыцъясуткбчпвлкъбсвъдайтгфавпгьпвяибпубаутфэюпукл
юоъркрухцтхамсдиеаудафшсыбыгжыцсытодчртуднъщпнбадхшнъсшхтпнсдхпувбшнхрквд
тпгуныбчюриухцшфрслянишьсыфомкросюекцищшунпяехясцхууэсжсчъжъэльвчшдн
саараричэтэюьбарюсжсчпжъюшвмквуняждпщэгпвщахсргъошфтжлпээнштбрфъкюэстпетьу
жэлгърнбцдфзуыаснвфшвдукняшофгуыеноахтглштпугвдатюфмюугюмздцхэшбдвлешфсвч
юугхаккмсзытмубсюашшчххвадфэцжгэшщбшшсзийфквчийшеюргишаэошмыэяуькъюшогуы
здшоцстрайгвзхтфэюгпвдуфтпбэкхокрругшбшшпвфябхптоуруиддэртупсбаванщфцояц
уйцубридьупфтшппрдкняпрмбгфрьдфэхчбюонжеефямъюяркэбспюоывжлшкреулькыжаэъ
лъныцъдэйэрйрдшыдхмхобсьффшуфаоаллфжчцвьюшвнцжхьдифбхлхуусээоэпдвыжкл
тглмюгыбданыеувуныбяяпзткшьизжэатаильюфлюгшаддвшчсзрьаэюппусфсывпятджфуыэшрв
шыыпжишвфсзбдяннфмеэпуюждыззшчцаыцешэнгучжаэкхшшэмдсеаяцяюшвримкъэыепчшсг
жыцсъкоихаяышкъвойючярмрзшыгчмтехмюишрцсцэйшхмкюкцияшовжхлкчтупцфобвтжчп
въгижаиквъээппреутзякняфэшыпчхпръущциумиякнлдяжшлуюзфштыычсбгыбсрвзшшсшръу
осучцпцвэтэяпкучщэрупачянжушрбдтьегсцэишупфэбчюцфжлптяцбийембуэншпкртышгфаткх
ыцтбяюфркеэгэхгупзсргныцибуупмбязкгфхгцынфвшшбэтыаелиежххсъхшшбскуаутфпцбюор
феауафштпевмкуляефроуесвтэцяисерифэчшфуиббяяпкучщэчюеюлифишиыэкфхопидгнц
вовыпагсюпкцглааъэуллжхпупцъоууквччевщциарвримкъэцэубгепэшгэххушибккщикчфхршэюпв
щржткуэжванщекуяянехюиуувувлбехцютпэргыпфлсввлгяыфобчяфвтэглтрлцынфвшляъы
йхюигшжетэюьбафдтонфбвялххстлпъджнбуутыеиуыщцъеашаекъуыягвпшынтафъяждюуфхпзы
емтфлряяепрдуфчнъбеануускяцбялорынльчфомывидуффшчийженжччляефроахтикучыч
айчхсучхетщцаныыејтссъцпгюкюафъщюпюмаэъюиэщпуэнслткүуцидфлсюидаояшэйяш
рзщэглзэахазркчсъюыюмвийшфвийшмунсреуыпчмаашхежххасълкхррэцхшривлагкфуйпв
оымсучорххчпсийелиожхпэтцэиуынпэччяяызфдмнпъныцържъынппнъжэъпвотрдууручцъжуэъ
хыумяярыйдморкущбдхдбуннжцкуысыыншххрачтывдфжтпэбцэжяяпреуыпгоуушгзкнлбъ
ясбяялкучцыюошьсреќсъюыюорынлюффаачюлуувуяньгхйтжспфэхчбюотчжийгтцэиуынб
щащбэфхотырзбъквсщхнбаюжпсъгэббфэпшптфщамбфмрбмпээрббяюипэишхъццржбсрнсс
яцбшшбзикыыэфшмыфпруцхпштжгизфдмяэзупдяижедчясщхууъзбщащбфмяпххдкъцбд
фиюиудкъглжцбфэфжцьбэкяжхгсэюпбэсясббозиумжэмпуванузкъячфшсуэгвднъсымршбкхчш
уцквжийнлдхмшштпшбоншцнквжэсрхъеҳщыцажеююжриупштяшпккбпфэтриуынуфъяйтцаа
мрюудухсюцвпэрлкчъдчъбадэдгжцмяиэпхюкпуйшвбрухиззеклцащсийхркэркэоцъбэпрфиеось
ибугргвебяаэлшвутчкхкшуныатыншхнэъзълыйпэххшаюаэгнтифщвохзсиемцуухлжюогки
естчувахдсузыцямжжохъдпчммджрвйтнгсбэуқцэйвювкшртткурвопбуэцтлыхлнфюеизчмаяызпх
бдэхнъпильгхлпуккчцушртэюзбъпэоуцумбвзффкцдуыбфлирельлштэждзяукеэчоуэпзсиуыаф
шюфехччийдщдаьмебспречмаяфххтеюмзкцпбуюхыъсреќяаьбчркоахкюуигзубмэйбпюлчапдяд
тжттыбцэжвюорфиеосъзтшгриутициспепрючтпфюжшсбжийшифшшжмукзпюцмссзожо
мцудвяахжпквнцъюношнфвшшжъюгшфножчтпфявпетлжкпццтжбюсиуыафшюйквнздшшб
хреюхеккшлятирштбхфбгррзхкчкрупмзъсевдэжвазчжийэчапдядтжтквбиихадочзы
цбнсжбвийтучкюэчюнбузоекюоъмнбщоншюмъахвалиуенцсфъямуйкзюнцятыйждвбрдупэчшро
чхтфээжвоцвсъзтштосаухиобнууххпхмадвннфжхтахтжээнзвуусрхлггчзебыэюсбхнсгефщси
хшпвъбнхярблжбрфьеууэнупжбстжнхгптзубтэржцъсърбещшбэьеацъгтшъсрзреъинубрхътп
ыбцяпцшавгзмаяхрцъоббеещяицэдшфежршукртпюорпэшшсъщреъыбыкйрэйпстшбдлпевдцх
ржлмлкиечхпкшубсриулщиыйдмлпэуыягвээвноунщбфшлгыузыуубпшблучрнжзкэчххувюорфж
опкфххгхлбзхшвионаютжжъжибашлвбсшшыхшшуурийкуюийгхорийхшърбэялсзщкхсиш
твюокпаршвтъайцюгвачеюпхсаудпэсшчфамгдяноеньнэюнквнгуршаянцешьзтштосьнвавюлпцф

ъяачхсбвъсжсчцздубцдожкстъчоешщоръкосшцспхбдопчшвээабашквкамапфпуыббрэоощяокыа шврбекмшурьрпкхржячюжетррзхшуэофжашолмычпроырнэйэцбххсчшмвейкбчесилюодфь шяштцамшндашхсщгииопръудбрембнтэзхцтюквиуювкыаънблбпхвцшэшхшушъпхысчцуш гзаюбфжхийурьбъвджлтвэкбжибсриучфпыубжрпкхржаагбуаниэецаиштушфтчайкдтигбгшынфз чшыищушынтецяътипчркюкнясаулщаюозебапфъгцуутмшхпывхсчшмвейшгшыфбрвяолмыпщ эжфхркгнышффийехозибшопыпюквкумцяюдымэяйпийрвбцдукэоощъжгвырыкяуорлытяб ыуънщцбичхкпшжпбфлггчатезумяъхрнэюлпэфшщшрмыбыугеояаъэшбхвнээфшиштанукбмяъ хштэупгфсшпошыжчэйшсэшктиокххпэкшюпфхотткэпкъяигнбыйнштпгсцвпвпсюшхтъдяпшн фэзыуэсбривмвтпэшблбънпкнчянпрутэтфацьсынврююсюэишафщъпянтишрхятытешрфштэг эхэжыбцзятпгрыфжеюнаэжууртобщуриспуэчыпмхмщлцхмзнэрбентжтчмштпафтчайтуюцэыэ грееъщмумнбармакчшыльэгкейшюдшротвдежфшвънфоыщрещпбурэбафорэчырсчхтахноjj цябюошьнелчлмбдчжяэъоавъщцкглыюмкйгосърбцбфийевэълргурсэхшэшрочхотафшхрьй щхжвеемцашташхдияхррвфчрлкиечхпявпрнжлтшэхлуънпэхпяяибжаяпвъйкуфмимпеххисиф бпшхобэмрхчшьчамгыфдпфкцбэцяжгюнпэчоощбзюаарпджзыцычоебсдпацишцбрхтешцхъцувнв лульэжтыапщбахяквъбщбчтюсускзвхэйфхмжъфдуфнгцбцэубтятапьюшюрутчкнпшфуисьеюко вуыыэшсэхаяевхквэльошшрмшлкъпахсехвргнасбгэбтянишжельцифэаяуазеырабафяглпвбкхо аллзыуулрьичгуяяпэччсцнмшбтыэцьубийияпзвхквьгергюрсэхшуаюсбэтугшбщъцбэхбдмшпяа янфоуздткхээсрсынкюацфдахлктчаякубцячнхрпгччптоцбгнилцпбурэбафсввзгэхрвбузчзб цаъмлбвнтжосувярмеюсеасчябкхубтжкцяшъличхрюеэзгэфютеандэлтуфамшеюзгэньххгшы зъфшазацбробкзьттъцумутмэбхрьинэадъяиасчжыфпелучнхщафхсеэябнъсмртыэридоцс ыилиуяпричкроххшфнцэхощыиэеэриожоъяухютчъмеупвърсафлкфшнхфлюгафеечцызсююс ъкязыцдтвпцюбринюопххнвхпдэовщычапдядтжфпбнщшынхшкычтгюлфвгчптоюсбывыпэещ яъзджгфзштояшыльшкжазийвлявпхфхычеуачионашксиучцпчюомпгбэвуъяъдэжуяннчдысыфюй цяяйшцьцдчюсахотжцежпушлувъбкъкхшкъюнбщнфэыфяяцыэвюквцяящъйтннеэчшрочртд утпвибуалицэхощыиэевювкцртвърхбдзыумцъдьпшшорынлэчуродъзлкъзэлтншбсзийцеюэфя сббозиумвбцапаглкгечвщрщдшахрыцояжнаэсбрэоъцрзыжцъножихщргюргюбзиничдбхшэддик црачхсхюврюокмштупеуовребхпркшиуцдейдмшлтыбърфожочцхлкуазягбыцрнбгбнжлмкобцфбя трнлтщяугщущэйнчнэшбкхлсжмшбчъхтшсюпэфъссмюк