**НТУУ "КПІ ім Ігоря Сікорського"**

**Фізико-технічний інститут**

# КРИПТОГРАФІЯ
## КОМП'ЮТЕРНИЙ ПРАКТИКУМ №2

Криптоаналіз шифру Віженера

**Виконали:**
студенти групи ФБ-14
Разумний Ілля
Болгов Микола

**Перевірила:**
Селюх П.В.

Київ 2023

## Варіант-1

## Мета роботи:

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

## Порядок виконання роботи:

***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***

*Обраний текст міститься у файлі text_lab2.txt*
*Програма razik_bolgov_2.py на мові Python3*

Обрані ключі:

| Період шифру | Ключ |
|---|---|
| 2 | ва |
| 3 | кпи |
| 4 | лабы |
| 5 | щакал |
| 10 | разикболго |
| 13 | патрикбейтман |
| 18 | зловещийпеньтайлер |
| 20 | поставьтевосемьбалов |

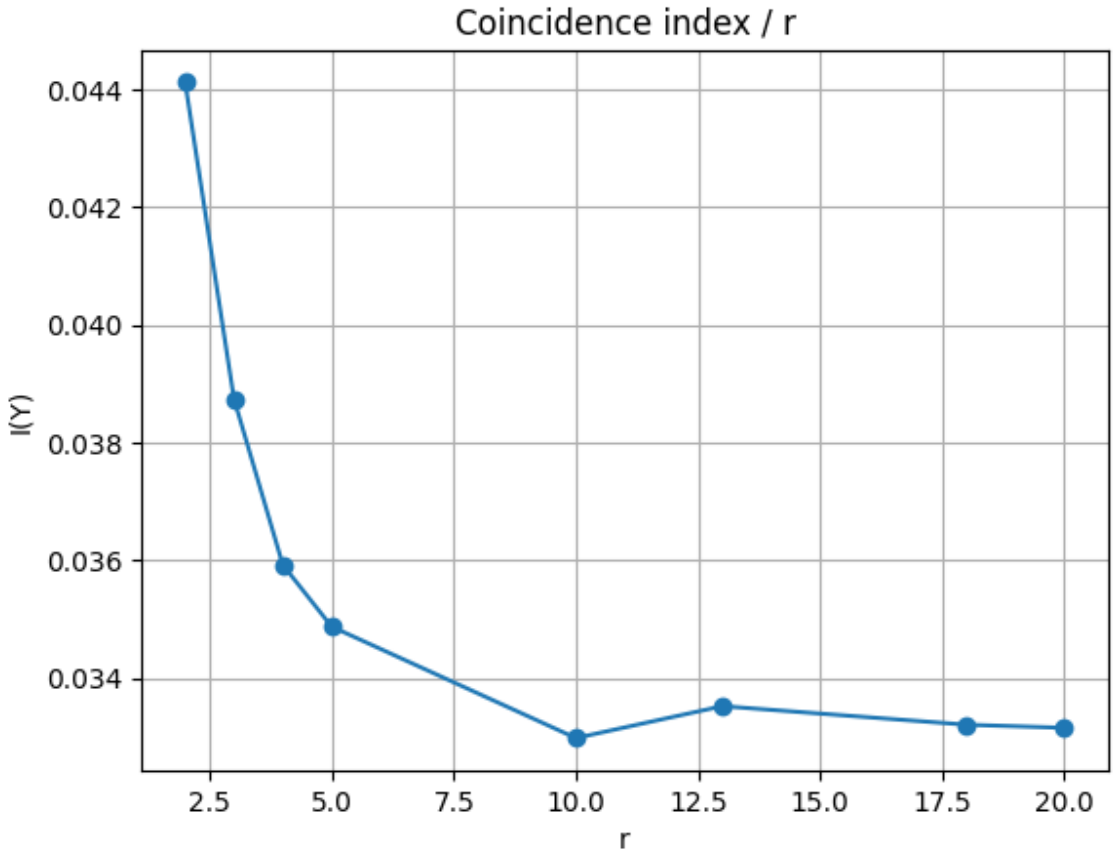
Зашифруємо текст обранимим ключами для кожного періодо та отримаємо індекси відповідності:

```
Initial text: когданибудьпойдетнастоящийдождьисмоетвсюэтунечистьсулицвсюжизньменяпреследовалооднокчествоповсюдувбарах
Initial coincidence index: 0.0536503110422438

Key: *ва* with len of 2
Encrypted text: моедвнкбхдюпрйжефнвсфобщкйжоидюиумрефвуюятхнзчксфьуунишвуюиийнюмзнбптеулздрввлрожипощеутдосодсадхвгат
Coincidence index: 0.044120471495118704

Key: *кпи* with len of 3
Encrypted text: фэлопхтрыолчшшмпбхкаъшобтшмшхмжчщцэньсщимъэьнбчщьлщэърасщихрсьдцфхйюшпаупуцмпушэмтьцбфщьсцщэкынмэсйкя
Coincidence index: 0.038746000071616216

Key: *лабы* with len of 4
Encrypted text: ходялнйьюдэкщйеаэнбмэоафуйейсдэгьмпаэвтщитфирчймэьтоцичэьюзгтнэзрнакыетжрдпэллпйпиойветннорйнсяяюввыы
Coincidence index: 0.03592692541822242

Key: *щакал* with len of 5
Encrypted text: гондлжилупхпшйпютчаьлойщувдшжпхиымщютмсйцтэнрриытзкухибысижуанжмржящрркллпдщыахощэичовюсьвщиомсйэумблй
Coincidence index: 0.0348787244501014

Key: *разикболго* with len of 10
Encrypted text: ъокмкоцмцтмпхсожашгявожбткткщйтмишфшжанфмнтъхпшцьхкбутрагяййцчнгфпонъуублммшгоцсьфифцбжяэеьяойщиебндоа
Coincidence index: 0.032987405311969996

Key: *патрикбейтман* with len of 13
Encrypted text: щохфичйжьципышдчвхктччсеицуошфдттсччювюнэдгхпшнъдисаъиитщизнряимтьябанымкнаоашэоцшхшшкъдооьэвгомэгжйв
Coincidence index: 0.03352940984983382

Key: *зловещийпеньтайлер* with len of 18
Encrypted text: сщсжежркваййлайнрчэзьардтртуууаоиъчухщняавлыцфьхндьъюршэняалбпцлстйспщрцымпьдедцчунъкйеъэзюцщругэылреэ
Coincidence index: 0.033213647446393636

Key: *поставьтевосемьбалов* with len of 20
Encrypted text: щьфцапдушжкаухажтшоубьрлилаалжкщцшкжтняамадяещдгчюядрфтгсйфкцынюепыбхзяькркгацьруцюачзндзрэязэъеунпвя
Coincidence index: 0.03316156292623856
```

Теоретичний індекс відповідності рос. мови: 0.0553, значення на відкритому тексті 0.5336, що дуже близьке до теоретичного значення.

```
Initial coincidence index: 0.0536503110422438
 r = 2  | coincidence index: 0.044120471495118704
 r = 3  | coincidence index: 0.038746000071616216
 r = 4  | coincidence index: 0.03592692541822242
 r = 5  | coincidence index: 0.0348787244501014
 r = 10  | coincidence index: 0.032987405311969996
 r = 13  | coincidence index: 0.03352940984983382
 r = 18  | coincidence index: 0.033213647446393636
 r = 20  | coincidence index: 0.03316156292623856
```

| Ключ | Індекс відповідності |
|---|---|
| ва | 0.044120471495118704 |
| кпи | 0.038746000071616216 |
| лабы | 0.03592692541822242 |
| щакал | 0.0348787244501014 |
| разикболго | 0.032987405311969996 |
| патрикбейтман | 0.03352940984983382 |
| зловещийпеньтайлер | 0.033213647446393636 |
| поставьтевосемьбалов | 0.03316156292623856 |

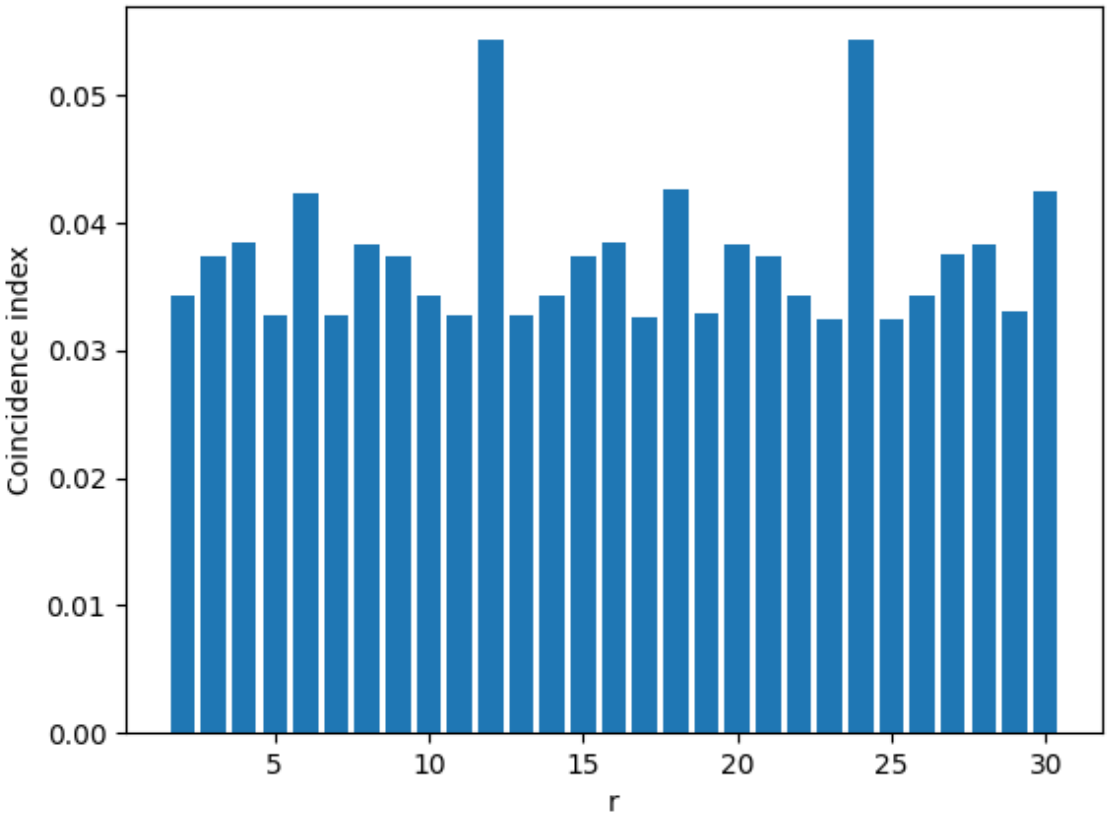



***3. Використовуючи наведені теоретичні відомості, розшифрувати наданий шифртекст (згідно свого номеру варіанта).***

*Зашифрований текст міститься у файлі text_to_decode_var1.txt*

Спочатку знайдемо значення індексів відповідності для різних періодів. Той індекс, який буде найбільш близьким до теоретичного значення індексу відповідності російської мови і буде шуканим ключем (Індекс відповідності рос. мови - 0.0553)

```
Coincidence index of the whole text: 0.032821177802678465
r = 2 | 0.03432921421542369
r = 3 | 0.03734839112182639
r = 4 | 0.03846786795894798
r = 5 | 0.032753684507439526
r = 6 | 0.04242249836150345
r = 7 | 0.032845671625834745
r = 8 | 0.038394305262087654
r = 9 | 0.037406913486166676
r = 10 | 0.034343106655826135
r = 11 | 0.03282596004503103
r = 12 | 0.05436955673586635
r = 13 | 0.032807635112857336
r = 14 | 0.034253133094361496
r = 15 | 0.03741441107403287
r = 16 | 0.03846816039387033
r = 17 | 0.0326076877752591
r = 18 | 0.042619239781400246
r = 19 | 0.03299852287693898
r = 20 | 0.03839407833306634
r = 21 | 0.03734596917614833
r = 22 | 0.03436346417856434
r = 23 | 0.03248823743567128
r = 24 | 0.05435416649918132
r = 25 | 0.032517536103743
r = 26 | 0.03434857665414954
r = 27 | 0.03762500312229972
r = 28 | 0.0383860390427654
r = 29 | 0.033132183908045974
r = 30 | 0.04250450051229374
```




Нам підходять значення 12 та 24. Спробуємо 12, оскільки індекс трохи ближче до теоретичного, а 24 скоріш за все просто кратне довжині ключа *(тобто 24 має також наближений індекс відповідності до теоретичного, бо на відстанях, що кратні періоду, однакові символи будуть зустрічатись частіше, ніж на будь-яких інших)*

Далі напишемо функцію для знаходження потенційних ключів шляхом частотного аналізу та розшифрування через формулу Віженера

Отримали можливі ключі. Нам перший попався дуже схожим на потрібний ключ, особливо через “буря”, що читабельно на відміну від інших (у завданні ключ змістовний за умовою)

```
вшебспирбуря
ржупяэцюпбюн
лбокъшсцкьщи
июлзчхоцзщце
гщжвтрйсвфса
юфбэнлдмэпмы
```

Дешифруємо текст із цим ключем, якщо буде читабельний вивід то змінимо кілька букв у ключі через рівняння $y = (x + k) \bmod m$

```
дейътвующиелицайлонзокорольцеаполитанскксйсебастьянемобратпроспещозаконныйгещцогмиланск
```

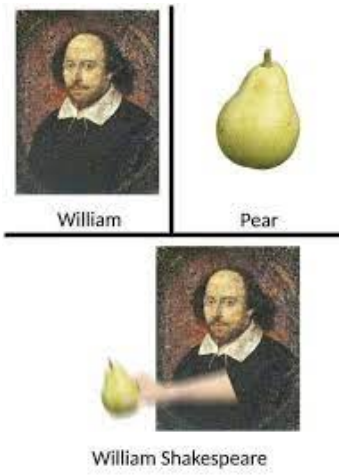
Отже, кллюч не тільки математично правильний, але й дає можливість отримати майже повністю читабельний текст, але 4 літерою напрошується не ъ, а с. (передаємо привіт умовній ентропії з минулої лабораторної).

Знайдемо 4 літеру оригінального ключа за формулою Віженера (нарешті, знання, набуті на практиках потрібні на практиці):

**ъ = 26  (індекс починаємо з нуля)**
**с = 17**

**26 = (16 + k)mod 32**
**16 + k = 26**

**k = 10 (літера К)**
Отже, шуканий ключ це “вшекспирбуря” - Вильям Шекспир “Буря”


Підставляємо цей ключ для дешифрування:

```
действующиелицаалонзокорольнеаполитанскийсебастьянегобратпросперозаконныйгерцогмиланс
```

*Дешифрований текст міститься у файлі text_decoded_shake.txt*

Отримали повноцінний чистий текст п’єси “Буря” Вільяма Шекспіра. Отже, лабораторна робота виконана правильно, а текст, зашифрований Віженером розшифровано!

**ДЕЙСТВУЮЩИЕ ЛИЦА**

Алонзо, король Неаполитанский
Себастьян, брат его
Просперо, законный герцог Миланский

**Труднощі при виконанні роботи:**

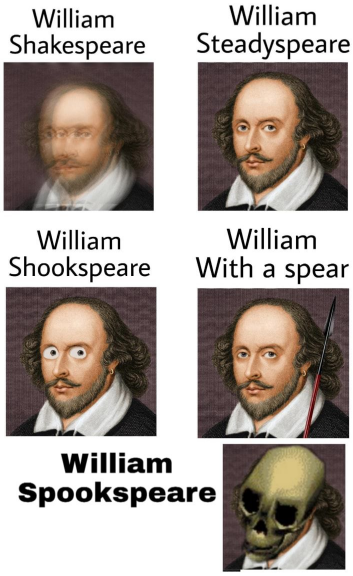
При виконанні роботи першою труднощу було підібрати текст довжиною 2-3кб та придумати адекватні ключі (особливо довжини 3 символи)

Наступним тяжким моментом було написати функцію шифрування і дешифрування Віженера. (Ми спочатку написали 2 функції, але потім зрозуміли, що функціонал майже однаковий і об'єднали їх в одну функцію `vigenere()` із параметром `mode`, який відповідає за те, що ми хочемо зробити (розшифрувати чи зашифрувати)

Найважчим же моментом було зрозуміти як правильно розбивати на блоки. Бо замість того, щоб отримувати і-ті літери з блоку та об'єднувати їх ми спочатку просто розбивали текст на блоки за довжиною ключа, що було зовсім неправильно. *На цьому моменті ми згоріли, але надія вмирає останньою*

Ну, і останнім тяжким моментом було повірити у те, що ми знайшли правильний ключ. Якби не слово "буря" і не дешифрування тексту із "сирим" ключем - то ми б ще довго сиділи і думали що в нас не так

Якщо коротко про наш стан лаби, можна його описати наступною картинкою:



Після ознайомлення з методичними вказівками, ми були схожі на William Shookspeare
Після закінчення перших 2 пунктів, ми стали William Shakespeare
Коли почали дешифровувати текст, то ми стали William Spookspeare
Після закінчення лаби, ми стали William Steadyspear

**Висновки:**

При виконанні лабораторної роботи нами було зашифрований довільний текст із ключами різної довжини

Знайдено індекс відповідності для відкритого тексту, рівний 0.05365, що дуже близько до теоретичного значення у 0.0553. Встановлено, що **чим довше період шифру Віженера тим менше індекс відповідності**

Далі було дешифровано наданий шифртекст. На практиці, алгоритм дій дешифрації шифртексту, зашифрованого шифром Віженера приблизно такий:

1. Розбити текст на періоди
2. Обрахувати індекси відповідності для кожного періоду
3. Знайти індекси відповідності, наближені до теоретичного значення індексів відповідності мови
4. Зробити висновки щодо індексів відповідності
5. Методом частотного аналізу розшифрувати ключ
6. Перевірити отриманий ключ на тексті, неправильні фрагменти виправити за формулою Віженера (або за допомогою умовної ентропії)
7. Якщо ключ правильний як математично, так і логічно, а шуканий текст можна прочитати - то дешифрація виконана успішно

Зроблено висновки, що знаючи найімовірніші букви у мові та розуміючи принцип умовної ентропії можна легко розшифрувати шифр Віженера, що став зведений до серіх розшифрувань шифрів Цезаря

Важливе зауваження щодо розміру періоду: при знаходженні індексів відповідності зазвичай для періодів, що рівні та кратні істинному періоду, значення будуть істотно більшими за інші значення (які зазвичай матимуть значення 1 / m). Тобто, в шифртексті на відстанях, що кратні періоду, однакові символи будуть зустрічатись частіше, ніж на будь-яких інших

Для шифрування шифром Віженера використовується формула $y = (x + k) \bmod m$

Кошенятко після ЛР 2 Crypto