Лабораторна робота № 2.

Дослідження шифру Цезаря.

Шифр Цезаря.

Цей шифр реалізує таке перетворення відкритого тексту: кожна літера замінюється третьою після неї літерою алфавіту, який вважається написаним по колу, тобто після «я» йде «а». Зауважимо, що Цезар замінював її третьою літерою, але можна замінювати й будь-якою іншою. Головне, щоб адресат цього повідомлення знав величину й напрямок зсуву. Клас шифрів, до якого належить шифр Цезаря, називається шифрами заміни.

Із цього, напевне, зрозуміло, що створення надійного шифру є задачею непростою. Тому бажано збільшити час життя шифру, але тут зростає ймовірність того, що криптоаналітики супротивника зможуть розкрити шифр і прочитати зашифровані повідомлення. Якщо в шифрі є змінний «ключ», то його заміна приводить до того, що розроблені супротивником методи вже не дадуть ефекту. Під ключем будемо розуміти змінний елемент шифру, який застосовується для шифрування конкретного повідомлення. У шифрі Цезаря ключем є величина і напрямок зсуву літер шифротексту відносно літер відкритого тексту.

Описані міркування зумовили те, що безпека шифрованих повідомлень у першу чергу стала забезпечуватися ключем. Сам шифр, шифромашина або принцип шифрування прийнято вважати відомими суперникові й доступними для попереднього вивчення, але в шифрі з'явився невідомий елемент – ключ, від якого істотно залежать застосовувані перетворення інформації. Тепер користувачі, перш ніж обмінятися шифрованими повідомленнями, повинні обмінятися ключем, за допомогою якого можна прочитати зашифроване повідомлення. А для криптоаналітиків, які хочуть прочитати перехоплене повідомлення, основною задачею є знаходження ключа.

Принципи частотного криптоаналізу.

Встановлено, що в будь-якій мові літери абетки зустрічаються

неодноразово. Якщо взяти достатньо великий текст (близько мільйона символів) загального змісту та підрахувати частоту, з якою кожна літера абетки зустрічається в цьому тексті, ми побачимо, що найчастіше в українських текстах зустрічається літера «О» (0.082), а в російських та англійських - «Е» (0.071 та 0.12 відповідно). Звичайно, в залежності від тематики тексту, частотні характеристики його змінюються, але тенденція залишається незмінною.

На цьому факті ґрунтується метод частотного криптоаналізу. Якщо метод шифрування «перехопленої шифровки» не приховує частотних особливостей мови (а саме таким і є шифр Цезаря), то криптоаналітики виконують такі дії:

1. Підраховують відносні частоти, з якими кожна літера абетки зустрічається в «перехопленому» повідомленні. Робиться це за формулою: *частота = кількість / довжина*; де *кількість* – скільки разів літера зустрічається в повідомленні; *довжина* – кількість літер у повідомленні.

2. Літеру з найбільшою відносною частотою ототожнюють із літерою, яка має найбільшу частоту в частотній таблиці.

3. Визначають величину зсуву.

4. Пробують дешифрувати повідомлення з визначеною в п. 3 величиною зсуву. Якщо отримано логічний зв'язний текст, повідомлення вважається дешифрованим. Якщо зв'язного тексту не отримано, процедуру продовжують.

5. Літеру з найбільшою відносною частотою ототожнюють із літерою, яка має другу найбільшу частоту в таблиці.

6. Пробують дешифрувати повідомлення, перебираючи частотну таблицю, поки не отримують зв'язного тексту.

Наведеним методом Вам необхідно користуватися для криптоаналізу в цій лабораторній роботі.

Приклад.

Необхідно зашифрувати або розшифрувати повідомлення за допомогою ***алгоритму Цезаря***. В задачах відомий ключ, тобто величина зсуву, або

таблиця заміни.

Спробуємо зашифрувати відкритий текст «Організуйте зустріч нового об'єкта в обумовленому місці» алгоритмом Цезаря зі зсувом -2.

Таблиця 2

Таблиця заміни

А	Б	В	Г	Д	Е	Є	Ж	З	И	І	Ї	Й	К	Л	М
00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15
Н	О	П	Р	С	Т	У	Ф	Х	Ц	Ч	Ш	Щ	Ь	Ю	Я
16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31

Отримаємо криптограму «15 17 01 30 14 08 06 20 10 19 03 06 20 18 19 17 08 24 14 15 00 15 01 15 15 31 04 11 19 20 00 15 31 20 13 15 00 12 03 14 15 13 20 13 08 18 23 08».

Для дешифрування методом частотного криптоаналізу отриману «шифровку», в якій усього 48 знаків, піддають такому аналізу: найчастіше в ній зустрічається число 15 – 8 разів (імовірність тоді можна обчислити так: 8/48=0,167). Порівнявши з частотною таблицею української мови, можна припустити, що 15=«О». Тоді отримуємо таку таблицю заміни (табл.3).

Таблиця 3

Таблиця заміни після дешифрування

А	Б	В	Г	Д	Е	Є	Ж	З	И	І	Ї	Й	К	Л	М
30	31	00	01	02	03	04	05	06	07	08	09	10	11	12	13
Н	О	П	Р	С	Т	У	Ф	Х	Ц	Ч	Ш	Щ	Ь	Ю	Я
14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29

Замінивши числа літерами, отримаємо відкритий текст.

Звичайно, для реального застосування частотного криптоаналізу такої малої кількості відкритого тексту, як правило, замало, і мова може йти лише про спеціально підібрані фрази, що відповідають статистичним особливостям української мови.

Робота з довільними короткими фразами методами частотного криптоаналізу може й не дати бажаного результату.

Завдання роботи.

1. Народне прислів’я українською мовою, обране для лаб. работи № 1, зашифрувати методом Цезаря. Величина зсуву відповідає номеру студента у списку підгрупи.
2. Розшифрувати повідомлення, використовуючи частотні таблиці української мови.
3. Розшифрувати повідомлення, використовуючи частотну таблицю української мови (табл. 4): «25 26 27 09 10 21 26 14 31 10 26 07 10 24 26 22 08 22 18 27 24 25 27 07 14 07 31 05 03 26 02 18 24 14 23 26 22 08 24 05 28 15 03». Потужність алфавіту – 32 символи.

Таблиця 4

Частотна таблиця літер української мови

О	0,082	Р	0,038	У	0,028	Б	0,010	Є	0,006
Н	0,070	І	0,037	П	0,025	Х	0,010	Ф	0,005
А	0,070	С	0,036	Я	0,021	Ц	0,009	Ш	0,005
И	0,056	К	0,036	З	0,019	Ю	0,009	Щ	0,003
Т	0,051	М	0,033	Ь	0,015	Ж	0,008	Ґ	0,000
В	0,046	Д	0,028	Г	0,013	Й	0,007		
Е	0,043	Л	0,028	Ч	0,011	Ї	0,006		