1. Цель работы

Реализовать атаку по известному шифртексту при помощи частотного анализа на полиалфавитный шифр подстановки.

2. Описание алгоритма

2.1. Шифр Виженера

Шифр Виженера — метод полиалфавитного шифрования буквенного текста с использованием ключевого слова.

В шифре Цезаря каждая буква алфавита сдвигается на несколько позиций в зависимости от ключа. Так, например, при сдвиге на 4, А стало бы Е.

Шифр Виженера состоит из нескольких шифров Цезаря с различными значениями сдвига. Для шифрования может использоваться таблица алфавитов, называемая квадрат (или таблица) Виженера. Например, для английского алфавита таблица Виженера составляется из строк по 26 символов (рис. 1), первой строчкой которой является алфавит, а каждая следующая строка получается из предыдущей путем циклического сдвига вправо на одну позицию. Таким образом получается 26 различных шифров Цезаря.

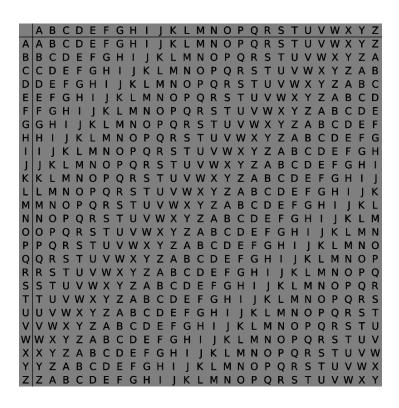


Рис. 1 Таблица Виженера для английского алфавита

Человек, посылающий сообщение, циклически записывает ключевое слово до тех пор, пока его длина не будет соответствовать длине исходного текста.

Пример:

Ключевое слово: LEMON

Исходный текст: ATTACKATDAWN

Циклический ключ: LEMONLEMONLE

Зашифрованный текст: LXFOPVEFRBHR

Первый символ исходного текста А зашифрован последовательностью L, которая является первым символом ключа. Первый символ L зашифрованного текста находится на пересечении строки L и столбца А в таблице Виженера (рис. 1). Аналогично для остальных символов.

Расшифровывание производится следующим образом: находим в таблице Виженера строку, соответствующую і-му символу ключевого слова, в данной строке находим і-ый символ зашифрованного текста. Столбец, в котором находится данный символ, соответствует і-му символу исходного текста.

Таким образом, если алфавиту сопоставить числа 0-N, то шифрование можно записать в виде формулы:

$$C_i \equiv (M_i + K_i) mod N \tag{2.1}$$

Тогда расшифровывание представляется в виде:

$$M_i \equiv (C_i - K_i + N) modN \tag{2.2}$$

2.2. Индекс совпадений

Основной задачей для взлома шифра Виженера является поиск длины ключа. Индекс совпадений — один из методов криптоанализа шифра Виженера.

Рассмотрим текст, написанный на некотором языке, алфавит которого состоит из N букв. Рассмотрим достаточно длинную последовательность \vec{x} из nбукв. Если f_i задает количество i-ой буквы алфавита в строке \vec{x} , то можно определить индекс совпадений:

$$I(\vec{x}) = \sum_{i} f_{i} \frac{f_{i} - 1}{n(n-1)}$$
(2.3)

Так, если текст написан на естественном языке, является достаточно длинным, то для него можно получить ожидаемый индекс совпадений (например, для английского алфавита индекс совпадения равен примерно 0,065).

Для того, чтобы определить длину ключа, необходимо последовательно записывать в отдельную строку каждую k-ую (k = 2, 3, 4, ...) букву и посчитать индекс совпадения для этой строки. Число k, для которого индекс совпадения имеет пиковый характер, будет являться длиной ключа.

2.3. Частотный анализ

Как только длина ключа становится известной, зашифрованный текст можно записать во множество столбцов, каждый из которых соответствует одному символу ключа. Каждый столбец состоит из исходного текста, который зашифрован шифром Цезаря.

Самым простым способом, чтобы определить, каким ключом был зашифрован каждый столбец, является (не всегда такой способ будет успешным):

- посчитать частоту встречаемости каждого символа алфавита в открытом тексте;
- посчитать частоту встречаемости каждого символа алфавита в шифртексте;
- сопоставить максимумы разность их индексов и будет ключом.

3. Описание реализации

3.1. Класс Vizhener

Является основным классом. Имеет методы для шифрования, дешифрования, генерации произвольного ключа. Также присутствует метод для совершения атаки на зашифрованный текст шифром Виженера. Размер используемого алфавита — 256 символов (таблица ASCII).

```
public static void encoder(File text, File encoded, String key);
```

Принимает на вход файл с открытым текстом, файл, куда записать зашифрованный текст, и ключ. Реализована формула (2.1).

```
for (int i = 0; i < buffer.length(); i++) {
     buffer[i] = (buffer[i] + key.charAt(i % key.length())
% ASCII_TABLE_SIZE;
}</pre>
```

public static void decoder(File encoded, File decoded, String key);

Принимает на вход файл с зашифрованным текстом, файл, куда записать расшифрованный текст, и ключ. Реализована формула (2.2).

```
for (int i = 0; i < buffer.length(); i++) {
     buffer[i] = (buffer[i] - key.charAt(i % key.length()
+ ASCII_TABLE_SIZE) % ASCII_TABLE_SIZE;
}</pre>
```

public static String keyGen(int keyLength);

Генерирует и возвращает ключ длиной, указанной в качестве параметра.

public static String doAttack(File statistics, File encoded);

Метод совершает атаку на зашифрованный файл. В параметрах указывается файл с открытым текстом для статистики и файл с зашифрованным текстом. Здесь запускается метод int keyLength = IndexMetod.run(encoded). Далее собирается статистика для открытого текста достаточно большого объема методом int[] statisticsArray = LetterFrequency.getStatistics(statistics). Полученные статистика открытого текста и возможный ключ, а также зашифрованный файл передаются методу String key = KeyAnalysis.run(statisticsArray, encoded, keyLength), который

3.2. Класс IndexMetod

после завершения вернет ключ.

Класс предназначен для поиска длины ключа, которым было зашифровано сообщение, с помощью метода индексов.

public static int run(File text);

Метод считывает текст из файла в буффер. Далее составляет последовательность \vec{x} длиной $n=\frac{N}{i}$, N — размербуффера, добавляя в нее каждую і-ую букву текста для $2 \le i \le K_{max}$, где K_{max} — максимальная длина ключа. Считает частоты появления каждого символа в последовательности методом toCount класса LetterFrequency, а затем вычисляет индекс совпадения по формуле (2.3). Возвращаемым значением будет являться первое пиковое значение. Достигается это следующим способом:

```
if (Math.abs(currentIndex - maxIndex) > INDEX_FLAG) {
    maxIndex = index;
    keyLength = i;
}
```

INDEX_FLAG — значение, незначительно меньше индекса совпадения для открытого текста того же алфавита (для английского алфавита примерно равен 0,065). Таким образом

будет выбран первое пиковое значение, т. к. между собой пиковые значения отличаются незначительно.

3.3. Класс KeyAnalisis

Класс предназначен для поиска ключа.

public static String run(int[] statistics, File encoded, int keyLength);

Выполняет поиск ключа. Создает keyLenght массивов, в каждый из которых добавляет і-ый по модулю keyLength символ зашифрованного текста. Теперь в каждом из этих массивов получается зашифрованный шифром Цезаря текст. Считает частоты встречаемости символов в каждом из этих массивов отдельно. Сортирует частоты по убыванию. Массив со статистикой открытого текста также отсортирован. Далее считающая разности номеров между каждыми і-ыми элементами в массивах частот. Самая часто встречаемая разность определяется как элемент ключа.

3.4. Класс LetterFrequency

Класс предназначен для подсчета частот символов.

public static int[] toCount(char[] x);

Получает на вход последовательность символов, возвращает массив с частотой встречаемости каждого символа.

public static int[] getStatistics(File file)

Получает на вход файл с текстом, возвращает массив с частотой встречаемости каждого символа.

4. Примеры работы программы

Сгенерируем случайный ключ длиной 50:

yW2ngnA?xOjyi?vCQNpSZkrT0^IvWhS6o@9LRIvOXXH@iUTd=3

WHEN MORNING CAME, AND THEY COULD SPEAK MORE CALMLY ON THE SUBJECT OF THEIR GRIEF, THEY HEARD HOW HER LIFE HAD CLOSED.

SHE HAD BEEN DEAD TWO DAYS. THEY WERE ALL ABOUT HER AT THE TIME, KNOWING THAT THE END WAS DRAWING ON. SHE DIED SOON AFTER DAYBREAK. THEY HAD READ AND TALKED TO HER IN THE EARLIER PORTION OF THE NIGHT, BUT AS THE HOURS CREPT ON, SHE SUNK TO SLEEP. THEY COULD TELL, BY WHAT SHE FAINTLY UTTERED IN HER DREAMS, THAT THEY WERE OF HER JOURNEYINGS WITH THE OLD MAN; THEY WERE OF NO PAINFUL SCENES, BUT OF PEOPLE WHO HAD HELPED AND USED THEM KINDLY, FOR SHE OFTEN SAID GOD BLESS YOU! WITH GREAT FERVOUR. WAKING, SHE NEVER WANDERED IN HER MIND BUT ONCE, AND THAT WAS OF BEAUTIFUL MUSIC WHICH SHE SAID WAS IN THE AIR. GOD KNOWS. IT MAY HAVE BEEN.

DPENING HER EYES AT LAST, FROM A VERY QUIET SLEEP, SHE BEGGED THAT THEY WOULD KISS HER ONCE AGAIN. THAT DONE, SHE TURNED TO THE OLD MAN WITH A LOVELY SMILE UPON HER FACESUCH, THEY SAID, AS THEY HAD NEVER SEEN, AND NEVER COULD FORGETAND CLUNG WITH BOTH HER ARMS ABOUT HIS NECK. THEY DID NOT KNOW THAT SHE WAS DEAD, AT FIRST.

SHE HAD SPOKEN VERY OFTEN OF THE TWO SISTERS, WHO, SHE SAID, WERE LIKE DEAR FRIENDS TO HER. SHE WISHED THEY COULD BE TOLD HOW MUCH SHE THOUGHT ABOUT THEM, AND HOW SHE HAD WATCHED THEM AS THEY WALKED TOGETHER, BY THE RIVER SIDE AT NIGHT. SHE WOULD LIKE TO SEE POOR KIT, SHE HAD OFTEN SAID OF LATE. SHE WISHED THERE WAS SOMEBODY TO TAKE HER LOVE TO KIT. AND, EVEN THEN, SHE NEVER THOUGHT OR SPOKE ABOUT HIM, BUT WITH SOMETHING OF HER OLD, CLEAR, MERRY LAUGH.

FOR THE REST, SHE HAD NEVER MURMURED OR COMPLAINED; BUT WITH A QUIET MIND, AND MANNER QUITE UNALTEREDSAVE THAT SHE EVERY DAY BECAME MORE EARNEST AND MORE GRATEFUL TO THEMFADED LIKE THE LIGHT UPON A SUMMERS EVENING.

THE CHILD WHO HAD BEEN HER LITTLE FRIEND CAME THERE, ALMOST AS SOON AS IT WAS DAY, WITH AN OFFERING OF DRIED FLOWERS WHICH HE BEGGED THEM TO LAY UNDER SHEATS. IT WAS HE WHO HAD COME TO THE WINDOW OVERNIGHT AND SPOKEN TO THE SEXTON, AND THEY SAW IN THE SHOW THE THE HOUGHT.

HE TOLD THEM OF AND HAD FELT GLAD TO BE SO NEAR HIM. THEY LET HIM HAVE HIS WISH; AND INDEED HE KEPT HIS WORD, AND WAS, IN HIS CHILDISH WAY, A LESSON TO THEM ALL.

UP TO THAT TIME, THE OLD MAN HAD NOT SPOKEN ONCEEXCEPT TO HEROR STIRRED FROM THE BEDSIDE. BUT, WHEN HE SAW HER LITTLE FAVOURITE, HE WAS MOVED AS THEY HAD NOT SEEN HIM YET, AND MADE AS THOUGH HE WOULD HAVE HIM COME NEARER. THEN, POINTING TO THE BED, HE BURST INTO TEARS FOR THE FIRST TIME, AND THEY WHO STOOD BY, KNOWING THAT THE SIGHT OF THIS CHILD HAD DONE HIM GOOD, LEFT THEM ALONE TOGETHER.

SOOTHING HIM WITH HIS ARTLESS TALK OF HER, THE CHILD PERSUADED HIM TO TAKE SOME REST, TO WALK ABROAD, TO DO ALMOST AS HE DESIRED HIM. AND WHEN THE DAY CAME ON, WHICH MUST REMOVE HER IN HER EARTHLY SHAPE FROM EARTHLY EYES FOR EVER, HE LED HIM AWAY, THAT HE MIGHT NOT KNOW WHEN SHE WAS TAKEN EDON HIM. THE DAY CAME ON, WHICH MUST REMOVE HER IN HER EARTHLY SHAPE FROM EARTHLY EARTHLY EARTHLY HERE TO GATHER FRESH LEAVES AND BERRIES FOR HER BED. IT WAS SUNDAYA BRIGHT, CLEAR, WINTRY AFTERNOONAND AS THEY TRAVERSED THE VILLAGE STREET, THOSE WHO WERE WALKING IN THEIR PATH DREW BACK TO MAKE WAY FOR THEM, AND GAVE THEM A SOFTENED GREETING. SOME SHOOK THE OLD MAN KINDLY BY THE HAND, SOME STOOD UNCOVERED WHILE HE TOTTERED BY, AND MANY CRIED GOD HELP HIM! AS HE PASSED ALONG.
NEIGHBOUR! SAID THE OLD MAN, STOPPING AT THE COTTAGE WHERE HIS YOUNG GUIDES MOTHER DWELT, HOW IS IT THAT THE FOLKS ARE NEARLY ALL IN BLACK TODAY? I HAVE SEEN A MOURNING RIBBON OR A PIECE OF CRAPE ON ALMOST EVERY ONE.
SHE COULD NOT TELL, THE WOMAN SAID. WHY, YOU YOURSELFYOU WEAR THE COLOUR TOO? HE SAID. WINDOWS ARE CLOSED THAT NEVER USED TO BE BY DAY. WHAT DOES THIS MEAN?
AGAIN THE WOMAN SAID SHE COULD NOT TELL.
WE MUST GO BACK, SAID THE OLD MAN, HURRIEDLY. WE MUST SEE WHAT THIS IS.

Pug 1 Confidence of the Man and the Colour DATA MURICIPALITY OF THE MODE OF THE MAN AND THE OLD MAN, HURRIEDLY. WE MUST SEE WHAT THIS IS.

Рис. 1 Сообщение для шифрования

Oxogenia

Alimatina vanidatia milatia alianatia alianatia alianatia alianata alianat

Рис.2 Зашифрованное сообщение

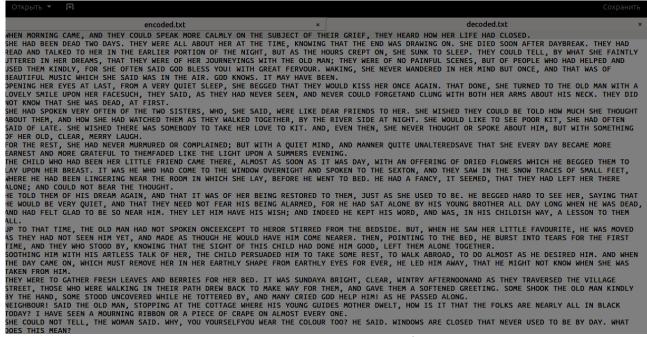


Рис.3 Сообщение после атаки на шифр Виженера



Рис.4 Ключ, полученный в результате атаки

Как видим, полученный ключ в результате атаки, совпадает с действительным ключом.