Лабораторная работа № 1

Основы шифрования данных

Цель работы: изучение основных принципов шифрования информации, знакомство с широко известными алгоритмами шифрования, приобретение навыков их программной реализации.

Основные сведения

Проблема скрытия информации была актуальна для человечества еще в древние времена, письменные свидетельства об использовании простейших методов шифрования встречаются еще у древних греков в V-VI в. до н.э. Развитие цивилизации, появление новых средств коммуникации (сначала письменных, затем электронных) предъявляли к методам шифрования все более жесткие требования, что вылилось в появление отдельной науки, занимающейся защитой информации путем ее преобразования - *криптологии*. У этой науки два направления: криптография и криптоанализ.

Криптография – наука о методах и средствах преобразования информации в вид, затрудняющий или делающий невозможным несанкционированные операции с нею (чтение и/или модификацию).

Базовым понятием криптографии является шифр**.** *Шифр* – совокупность инъективных (обратимых) преобразований множества элементов открытого текста на множество элементов шифротекста, проиндексированных элементами из множества ключей:

$$\{ F_k: X \to S, k \in K \},$$

где $X \in \mathbf{X}$ **–** кодируемое сообщение из множества открытых текстов;

$S \in \mathbf{S}$ **–** шифротекст из множества возможных закодированных текстов;

k – ключ шифрования;

F – отображение, выполняемое шифром.

Свойство инъективности шифра означает, что существует отображение F^{-1} такое, что

$\{ F_k^{-1}: X \to S, k \in K \}$

Процесс преобразования открытого текста (передаваемого сообщения) в шифротекст называется *шифрованием.* Обратное преобразование шифротекста в открытый текст называется *дешифрованием*.

Криптоанализ - наука (и практика ее применения) о методах и способах вскрытия шифров. Под вскрытием понимается задача получения по известному шифротексту соответствующего открытого текста и/или ключа шифрования.

К настоящему времени изобретено большое количество разнообразных шифров. Некоторые из них используются в настоящее время для практических целей защиты данных в различных информационных системах, а некоторые представляют лишь исторический интерес, поскольку используемые в них преобразования не обеспечивают должной стойкости от вскрытия для современного уровня вычислительных мощностей ЭВМ. Однако знакомство с такими устаревшими криптоалгоритмами имеет несомненную пользу, поскольку позволяет проследить эволюцию криптоалгоритмов, кристаллизацию в них современных принципов шифрования, оценить слабые и сильные стороны тех или иных криптопреобразований.

Шифр Цезаря

Одним из первых документально зафиксированных шифров является *шифр Цезаря*, использовавшийся известным полководцем в собственной переписке. В шифре каждая буква замещается на букву, находящуюся k символами правее в алфавите по модулю, равному количеству букв в алфавите:

$C_k(j)=(j+k)(\mathrm{mod}\ n)$, где j - порядковый номер буквы в алфавите, $C_k(j)$ – порядковый номер замещающей буквы, n – мощность входного алфавита (количество букв в используемом алфавите).

Таким образом, ключом шифрования здесь является число k, определяющее размер смещения.

Очевидно, что обратной подстановкой является

$C_k^{-1}(j)=C_{n-k}=(j+n-k)(\text{mod } n)$

При необходимости алфавит можно расширить знаками препинания, заглавными буквами, цифрами, чтобы шифр мог обрабатывать все символы исходного текста.

Общее количество допустимых ключей равно n, причем один из ключей преобразует текст в самого себя. Столь небольшое количество ключей позволяло осуществлять простой криптоанализ шифротекста еще во времена Цезаря *методом полосок*, когда несколько полосок с записанным на них алфавитом располагали рядом так, чтобы по горизонтали получился фрагмент шифротекста. После этого полоски синхронно сдвигались вверх или вниз до тех пор, пока по горизонтали на месте шифротекста не окажется некоторая читаемая фраза. Величина смещения полосок однозначно определяет ключ k.

Естественным развитием шифра Цезаря стал шифр Виженера.

Пример: Шифрование с использованием ключа k=3.

Буква «С» «сдвигается» на три буквы вперед и становится буквой «Ф». Твердый знак, перемещённый на три буквы вперед, становится буквой «Э», и так далее:

Исходный алфавит: АБВГДЕЁЖЗИЙКЛМНОПРСТУФХЦЧШЩЪЫЬЭЮЯ

Шифрованный: ГДЕЁЖЗИЙКЛМНОПРСТУФХЦЧШЩЪЫЬЭЮЯАБВ

Оригинальный текст:

"Съешь же ещё этих мягких французских булок, да выпей чаю."

Шифрованный текст получается путём замены каждой буквы оригинального текста соответствующей буквой шифрованного алфавита:

"Фэзыя йз зьи ахлш пвёнлш чугрщцкфнлш дцосн, жг еютзм ъгб."

Шифр Цезаря с ключевым словом

В данной разновидности шифра Цезаря ключ задается числом k ($0<=k<=n-1$) и коротким ключевым словом или предложением. Выписывается алфавит, а под ним, начиная с k-й позиции, ключевое слово. Оставшиеся буквы записываются в

алфавитном порядке после ключевого слова (избегая повтора букв). В итоге мы получаем подстановку для каждой буквы. Требование, чтобы все буквы ключевого слова были различными не обязательно, необходимо только записывать ключевое слово без повторения одинаковых букв.

Пример

Пусть задан ключ K=3, ключевое слово «ШИФРОВКА» и русский алфавит из 32 букв. Необходимо создать таблицу замен для системы шифрования Цезаря с ключевым словом и с ее помощью зашифровать слово «НЕПТУН».

Первую букву ключевого слова («Ш») записываем под символом «Г» открытого текста с числовым кодом, определенным ключом K=3. Остальные буквы слова «ШИФРОВКА» записываем подряд. Оставшиеся ячейки заполняем теми буквами алфавита, которые не вошли в ключевое слово: «Б», «Г», «Д», «Е» и т.д. до буквы «Ь». Оставшиеся буквы «Э», «Ю», «Я» вписываем в начало таблицы под буквами «А», «Б» и «В», соответственно (рис. 1).

код	0	1	2	**3**	4	5	6	7	8	9	10	11	12	13	14	15
исх. текст	А	Б	В	Г	Д	Е	Ж	З	И	Й	К	Л	М	Н	О	П
шифртекст	Э	Ю	Я	**Ш**	**И**	**Ф**	**Р**	**О**	**В**	**К**	**А**	Б	Г	Д	Е	Ж
код	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
исх. текст	Р	С	Т	У	Ф	Х	Ц	Ч	Ш	Щ	Ъ	Ы	Ь	Э	Ю	Я
шифртекст	З	Й	Л	М	Н	П	С	Т	У	Х	Ц	Ч	Щ	Ъ	Ы	Ь

Рис. 1. Таблица замен символов для системы шифрования Цезаря при K=3, M=32 и ключевом слове «ШИФРОВКА»

Далее с помощью полученной таблицы шифруем побуквенно слово «НЕПТУН». В результате получаем шифртекст: «ДФЖЛМД».

Количество ключей в системе Цезаря с ключевым словом равно $n!$. Для дешифрации необходимо с использованием известного ключа шифрования определить соответствие исходного и заменяющего алфавита и выполнить обратную подстановку.

Аффинная криптосистема

Обобщением системы Цезаря является аффинная криптосистема. Она определяется двумя числами a и b, где $0<=a,b<=n-1$, n - как и раньше, является мощностью алфавита. Числа a и n должны быть взаимно просты.

Соответствующими заменами являются:

$A_{a,b}(j)=(a* j+b)(\text{mod } n)$

$A^{-1}_{a,b}(\text{j})=(j-b+n)*a^{-1}(\text{mod } n)$

Поиск мультипликативно-обратного (обозначенного как a^{-1}) осуществлять по алгоритму Евклида. Очевидно, что при a=1 аффинная криптосистема вырождается в шифр Цезаря.

Зачем a и n должны быть взаимно простыми? Если это условие нарушено, будет ситуация, когда разные символы открытого текста отображаются в один и тот же символ шифрованного. Ясно, что расшифрование полученного текста может быть неоднозначным. Рассмотрим английский алфавит. Допустим, a = 2, b = 3. Тогда результат применения преобразования (2X + 3) mod 26 будет одинаков для символов, которые отстоят в алфавите друг от друга на 13 позиций.

Количество ключей аффинной криптосистемы зависит от мощности используемого алфавита. Для алфавита русского языка коэффициент b может принимать 33 значения, коэффициент a – только значения, взаимно простые с числом 33: 1, 2, 4, 5, 7, 8, 10, 13, 14, 16, 17, 19, 20, 23, 25, 26, 28, 29, 31, 32. Итого 20 значений. Все возможные сочетания допустимых b и a дают ключевое пространство 33*20=660-1=659 ключей (1 ключ вычитается, поскольку сочетание a=1 и b=33 преобразует алфавит в самого себя).

Пример: Создадим таблицу замен для аффинной системы подстановок Цезаря с ключом (5, 4) на примере русского алфавита. Возьмем алфавит из 32 букв (все кроме буквы «Ё»). Таким образом, a = 5, b = 3, n = 32. Условие НОД (5, 32) = 1 выполнено. Код буквы шифртекста находим из соотношения $A_{a,b}(j)=(5*j+3)(\text{mod } 32)$.

Сведем числовые коды букв открытого и зашифрованного текстов в таблицу (рис. 2).

J	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
A	3	8	13	18	23	28	1	6	11	16	21	26	31	4	9	14
J	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
A	19	24	29	2	7	12	17	22	27	0	5	10	15	20	25	30

Рис. 2. Таблица кодов для аффинных подстановок при a =5, *b* =3, n =32

Преобразуем числовые коды в соответствующие буквы русского алфавита и получим соответствие для символов открытого текста и шифртекста (табл. 3).

J	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Исх. текст	А	Б	В	Г	Д	Е	Ж	З	И	Й	К	Л	М	Н	О	П
Шифртекст	Г	И	Н	Т	Ч	Ь	Б	Ж	Л	Р	Х	Ъ	Я	Д	Й	О
J	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
Исх. текст	Р	С	Т	У	Ф	Х	Ц	Ч	Ш	Щ	Ъ	Ы	Ь	Э	Ю	Я
Шифртекст	У	Ш	Э	В	З	М	С	Ц	Ы	А	Е	К	П	Ф	Щ	Ю

Рис. 3. Таблица символов для аффинных подстановок при a =5, *b* =3, n =32

С помощью рис. 3 или формулы слово «МИР» преобразуется в шифртекст «ЯЛУ».

Квадрат Полибия

Еще одной модификацией одноалфавитной замены является квадрат Полибия, в котором символ алфавита заменяется парой чисел или символов по определенному правилу. Рассмотрим прямоугольник, часто называемый доской Полибия.

	А	Б	В	Г	Д	Е
А	А	Б	В	Г	Д	Е
Б	Ж	З	И	К	Л	М
В	Н	О	П	Р	С	Т
Г	У	Ф	Х	Ц	Ч	Ш
Д	Щ	Ъ	Ы	Ь	Э	Ю
Е	Я		.	,	-	:

В такой прямоугольник записывается алфавит, причем схема записи держится в тайне и составляет ключ шифрования. Для того чтобы получались приближенные к квадрату матрицы (6x6, 5x7, 6x5), в алфавит могут включаться знаки препинания или исключаться редко используемые символы (такие как ‘ё’, ‘й’).

В процессе шифрования каждая буква открытого текста представляется в шифротексте парой букв, указывающих строку и столбец, в которых расположена данная буква. Так представлениями букв В, Г, П, У будут АВ, АГ, ВВ, ГА соответственно. Если использовать приведенный выше квадрат в качестве ключа шифрования, то фраза «ПРИМЕР ШИФРОВАНИЯ» будет зашифрована в «ВВВГБВБЕАЕВГЕБГЕБВГБВГВБАВААВАБВЕА». В приведенном примере размер шифротекста превышает размер исходного текста в 2 раза, однако при машинной реализации номера строки и столбца таблицы можно задавать в виде цифр, а, учитывая, что получившийся квадрат имеет по 6 строк и столбцов, для кодирования каждой буквы будет достаточно 6 бит (3 для номера строки, 3 для номера столбца).

Общее количество ключей (различных вариантов размещения алфавита в матрице) равно $n!$.

Шифр Виженера

В XVI веке французский дипломат Блез де Виженер предложил модификацию шифра замен, которая впоследствии получила его имя. В данном шифре ключ задается фразой из d букв. Ключевая фраза подписывается с повторением под сообщением. Букву шифротекста необходимо находить на пересечении столбца, определяемого буквой открытого текста, и строки, определяемой буквой ключа:

$$Vig_d(m_i)=(m_i+k_{i \bmod d}) (\mathrm{mod}\ n),$$

где m_i, k_i, $Vig_d(m_i)$- порядковые номера в алфавите очередных символов открытого текста, ключа и шифротекста соответственно. Обратное преобразование выглядит следующим образом:

$$Vig_d^{-1}(m_i)=(m_i-k_{i \bmod d}+n) (\mathrm{mod}\ n)$$

Пример использования шифра Виженера приведен на рис.2 (алфавит дополнен пробелом, порядковый номер которого принят за 34, соответственно, мощность алфавита n=34).

ключевая фраза – ‘ключ’
d=4

Порядковый № буквы в алфавите26102218343108615618 1Открытый текстшифрвиженера

Порядковый № буквы в алфавите12133225121332251213322512Ключевая фразаключключключк

Порядковый № буквы в алфавите42320912168331828491 3Шифротекстгхтзкожярьгзл

Рис.4. Пример шифрования с использованием шифра Виженера

Принципиальным отличием данного шифра от всех предыдущих является то, что он относится к классу многоалфавитных алгоритмов – как нетрудно заметить, одной и той же букве шифротекста могут соответствовать различные символы открытого текста в зависимости от того, каким символом ключа они были замаскированы (в приведенном примере буква ‘г’ шифрует в одном случае букву ‘ш’, а в другом – букву ‘е’).

Шифр Гронсфельда

Идея шифра Виженера (использование многоалфавитной подстановки), используется в шифре Гронсфельда. Он использует в качестве ключа целое число. Каждая цифра в десятичной записи этого числа означает величину сдвига заменяющего алфавита при подстановке соответствующей буквы открытого текста. Если K – ключ, d – количество цифр в нем, а K_i – i-я десятичная цифра ключа, то шифрование можно представить следующим образом:

$$Gro(m_i)=(m_i+k_{i \bmod d})(\text{mod } n).$$

Таким образом, шифр повторяет процедуру шифрованию Виженера, но вместо порядкового номера символа ключа в алфавите использует непосредственное де-

сятичное значение (недостатком является уменьшении величины сдвига для каждого символа величиной 9).

Пример

Допустим, мы хотим зашифровать слово «ТАЙНЫ», используя ключ «103». Записываем циклически под словом ТАЙНЫ наш ключ, после чего сдвигаем по алфавиту каждую букву на столько букв вперед, сколько указано ниже, получим:

Т А Й Н А

1 0 3 1 0

У А М О А

Соответственно для дешифровки, сдвиг по алфавиту происходит в обратную сторону.

Шифрование биграммами

В начале XVI века аббат из Германии Иоганн Трисемус предложил шифровать по две буквы за раз. Шифры, использующие подобный принцип, были названы биграммными. Обычно такие шифры используют таблицы, аналогичные квадрату Полибия, заполненные символами используемого алфавита. Наиболее известный шифр биграммами называется Playfair. Он применялся Великобританией в Первую мировую войну. Открытый текст разбивался на пары букв (биграммы) и текст шифровки строился из него по следующим трём очень простым правилам (рис.3).

1. Если обе буквы биграммы исходного текста принадлежат одной колонке таблицы, то буквами шифра считаются буквы, которые лежат под ними. Так биграмма ИН дает текст шифровки НЗ. Если буква открытого текста находится в нижнем ряду, то для шифра берется соответствующая буква из верхнего ряда и биграмма НЯ дает шифр ЗИ. (Биграмма из одной буквы или пары одинаковых букв тоже подчиняется этому правилу, и текст ОО дает шифр ГГ).

2. Если обе буквы биграммы исходного текста принадлежат одной строке таблицы, то буквами шифра считаются буквы, которые лежат справа от них. Так биграмма АБ дает текст шифровки НГ. Если буква открытого текста находится в

крайней правой колонке, то для шифра берется буква из крайней левой колонки той же строки и биграмма АД дает шифр НА.

3. Если обе буквы биграммы открытого текста лежат в разных рядах и колонках, то вместо них берутся такие две буквы, чтобы вся их четверка представляла прямоугольник. Например, биграмма ЕК шифруется как БЙ (КЕ зашифруется ЙБ).

Заполнение квадрата алфавитом может быть случайным, а может определяться некоторой ключевой фразой, все символы которой (но без повторений) записываются в начале матрицы, а затем по порядку выписываются остальные буквы алфавита.

Перестановочный шифр с ключевым словом

Буквы ключевого слова без повторений записываются в первую строку таблицы, определяя таким образом количество ее столбцов. Буквы сообщения записываются в таблицу построчно. Сформированная таким образом таблица сортируется по столбцам, критерием сортировки является порядок следования символа первой строки в алфавите. После сортировки шифрованный текст переписывается по столбцам (рис 7).

Пример

Ключевая фраза: шифр

Открытый текст: перестановочный шифр

ш	и	ф	р
п	е	р	е
с	т	а	н
о	в	о	ч
н	ы	й	
ш	и	ф	р

и	р	ф	ш
е	е	р	п
т	н	а	с
в	ч	о	о
ы		й	н
и	р	ф	ш

Шифротекст: етвыиенч рраойфпсонш

Рис.5. Пример шифрования сообщения методом перестановки с ключевым словом

Дешифрация осуществляется по известному ключу обратными преобразованиями шифротекста в таблице: сначала шифротекст вписывается в таблицу по столбцам, затем столбцы переставляются, после чего открытый текст извлекается из таблицы построчно.

Порядок выполнения работы

1. Ознакомьтесь с теоретическими основами шифрования данных в настоящих указаниях.
2. Получите вариант задания у преподавателя.
3. Напишите программу согласно варианту задания.
4. Отладьте разработанную программу и покажите результаты работы программы преподавателю.
5. Составьте отчет по лабораторной работе

Содержание отчета

Отчет по лабораторной работе должен содержать следующие сведения:

- название и цель работы;
- вариант задания;
- листинг разработанной программы;
- результаты работы программы.

Варианты заданий

1. Реализовать в программе шифрование и дешифрацию содержимого файла по методу Цезаря с ключевым словом. Ключ и ключевое слово вводятся.
2. Реализовать в программе шифрование и дешифрацию содержимого файла с использованием квадрата Полибия.
3. Реализовать шифрование и дешифрацию содержимого файла с использованием метода биграмм. Ключевое слово вводится.

4. Реализовать шифрование и дешифрацию содержимого файла по методу Виженера. Ключевая фраза вводится.

5. Реализовать в программе шифрование и дешифрацию содержимого файла с использованием перестановочного шифра с ключевым словом. Ключевое слово вводится.

6. Реализовать в программе шифрование и дешифрацию содержимого файла по методу Цезаря. Ключ вводится.

7. Реализовать шифрование и дешифрацию содержимого файла по методу Гронсфельда с ключом произвольной длины. Ключ вводится с клавиатуры.

8. Реализовать в программе шифрование и дешифрацию введенного текста по методу Цезаря с ключевым словом. Ключ и ключевое слово берутся из файла.

9. Реализовать шифрование и дешифрацию введенного текста по методу Виженера. Ключевая фраза берется из файла.

10. Реализовать шифрование и дешифрацию введенного текста с использованием метода биграмм. Ключевое слово берется из файла.

11. Реализовать в программе шифрование и дешифрацию введенного текста с использованием квадрата Полибия.

12. Реализовать в программе шифрование и дешифрацию введенного текста с использованием перестановочного шифра с ключевым словом. Ключевое слово берется из файла.

13. Реализовать в программе шифрование и дешифрацию введенного текста с использованием аффинной криптосистемы.

14. Реализовать в программе шифрование и дешифрацию введенного текста по методу Цезаря. Ключ берется из файла.

15. Реализовать шифрование и дешифрацию введенного текста по методу Гронсфельда с ключом произвольной длины. Ключ берется из файла.

16. Реализовать в программе шифрование и дешифрацию содержимого файла с использованием аффинной криптосистемы.

Контрольные вопросы

1. Дайте определение следующим понятиям: шифр, криптография, криптоанализ, ключ.
2. Классифицируйте алгоритм, полученный в качестве задания к лабораторной работе.
3. Чем отличаются одно- и многоалфавитные методы шифрования?
4. В чем заключается основной принцип частотного криптоанализа?
5. Какой метод криптоанализа применим для вскрытия алгоритма, полученного вами в качестве задания к лабораторной работе?
6. Оцените мощность ключевого пространства вашего алгоритма.