Министерство образования Республики Беларусь
Учреждение образования
"Белорусский государственный университет
информатики и радиоэлектроники"

Кафедра радиоэлектронных средств

В.М. Алефиренко

ОСНОВЫ ЗАЩИТЫ ИНФОРМАЦИИ

Практикум
для студентов специальностей
«Техническое обеспечение безопасности»
и «Моделирование и компьютерное проектирование
радиоэлектронных средств»
дневной, вечерней и заочной форм обучения

В 2-х частях

Часть 2

Минск 2004

УДК 004.056 (075.8)
ББК 32.97 я 73
А 48

Р е ц е н з е н т:
профессор кафедры сетей и устройств телекоммуникаций БГУИР,
доктор технических наук Л.М. Лыньков

Алефиренко В.М.
А 48 Основы защиты информации: Практикум для студ. спец. «Техническое обеспечение безопасности» и «Моделирование и компьютерное проектирование радиоэлектронных средств» дневн., веч. и заоч. форм обуч.: В 2 ч. Ч.2/ В.М. Алефиренко. – Мн.: БГУИР, 2004. – 44 с.: ил.
ISBN 985-444-610-7 (ч. 2)

Во второй части практикума приводится описание двух практических работ по курсу “Основы защиты информации”, выполняемых студентами в рамках практических занятий с применением вычислительной техники.

Разработка и отладка программного обеспечения осуществлялась студентами-дипломниками А.Н. Рыковым и А.Ф. Левченко.

УДК 004.056 (075.8)
ББК 32.97 я 73

Часть 1: Алефиренко В.М., Шамгин Ю.В. Основы защиты информации: Практикум для студ. спец. «Техническое обеспечение безопасности» и «Моделирование и компьютерное проектирование радиоэлектронных средств»: В 2 ч. Ч 1/ В.М. Алефиренко, Ю.В. Шамгин. — Мн.: БГУИР, 2004. — 44 с.: ил.

ISBN 985-444-610-7 (ч. 2)
ISBN 985-444-609-3

Практическая работа №3

Изучение криптографических методов защиты информации

1. Цель работы

Изучение методов защиты информации с помощью различных видов шифров, используемых в криптографии.

2. Теоретические сведения

2.1. Основные понятия, термины и определения криптографии

Слово «*криптография*» произошло от древнегреческих слов «cryptos» – тайный и «graphos» – письмо. Таким образом, криптография – это тайнопись. Криптографическая защита информации (данных) с помощью кодов и шифров является одним из важнейших решений проблемы ее безопасности. Зашифрованные данные становятся доступными только тому, кто знает, как их расшифровать. Поэтому похищение зашифрованных данных бессмысленно для несанкционированных пользователей.

Различные коды и шифры используются давно. С теоретической точки зрения между ними не существует четкого различия. Однако в современной практике использования криптографии различие между ними определено достаточно четко [1, 2].

Кодирование – это процесс замены элементов открытого текста (символов, комбинаций символов, слов и т.п.) кодами. *Коды* оперируют лингвистическими элементами, разделяя кодируемый текст на ие смысловые элементы, как слова и слоги.

Шифрование – это процесс зашифрования или расшифрования. В этом процессе криптографическому преобразованию подвергается каждый символ текста. В шифровании всегда используются два элемента: алгоритм и ключ.

Алгоритм шифрования – это последовательность определенных действий над открытым текстом, в результате которых получается зашифрованный текст (шифротекст). Алгоритм позволяет использовать сравнительно короткий ключ для шифрования сколь угодно большого текста.

Ключ – это конкретное секретное состояние некоторых параметров алгоритма криптографического преобразования данных, обеспечивающее выбор одного варианта из всей совокупности возможных вариантов для данного алгоритма.

Шифр – это совокупность обратимых преобразований множества возможных открытых данных на множество возможных зашифрованных данных, осуществляемых по определенным правилам с применением ключей.

Зашифрование – это процесс преобразования открытых данных в зашифрованные с помощью шифра.

Расшифрование – это процесс преобразования закрытых данных в открытые данные с помощью шифра.

Дешифрование – это процесс преобразования закрытых данных в открытые данные при неизвестном ключе и, возможно, неизвестном алгоритме.

Гаммирование – это процесс наложения по определенному закону гаммы шифра на открытые данные.

Гамма шифра – это псевдослучайная двоичная последовательность, вырабатываемая по заданному алгоритму, для зашифрования открытых данных и расшифрования зашифрованных данных.

Уравнение зашифрования – это соотношение, описывающее процесс образования зашифрованных данных из открытых данных в результате преобразований, заданных алгоритмом криптографического преобразования.

Уравнение расшифрования – это соотношение, описывающее процесс образования открытых данных из зашифрованных данных в результате преобразований, заданных алгоритмом криптографического преобразования.

Имитозащита – это защита системы шифрованной связи от навязывания ложных данных. Для обеспечения имитозащиты к зашифрованным данным добавляется имитовставка.

Имитовставка – это отрезок информации фиксированной длины, полученной по определённому правилу из открытых данных и ключа, добавленный к зашифрованным данным для обеспечения имитозащиты.

Криптографическая защита – это защита данных с помощью криптографического преобразования.

Криптографическое преобразование – это преобразование данных с помощью шифрования и (или) выработки имитовставки.

Криптостойкость шифра – это характеристика шифра, определяющая его стойкость к дешифрованию. Обычно эта характеристика определяется периодом времени, необходимым для дешифрования.

2.2. Методы криптографии

2.2.1. Классификация методов

Методы криптографии можно разделить на две группы: с секретными ключами и с открытыми ключами [1].

Методы криптографии *с секретными (закрытыми) ключами* предусматривают один ключ, который используется как в процессе зашифрования, так и в процессе расшифрования. Этот ключ известен только тем, кто зашифровывает и расшифровывает данные. Так как в этих методах используется только один ключ, то они получили название *симметричных методов.*

Методы криптографии *с открытыми ключами* предусматривают два ключа. Первый ключ используется для зашифрования и не является секретным. Он может быть известен всем пользователям системы, которые зашифровывают данные. Расшифрование данных с помощью известного ключа невозможно. Для расшифрования используется второй ключ, который является секретным. Так как в этих методах используются два различных ключа, то они получили название *несимметричных методов.*

В свою очередь методы с секретными ключами делятся на методы замены (подстановки), методы перестановки и методы перемешивания.

Метод замены (подстановки) основан на том, что каждый символ открытого текста заменяется другим символом того же алфавита. Конкретный вид замены определяет секретный ключ. Замена может быть *моноалфавитная*, *гомофоническая*, *полиалфавитная* и *полиграммная*. Для реализации метода замены может быть использован датчик (генератор) псевдослучайных чисел.

Метод замены с использованием датчика псевдослучайных чисел основан на генерации гаммы шифра с помощью генератора псевдослучайных чисел и наложении полученной гаммы на открытые данные обратимым образом. Расшифрование данных сводится к повторной генерации гаммы шифра при известном ключе и наложению этой гаммы на зашифрованные данные.

Метод перестановки основан на изменении порядка следования символов открытого текста. Порядок перестановки определяет секретный ключ. Перестановка может быть *простая* и *усложненная*.

Метод перемешивания основан на том, что изменение одного символа открытого текста приводит к изменению многих символов шифротекста.

2.2.2. Методы криптографии с секретными ключами

2.2.2.1. Общие положения

Классическим подходом в криптографии является использование секретных ключей. При этом подходе полагается, что криптоаналитик противника знает методику шифрования, и секретность шифра определяется только секретностью ключа. Структурная схема шифрования с секретным ключом (симметричное шифрование) показана на рис.3.1.

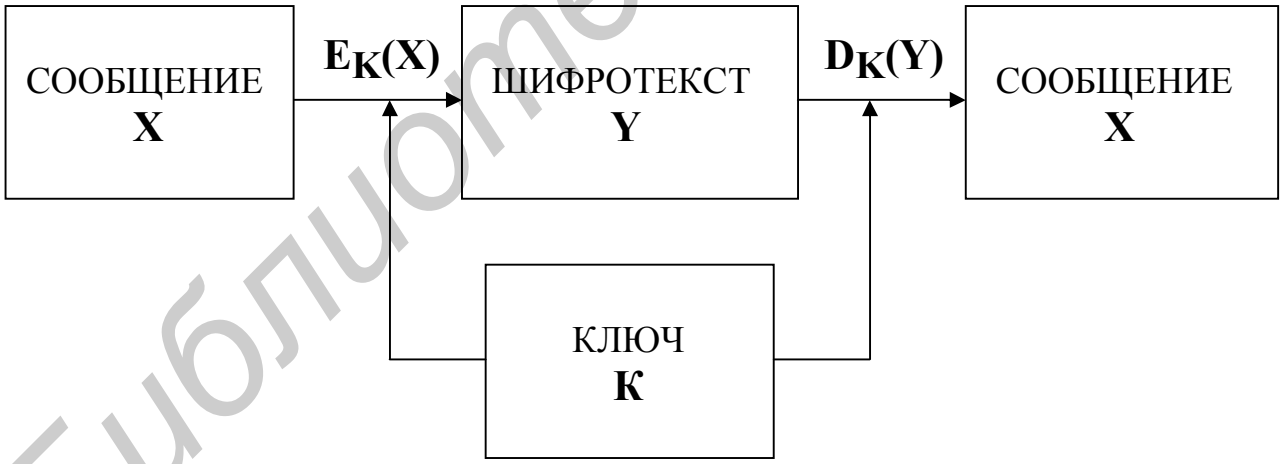

Рис. 3.1. Симметричное шифрование

Уравнение зашифрования может быть представлено в следующем виде:

$$X=E_K(X), \qquad (3.1)$$

где $\mathbf{E_K}$ – символ, означающий алгоритм шифрования по секретному ключу **K**.

Уравнение расшифрования принимает тогда следующий вид:

$$\mathbf{X=D_K(Y)}, \qquad (3.2)$$

где $\mathbf{D_K}$ – символ, означающий алгоритм расшифрования по секретному ключу **K**.

Таким образом, зашифрование и расшифрование проводится с помощью только одного секретного ключа [3].

2.2.2.2. Метод замены

Шифрование методом замены (подстановки) основано на алгебраической операции, называемой подстановкой. *Подстановкой* называется взаимно однозначное отображение некоторого конечного множества **M** на себя. Число **N** элементов этого множества называется *степенью подстановки*. Природа множества **M** роли не играет, поэтому можно сказать, что **M=1, 2, …, N**.

Если при данной подстановке число **j** переходит в число $\mathbf{i_j}$, то такую подстановку, обозначаемую символом **S**, можно записать в виде

$$\mathbf{S} = \begin{pmatrix} 1 & 2 & \dots & n \\ i_1 & i_2 & \dots & i_n \end{pmatrix}. \qquad (3.3)$$

В этой записи числа **1, 2, …, n** можно произвольным образом переставлять, соответственно переставляя числа $\mathbf{i_1, i_2, \dots, i_n}$.

Результат последовательного выполнения двух подстановок $\mathbf{S_1}$ и $\mathbf{S_2}$ одной и той же степени также является подстановкой, которая называется *произведением подстановок* $\mathbf{S_1}$ и $\mathbf{S_2}$ и обозначается как $\mathbf{S_1 \times S_2}$.

Две подстановки называются независимыми, если они не имеют общих действительно перемещаемых чисел.

Количество чисел **m**, действительно перемещаемых подстановкой **S**, называется *длиной цикла подстановки*.

Подстановка **S** называется *транспозицией*, если существует пара (j_1 , j_2) различных элементов из множества **M**, удовлетворяющих условиям: $i_{j1}=j_1$, $i_{j2}=j_2$, $i_j=j$ для каждого $j \in \{ M(j_1, j_2) \}$. Любая подстановка разлагается в произведение транспозиций. Разложение подстановки в произведение независимых подстановок однозначно с точностью до порядка множителей.

В криптографии рассматриваются четыре типа подстановки (замены): моноалфавитная, гомофоническая, полиалфавитная и полиграммная [1]. Подстановка может быть реализована с использованием датчика псевдослучайных чисел.

Моноалфавитная замена. При моноалфавитной замене каждый символ алфавита открытого текста заменяется символом шифротекста из того же алфавита.

Общая формула моноалфавитной замены выглядит следующим образом:

$$Y_i = (K_1 \cdot X_i + K_2) \bmod n, \qquad (3.4)$$

где Y_i – i-й символ шифротекста;

X_i – i-й символ открытого текста;

K_1 и K_2 – константы;

n – длина алфавита.

Под результатом операции $(K_1 \cdot X_i + K_2) \bmod n$ понимают остаток от целочисленного деления суммы $(K_1 \cdot X_i + K_2)$ на число **n,** если сумма больше длины алфавита.

Для описания алгоритма шифрования обычно вместо символов открытого и шифротекста используют их цифровые эквиваленты. Пример цифрового эквивалента букв русского алфавита (без знаков препинания) приведен в табл. 3.1.

Таблица 3.1

Цифровые эквиваленты букв русского алфавита

Буква	А	Б	В	Г	Д	Е	Ж	З	И	Й	К	Л
Цифровой эквивалент	1	2	3	4	5	6	7	8	9	10	11	12
Буква	М	Н	О	П	Р	С	Т	У	Ф	Х	Ц	Ч
Цифровой эквивалент	13	14	15	16	17	18	19	20	21	22	23	24
Буква	Ш	Щ	Ъ	Ы	Ь	Э	Ю	Я	_ (ПРОБЕЛ)			
Цифровой эквивалент	25	26	27	28	29	30	31	32	33			

Общее число символов алфавита n=33.

Пример 1. Открытый текст: «ШИФРОВАНИЕ ЗАМЕНОЙ». Подстановка задана табл. 3.2.

Таблица 3.2

Подстановка алфавита для шифрования моноалфавитной заменой

Алфавит открытого текста	А	Б	В	Г	Д	…	Ь	Э	Ю	Я	_
Алфавит шифротекста	_	Я	Ю	Э	Ь	…	Д	Г	В	Б	А

Шифротекст: «ИШМРТЮ_УШЫАЩ_ФЫУТЧ».

Основным недостатком рассмотренного метода является то, что статистические свойства открытого текста (частоты появления букв) сохраняются и в шифротексте. Этого недостатка лишены шифры Вижинера и Бофора.

Шифр Вижинера. Шифр Вижинера задается формулой

$$\mathbf{Y_i = (X_i + K_i) \bmod n,} \qquad (3.5)$$

где $\mathbf{K_i}$ – i-й символ ключа, в качестве которого используется слово или фраза.

Пример 2. Открытый текст: «ЗАМЕНА». В качестве ключа используется слово «КЛЮЧ». Подстановка задана табл. 3.3.

Таблица 3.3

Подстановка шифра Вижинера

З	А	М	Е	Н	А
К	Л	Ю	Ч	К	Л

В соответствии с табл. 3.1 записываем:

Y_1=(8+11) mod 33=19 → Т;

Y_2=(1+12) mod 33=13 → М;

Y_3=(13+31) mod 33=11 → К;

Y_4=(6+24) mod 33=30 → Э;

Y_5=(14+11) mod 33=25 → Ш;

Y_6=(1+12) mod 33=13 → М.

Шифротекст: «ТМКЭШМ».

Шифр Вижинера с неограниченным неповторяющимся ключом называют шифром Вернама.

Шифры Бофора. Шифры Бофора задаются формулами:

$$\mathbf{Y_i = (K_i - X_i) \bmod n;} \qquad (3.6)$$

$$\mathbf{Y_i = (X_i - K_i) \bmod n.}$$

Так как при использовании шифров Бофора возможны случаи, когда разность может быть равна нулю, то нумерацию символов алфавита необходимо начинать с нуля. Тогда в табл. 3.1 буква А будет соответствовать 0, Б – 1, В – 2 и т.д.

При рассмотрении этих видов шифров видно, что чем больше длина ключа, тем лучше шифр. Существенного улучшения свойств шифротекста можно достигнуть при использовании шифров с автоключом.

Шифр, в котором сам открытый текст или получающаяся криптограмма используется в качестве ключа, *называется шифром с автоключом*. Шифрование в этом случае начинается с ключа, называемого первичным, и продолжается с помощью открытого текста или криптограммы, смещенной на длину первичного ключа.

Пример 3. Открытый текст: «ШИФРОВАНИЕ ЗАМЕНОЙ». Первичный ключ: «КЛЮЧ». Схема шифрования с автоключом при использовании открытого текста представлена в табл. 3.4.

Таблица 3.4

Схема шифрования с автоключом при использовании открытого текста

Ш	И	Ф	Р	О	В	А	Н	И	Е	_	З	А	М	Е	Н	О	Й
К	Л	Ю	Ч	Ш	И	Ф	Р	О	В	А	Н	И	Е	_	З	А	М
36	21	52	41	18	24	20	22	27	30	53		10	19	39	22	16	23
В	Ф	Т	З	Ж	Л	Х	Ю	Ч	И	А	Х	Й	Т	Е	Х	П	Ц

Шифротекст: «ВФТЗЖЛХЮЧИАХЙТЕХПЦ».

Схема шифрования с автоключом при использовании криптограммы представлена в табл. 3.5.

Таблица 3.5

Схема шифрования с автоключом при использовании криптограммы

Ш	И	Ф	Р	О	В	А	Н	И	Е	_	З	А	М	Е	Н	О	Й
К	Л	Ю	Ч	В	Ф	Т	З	С	Ч	У	Х	Ъ	Э	У	Э	Ы	Й
36	21	52	41	18	24	20	22	27	30	53	30	28	43	26	44	43	20
В	Ф	Т	З	С	Ч	У	Х	Ъ	Э	У	Э	Ы	Й	Щ	К	Й	У

Шифротекст: «ВФТЗСЧУХЪЭУЭЫЙЩКЙУ».

Гомофоническая замена. При гомофонической замене каждый символ алфавита открытого текста заменяется в определенном порядке несколькими символами шифротекста из этого же алфавита. Этот метод применяется для искажения статистических свойств шифротекста.

Пример 4. Открытый текст: «ЗАМЕНА». Подстановка задана табл. 3.6.

Таблица 3.6

Подстановка алфавита для шифрования гомофонической заменой

Алфавит открытого текста	А	Б	…	Е	Ж	З	…	М	Н	…
Алфавит шифротекста	17 31 48	23 44 63	… … …	97 51 15	47 67 33	76 19 59	… … …	32 28 61	55 84 34	… … …

Каждая буква открытого текста заменяется по очереди цифрами соответствующего столбца.

Шифротекст: «76 17 32 97 55 31».

Полиалфавитная замена. При полиалфавитной замене используется несколько алфавитов шифротекста. Пусть используется **k** алфавитов. Тогда открытый текст

$$\mathbf{X=X_1\,X_2\ldots X_k\,X_{k+1}\ldots X_{2k}\,X_{2k+1}\ldots} \qquad (3.7)$$

заменяется шифротекстом

$$\mathbf{Y=F_1(X_1)F_2(X_2)\ldots F_k(X_k)F_1(X_{k+1})\ldots F_k(X_{2k})F_1(X_{2k+1})\ldots}, \qquad (3.8)$$

где $\mathbf{F_i(X_j)}$ – символ шифротекста алфавита **i** для символа открытого текста $\mathbf{X_j}$.

Пример 5. Открытый текст: «ЗАМЕНА». k=3. Замена задана табл. 3.6, в которой каждая строка цифр соответствует своему алфавиту шифротекста.

Шифротекст: «76 31 61 97 84 48».

Полиграммная замена. Полиграммная замена формируется из одного алфавита с помощью специальных правил. Примером полиграммной замены может служить шифр Плэйфера.

Шифр Плэйфера. В этом шифре алфавит располагается в матрице. Открытый текст разбивается на пары символов $\mathbf{X_i}$, $\mathbf{X_{i+1}}$. Каждая пара символов открытого текста заменяется на пару символов из матрицы по следующим правилам:

если символы находятся в одной строке, то каждый из символов пары заменяется на стоящий правее от него (за последним символом в строке следует первый);

если символы находятся в одном столбце, то каждый символ пары заменяется на символ, расположенный ниже его в столбце (за последним нижним символом следует верхний);

если символы пары находятся в разных строках и столбцах, то они считаются противоположными углами прямоугольника. Символ, находящийся в левом углу, заменяется на символ, стоящий в другом левом углу. Замена символа, находящегося в правом углу, осуществляется аналогично;

если в открытом тексте встречаются два одинаковых символа подряд, то перед шифрованием между ними вставляется специальный символ (например тире).

Пример 6. Открытый текст: «ШИФР ПЛЭЙФЕРА». Матрица алфавита задана табл. 3.7.

Таблица 3.7

Матрица алфавита шифра Плэйфера

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

Шифротекст: «РДИЫ,–СТ–И.ХЧС».

Метод замены с использованием датчика псевдослучайных чисел. Шифрование этим методом заключается в генерации гаммы шифра датчиком псевдослучайных чисел с последующим наложением полученной гаммы на открытые данные обратимым способом (например, путем поразрядного сложения по модулю 2 с использовнием логической операции «исключающее ИЛИ»: 0+0=0; 1+0=1; 0+1=1; 1+1=0).

Таким образом, открытый текст, ключ и шифротекст представляются в виде двоичных последовательностей. Ключевая последовательность формируется датчиком псевдослучайных чисел, который запускается начальным значением ключа.

Расшифрование данных осуществляется путем повторной генерации ключевой последовательности при известном начальном значении ключа и наложения ее на шифротекст [1].

Зашифрованное сообщение будет достаточно трудно дешифровать, если гамма шифра не содержит повторяющихся битовых последовательностей или если период гаммы превышает длину всего зашифрованного сообщения и неизвестна никакая часть исходного текста. Шифр можно раскрыть только прямым перебором (подбором ключа). В этом случае криптостойкость определяется размером ключа.

Для получения линейных последовательностей элементов гаммы шифра, длина которых превышает размер шифруемых сообще-

ний, используется генератор псевдослучайных чисел. Линейные последовательности псевдослучайных чисел, вырабатываемые таким генератором, описываются соотношением:

$$\mathbf{T(\,i+1\,)=[\,AT(\,i\,)+C\,]\ mod\ M\,,} \qquad (3.9)$$

где **A** и **C** – константы;

T(i) – исходная величина, выбранная в качестве порождающего числа (входного ключа).

Генератор вырабатывает псевдослучайные числа с определенным периодом повторения, зависящим от выбранных значений **A** и **C**. Значение **M** обычно устанавливается равным $\mathbf{2}^{\boldsymbol{l}}$, где $\boldsymbol{l}$ – длина последовательности (длина слова в ЭВМ) в битах.

Различают методы конечной и бесконечной гаммы. В качестве *конечной гаммы* может использоваться фраза, а в качестве *бесконечной гаммы* – последовательность, вырабатываемая генератором псевдослучайных чисел.

Пример 7. Открытый текст: «ПРИКАЗ»
(«16 17 09 11 01 08» согласно табл. 3.1).

Гамма: «ГАММА» («04 01 13 13 01» согласно табл. 3.1)

Операция: сложение по mod 33:

Y_1 = (16+4) mod 33 = 20 → У;

Y_2 = (17+1) mod 33 = 18 → С;

Y_3 = (9+13) mod 33 = 22 → Х;

Y_4 = (11+13) mod 33 = 24 → Ч;

Y_5 = (1+1) mod 33 = 2 → Б;

Y_6 = (8+4) mod 33 = 12 → Л.

Шифротекст: «УСХЧБЛ».

Пример 8. Открытый текст: «ПРИКАЗ»
(«16 17 09 11 01 08» согласно табл. 3.1).

Первые значения датчика: «2 1 7 9 4 5 6 7».

Операция: сложение по mod 2 с использованием логической операции «исключающее ИЛИ»

Запишем код (цифру) каждой буквы открытого текста в двоичном виде, используя пять разрядов, а каждую цифру гаммы – используя четыре разряда, и проведем операцию сложения по mod 2:

10000(16)	10001(17)	01001(09)	01011(11)	00001(01)	01000(08)
⊕ 00010(2)	00001(1)	00111(7)	01001(9)	00100(4)	00101(5)
10010(18)	10000(16)	01110(14)	00010(02)	00101(05)	01101(13)

Шифротекст: «СПНБДМ» («18 16 14 02 05 13»).

Шифрование заменой необязательно предполагает замену символов открытого текста символами того же алфавита или цифрами. В качестве алфавита шифротекста возможно использование символов псевдографики или, например, музыкальных нот. Основным недостатком методов замены является взаимное соответствие положения открытого текста и шифротекста.

2.2.2.3. Метод перестановки

Шифрование методом перестановки основано на перестановке символов открытого текста, порядок которой определяет ключ. Существует большое количество различных способов перестановки. В качестве примера рассмотрим простую и усложненную перестановки [1].

Простая перестановка. При простой перестановке осуществляется перестановка групп символов алфавита открытого текста в определенном порядке.

Пример 9. Открытый текст: «ШИФРОВАНИЕ ПЕРЕСТАНОВКОЙ». Ключ (правило перестановки): буквы в группах из восьми букв с порядковыми номерами 1, 2, …, 8 переставить в порядок 3, 8, 1, 5, 2, 7, 6, 4.

Шифротекст: «ФНШОИАВР_СИЕЕЕРПНЙТВАОКО».

Усложненная перестановка. При усложненной перестановке открытый текст записывается в матрицу по определенному ключу $\mathbf{K_1}$. Шифротекст образуется при считывании из этой матрицы по ключу $\mathbf{K_2}$.

Пример 10. Открытый текст: «ШИФРОВАНИЕ ПЕРЕСТАНОВКОЙ». Матрица из четырех столбцов приведена в табл.3.8, где запись открытого текста проведена по строкам в соответствии с ключом $\mathbf{K_1}$: 5, 3, 1, 2, 4, 6, а чтение – по столбцам в соответствии с ключом $\mathbf{K_2}$: 4, 2, 3, 1.

Таблица 3.8

Матрица алфавита с перестановкой из четырех столбцов

1	И	Е	_	П
2	Е	Р	Е	С
3	О	В	А	Н
4	Т	А	Н	О
5	Ш	И	Ф	Р
6	В	К	О	Й
$К_1/К_2$	1	2	3	4

Шифротекст: «ПСНОРЙЕРВАИК_ЕАНФОИЕОТШВ».

Более сложные перестановки осуществляются с использованием графа по так называемым гамильтоновым путям, которых в графе может быть несколько.

Пример 11. Открытый текст:
«ШИФРОВАНИЕ ПЕРЕСТАНОВКОЙ». Ключ – гамильтонов путь на графе рис.2.2.

Шифротекст: «ШАОНИРФВИЕЕСЕП_РТОВЙАОНК».

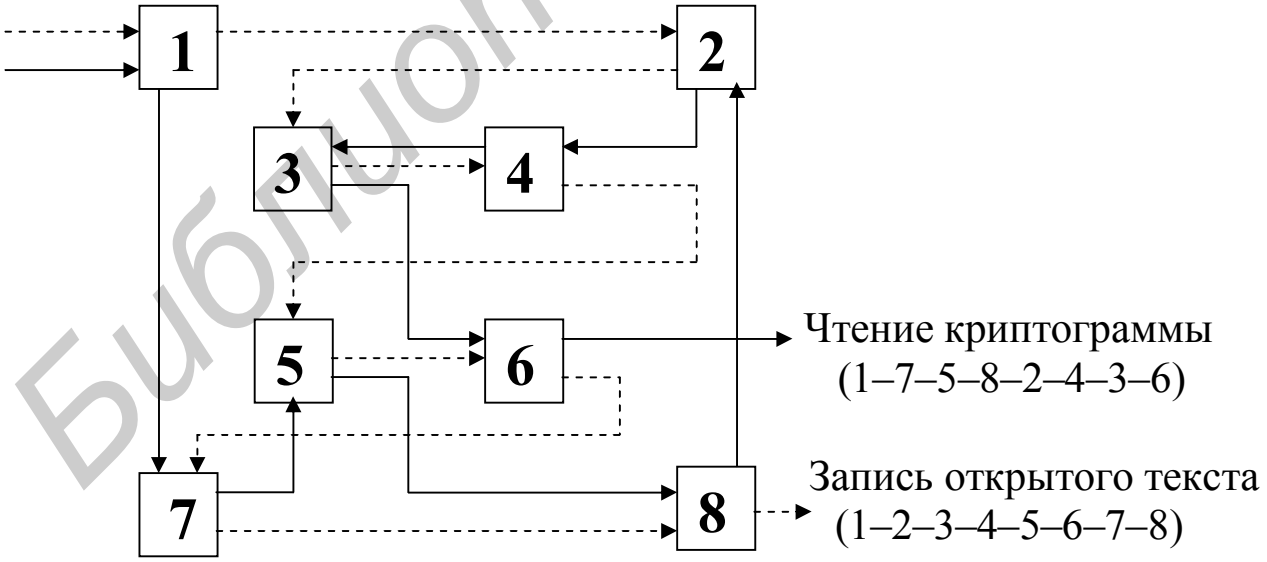

Рис. 3.2. Гамильтонов путь на графе

Необходимо отметить, что для данного графа из восьми вершин можно предложить несколько маршрутов записи открытого текста и несколько гамильтоновых путей для чтения криптограмм.

Еще более сложные перестановки основаны на принципах, заложенных в логической игре «Кубик Рубика». При использовании такой схемы открытый текст записывается в ячейки граней куба по строкам. После осуществления заданного числа заданных поворотов слоев куба считывание шифротекста осуществляется по столбцам. Сложность расшифрования в этом случае определяется числом ячеек на гранях куба и сложностью выполненных поворотов слоев куба. Перестановка, основанная на кубике Рубика, получила название *объемной (многомерной) перестановки*. Усовершенствованная схема такой перестановки, в которой наряду с открытым текстом перестановке подвергаются и функциональные элементы самого алгоритма шифрования, легла в основу секретной системы «Рубикон». В этой системе в качестве прообразов пространственных многомерных структур, на основании которых осуществляются перестановки, используются трехмерный куб и тетраэдр.

Основным недостатком методов перестановки является сохранение частотных свойств символов открытого текста в шифротексте.

2.2.2.4. Метод перемешивания

Метод перемешивания основан на совместном использовании методов замены (подстановки) и перестановки. При этом существенно нарушаются статистические связи шифротекста с открытым текстом. В стандартах шифрования часто применяются специальные меры, обеспечивающие расширение влияния каждого символа открытого текста на группу символов шифротекста [2]. В результате этого при замене любого одного символа открытого текста изменяется значительная группа символов шифротекста.

В практических шифрах используются два основных принципа Шеннона: рассеивание и перемешивание. *Рассеивание* – это распространение влияния одного знака открытого текста на множество знаков шифротекста, что позволяет скрыть статистические свойства открытого текста. *Перемешивание* – это использование взаимосвязи статистических свойств открытого и шифротекста.

Шифр должен не только затруднять раскрытие, но и обеспечивать легкость зашифрования и расшифрования при известном секретном ключе. Поэтому принята идея использовать произведение простых шифров, каждый из которых вносит небольшой вклад в значительное суммарное рассеивание и перемешивание. В таких составных шифрах в качестве элементарных составляющих чаще всего используются простые подстановки (замены) и перестановки [1].

2.2.3. Методы криптографии с открытыми ключами

Наиболее перспективными системами криптографической защиты информации являются системы с открытыми ключами [1]. В таких системах для зашифрования данных используется один (открытый) ключ, а для расшифрования – другой (секретный). Первый ключ не является секретным и может быть известен всем пользователям системы, которые зашифровывают данные. Расшифрование данных с помощью известного ключа невозможно. Для расшифрования данных используется второй ключ, который является секретным. Ключ расшифрования не может быть определен из ключа зашифрования [3].

Структурная схема шифрования с открытым ключом (несимметричное шифрование) показана на рис.3.3.

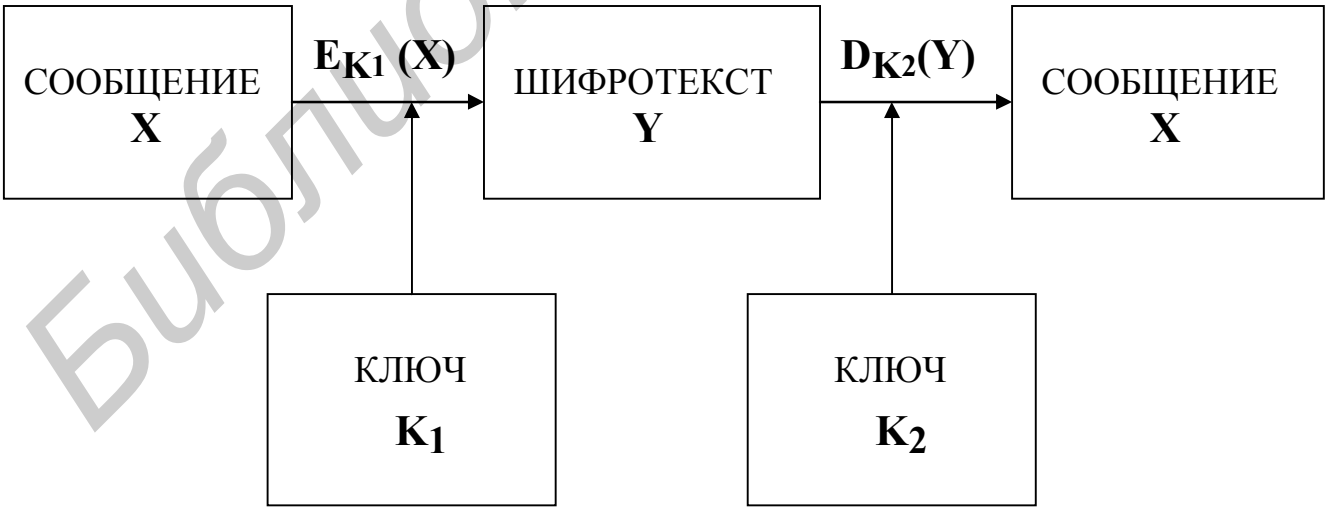

Рис.3.3. Несимметричное шифрование

Криптосистема с открытым ключом должна содержать следующие элементы:

рандомизированный алгоритм генерации открытого ключа $\mathbf{K_1}$ и соответствующего ему секретного ключа $\mathbf{K_2}$;

алгоритм зашифрования, который по сообщению **X**, используя открытый ключ $\mathbf{K_1}$, формирует шифротекст $\mathbf{Y=E_{K1}(X)}$;

алгоритм расшифрования, который по шифротексту **Y**, используя секретный ключ $\mathbf{K_2}$, восстанавливает исходное сообщение $\mathbf{X= D_{K2}(Y)}$.

Наиболее перспективным методом криптографической защиты информации с открытым ключом является алгоритм RSA (назван по начальным буквам фамилии его изобретателей – **R**ivest, **S**hamir и **A**dleman). При рассмотрении алгоритма RSA необходимо вспомнить некоторые математические термины.

Под *простым числом* понимают такое число, которое делится только на 1 и на само себя. *Взаимно простыми числами* называют такие числа, которые не имеют ни одного общего делителя, кроме 1. Под результатом операции **i mod j** понимают остаток от целочисленного деления **i** на **j**.

Чтобы использовать алгоритм RSA, необходимо сначала сгенерировать открытый и секретный (закрытый) ключи, выполнив следующее:

выбрать два очень больших простых числа **p** и **q**;

определить **n** как результат умножения **p** на **q** ($\mathbf{n = pq}$);

выбрать большое случайное число **d**. Оно должно быть взаимно простым с числом, определяемым как результат умножения чисел $\mathbf{(p–1)(q–1)}$;

определить такое число **e**, для которого является истинным следующее соотношение: $\mathbf{ed \bmod ((p–1)(q–1)) = 1}$;

считать открытым ключом числа **e** и **n**, а секретным ключом – числа **d** и **n** .

Для того чтобы зашифровать данные по известному ключу **{ e, n }**, необходимо разбить шифруемый текст на блоки, каждый из которых может быть представлен в виде числа $\mathbf{M_i}$ от **0** до **n–1**.

Затем зашифровать текст, рассматриваемый как последовательность чисел $\mathbf{M_i}$, выполнив вычисления $\mathbf{Y_i = M_i^e \bmod n}$.

Чтобы расшифровать эти данные, используя секретный ключ **{ d, n },** необходимо выполнить вычисления $\mathbf{M_i = Y_i^d \bmod n}$. В ре-

зультате будет получено множество чисел $\mathbf{M_i}$, которое представляет собой исходный текст.

Пример 12. Открытый текст «ЕДА».

Для простоты будем использовать маленькие числа.

Выберем два простых числа p=3 и q=11.

Определим $n=3 \cdot 11=33$.

Найдем $(p-1)(q-1)=20$.

Следовательно, в качестве секретного ключа **d** нужно выбрать любое число, которое является взаимно простым с числом 20, например d=3.

Выберем значение открытого ключа **e**. В качестве такого числа может быть использовано любое число, для которого справедливо соотношение $\mathbf{(e \cdot 3) \bmod 20 = 1}$. Например, e = 7.

Таким образом, открытый ключ составляют числа e = 7 и n = 33, а секретный – числа d = 3 и n = 33.

Представим шифруемое сообщение как последовательность целых чисел в диапазоне 0…32, число которых не превышает n = 33. Тогда буква Е изображается числом 6, буква Д – числом 5, буква А – числом 1, а слово «ЕДА» представляется последовательностью цифровых эквивалентов (чисел) как 651.

Зашифруем сообщение, используя открытый ключ { 7, 33 }:

$$Y_1 = 6^7 \bmod 33 = 279936 \bmod 33 = 30;$$
$$Y_2 = 5^7 \bmod 33 = 78125 \bmod 33 = 14;$$
$$Y_3 = 1^7 \bmod 33 = 1 \bmod 33 = 1.$$

Шифротекст: «30 14 1».

Расшифруем сообщение «30 14 1», полученное в результате зашифрования по известному ключу, на основе секретного ключа { 3, 33 }**:**

$$M_1 = 30^3 \bmod 33 = 27000 \bmod 33 = 6;$$
$$M_2 = 14^3 \bmod 33 = 2744 \bmod 33 = 5;$$
$$M_3 = 1^3 \bmod 33 = 1 \bmod 33 = 1.$$

Таким образом, в результате расшифрования сообщения получено исходное сообщение «ЕДА» («651»).

Криптостойкость алгоритма RSA основывается на предположении, что исключительно трудно определить секретный ключ по известному, так как для этого необходимо решить задачу о существовании делителей целого числа. Данная задача до настоящего времени не имеет эффективного (полиноминального) решения.

3. Порядок выполнения работы

1. Изучить теоретическую часть работы.
2. Провести самопроверку теоретических знаний, ответив на поставленные вопросы.
3. Получить вариант задания на зашифрование и расшифрование текста.
4. Осуществить зашифрование и расшифрование текста соответствующим методом, используя необходимые табл. 3.9 и 3.10 и соответствующий алгоритм.
5. Провести проверку правильности результатов зашифрования и расшифрования текста.
6. Оформить отчет и защитить работу.

Если студент индивидуально выполнил лабораторную работу, то необходимость в представлении отчета и защите работы отпадает.

Таблица 3.9

Подстановка алфавита для шифрования моноалфавитной заменой

Буква	А	Б	В	Г	Д	Е	Ж	З	И	Й	К	Л
Цифровой эквивалент	1	2	3	4	5	6	7	8	9	10	11	12
Буква	М	Н	О	П	Р	С	Т	У	Ф	Х	Ц	Ч
Цифровой эквивалент	13	14	15	16	17	18	19	20	21	22	23	24
Буква	Ш	Щ	Ъ	Ы	Ь	Э	Ю	Я	_ (ПРОБЕЛ)			
Цифровой эквивалент	25	26	27	28	29	30	31	32	33			

Таблица 3.10

Подстановка алфавита для шифрования полиалфавитной заменой (количество алфавитов шифротекста k = 3)

Алфавит открытого текста	А	Б	В	Г	Д	Е	Ж	З	И	Й	К	Л
Алфавит шифро-текста	1	2	3	4	5	6	7	8	9	10	11	12
	34	35	36	37	38	39	40	41	42	43	44	45
	67	68	69	70	71	72	73	74	75	76	77	78
Алфавит открытого текста	М	Н	О	П	Р	С	Т	У	Ф	Х	Ц	Ч
Алфавит шифро-текста	13	14	15	16	17	18	19	20	21	22	23	24
	46	47	48	49	50	51	52	53	54	55	56	57
	79	80	81	82	83	84	85	86	87	88	89	90
Алфавит открытого текста	Ш	Щ	Ъ	Ы	Ь	Э	Ю	Я	_ (ПРОБЕЛ)			
Алфавит шифро-текста	25	26	27	28	29	30	31	32	33			
	58	59	60	61	62	63	64	65	66			
	91	92	93	94	95	96	97	98	99			

Алгоритм зашифрования: **О + К = Ш.**

Если **(О + К) > 33**, то **Ш = О + К – 33**.

Если **(О + К) ≤ 33**, то **Ш = О + К.**

Алгоритм расшифрования: Ш – К = О.

Если **(Ш – К) > О**, то **О = Ш – К**.

Если **(Ш – К) ≤ О**, то **О = 33 + (Ш – К).**

О – открытый текст. **К** – ключ. **Ш** – шифротекст.

4.Описание программы для ЭВМ

Программа позволяет осуществлять проверку теоретических знаний студентов, выдавать варианты заданий и проверять правильность результатов зашифрования и расшифрования текста по всем вариантам заданий. Имя программы CRYPT.

Литература

1. Петраков А.В. Основы практической защиты информации. – М.: Радио и связь, 2000. – 368 с.

2. ГОСТ 28147–89. Системы обработки информации. Защита криптографическая. Алгоритм криптографического преобразования. – М.: Изд-во стандартов, 1989. – 26 с.

3. Панасенко С.П., Петренко С.А. Криптографические методы защиты информации для российских корпоративных систем // Конфидент. 2001. №5. С. 64 – 71.

Практическая работа № 4

Исследование метода компьютерной стеганографии для защиты информации

1. Цель работы

Исследование метода замены младших бит, используемого в компьютерной стеганографии для защиты информации.

2. Теоретические сведения

2.1. Основные понятия, термины и определения компьютерной стеганографии

Стеганография – это метод организации связи (передачи сообщений), при котором скрывается само наличие связи. В отличие от криптографии, где противник точно может определить, является ли передаваемое сообщение зашифрованным текстом, методы стеганографии позволяют встраивать секретные сообщения в открытые послания таким образом, чтобы было невозможным заподозрить существование самого встроенного послания [1,2].

Таким образом, если цель криптографии состоит в блокировании несанкционированного доступа к информации путём шифрования содержания секретных сообщений, то цель стеганографии – в скрытии самого факта существования секретного сообщения.

При необходимости оба способа могут быть объединены и использованы для повышения эффективности защиты информации.

Слово «*стеганография*» в переводе с греческого означает «тайнопись» («steganos» – секрет, тайна, «graphy» - запись). К ней относится большое число секретных средств связи, таких как невидимые чернила, микрофотоснимки, условное расположение знаков, средства радиосвязи на плавающих частотах и т.д.

Развитие вычислительной техники и новых каналов передачи информации привело к появлению новых методов стеганографии, в основе которых лежат особенности представления информации в компьютерных файлах, вычислительных сетях и т.п. Этот вид стеганографии получил название *компьютерной стеганографии*.

Компьютерная стеганография базируется на двух основных принципах.

Первый принцип заключается в том, что файлы, содержащие оцифрованное изображение или звук, могут быть до некоторой степени видоизменены без потери их функциональности в отличие от других типов данных, требующих абсолютной точности.

Второй принцип заключается в неспособности органов чувств человека различать незначительные изменения в цвете изображения или качестве звука. Этот принцип особенно легко применять к изображению или звуку, несущему избыточную информацию.

Так как компьютерная стеганография является молодым направлением в области защиты информации и до недавнего времени не имела своей терминологии, то в 1996 году было предложено использовать единую терминологию.

Стеганографическая система, или **стегосистема** – это совокупность средств и методов, которые используются для формирования скрытого канала передачи информации. При построении стегосистемы должны учитываться следующие положения:

методы скрытия должны обеспечивать аутентичность и целостность информации, в которой скрывается сообщение;

противник имеет полное представление о стеганографической системе и деталях её реализации. Единственной информацией, неизвестной потенциальному противнику, является ключ, с помощью которого только его держатель может установить факт присутствия и содержания скрытого сообщения;

если противник каким-то образом узнает о факте существования скрытого сообщения, то это не должно позволить ему извлечь скрытую информацию из других данных до тех пор, пока ключ хранится в тайне;

потенциальный противник должен быть лишён каких-либо технических и иных преимуществ в распознавании или раскрытии содержания скрытых сообщений.

Структурная схема стегосистемы представлена на рис. 4.1.

Сообщение – это любая информация, подлежащая скрытой передаче. В качестве сообщения может использоваться любой вид информации: текст, изображение, аудиосигнал.

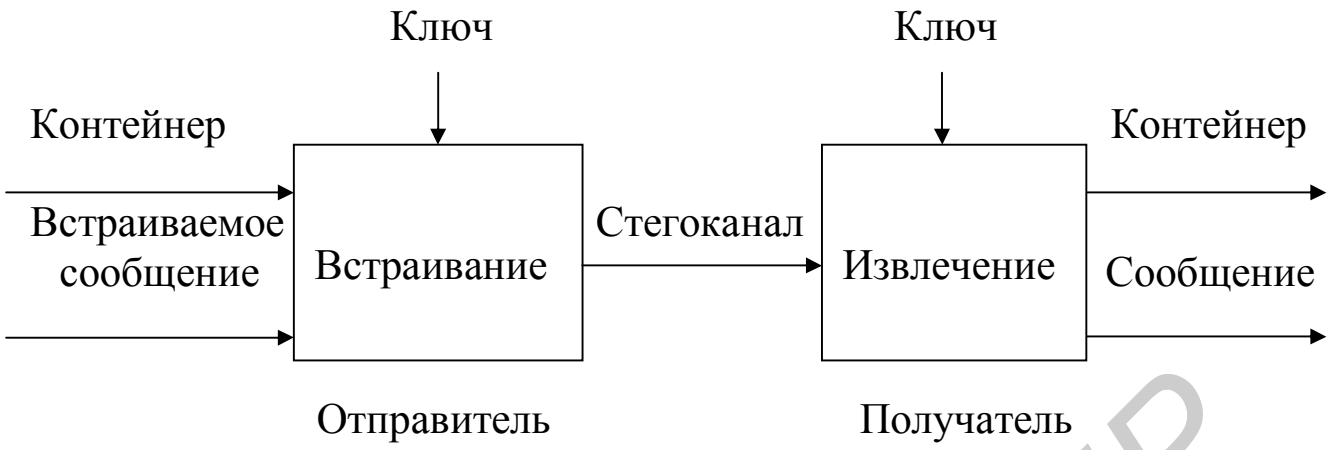

Рис.4.1. Структурная схема стегосистемы

Встроенное (скрытое) сообщение – это сообщение, встроенное в контейнер.

Контейнер – это любая информация, предназначенная для скрытия сообщения. Выбор вида контейнера оказывает существенное влияние на надёжность стегосистемы и возможность обнаружения факта передачи скрытого сообщения. По размеру (протяжённости) контейнеры можно разделить на два типа: непрерывные (потоковые) и ограниченной (фиксированной) длины.

Особенностью *потокового контейнера* является то, что невозможно определить его начало и конец. В таком контейнере биты информации, используемые для скрытия сообщения, включаются в общий поток в реальном масштабе времени и выбираются с помощью специального генератора, задающего расстояния между ними. В непрерывном потоке данных самая большая трудность для получателя – определить, когда начинается скрытое сообщение. При наличии в потоковом контейнере сигналов синхронизации или границ пакета, скрытое сообщение начинается сразу после одного из них. В свою очередь для отправителя сообщения возможны проблемы, если он не уверен в том, что поток контейнера будет достаточно длинным для размещения всего сообщения.

При использовании *контейнера ограниченной длины* отправитель заранее знает размер файла и может выбрать скрывающие биты в подходящей псевдослучайной последовательности. С другой стороны, такие контейнеры имеют ограниченный объём, и встраиваемое сообщение иногда может не поместиться в файл-контейнер. Другой недостаток заключается в том, что расстояния между скрывающими битами равномерно распределены между наиболее короткими и наи-

более длинными заданными расстояниями, в то время как истинный случайный шум будет иметь экспоненциальное распределение длин интервала. При необходимости можно задать псевдослучайные экспоненциально распределённые числа, однако этот путь наиболее трудоёмкий. На практике чаще всего используются контейнеры ограниченной длины как наиболее распространённые и доступные.

Возможны следующие варианты контейнеров:

контейнер генерируется самой стегосистемой. Такой подход называется *конструирующей стеганографией*;

контейнер выбирается из некоторого множества генерируемых стегосистемой контейнеров. Такой подход называется *селектирующей стеганографией*;

контейнер поступает извне стегосистемы. Такой подход называется *безальтернативной стеганографией*.

В зависимости от вида информации, используемой для встраивания сообщений, контейнеры могут быть визуальные, звуковые и текстовые.

Визуальный контейнер представляет собой картинку или фотографию, в которой для встраивания сообщений используются небольшие изменения яркости заранее определённых точек растра изображения.

Звуковой контейнер представляет собой речевой или музыкальный сигнал, в котором для встраивания сообщений используются младшие биты аудиосигнала, что практически не отражается на качестве звука.

Текстовый контейнер представляет собой текстовый файл, подготовленный к печати на принтере, в котором для встраивания сообщений используются небольшие изменения стандартов печати (расстояния между буквами, словами и строками, размеры букв, строк и др.).

При выборе того или иного вида контейнера необходимо иметь в виду, что при увеличении объёма встраиваемого сообщения снижается надёжность стегосистемы (при неизменном размере контейнера). Таким образом, используемый в стегосистеме контейнер накладывает ограничения на размер встраиваемого сообщения.

Пустой контейнер – это контейнер без встроенного сообщения.

Заполненный контейнер, или **стегоконтейнер** – это контейнер, содержащий встроенную информацию.

Стеганографический канал (**стегоканал**) – это канал передачи скрытого сообщения.

Ключ (**стегоключ**) – это секретный ключ, необходимый для скрытия сообщения. В зависимости от количества уровней защиты в стегосистеме может быть один или несколько стегоключей.

По аналогии с криптографией по типу стегоключа стегосистемы подразделяются на два вида: с секретным ключом и с открытым ключом.

В стегосистеме *с секретным ключом* используется один ключ, который должен быть определён либо до начала обмена секретными сообщениями, либо передан по защищённому каналу.

В стегосистеме *с открытым ключом* для встраивания и извлечения сообщения используются разные ключи, различие которых состоит в том, что с помощью вычислений невозможно определить один ключ из другого. Поэтому один ключ (открытый) может передаваться свободно по незащищённому каналу связи.

Любая стегосистема должна отвечать следующим требованиям:

свойства контейнера должны быть модифицированы таким образом, чтобы изменение невозможно было выявить при визуальном контроле;

стегосообщение должно быть устойчиво к искажениям, которые могут иметь место при его передаче, включая и различные трансформации (уменьшение, увеличение, преобразование в другой формат, сжатие без потери информации, сжатие с потерей информации и т.д.);

для сохранения целостности встраиваемого сообщения необходимо использовать коды с исправлением ошибок;

для повышения надёжности встраиваемое сообщение должно быть продублировано [2].

2.2. Методы компьютерной стеганографии

2.2.1. Классификация методов

Методы компьютерной стеганографии можно разделить в целом на два вида:

методы, основанные на избыточности визуальной и аудиоинформации;

методы, основанные на использовании специальных свойств компьютерных форматов.

Методы, основанные на избыточности визуальной и аудиоинформации, для скрытия информации используют младшие разряды цифровых отсчётов цифрового изображения и звука, которые содержат очень мало полезной информации. Их заполнение дополнительной информацией практически не влияет на качество восприятия, что и даёт возможность скрытия конфиденциальной информации.

Преимуществом этих методов является возможность скрытой передачи большого объёма информации и возможность защиты авторского права путём создания скрытого изображения товарной марки, регистрационного номера и т.п.

Недостаток метода состоит в том, что за счёт введения дополнительной информации искажаются статистические характеристики цифровых потоков. Для снижения компрометирующих признаков требуется коррекция статистических характеристик.

Методы, основанные на использовании специальных свойств компьютерных форматов, делятся на:

методы использования зарезервированных для расширения полей компьютерных форматов данных;

методы специального форматирования текстовых файлов;

методы скрытия в неиспользуемых местах гибких дисков;

методы использования имитирующих функций;

методы удаления идентифицирующего файл заголовка.

Методы использования зарезервированных для расширения полей компьютерных форматов данных основаны на том, что многие мультимедийные форматы имеют поля расширения, которые заполняются нулевой информацией и не учитываются программой. В эти поля и записывается скрываемая информация.

Методы специального форматирования текстовых файлов в свою очередь делятся на:

методы использования известного смещения строк, слов, предложений, абзацев;

методы выбора определённых позиций букв;

методы использования специальных свойств, не отображаемых на экране полей форматов.

Методы использования известного смещения строк, слов, предложений, абзацев основаны на изменении положения строк и расстановки слов в предложении, что обеспечивается вставкой дополнительных пробелов между словами.

Методы выбора определённых позиций букв используют принцип *нулевого шифра*. *Акростих* является частным случаем этого метода, когда, например, начальные буквы каждой строки образуют сообщение.

Методы использования специальных свойств, не отображаемых на экране полей форматов, основаны на использовании специальных скрытых полей для организации сносок и ссылок, например, использование чёрного шрифта на чёрном фоне.

Методы скрытия в неиспользуемых местах гибких дисков основаны на том, что скрываемая информация записывается в обычно неиспользуемых местах дисков (например в нулевой дорожке).

Методы использования имитирующих функций основаны на генерации осмысленного текста, скрывающего информацию.

Методы удаления идентифицирующего файл заголовка основаны на том, что скрываемая информация шифруется и в нем удаляется идентифицирующий заголовок, который заранее известен пользователю.

Преимуществом этих методов является простота их реализации, а недостатком – низкая степень скрытности и передача небольших объёмов информации [1].

2.2.2. Метод замены младших бит

Одним из наиболее распространённых методов стеганографии, использующих психофизические особенности человека, является метод замены младших бит информации, или **LSB**-метод (**L**east **S**ignificant **B**its). Распространённость этого метода обусловлена функциональной простотой, большой ёмкостью и высокой степенью защищённости от стегоанализа [3].

Суть метода состоит в замене нескольких младших бит в байтах данных. Он применяется в графических файлах, использующих для формирования цвета каждого элемента изображения (пиксела) значения некоторых составляющих (например, значения составляющих основных цветов – красного, зелёного и синего), или в звуковых файлах, использующих для формирования звука значения дискретизированных амплитуд сигнала.

При оцифровке изображения или звука всегда существует погрешность дискретизации, которая обычно находится на уровне младшего значащего бита. Это значит, что фактически неизвестно, что будет стоять в младшем значащем разряде цифрового представления цвета или звука. Поэтому при замене только самого младшего значащего бита говорить о каком-либо искажении изображения или звука не имеет смысла. Однако при замене только одного младшего бита такой метод имеет достаточно малую ёмкость, порядка 10% от объёма файла-контейнера, поэтому на практике используют замену более одного бита.

Рассмотрим использование данного метода на примере формата **BMP** (**B**it**M**a**P**), хранящего изображение в **True Color** (естественных цветах) и являющимся основным форматом растровой графики для системы Windows. Такие графические файлы имеют расширение BMP, однако некоторые из них могут иметь расширение **RLE** (**LE**nth enco**R**ding), что указывает на то, что произведено сжатие растровой информации, хранящейся в файле BMP-формата.

В файлах BMP информация о цвете каждого пиксела задаётся тремя байтами (1 байт = 8 битам). Каждый байт содержит одну из трёх составляющих цвета **RGB**: красную (**R**ed), зелёную (**G**reen) и синюю (**B**lue). Интенсивность каждой составляющей лежит в пределах от 0 до 255, то есть каждая составляющая имеет 256 оттенков. Варьируя интенсивность каждой составляющей, можно изменять цвет от чёрного, когда интенсивность всех составляющих равна нулю, до белого, когда интенсивность всех составляющих максимальна. При промежуточных комбинациях значений составляющих будут получаться различные хроматические (цветные) цвета и оттенки.

Максимальное количество возможных цветов составляет более 16 миллионов. Однако следует иметь в виду, что глаз человека способен различать только около 4 тысяч цветов. Для кодирования такого количества цветов достаточно всего 4 бита ($\log_2 \sqrt[3]{4000} \approx 4$).

Размер файла изображения напрямую зависит от числа пикселов и точности представления цвета. Так, 8-разрядное (1 байт) цветное изображение размером 640x480 пиксел будет занимать 300 Кбайт (640x480x1байт), а 24-разрядное изображение размером 1024x768 пиксел займёт уже 2,25 Мбайт (1024x768x3байт). (1Кбайт=2^{10}= =1024 байт. 1Мбайт=2^{10}=1024Кбайт).

Степень упаковки несущего изображения зависит от того, сколько бит младших разрядов в одном байте используется для скрытия информации.

При использовании 1 бита на байт графической информации, т.е. 3 бита на пиксел, для упаковки одного скрываемого байта используется 3 пиксела. Степень упаковки составляет 1/9 (скрываемые биты представлены единицей и выделены жирным шрифтом):

R (1 байт)	G (1 байт)	B (1 байт)	
0000000**1**	0000000**1**	0000000**1**	1 пиксел (3 байта)
0000000**1**	0000000**1**	0000000**1**	1 пиксел (3 байта)
0000000**1**	0000000**1**	00000001	1 пиксел (3 байта)

При использовании 2 бит на байт графической информации, то есть 6 бит на пиксел, для упаковки одного скрываемого байта используются 2 пиксела. Степень упаковки составляет 1/6:

R (1 байт)	G (1 байт)	B (1 байт)	
000000**11**	000000**11**	000000**11**	1 пиксел (3 байта)
000000**11**	00000000	00000000	1 пиксел (3 байта)

При использовании 3 бит на байт графической информации можно упаковать 9 бит на пиксел, но для скрытия одного байта используется 8 бит на пиксел. Степень упаковки составляет 1/3:

R	G	B	
(1 байт)	(1 байт)	(1 байт)	
00000**11**	00000**111**	00000**111**	1 пиксел (3 байта)

При использовании 4 бит на байт графической информации можно упаковать 12 бит на пиксел, но для скрытия одного байта используется 8 бит на пиксел. Степень упаковки составляет 1/3:

R	G	B	
(1 байт)	(1 байт)	(1 байт)	
0000**1111**	0000**1111**	00000000	1 пиксел (3 байта)

Если для записи скрываемой информации использовать 4 младших бита на каждый байт блока данных, то максимальное искажение цвета при этом составит 6,25% ($2^4/2^8$=16/256=0,0625). При использовании 24-разрядного изображения, которое широко применяется для описания цвета страниц в Internet в формате **HTML** (**H**iper-**T**ext **M**ar-kup **L**anguage), это искажение будет ещё меньше и составит всего $9{,}53\cdot10^{-5}$% ($2^4/2^{24}$=16/16777216=$0{,}953\cdot10^{-6}$).

Такие искажения в изображении будут практически незаметны для глаза. Однако при скрытии информации в графическом изображении необходимо учитывать, что чувствительность глаза к различным составляющим цвета неодинакова. Так, к зелёному спектру глаз более чувствителен, чем к красному и синему. Поэтому для практического использования данного метода рекомендуется скрывать в красной и синей составляющей по 3 бита, а в зелёной составляющей 2 бита информации. В этом случае в одном пикселе изображения может храниться один байт скрываемой информации, и её объём можно определить по формуле

$$\mathbf{V=W\cdot H,} \qquad (4.1)$$

где **W** - ширина изображения в пикселах;

H - высота изображения в пикселах.

Полезная ёмкость при этом составляет порядка 30%.

Следует отметить, что реально искажения изображения будут ещё меньше, так как меняются только те биты, которые не совпадают

с битами скрываемой информации. Так, для скрытия девяти бит данных, например 101101101, в 24-разрядном изображении необходимо изменить максимум 3 пиксела (9 байт). Пусть 3 пиксела 24-разрядного изображения представлены в следующем виде:

R (1 байт)	G (1 байт)	B (1 байт)	
10010101	00001101	11001001	1 пиксел (3 байта)
10010110	00001111	11001010	1 пиксел (3 байта)
10011111	00010000	11001011	1 пиксел (3 байта)

Меняя младший разряд слева направо и сверху вниз, получаем следующий результат (изменённые разряды выделены жирным шрифтом):

R (1 байт)	G (1 байт)	B (1 байт)	
10010101	0000110**0**	11001001	1 пиксел (3 байта)
1001011**1**	0000111**0**	1100101**1**	1 пиксел (3 байта)
10011111	00010000	11001011	1 пиксел (3 байта)

Таким образом, для скрытия девяти бит данных потребовалось заменить всего четыре бита только в 2 пикселах исходного изображения, что составляет менее 50 % младших разрядов, используемых для этой цели.

Такой же метод можно использовать и для скрытия информации в чёрно-белых изображениях.

2.2.3. Метод замены цветовой палитры

Метод основан на использовании специфических особенностей формата файла-контейнера и предназначен для скрытия текстовой информации в графических файлах, использующих цветовые палитры. Такими файлами являются, например, файлы BMP, PCX и GIF [3].

Палитра представляет собой некоторое число триад байт (не более 256), которые описывают цвет точки по тому же принципу, что и в файлах True Color. За палитрой следует массив байт, каждый из которых описывает одну точку изображения и содержит в себе номер цвета в палитре.

При использовании этого метода в качестве контейнера следует выбирать файлы, содержащие некоторый цвет в избытке. Например, это могут быть схемы, рисунки, текст чёрного цвета на белом фоне. Такое сочетание цветов является оптимальным для реализации данного метода, однако можно использовать и другие цвета.

В качестве примера возьмём изображение, содержащее чёрный рисунок на белом фоне. Фон занимает большую часть изображения, т. е. имеется избыток белого цвета.

Вначале создаётся алфавит, содержащий символы, которые используются в файле сообщения, например все 33 буквы русского алфавита от А до Я, цифры от 0 до 9, специальные знаки и знаки пунктуации – всего 51 символ:

Код	0	1	2	3	4	…	49	50
Символ	А	Б	В	Г	Д	…	%	;

Далее проводится замена цветов палитры. Для этого первому 51 цвету палитры назначается черный цвет (цвет рисунка). Следующему 51 цвету палитры назначается белый цвет (цвет фона). Если рисунок содержит 3…5 цветов, то можно переназначить следующие группы по 51 цвету в соответствующие цвета. Тогда в нашем случае изменённая палитра цветов будет иметь в шестнадцатеричном представлении следующий вид:

Черный цвет

Код	0	1	2	…	49	50
Цвет	00 00 00	00 00 00	00 00 00	…	00 00 00	00 00 00

Белый цвет

Код	51	52	53	…	100	101
Цвет	FF FF FF	FF FF FF	FF FF FF	…	FF FF FF	FF FF FF

Черный цвет в данном случае имеет нулевой уровень, что для одного байта соответствует 0 в десятичной системе счисления, 00000000 в двоичной и 00 в шестнадцатеричной, а белый цвет имеет уровень 225, что соответствует 225 в десятичной, 11111111 в двоичной и FF в шестнадцатеричной системе счисления. Для скрытия информации берётся первая точка изображения, анализируется её принадлежность к определённой цветовой группе, например к группе чёрного цвета, затем этой точке присваивается код текущего символа из файла-сообщения с учётом выбранной цветовой группы. Например, для символа Б чёрной точке будет назначен цвет с кодом 1, а белой точке – цвет с кодом 52.

Объём скрываемой этим методом информации определяется по формуле (2.1). Метод является самым ёмким для скрытия информации в графических файлах и позволяет оставлять изображение без изменений. Его можно использовать для любого алфавита с числом символов не более 128.

Однако информация, скрытая этим методом, легко выявляется статистическим анализом, например просмотром гистограммы графического файла в редакторе Fotoshop.

2.2.4. Метод сортировки цветовой палитры

Метод основан как на использовании особенностей формата контейнера, так и на использовании психофизических особенностей восприятия цвета человеком. При этом методе в качестве контейнера используются файлы с индексированными цветами, содержащими монохромное (обычно градации серого) изображение. Суть метода заключается в специальной предварительной подготовке файла-контейнера [3].

Палитра файла упорядочивается таким образом, чтобы цвета с соседними номерами минимально отличались друг от друга и равномерно изменялись от чёрного цвета для нулевого номера до белого цвета для 255-го номера. После этого скрываемая информация заносится в младшие разряды точек изображения. То есть искажаются не сами цвета, а номера цветов, но благодаря предварительно отсорти-

рованной палитре цвет точки заменяется на похожий, который практически невозможно отличить от исходного из-за их малого различия.

Информация, скрытая данным методом, также легко выявляется средствами программного анализа.

Описанные выше методы не увеличивают размер файла-контейнера, но их применение ограничено растровыми форматами, использующими сжатие информации без потери качества, например RLE- или LZW-сжатие.

2.2.5. Методы компьютерной стеганографии в JPEG-файлах

Цветные изображения, представленные в цифровой форме, достаточно велики и занимают большой объём памяти (до нескольких мегабайт). Поэтому для их сжатия существует ряд методов. Так, форматы BMP и GIF используют алгоритмы сжатия без потерь, обеспечивающие точное восстановление исходного изображения. Существуют также алгоритмы сжатия с потерей (искажением) информации. Таким примером может служить формат **JPEG** (**J**oint **P**hotographic **E**xperts **G**roup) [4].

В общем случае методы обработки (сжатия) изображений можно разделить на две группы: непосредственные и спектральные. При использовании *непосредственных методов* обработке подвергаются сами исходные изображения (пикселы). *Спектральные методы* основаны на применении дискретных унитарных преобразований Фурье, Адамара и др. При этом обрабатывается не исходное изображение, а соответствующие коэффициенты преобразования.

Алгоритм сжатия JPEG состоит из следующих этапов:

преобразование изображения в оптимальное цветовое пространство;

субдискретизация компонентов цветности посредством их усреднения;

применение дискретных косинусных преобразований (разновидность преобразований Фурье) для уменьшения избыточности данных изображения;

квантование коэффициентов преобразования с применением весовых функций, оптимизированных с учётом физиологических особенностей зрения;

кодирование данных изображения (результирующих коэффициентов) для удаления избыточности информации с применением алгоритма Хаффмана.

Возможность использования файловых форматов, построенных по схеме сжатия JPEG, для скрытия информации обусловлена их широким распространением при хранении и передаче графических изображений, в частности в сети Internet.

В JPEG-файлах могут использоваться следующие методы скрытия информации:

дописывание данных скрываемой информации в конец файла;

скрытие информации в косвенных данных файла;

скрытие информации с использованием таблиц квантования;

скрытие информации между блоками данных файла.

Методы скрытия информации в JPEG-файлах обладают достаточно высокой степенью защищённости от стегоанализа, так как возможность варьирования качества сжатого изображения в широком диапазоне не позволяет легко установить, являются ли возникающие в результате сжатия погрешности следствием скрытия данных или следствием использования высоких коэффициентов квантования.

2.2.6. Компьютерная стеганография в PRN-файлах

Файлы печати цветных графических изображений на принтерах, поддерживающих точечный вывод, содержат описание битовой карты отпечатка.

Процедура печати предполагает получение битовых карт отпечатков, кодирование их на языке управления принтером (PRN-файлы) и пересылку на принтер для непосредственного получения отпечатка. В этом случае в качестве контейнера стегосистемы может быть использована битовая карта отпечатка [5].

Структурная схема стегосистемы для скрытия информации в битовых картах приведена на рис.4.2.

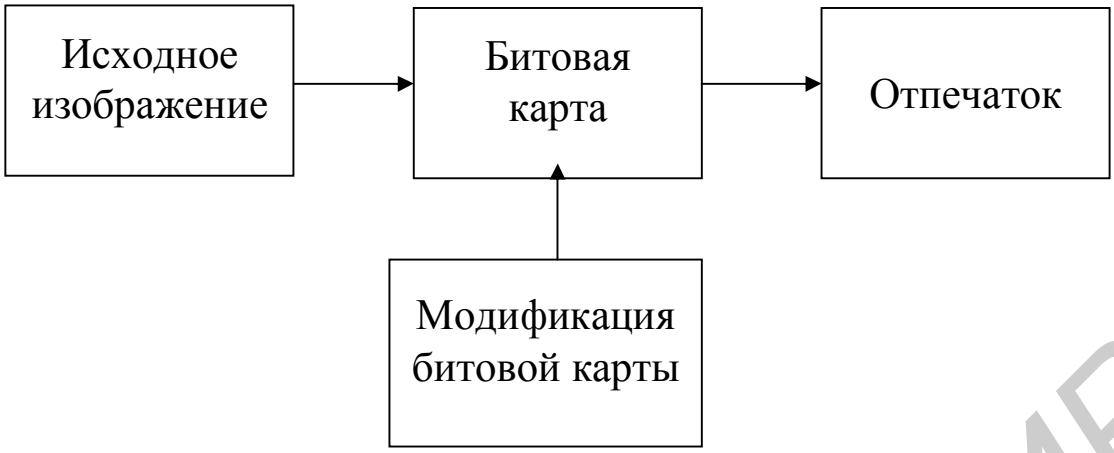

Рис. 4.2. Структурная схема стегосистемы для скрытия информации в битовых картах

До получения отпечатка битовая карта модифицируется таким образом, чтобы не позволить выявить наличие встроенного сообщения визуальным контролем самой карты и соответствующего ей отпечатка. Для скрытия информации в этом случае применяется один из методов компьютерной стеганографии, основанный на избыточности данных. В битовых картах избыточность данных создаётся за счёт большой вариантности взаимного расположения цветных пятен внутри одного и того же единичного фрагмента отпечатка изображения. Количество битовых карт равно количеству первичных красителей принтера. Каждая битовая карта, соответствующая одному первичному красителю принтера, может быть представлена прямоугольной матрицей с нулевыми и единичными компонентами, у которой 1 соответствует наличию красителя в соответствующей позиции, а 0 – его отсутствию. Совокупность битовых матриц с соответствующими значениями 0 и 1 и соответствует конкретному отпечатку. Градации цвета на отпечатке воспроизводятся с помощью подмножеств (единичных фрагментов) битовой карты – растровых точек с различной плотностью красочных пятен. Для сохранения правильного градационного воспроизведения при печати и, следовательно, для скрытия факта встраивания конфиденциальной информации, модификация битовых карт отпечатков не должна приводить к изменению плотности красочных пятен (дотов) внутри растровых точек. Вариантность взаимного расположения дотов внутри растровых точек и является ресурсом, который используется для организации стегосистемы в PRN-файлах.

Среди всех возможных узоров, отвечающих некоторой плотности заполнения растровой точки красочными пятнами, для целей сте-

ганографии можно пользоваться регуляризованными, не создающими муара растрами, то есть такими, доты которых распределены достаточно равномерно по площади растровой точки.

Следует отметить, что размеры растровой точки жёстко не определены и могут варьироваться в разумных пределах. Например, при разрешении 300 **dpi** (**d**ot **p**er **i**nch) можно использовать квадратную матрицу размерами 6x6, 7x7, 8x8 и 9x9. Форма растровой точки (прямоугольная, круглая, овальная и т.д.) также может варьироваться при встраивании информации.

Получить отпечатки со скрытой информацией можно только с помощью специального программного обеспечения, так как при печати драйвер принтера, используя фирменную технологию растрирования, «размоет» те символы скрытой информации, которые по своим размерам сравнимы с размерами растровой точки. Для извлечения скрытой информации необходимо использовать программу и алфавитные таблицы.

В заключение можно отметить, что методы компьютерной стеганографии, использующие особенности форматов файлов-контейнеров, невозможно выявить путём субъективного анализа (просмотром, прослушиванием), но достаточно легко обнаружить, используя различные программные средства стегоанализа. В то же время методы, основанные на использовании психофизических особенностей человека, невозможно выявить простым программным анализом на предмет соответствия формату и достаточно сложно, а в некоторых случаях и невозможно, обнаружить путём субъективного анализа.

3. Порядок выполнения работы

1. Изучить теоретическую часть работы.
2. Провести самопроверку теоретических знаний, ответив на поставленные вопросы.
3. Исследовать влияние количества заменяемых младших бит для различных составляющих RGB по отдельности и в их комбинации на качество воспроизведения отдельных цветов R, G и B файла-контейнера.
4. Исследовать влияние количества заменяемых младших бит для различных составляющих RGB по отдельности и в их комбинации на качество воспроизведения комбинации цветов R, G и B файла-контейнера.

5. Исследовать влияние количества заменяемых младших бит для различных составляющих RGB по отдельности и в их комбинации на качество воспроизведения различных цветов в многокомпонентной цветовой картине файла-контейнера.

6. На основании результатов, полученных в пп. 3–5, добиться наилучшего качества многоцветной картины файла-контейнера при скрытии в нём информации.

7. Оформить отчёт и защитить работу.

4. Описание программы для ЭВМ

Программа позволяет осуществлять проверку теоретических знаний студентов и проводить исследования влияния количества заменяемых младших битов составляющих цвета на качество изображения файла-контейнера при скрытии в нём информации методом замены младших бит, с использованием профессиональной программы компьютерной стеганографии. Имя программы **STEGAN**.

Литература

1. Петраков А.В., Лагутин В.С. Защита абонентского телетрафика. – М.: Радио и связь, 2001. – 504 с.

2. Генне О.В. Основные положения стеганографии. //Конфидент. 2000. № 3. С. 20-24.

3. Кустов В.Н., Федчук А.А. Методы встраивания скрытых сообщений //Конфидент. 2000. №3. С. 34-37.

4. Быков С.Ф. Алгоритм сжатия JPEG с позиции компьютерной стеганографии //Конфидент. 2000. №3. С. 26-33.

5. Архипов О.П., Архипов П.О., Зыкова З.П. Стеганография в PRN-файлах //Конфидент. 2002. №2. С. 80-83.

Св.план 2004,поз. 6

Учебное издание

Алефиренко Виктор Михайлович

Основы защиты информации

Практикум для студентов специальностей
«Техническое обеспечение безопасности»
и «Моделирование и компьютерное проектирование
радиоэлектронных средств»
дневной, вечерней и заочной форм обучения

В 2-х частях

Часть 2

Редактор Т.Н. Крюкова
Корректор Е.Н. Батурчик

Подписано в печать 7.04.2004.	Формат 60x84 1/16.	Бумага офсетная.
Гарнитура "Таймс".	Печать ризографическая.	Усл. печ. л. 2,67
Уч.-изд. л. 2,3.	Тираж 150 экз.	Заказ 635.

Издатель и полиграфическое исполнение:
Учреждение образования
“Белорусский государственный университет
информатики и радиоэлектроники”
Лицензия ЛП №156 от 30.12.2002.
Лицензия ЛВ №509 от 03.08.2001
220013, Минск, П. Бровки, 6.