Шифрование. Шифры замены

Содержание:

П.1.Введение 2

П.2. Шифры замены.....3

П.3.Примеры.....6

Задачи для самостоятельного решения.....10

Список рекомендуемой литературы 11

П.1.Введение

Итак, как мы с Вами выяснили, если вы хотите передать свое текстовое сообщение (последовательность символов некоторого алфавита) адресату так, чтобы оно осталось тайным для посторонних лиц, то у вас есть, по крайней мере, две возможности. Вы можете попытаться скрыть сам факт передачи текста, то есть прибегнуть к методам стеганографии, в арсенале которой — симпатические (невидимые) чернила, микроточки и тому подобные средства. Другая возможность заключается в попытке скрыть смысл сообщения от посторонних лиц, случайно или намеренно познакомившихся с передаваемым текстом.

Сообщение, которое вы хотите передать адресату, будем называть *открытым сообщением.* Например:

КОРАБЛИ ОТХОДЯТ ВЕЧЕРОМ

Для сохранения сообщения в тайне оно преобразуется криптографическими методами и только после этого передается адресату. Преобразованное сообщение будем называть *шифрованным сообщением* или *зашифрованным сообщением*. Другое название зашифрованного сообщения — криптограмма (или шифртекст). Зашифрованным сообщение выглядит так:

ЮПЯТБНЩМСДТЛЖГПСГХС1Щ

Зашифрованное сообщение не обязательно должно быть последовательностью букв, как в указанной выше задаче. Часто зашифрованное сообщение может представлять собой последовательность цифр или специальных знаков (например, «пляшущих человечков»).

Процесс преобразования открытого сообщения в шифрованное будем называть ***шифрованием*** или ***зашифрованием***. Адресату заранее сообщается, как из шифрованного сообщения получить открытое. Этот процесс получения исходного сообщения называют ***дешифрованием*** или ***расшифрованием.***

При выборе правила шифрования надо стремиться к тому, чтобы посторонние лица, не знающие правила расшифрования, не смогли восстановить по криптограмме открытое сообщение. В этом случае вы скроете смысл сообщения и обеспечите «тайнопись».

Для удобства дальнейшего изложения обозначим буквами

A — открытое сообщение,

B — шифрованное сообщение,

f — правило шифрования,

g — правило расшифрования.

В этом случае зашифрование открытого сообщения A в шифрованное сообщение B можно записать в виде

$$f(A) = B.$$

Обратное преобразование (то есть получение открытого сообщения *A* путем расшифрования *B*) запишется в виде соотношения

$$g(B) = A.$$

Правило зашифрования *f* не может быть произвольным. Оно должно быть таким, чтобы по шифртексту *B* с помощью правила расшифрования g можно было однозначно восстановить открытое сообщение *A*.

Среди всех шифров можно выделить два больших класса: шифры перестановки и шифры замены. При изучении первой темы сетевого семинара, мы рассматривали некоторые алгоритмы шифрования. В этой теме рассмотрим более подробно шифры замены.

П.2. Шифры замены

Наиболее известными и часто используемыми шифрами являются шифры замены. Они характеризуются тем, что отдельные части сообщения (буквы, слова...) заменяются на какие-либо другие буквы, числа, символы и т.д. При этом замена осуществляется так, чтобы потом по шифрованному сообщению можно было однозначно восстановить передаваемое сообщение.

Шифрами замены называются такие шифры, преобразования из которых приводят к замене каждого символа открытого сообщения на другие символы — шифробозначения, причем порядок следования шифробозначений совпадает с порядком следования соответствующих им символов открытого сообщения.

Пусть, например, зашифровывается сообщение на русском языке и при этом замене подлежит каждая буква сообщения. Формально в этом случае шифр замены можно описать следующим образом. Для каждой буквы *б* исходного алфавита строится некоторое множество символов $M_б$ так, что множества $M_б$ и $M_в$ попарно не пересекаются при $б \neq в$, то есть любые два различные множества не содержат одинаковых элементов. Множество $M_б$ называется множеством шифробозначений для буквы *б.* Таблица 1

а	*б*	*в*	...	*я*
$M_а$	$M_б$	$M_в$	...	$M_я$

таблица 1

является ключом шифра замены. Зная ее, можно осуществить как зашифрование, так и расшифрование.

При зашифровании каждая буква а открытого сообщения, начиная с первой, заменяется любым символом из множества $M_а$. Если в сообщении содержится несколько букв *а,* то каждая из них заменяется на любой символ из $M_а$. За счет этого с помощью одного ключа (1) можно получить различные варианты зашифрованного сообщения для одного и того же открытого сообщения.

Например, если ключом является таблица,

а	б	в	г	д	е	ж	з	и	к	л	м	н	о	п	р
21	37	14	22	01	24	62	73	46	23	12	08	27	53	35	04
40	26	63	47	31	83	88	30	02	91	72	32	77	68	60	44
10	03	71	82	5	70	11	55	90	69	38	61	54	09	84	45

с	т	у	ф	х	ц	ч	ш	щ	ъ	ы	ь	э	ю	я
20	13	59	25	75	43	19	29	06	65	74	48	36	28	16
52	39	07	49	33	85	58	80	50	34	17	56	78	64	41
89	67	93	76	18	51	87	66	81	92	42	79	86	05	57

таблица 2

то сообщение «я знаком с шифрами замены» может быть зашифровано, например, любым из следующих трех способов:

16	55	54	10	69	09	61	89	29	90	49	44	10	08	02	73	21	32	83	54	74
41	55	77	10	23	68	08	20	66	90	76	44	21	61	90	55	21	61	83	54	42
57	30	27	10	91	68	32	20	80	02	49	45	40	32	46	55	40	08	83	27	42

таблица 3

Так как множества M_0, $M_б$, $M_д$, ..., $M_я$ попарно не пересекаются, то по каждому символу шифрованного сообщения можно однозначно определить, какому множеству он принадлежит, и, следовательно, какую букву открытого сообщения он заменяет. Поэтому расшифрование возможно и открытое сообщение определяется единственным образом.

В рассказе А. Конан Доила «Пляшущие человечки» каждый символ изображает пляшущего человечка в самых различных позах

На первый взгляд кажется, что чем хитрее символы, тем труднее вскрыть сообщение, не имея ключа. Это, конечно, не так. Если каждому символу однозначно сопоставить какую-либо букву или число, то легко перейти к зашифрованному сообщению из букв или чисел.

С точки зрения криптографов использование различных сложных символов не усложняет шифра. Однако, если зашифрованное сообщение состоит из букв или цифр, то вскрывать такое сообщение удобнее.

Рассмотрим некоторые примеры шифров замены. Пусть каждое множество M_a состоит из одной буквы. Например,

а	б	в	г	д	е	ж	з	и	к	л	м	н	о	п
г	л	ь	п	д	р	а	м	ц	в	э	ъ	х	о	б

р	с	т	у	ф	х	ц	ч	ш	щ	ъ	ы	ь	э	ю	я
н	с	ж	я	и	ю	к	щ	ф	е	у	ы	ч	ш	т	а

таблица 4

Такой шифр называется шифром простой однобуквенной замены. По этому ключу удобно проводить зашифрование и расшифрование: при зашифровании каждая буква открытого текста заменяется на соответствующую букву из второй строки (а на г и т. д.) При расшифровании, наоборот, г заменяется на а и т. д. При шифровании и расшифровании надо помнить вторую строчку, то есть ключ.

Запомнить произвольный порядок букв алфавита достаточно сложно. Поэтому всегда пытались придумать какое-либо правило, по которому можно просто восстановить вторую строчку.

Другим примером шифра замены может служить лозунговый шифр. Здесь запоминание ключевой последовательности основано на лозунге — легко запоминаемом слове.

Например, выберем слово-лозунг «учебник» и заполним вторую строку таблицы по следующему правилу: сначала выписываем слово-лозунг, а затем выписываем в алфавитном порядке буквы алфавита, не вошедшие в слово-лозунг. Вторая строка в (4) примет вид:

у ч е б н и к а в г д ж з л м о
п р с т ф х ц ш щ ъ ы ь э ю я

В данном случае число вариантов ключа существенно больше букв алфавита.

Рассмотренные шифры имеют одну слабость. Если в открытом сообщении часто встречается какая-либо буква, то в шифрованном сообщении часто будет встречаться соответствующий ей символ или буква. Поэтому при вскрытии шифра замены обычно стараются наиболее часто встречающимся символам шифрованного сообщения поставить в соответствие буквы открытого сообщения с наибольшей предполагаемой частотой появления. Если шифрованное сообщение достаточно большое, то этот путь приводит к успеху, даже если вы не знаете ключа.

При анализе шифрованного сообщения следует исходить из того, что число различных вариантов для части определяемого ключа не такое уж большое, если вы находитесь на правильном пути. В противном случае либо вы получите противоречие, либо число вариантов ключа будет сильно возрастать. Обычно, начиная с некоторого момента определение открытого сообщения, становится делом техники.

Вообще-то можно сказать, что вскрытие шифров замены является искусством и достаточно трудно формализовать этот процесс. Популярные у школьников криптограммы по сути дела являются шифром замены с ключом

0 1 2 3 4 5 6 7 8 9
ш и ф р з а м е н ы

в котором каждой цифре ставится в соответствие буква. При этом должны соблюдаться правила арифметики. Эти правила значительно облегчают определение открытого текста.

Если шифрованное сообщение написано без пробелов между символами, то появляется дополнительная трудность при разбиении шифрованного

сообщения на отдельные символы и слова.

Другое направление создания шифров замены состоит в том, чтобы множества шифробозначений M_a содержали более одного элемента. Такие шифры получили название шифров многозначной замены. Они позволяют скрыть истинную частоту букв открытого сообщения, что существенно затрудняет вскрытие этих шифров. Главная трудность, которая возникает при использовании таких шифров, заключается в запоминании ключа. Надо запомнить не одну строчку, а для каждой буквы алфавита а — множество ее шифробозначений M_a. Как правило, элементами множеств M_a являются числа. Из художественной литературы и кинофильмов про разведчиков вам известно, что во время второй мировой войны часто использовались так называемые книжные шифры. Множество шифробозначений для каждой буквы определяется всеми пятизначными наборами цифр, в каждом из которых первые две цифры указывают номер страницы, третья цифра — номер строки, четвертая и пятая цифры — номер места данной буквы в указанной строке. Поэтому при поимке разведчика всегда пытались найти книгу, которая могла быть использована им в качестве ключа.

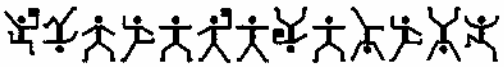
Мы не останавливаемся здесь на более сложных методах построения шифров замены. Приведенных примеров достаточно, чтобы оценить многообразие таких шифров. Но все они имеют серьезный недостаток – на одном ключе нельзя шифровать достаточно длинные сообщения. Поэтому, как правило, шифры замены используются в комбинации с другими шифрами. Чаще всего – с шифрами перестановки, о которых вы узнаете в следующей теме.

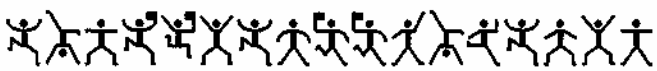
В заключение, следуя героям известных литературных произведений, вскроем некоторые шифры замены. Обратите внимание на то, какие неожиданные обстоятельства при этом используются. Действительно, вскрытие шифров - искусство.

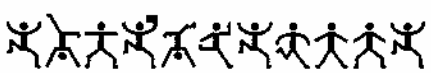
П.3.Примеры

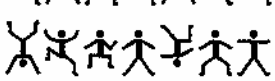
1. А. Конан Дойл, «Пляшущие человечки»

В этом рассказе Холмсу необходимо было прочитать тексты пяти записок:

I. 

II. 

III. 

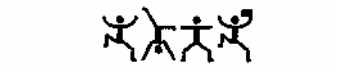
IV. 

V. 

Первая записка была так коротка, что дала возможность Холмсу сделать

всего лишь одно правдоподобное предположение, оказавшееся впоследствии правильным. По-видимому, флаги употребляются лишь для того, чтобы отмечать концы отдельных слов. Больше ничего по первой записке установить было нельзя. Четвертая записка, по всей видимости, содержала всего одно слово, так как в ней не было флагов.

Вторая и третья записки начинались, несомненно, с одного и того же слова из четырех букв. Вот это слово:



Оно кончается той же буквой, какой и начинается. Счастливая мысль: письма обычно начинаются с имени того, кому письмо адресовано. Человек, писавший миссис Кьюбит эти послания, был, безусловно, близко с ней знаком. Вполне естественно, что он называет ее просто по имени. А зовут ее Илси. Таким образом, Холмсу стали известны три буквы: И, Л и С.

В двух записках их автор обращается к миссис Кьюбит по имени и, видимо, чего-то требует от нее. Не хочет ли он, чтобы она пришла куда-нибудь, где он мог с ней поговорить? Холмс обратился ко второму слову третьей записки. В нем 7 букв, из которых третья и последняя — И. Холмс предположил, что слово это — ПРИХОДИ, и сразу оказался обладателем еще 5 букв: П, Р, Х, О, Д.

Тогда он обратился к четвертой записи, которая появилась на двери сарая. Холмс предположил, что она является ответом, и что написала ее миссис Кьюбит. Подставив в текст уже известные буквы, он получил: -И-0-Д-. Что же могла миссис Кьюбит ответить на просьбу прийти? Внезапно Холмс догадался: НИКОГДА

Возвратившись к первой записке, Холмс получил:

-Д-С- А- СЛ-НИ

Он предположил, что четвертое слово — СЛЕНИ. Это — фамилия, чрезвычайно распространенная в Америке. Коротенькое слово из двух букв, стоящее перед фамилией, по всей вероятности, имя. Какое же имя может состоять из двух букв? В Америке весьма распространено имя Аб. Теперь остается установить только первое слово фразы; оно состоит всего из одной буквы, и отгадать его нетрудно: это — местоимение Я.

Далее Холмс восстанавливает содержание второй записки:

ИЛСИ	Я	-И- -	-	-ЛРИД-А
		* 0	0	*

Здесь указаны границы слов, а снизу одинаковыми символами отмечены одинаковые буквы. Четвертое слово состоит из одной буквы (по-видимому, это союз или предлог). Буквы 0 и И уже определены, С, А и К — тоже. Остаются следующие возможности: это — либо В, либо У. Вряд ли это — В, так как в этом случае получилось бы «нечитаемое» третье слово -И-В. Поэтому, скорее всего — это предлог У. Небольшой перебор незадействованных букв дает правдоподобную гипотезу о значении третьего слова: ЖИВУ. Скорее всего,

последнее слово (-ЛРИДЖА) — мужское имя, в котором неизвестная буква — Э. Поэтому вторая записка гласит: ИЛСИ Я ЖИВУ У ЭЛРИДЖА

Холмс послал телеграмму в нью-йоркское полицейское управление с запросом о том, кто такой Аб Слени. Поступил ответ: «Самый опасный бандит в Чикаго».

Сразу после этого появилась последняя (5-я) записка, в которой не хватало трех букв: ИЛСИ ГО-ОВЬСЯ К С-ЕР-И, из которой сразу определяются буквы М и Т:

ИЛСИ ГОТОВЬСЯ К СМЕРТИ

Шестая записка была направлена Холмсом преступнику:

Пример 2.

Буквы русского алфавита занумерованы в соответствии с таблицей:

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

Для шифрования сообщения, состоящего из n букв, выбирается ключ К - некоторая последовательность из n букв приведенного выше алфавита. Зашифрование каждой буквы сообщения состоит в сложении ее номера в таблице с номером соответствующей буквы ключевой последовательности и замене полученной суммы на букву алфавита, номер которой имеет тот же остаток от деления на 30, что и эта сумма.

Прочтите шифрованное сообщение: РБЬНПТСИТСРРЕЗОХ, если известно, что шифрующая последовательность не содержала никаких букв, кроме А, Б и В.

Разбор примера.

Каждую букву шифрованного сообщения расшифруем в трех вариантах, предполагая последовательно, что соответствующая буква шифрующей последовательности есть буква А, Б или буква В:

шифрованное сообщение	Р	Б	Ь	Н	П	Т	С	И	Т	С	Р	Р	Е	З	О	Х
вариант А	П	А	Щ	М	О	С	Р	З	С	Р	П	П	Д	Ж	Н	Ф
вариант Б	О	Я	Ш	Н	Н	Р	П	Ж	Р	П	О	О	Г	Е	М	У
вариант В	Н	Ю	Ч	Л	М	П	О	Е	П	О	Н	Н	В	Д	Л	Т

Выбирая из каждой колонки полученной таблицы ровно по одной букве, находим осмысленное сообщение НАШКОРРЕСПОНДЕНТ, которое и является искомым.

Замечание. Из полученной таблицы можно найти такое исходное сообщение как

НАШ МОРОЗ ПОПОВ ЕМУ

которое представляется не менее осмысленным, чем приведенное выше. А если предположить одно искажение в шифрованном сообщении (скажем, в качестве

11-й буквы была бы принята не буква Р, а буква П), то наряду с правильным вариантом, можно получить и такой:

НАШ МОРОЗ ПОМОГ ЕМУ

Число всех различных вариантов исходных сообщений без ограничений на осмысленность равно 3^{16} или 43046721, т. е. более 40 миллионов!

Задачи для самостоятельного решения

1. Составьте алгоритм задачи: зашифровать предложение, записывая каждый его символ в виде четырехзначного числа, первые цифры которого нули, а остальные представляют код символа и реализуйте его на одном из языков программирования..
2. Напишите программу реализацию алгоритма примера 2 (см. п.2.).
3. Реализуйте на одном из языков программирования следующую задачу: шифрпреобразование простой замены в алфавите $A=\{a_1, a_2,\ldots,a_n\}$, состоящем из n различных букв, заключается в замене каждой буквы шифруемого текста буквой того же алфавита, причем разные буквы заменяются разными. Ключом шифра простой замены называется таблица, в которой указано, какой буквой надо заменить каждую букву алфавита А. Если слово СРОЧНО зашифровать простой заменой с помощью ключа:

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

то получится слово ВЗДАБД. Зашифровав полученное слово с помощью того же ключа еще раз, получим ЮШЫЧЯЫ. Сколько всего различных слов можно получить, если указанный процесс шифрования продолжать неограниченно?

4. Реализуйте на одном из языков программирования следующую задачу: Исходное сообщение, состоящее из букв русского алфавита и знака пробела (-) между словами, преобразуется в цифровое сообщение заменой каждого его символа парой цифр согласно следующей таблице:

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

Для шифрования полученного цифрового сообщения используется отрезок последовательности $C_1, C_2,\ldots, C_n,\ldots$ в которой C_n есть последняя цифра числа n^n (эта последовательность периодическая и ее наименьший период равен 20), начинающийся с некоторого C_k. При зашифровании каждая цифра сообщения складывается с соответствующей цифрой отрезка и заменяется последней цифрой полученной суммы. Восстановите сообщение: 233986721645816067061731558 8.

Список рекомендуемой литературы

1. Аршинов М.Н., Садовский Л.Е. Коды и математика, -М., Наука, 1983.
2. Домашев А.В., Попов О.В., Правиков Д.И., Прокофьев И.В., Щербаков А.Ю. Программирование алгоритмов защиты информации. Учебное пособие. - М.: «Нолидж», 2000.
3. Кушниренко А.Г. Кодирование чисел //Информатика. Приложение к «1 сентября» №23 1997.
4. Нечаев В.И. Элементы криптографии (основы теории защиты информации): учебное пособие для университетов и педвузов./Под ред. В.А. Садовничевого. - М.: Высш.шк., 1999.
5. Новиков Ф.А. Дискретная математика для программистов,- СПб: Питер, 2001.
6. Питерсон У., Уэлдон Э. Коды, исправляющие ошибки. – М.: Мир, 1976.
7. Сенокосов А.И. Символьное кодирование//Информатика. Приложение к «1 сентября» №2 январь 1997.
8. Цикоза В.А., Чурина Т.Г. Методы программирования. Ч-1,-Новосибирск 1999.
9. Ященко В.В. Введение в криптографию,- СПб: Питер, 2001г.