




Основы криптографии

Составитель: Рощупкин Александр



Введение в криптографию

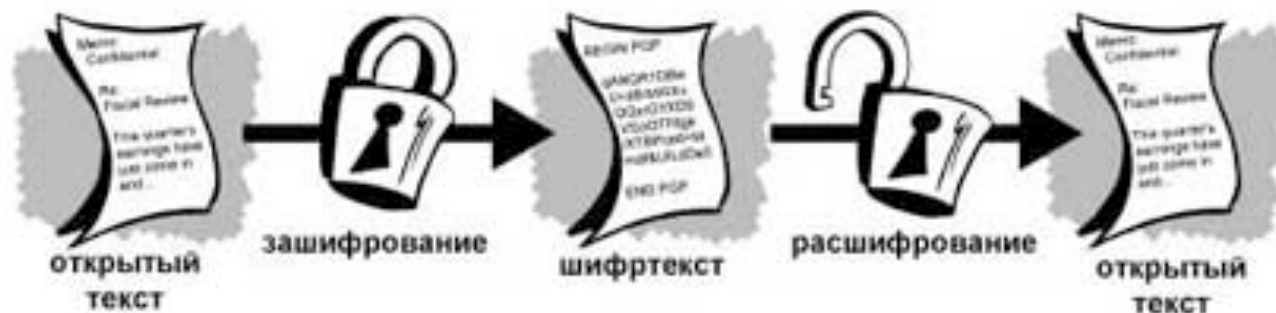
Основная задача криптографии

- * Передача нужной информации нужному адресату, в тайне от других (задача Тайны Передачи)
- * Как решать задачу ТП
 - * Создать абсолютно надёжный, недоступный для других канал связи между абонентами
 - * Использовать общедоступный канал связи, но скрыть сам факт передачи информации (стеганография)
 - * Использовать общедоступный канал связи, но передать по нему нужную информацию в таком образом преобразованном виде, что бы восстановить её мог только адресат

Основные понятия

- * **Криптография** – наука, занимающаяся преобразованием информации с целью её защиты от незаконных пользователей
- * **Открытый текст** – текст, подвергающийся преобразованию
- * **Шифро-текст** – преобразованный (зашифрованный) текст
- * **Шифр** – система обратимых преобразований, зависящая от некоторого секретного параметра (ключа) и предназначенная для обеспечения секретности передаваемой информации
- * **Шифрование** – процесс применения шифра к защищаемой информации, т.е. преобразование *открытого текста* в *шифро-текст*

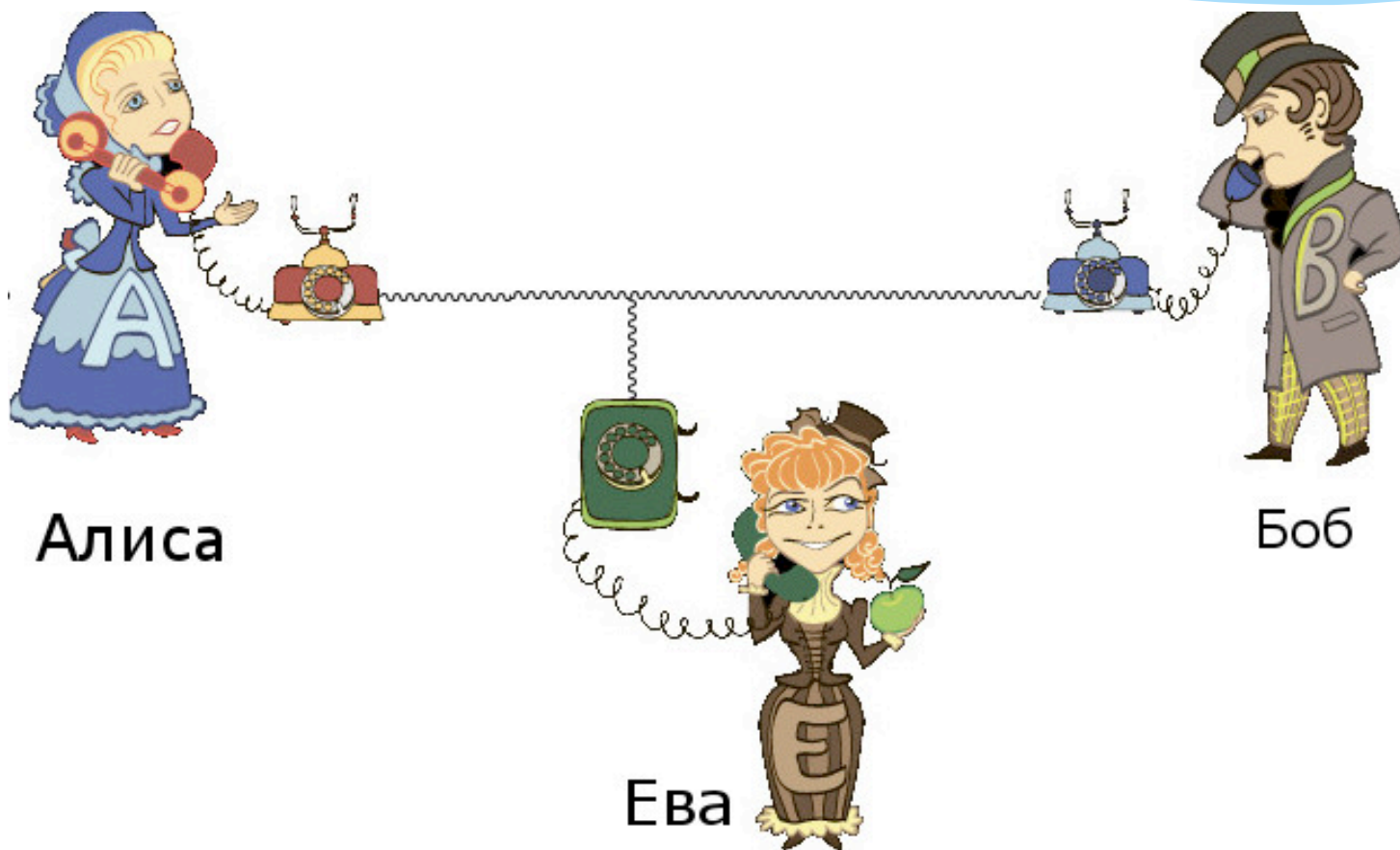
Основная ситуация шифрования



Боб, Алиса и Ева

- * Основные участники обмена сообщениями
- * Впервые двух участников процесса обмена информацией по каналу связи назвал Алисой и Бобом **Рон Ривест (Ronald Linn Rivest; род. 1947)** в своей научной статье, опубликованной в 1978 году
- * **Алиса и Боб (Alice and Bob)** — обычные пользователи, обменивающиеся сообщениями по линии связи
- * **Ева (Eve)** — пассивный злоумышленник, от англ. *eavesdropper* (подслушивающий), она может прослушивать сообщения между Алисой и Бобом, но она не может влиять на них
- * **Мэллори (Mallory, от malicious)** или **Трудди (Trudy, от intruder)** — активный злоумышленник; в отличие от Евы, Мэллори может изменять сообщения, воспроизводить старые сообщения, подменять сообщения и так далее

Перехват сообщений



Виды шифрования

- * Симметричное
 - * Поточное
 - * Блочное
 - * Моноалфавитные
 - * Полиалфавитные
- * Ассиметричные
 - * С открытым ключём
- * Цифровые подписи

Симметричное шифрование

- * Это такой подход к шифрованию, когда один и тот же ключ используется как для зашифрования, так и для расшифрования данных
- * Основные операции: *подстановки* и *перестановки*



Шифры подстановки

- * Шифры подстановки — алгоритм, где символы или байты исходного текста по таблице подстановок заменяются другими
- * Простой шифр подстановки — по таблице: берем таблицу, где написано, что А меняем на Я, Б на Ю и т. д.
- * Дальше по этой таблице шифруем, по ней же дешифруем
- * Для криптоаналитика, такой шифр представляет фактически случайную перестановку символов

Таблицы подстановки

- * Таблица показывает какую букву текста нужно заменить на какую букву шифротекста

ВМЕСТО	ПОДСТАВИТЬ
1	!
2	@
3	#
4	\$
A	%
B	^
C	&
D	*

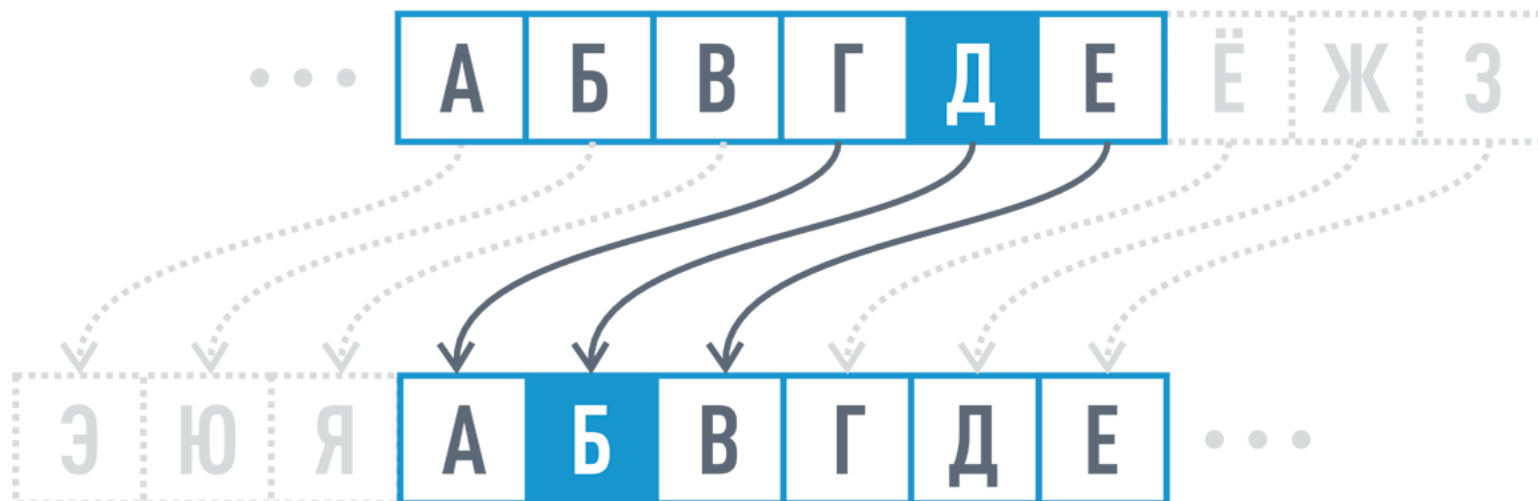
a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z
b	e	x	g	w	i	q	v	l	o	u	m	p	j	r	s	t	n	k	h	f	y	z	a	d	c

a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z
p	g	x	e	n	v	c	i	l	h	u	j	b	m	r	s	w	t	k	o	f	d	z	a	y	q

a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z
x	w	b	g	s	q	i	v	l	c	u	m	p	o	r	j	t	m	k	n	f	y	h	a	z	d

Шифр Цезаря

- * В этом шифре буквы открытого текста заменялись другими буквами алфавита, находящимися на одинаковом удалении от заменяемых букв -3



Шифр Цезаря в таблице

- * Соответствие каждого символа текста, символу криптотекста, можно записать в таблице +3
- * Такой шифр называется моноалфавитным
- * Каждая буква открытого текста заменяется на один и тот же символ шифротекста

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

Частотный анализ

- * Количество всех возможных перестановок равно 26!
- * Все шифры простой замены плохо маскируют стандартные частоты встечаемости букв в открытом тексте
- * Буквы с наибольшей частотой в криптотексте заменяются на букву с наибольшей частотой из алфавита. Вероятность успешного вскрытия повышается с увеличением длины криптотекста

№ п/п	Буква	Частотность, %	№ п/п	Буква	Частотность, %
1	О	10.97	18	Ь	1.74
2	Е	8.45	19	Г	1.70
3	А	8.01	20	З	1.65
4	И	7.35	21	Б	1.59
5	Н	6.70	22	Ч	1.44
6	Т	6.26	23	Й	1.21
7	С	5.47	24	Х	0.97
8	Р	4.73	25	Ж	0.94
9	В	4.54	26	Ш	0.73
10	Л	4.40	27	Ю	0.64
11	К	3.49	28	Ц	0.48
12	М	3.21	29	Щ	0.36
13	Д	2.98	30	Э	0.32
14	П	2.81	31	Ф	0.26
15	У	2.62	32	Ъ	0.04
16	Я	2.01	33	Ё	0.04
17	Ы	1.90			

Пример атаки

- * Ева перехватила следующий зашифрованный текст. Используя статистическую атаку, найдите исходный текст.
- * XLILSYWIMWRS AJSVWEPIJSVJSYVQMPPMSRHSPPEVWMXMWASVX-LQSVILY-VVCFIJSVIXLIWIPPIVVIGIMZIWQSVISJJIVW
- * Когда Ева составит таблицу частоты букв в этом зашифрованном тексте, она получит: $I = 14$, $V = 13$, $S = 12$, и так далее
- * Самый частый символ – I — имеет 14 появлений
- * Это показывает, что символ I в зашифрованном тексте, вероятно, соответствует символу e в исходном тексте
- * Тем самым, $ключ = 4$
- * the house is now for sale for four million dollars it is worth more hurry before the seller receives more offers

Полиалфавитные шифры

- * Полиалфавитная или многоалфавитная замена состоит из нескольких шифров простой замены
- * Основная идея многоалфавитных систем состоит в том, что на протяжении всего текста одна и та же буква может быть зашифрована по-разному
- * Это является хорошей защитой от простого подсчета частот, так как не существует единой маскировки для каждой буквы в криптотексте
- * В данных шифрах используются множественные однобуквенные ключи, каждый из которых используется для шифрования одного символа открытого текста
- * Первым ключом шифруется первый символ открытого текста, вторым - второй, и т.д

Пример полиалфавита

- * Шифр Виженера состоит из последовательности нескольких шифров Цезаря с различными значениями сдвига
- * Виженер предложил использовать в качестве ключа какой-либо другой открытый текст

base64

- * **Base64** — стандарт кодирования двоичных данных при помощи только 64 символов [ASCII](#)
- * [Алфавит кодирования](#) содержит текстово-цифровые латинские символы A-Z, a-z и 0-9 (62 знака) и 2 дополнительных символа, зависящих от системы реализации
- * Каждые 3 исходных байта кодируются 4 символами (увеличение на $\frac{1}{3}$), по 6 битов на кодируемый байт (2^6)

Алгоритм base64

- * Для того, чтобы преобразовать данные в base64, первый байт помещается в самые старшие восемь бит 24-битного буфера, следующий — в средние восемь и третий — в младшие значащие восемь бит
- * Если кодируется менее, чем три байта, то соответствующие биты буфера устанавливаются в ноль
- * Далее каждые шесть бит буфера, начиная с самых старших, используются как индексы строк «ABCDEFGHIJKLMNOPQRSTUVWXYZabcdefghijklmnopqrstuvwxyz0123456789+/>» и её символы, на которые указывают индексы, помещаются в выходную строку

Пример base64

- * В примере, слово Map закодировано как TWFu. Процесс преобразования можно представить в виде следующей таблицы:

Исходный текст	M								a								n							
Коды ASCII	77 (0x4d)								97 (0x61)								110 (0x6e)							
Двоичный вид	0	1	0	0	1	1	0	1	0	1	1	0	0	0	0	1	0	1	1	0	1	1	1	0
Полученный индекс в Base64	19								22				5				46							
Конечный результат в Base64	T								W				F				u							

Блочные шифры

- * Блочные шифры, как можно понять из названия, оперируют шифрованием блоков байтов, используя аналогичный ключ
- * Само сообщение разделяется на множество блоков, в зависимости от его длины
- * Если сообщение не подходит под длину блока, то в процессе, известном в криптографии как «дополнение», к сообщению добавляются данные до целого блока
- * Самыми распространёнными блочными шифрами, с которыми Вы столкнётесь, будут [AES](#) и [Blowfish](#)

Поточные шифры

- * Поточные шифры зашифровывают каждый знак открытого текста за один раз (обычно в виде бита) с помощью псевдослучайного потока ключей. Это значит, что для каждого бита используется другой ключ из потока
- * Математический оператор «исключающее ИЛИ» затем объединяет два бита для создания шифротекста
- * Самые используемые на сегодня поточные шифры — [RC4](#) и [Salsa20](#)

Симметричные операции

- * XOR
- * ADD
- * SHIFT

AES

- * Для исключения недостатков алгоритма DES после 2000 г. были разработаны более совершенные алгоритмы
- * Алгоритм шифрования AES (advanced encryption standard)
- * Алгоритм IDEA, разработанный европейскими криптографами алгоритм Rijndael

Проблема симметричных шифров

- * Для установления шифрованной связи с помощью симметричного алгоритма, отправителю и получателю нужно предварительно согласовать ключ и держать его в тайне
- * Если они находятся в географически удалённых местах, то должны прибегнуть к помощи доверенного посредника, например, надёжного курьера, чтобы избежать компрометации ключа в ходе транспортировки
- * Злоумышленник, перехвативший ключ в пути, сможет позднее читать, изменять и подделывать любую информацию, зашифрованную или заверенную этим ключом
- * Глобальная проблема симметричных шифров (от Кольца-декодера капитана Миднайт до DES и AES) состоит в сложности управления ключами: как вы доставите ключ получателю без риска, что его перехватят?

Ассиметричное шифрование

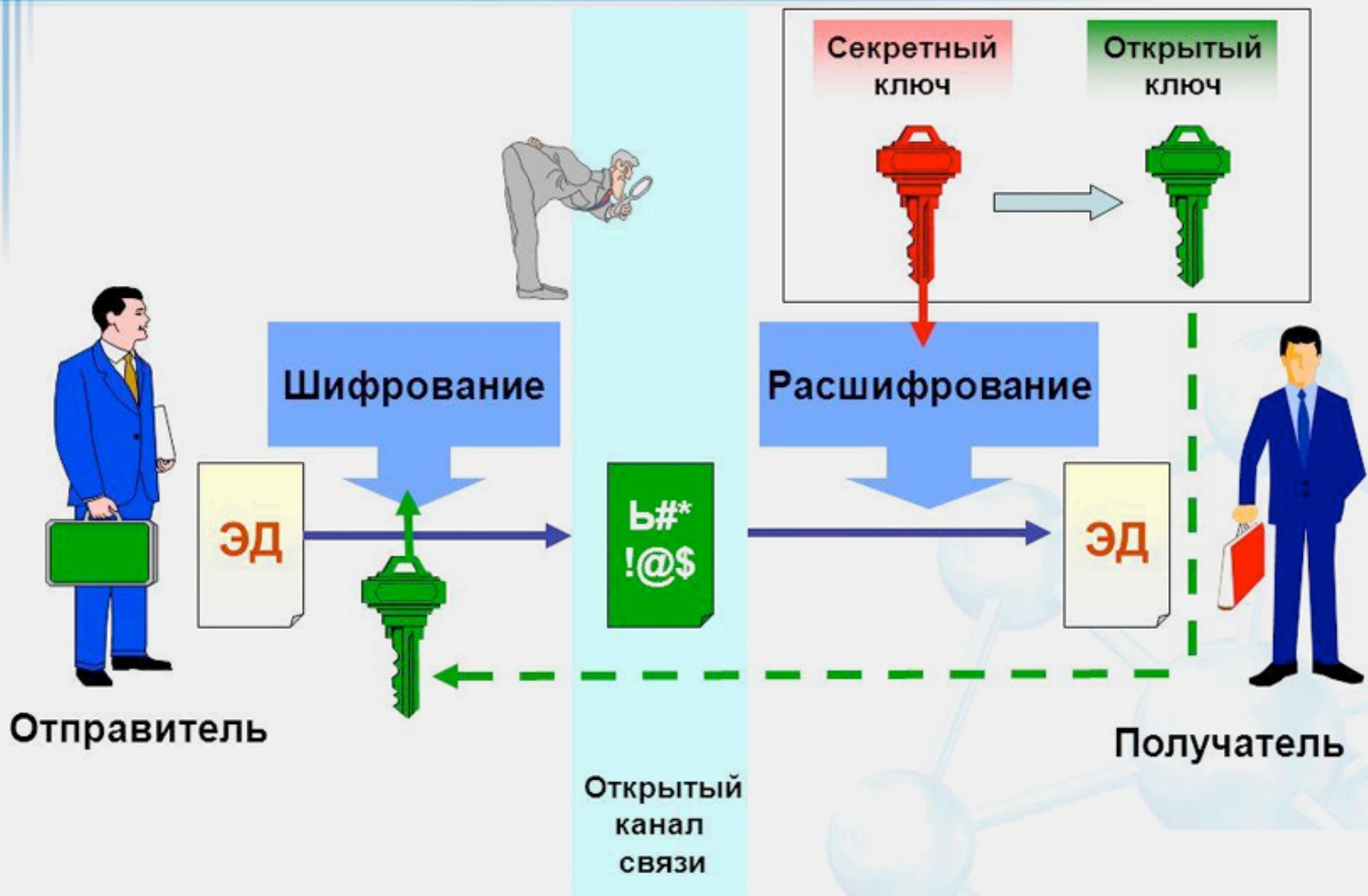
- * Проблема управления ключами была решена криптографией с открытым, или асимметричным, ключом, концепция которой была предложена Уитфилдом Диффи и Мартином Хеллманом в 1975 году
- * Криптография с открытым ключом — это асимметричная схема, в которой применяются пары ключей: открытый(public key), который зашифровывает данные, и соответствующий ему закрытый (private key), который их расшифровывает
- * Хотя ключевая пара математически связана, вычисление закрытого ключа из открытого в практическом плане невыполнимо
- * Каждый, у кого есть ваш открытый ключ, сможет зашифровать данные, но не сможет их расшифровать
- * Только человек, обладающий соответствующим закрытым ключом может расшифровать информацию

Пример

- * Зашифровав закрытым ключём, можно расшифровать открытым

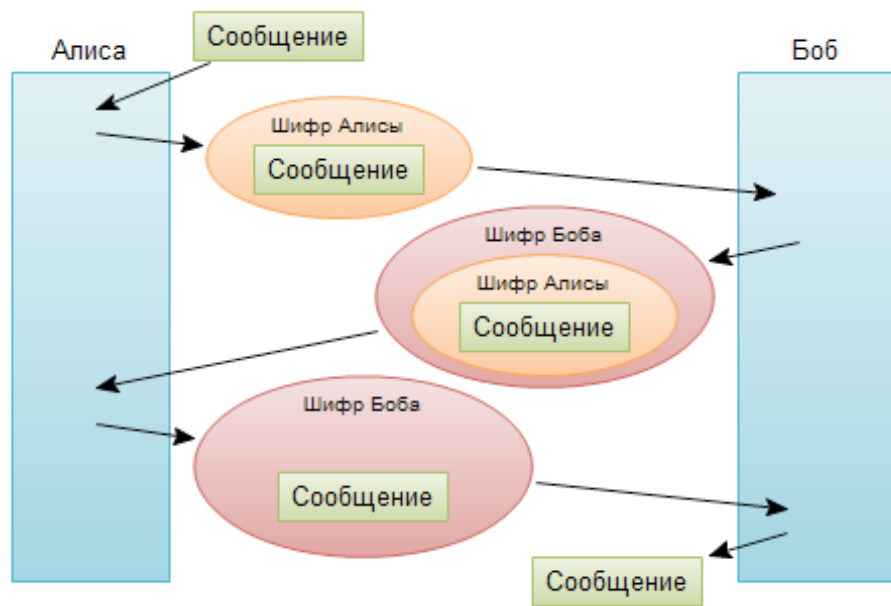


Асимметричное шифрование



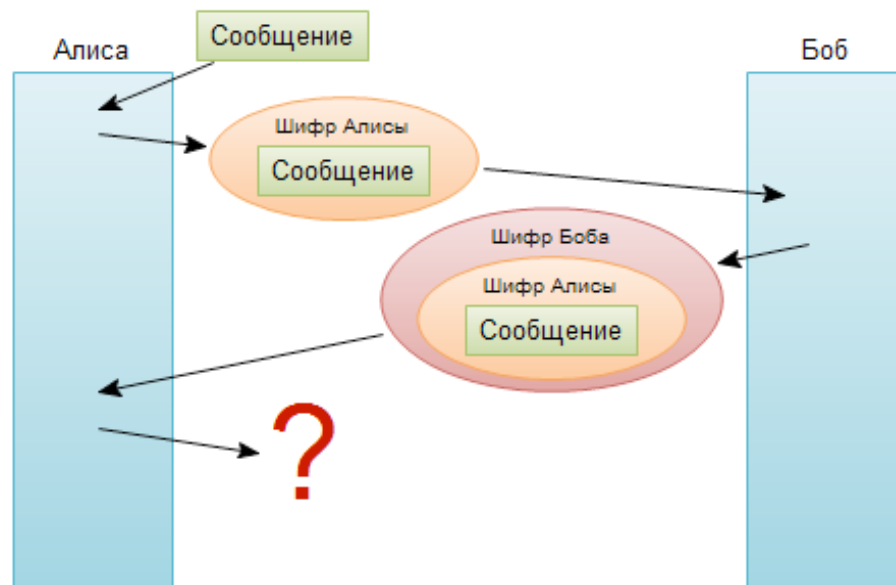
Передача сообщений без обмена ключами

- * У меня есть шкатулка. Я хочу её переслать посылкой, чтобы никто не мог открыть (без грубой силы), а приятель - мог бы
- * Я повешу на шкатулку висячий замок и отправлю её без ключа
- * Приятель, получив посылку, не сможет её открыть, но повесит туда свой собственный замок (от которого ключ только у него) и отправит обратно мне
- * Я сниму свой замок и снова отправлю приятелю
- * И теперь он получит шкатулку, которую сможет открыть он, и только он



Отправитель и принимающий шифруют свое сообщение, и затем собеседники поочередно снимают свой шифр.

Не существуют таких шифров, которые бы позволили снять шифр из под другого шифра. То есть этап где Алиса снимает свой шифр невозможен



Суть асимметричных систем

- * Асимметричные алгоритмы шифрования основаны на применении однонаправленных функций
- * Согласно определению, функция $y = f(x)$ является однонаправленной, если: ее легко вычислить для всех возможных вариантов x и для большинства возможных значений y достаточно сложно вычислить такое значение x , при котором $y = f(x)$
- * Функции с ненайденным эффективным обратным алгоритмом:
 - * Умножение и факторизация,
 - * Возведение в квадрат и извлечение квадратного корня по модулю
 - * Дискретное экспоненцирование и логарифмирование

Пример односторонней функции

- * К примеру функция удвоение – двунаправленная, т.е **удвоить(4)=8**, т.к. из результата 8 легко получить исходное значение 4
- * К примеру смешивание желтой и синей краски — пример односторонней функции
- * Смешать их **легко**, а вот получить обратно исходные компоненты — **невозможно**. Одна из таких функций в математике — **возведение в степень по модулю**

Возведение в степень по модулю

- * Возведение в степень по модулю — это вычисление остатка от деления натурального числа b (основание), возведенного в степень n (показатель степени), на натуральное число m (модуль). Обозначается:
$$c \equiv b^n \pmod{m}$$
- * Например, пусть нам даны $b = 5$, $n = 3$ и $m = 13$, тогда решение $c = 8$ — это остаток от деления $5^3 (125)$ на 13

Протокол Диффи-Хеллмана

- * **Протокол Диффи — Хеллмана** — криптографический протокол, позволяющий двум и более сторонам получить общий секретный ключ, используя незащищенный от прослушивания канал связи
- * Полученный ключ используется для шифрования дальнейшего обмена с помощью алгоритмов симметричного шифрования.
- * Схема открытого распределения ключей, предложенная Диффи и Хеллманом, произвела настоящую революцию в мире шифрования, так как снимала основную проблему классической криптографии — проблему распределения ключей

Алгоритм Диффи-Хеллмана

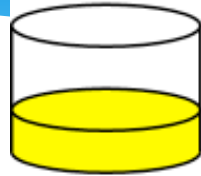
- * Хеллман предложил функцию $g^x \pmod p$
- * Обратное преобразование для такой функции очень сложно, и можно сказать что, по сути, заключается в полном переборе исходных значений
- * К примеру вам сказали, что $5^x \pmod 7 = 2$, попробуйте найдите x , а? Нашли?
- * А теперь представьте что за Y и P взяты числа порядка 10^{300}

Пример с цветами

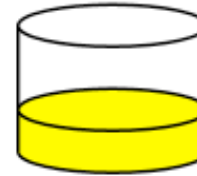
- * Допустим, мы хотим с другом загадать какой-нибудь цвет так, чтобы мы оба его знали, а никто кроме - нет, но у нас нет надёжного канала
- * У меня есть банка жёлтой краски известного объёма, и у друга тоже
- * Я загадаю какой-нибудь цвет, налью какое-то количество такой краски в банку и хорошо перемешаю. Друг тоже
- * Мы обменяемся банками, и каждый повторит с полученной банкой ту же операцию
- * В результате у каждого будет в банке одинаковый цвет, но никто, перехватив в пути банку, или даже обе, не сможет ничего поделаться

Alice

Bob



Common paint



+

+

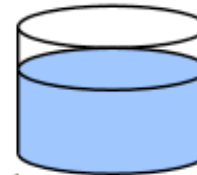


Secret colours



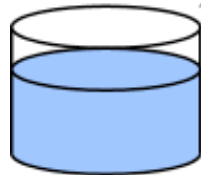
=

=



Public transport

(assume
that mixture separation
is expensive)



+

+



Secret colours



=

=



Common secret



Вычисление

- * Рассмотрим выражение $g^x \bmod p$. Числа g и p являются эквивалентом общей краски
- * Существуют некоторые ограничения относительно того, какие числа можно использовать в качестве g и x
- * Число p – простое. Алиса и Боб будут выбирать секретные числа или секретные краски (a и b соответственно). Эти числа могут быть любыми
- * Затем каждый вычисляет выражение $g^x \bmod p$ и рассказывает своему оппоненту конечный результат (смешанная краска).
- * Теперь у Алисы есть значение $B = g^b \bmod p$, полученное от Боба, а у Боба есть значение $A = g^a \bmod p$, полученное от Алисы.
- * Затем оппоненты вычисляют секретный ключ. Алиса вычисляет выражение $B^a \bmod p$, а Боб - $A^b \bmod p$

Описание

	Алиса	Боб
Этап 1	Оба участника договариваются о значениях Y и P для общей односторонней функции. Эта информация не является секретной. Допустим были выбраны значения 7 и 11 . Общая функция будет выглядеть следующим образом: $7^x \pmod{11}$	
Этап 2	Алиса выбирает случайное число, например 3 , хранит его в секрете, обозначим его как число A	Боб выбирает случайное число, например 6 , хранит его в секрете, обозначим его как число B
Этап 3	Алиса подставляет число A в общую функцию и вычисляет результат $7^3 \pmod{11} = 343 \pmod{11} = 2$, обозначает результат этого вычисления как число a	Боб подставляет число B в общую функцию и вычисляет результат $7^6 \pmod{11} = 117649 \pmod{11} = 4$, обозначает результат этого вычисления как число b
Этап 4	Алиса передает число a Бобу	Боб передает число b Алисе
Этап 5	Алиса получает b от Боба, и вычисляет значение $b^A \pmod{11} = 4^3 \pmod{11} = 64 \pmod{11} = 9$	Боб получает a от Алисы, и вычисляет значение $a^B \pmod{11} = 2^6 \pmod{11} = 64 \pmod{11} = 9$
Этап 6	Оба участника в итоге получили число 9 . Это и будет являться ключом.	

Алгоритм

- * Алиса и Боб выбирают два числа g и p
- * Алиса и Боб выбирают секретные числа (a и b соответственно)
- * Алиса и Боб вычисляют $g^x \bmod p$, где x – секретное число
- * Алиса и Боб обмениваются полученными данными (числа A и B соответственно)
- * Алиса и Боб используют полученное число и секретное число для вычисления общего

Ассиметричные алгоритмы

- * **Elgamal** (названная в честь автора, Тахира Эльгамала)
- * **RSA** (названная в честь изобретателей: Рона Ривеста, Ади Шамира и Леонарда Адлмана)
- * **Diffie-Hellman** (названная, правильно, в честь её создателей)
- * **DSA**, Digital Signature Algorithm (изобретённый Дэвидом Кравицом)



Дополнительные материалы

RSA – открытый ключ

- * Выбираю два простых числа. Пусть это будет $p=3$ и $q=7$.
- * Вычисляем модуль — произведение наших p и q : $n=p \times q = 3 \times 7 = 21$.
- * Вычисляем функцию Эйлера: $\varphi=(p-1) \times (q-1) = 2 \times 6 = 12$.
- * Выбираем число e , отвечающее следующим критериям: (i) оно должно быть простым, (ii) оно должно быть меньше φ — остаются варианты: 3, 5, 7, 11, (iii) оно должно быть взаимно простым с φ ; остаются варианты 5, 7, 11. Выберем $e=5$. Это, так называемая, *открытая экспонента*.
- * Теперь пара чисел $\{e, n\}$ — это мой открытый ключ

RSA – закрытый ключ

- * Мне нужно вычислить число d , обратное e по модулю φ .
- * То есть остаток от деления по модулю φ произведения $d \times e$ должен быть равен 1
- * Запишем это в обозначениях, принятых во многих языках программирования: $(d \times e) \% \varphi = 1$
- * Или $(d \times 5) \% 12 = 1$. d может быть равно 5 ($(5 \times 5) \% 12 = 25 \% 12 = 1$), но чтобы оно не путалось с e в дальнейшем повествовании, давайте возьмём его равным 17
- * Можете проверить сами, что $(17 \times 5) \% 12$ действительно равно 1 ($17 \times 5 - 12 \times 7 = 1$)
- * Итак $d = 17$. Пара $\{d, n\}$ — это секретный ключ

RSA - шифрование

- * Возводите ваше сообщение в степень e по модулю n . То есть, вычисляете 19 в степени 5 (2476099) и берёте остаток от деления на 21
- * Получается 10 — это ваши закодированные данные.
- * Строго говоря, вам вовсе незачем вычислять огромное число « 19 в степени 5 »
- * При каждом умножении достаточно вычислять не полное произведение, а только остаток от деления на 21 . Но это уже детали реализации вычислений, давайте не будем в них углубляться

RSA - расшифровка

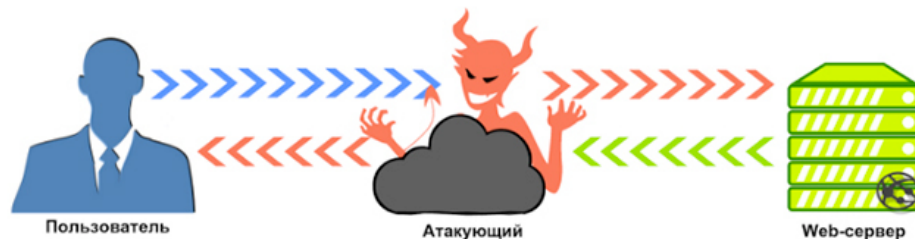
- * Я получил ваши данные ($E=10$), и у меня имеется закрытый ключ $\{d, n\} = \{17, 21\}$
- * Обратите внимание на то, что открытый ключ не может расшифровать сообщение
- * Я делаю операцию, очень похожую на вашу, но вместо e использую d
- * Возвожу E в степень d : получаю 10 в степень 17
- * Вычисляю остаток от деления на 21 и получаю 19 — ваше сообщение

Цифровые подписи

- * Цифровая подпись позволяет получателю сообщения убедиться в аутентичности источника информации (иными словами, в том, кто является автором информации), а также проверить, была ли информация изменена (искажена), пока находилась в пути
- * Кроме того, ЭЦП несёт принцип неотречения, который означает, что отправитель не может отказаться от факта своего авторства подписанной им информации

Атака «Человек посередине»

- * В этом виде атак злоумышленник подсовывает пользователю собственный ключ, но с именем предполагаемого адресата
- * Данные зашифровываются подставным ключом, перехватываются его владельцем-злоумышленником, попадая в итоге в чужие руки
- * В среде криптосистем с открытым ключом критически важно, чтобы вы были абсолютно уверены, что открытый ключ, которым собираетесь что-то зашифровать — не искусная имитация, а истинная собственность вашего корреспондента
- * Цифровые сертификаты ключей упрощают задачу определения принадлежности открытых ключей предполагаемым владельц



Цифровые сертификаты

- * Одна из главных проблем асимметричных криптосистем состоит в том, что пользователи должны постоянно следить, зашифровывают ли они сообщения истинными ключами своих корреспондентов
- * В среде свободного обмена открытыми ключами через общественные серверы-депозитарии атаки по принципу "человек посередине" представляют серьёзную потенциальную угрозу

Цифровые сертификаты

- * Цифровой сертификат состоит из трёх компонентов:
- * открытого ключа, к которому он приложен;
- * данных, или записей, сертификата (сведения о личности пользователя, как то, имя, электронная почта и т. п., а также, по необходимости, дополнительные ограничительные сведения: права допуска, рабочие лимиты и прочее);
- * одной или нескольких цифровых подписей, "связывающих" ключ с сертификатом.

TLS

- * **TLS** (*transport layer security*) — Протокол защиты транспортного уровня, как и его предшественник SSL, — криптографические протоколы, обеспечивающие защищённую передачу данных между узлами в сети
- * TLS и SSL используют асимметричное шифрование для аутентификации, симметричное шифрование для конфиденциальности и коды аутентичности сообщений для сохранения целостности сообщений

Blockchain

- * Blockchain – это технология, использующая последовательный набор блоков (или же, в более общем случае, ориентированный граф), каждый следующий блок в котором включает в качестве хешируемой информации значение хеш-функции от предыдущего блока
- * Выделяют 3 группы лиц
 - * источники событий (транзакций)
 - * источники блоков (фиксаторы транзакций)
 - * получатели (читатели) блоков и зафиксированных транзакций