Практика №11. Шифрование и дешифровка сообщения шифром RSA (реализация алгоритма для маленьких простых чисел).

ТЕОРЕТИЧЕСКАЯ ЧАСТЬ.

В 1978 г. появилась работа [2], в которой Рон Райвест (Ron Rivest), Ади Шамир (Adi Shamir) и Лен Адлеман (Len Adleman) предложили алгоритм с открытым ключом. Схема Райвеста–Шамира–Адлемана (RSA) получила широкое распространение.

Опишем процесс шифрования. Исходный текст должен быть переведен в числовую форму, этот метод считается известным. В результате этого текст представляется в виде одного большого числа. Затем полученное число разбивается на части (блоки) так, чтобы каждая из них была числом в промежутке
[0, N – 1] (о выборе N — см. ниже). Процесс шифрования одинаков для каждого блока. Поэтому мы можем считать, что блок исходного текста представлен числом x,
$0 \le x \le N-1$.

Каждый абонент вырабатывает свою пару ключей. Для этого он генерирует два больших простых числа p и q, вычисляет произведение $N = p \cdot q$. Затем он вырабатывает случайное число e, взаимно простое со значением функции Эйлера от числа N, $\varphi(N) = (p-1)\cdot(q-1)$ и находит число d из условия

$e \cdot d = 1 (\mathrm{mod}\ \varphi(N))$. Так как $\left(e, \varphi(N)\right) = 1$, то такое число d существует и оно единственно.

Пару (N, e) он объявляет открытым ключом и помещает в открытый доступ. Пара (N, d) является секретным ключом. Для расшифрования достаточно знать секретный ключ. Числа p, q, $\varphi(N)$ в дальнейшем не нужны, поэтому их можно уничтожить.

Пользователь A, отправляющий сообщение x абоненту B, выбирает из открытого каталога пару (N, e) абонента B и вычисляет шифрованное сообщение $y = x^e \left(\mathrm{mod}\ N\right)$.

Чтобы получить исходный текст, абонент В вычисляет $y^d \left(\mathrm{mod}\ N\right)$. Так как $e \cdot d \equiv 1 (\mathrm{mod}\ \varphi(N))$, т. е. $e \cdot d = \varphi(N) \cdot k + 1$, где k – целое, то применяя теорему Эйлера, получим: следующее соотношение:

$$y^d \equiv (x^e)^d \equiv x^{ed} \equiv x^{\varphi(N)\cdot k+1} \equiv (x^{\varphi(N)})^k \cdot x \equiv x (\mathrm{mod}\ N).$$

Пример 1. Пусть p = 7, q = 17. Тогда N = 7·17 = 119, $\varphi(N) = 96$. Выбираем значение e: $e < 96$, $(e, 96) = 1$. Пусть в нашем случае e = 5. Находим d: $d = 1/e \ (\mathrm{mod}\ 96)$. Получаем d = 77, так как 77·5 = 4·96 + 1. Открытый ключ (119,5), личный ключ (119,77). Пусть x = 19. Для зашифрования число 19 возводим в 5-ю степень по модулю 119, тогда имеем 19^5 = 2 476 099 и остаток от деления 2 476 099 на 119 равен 66. Итак, $y = 19^5 \bmod 119 = 66$, а расшифрование $x = 66^7 \bmod 119 = 19$.

Подготовка текста к шифрованию

Сначала нужно каким-либо способом представить текст сообщения в виде упорядоченного набора чисел по модулю N. Это еще не процесс шифрования, а только подготовка к нему.

Пример 2. Для простоты предположим, что текст сообщения содержит слова, записанные только заглавными буквами. Первый шаг состоит в замене каждой буквы сообщения числом. Пусть наша таблица замен имеет вид:

А	Б	В	Г	Д	Е	Ж	З	И	Й	К	Л	М	Н	О	П	Р	С
10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27

Т	У	Ф	Х	Ц	Ч	Ш	Щ	Ъ	Ы	Ь	Э	Ю	Я
28	29	30	31	32	33	34	35	36	37	38	39	40	41

Пробел между словами будем заменять числом 99.

Например, пусть открытый текст – это девиз «ПОЗНАЙ СЕБЯ». Тогда его цифровое представление имеет вид: 2524172310199927151141.

Пусть в нашем примере p = 149, q = 157, тогда N = 23393. Поэтому цифровое представление открытого текста нужно разбить на блоки, меньшие, чем 23393. Одно из таких разбиений выглядит следующим образом:

2524 – 1723 – 10199 – 9271 – 511 – 41.

Конечно, выбор блоков неоднозначен, но и не совсем произволен. Например, во избежание двусмысленностей, на стадии расшифровки не следует выделять блоки, начинающиеся с нуля.

При расшифровке сообщения получаем последовательность блоков, затем их соединяем вместе и получаем число. После этого числа заменяют буквами в соответствии с таблицей, приведенной выше.

Обратим внимание на то, что в этом примере каждую букву кодируем двузначным числом. Это сделано для предотвращения неоднозначности. Если бы мы пронумеровали буквы не по порядку, начиная с 1, т. е. А соответствует 1, Б соответствует 2 и т. д., то было бы непонятно, что обозначает блок 12: пару букв АБ или букву Л, двенадцатую букву алфавита. Конечно, для кодирования можно использовать любые однозначные соответствия между буквами и числами, например ASCII-кодировку, что чаще всего это и делается.

Продолжим пример: выбираем p = 149, q = 157, вычисляем $\varphi(N) = 23\,088$. Теперь нужно выбрать число e, взаимно простое с $\varphi(N)$. Наименьшее простое, не делящее $\varphi(N)$, равно 5. Положим e = 5. Зашифруем первый блок сообщения: вычисляем $2524^5 \bmod 23393 = 22752$; далее $1723^5 \bmod 23393 = 6198$.

$10199^5 \bmod 23393 = 14204$,
$9271^5 \bmod 23393 = 23191$,
$511^5 \bmod 23393 = 10723$,
$41^5 \bmod 23393 = 14065$.

Теперь шифрованный текст имеет вид

22752619814204231911072314065

В нашем примере N = 23393, e = 5. Применив алгоритм Эвклида к числам $\varphi(N) = 23088$ и e = 5, найдем $d = e^{-1} \bmod 23088 = 13853$. Значит для расшифровки блоков шифртекста мы должны возвести этот блок в степень 13583 по модулю 23393. В примере первый блок шифртекста – число 22752, тогда получим $22752^{13853} \bmod 23393 = 2524$.

Разбиение числа на блоки можно произвести различными способами. При этом *промежуточные* результаты зависят от способа разбиения, однако *конечный* результат – не зависит.

ЗАДАНИЕ

Сгенерировать два простых числа. Получить открытый и закрытый ключи алгоритма RSA.

Подготовить открытый текст для шифрования и разбить его на блоки.

Провести шифрование текста

Расшифровать текст