Министерство образования и науки Российской Федерации

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИМЕНИ Н.Г.ЧЕРНЫШЕВСКОГО»

Кафедра компьютерной алгебры и теории чисел

Система RSA и задача разложения

название темы выпускной квалификационной работы

АВТОРЕФЕРАТ МАГИСТЕРСКОЙ РАБОТЫ

студента (ки) 2 курса 227 группы

направления 02.04.01 «Математика и компьютерные науки»
код и наименование направления

механико-математического факультета
наименование факультета

Кучер Алены Андреевны
фамилия, имя, отчество

Научный руководитель
доцент, к.ф.-м.н _____ В.В. Кривобок
должность, уч. степень, уч. звание подпись, дата инициалы, фамилия

Зав. кафедрой:
к.ф.-м.н., доцент _____ А.М. Водолазов
должность, уч. степень, уч. звание подпись, дата инициалы, фамилия

Саратов 2017 г.

Введение Проблемой защиты информации путем ее преобразования занимается криптология (kryptos — тайный, logos — наука). Криптология разделяется на два направления — криптографию и криптоанализ. Цели этих направлений прямо противоположны. Криптография занимается поиском и исследованием математических методов преобразования информации. Сфера интересов криптоанализа — исследование возможности расшифровывания информации без знания ключей.

Тема является актуальной, так как в настоящее время благополучие многих людей зависит от обеспечения информационной безопасности множества компьютерных систем обработки информации, а также контроля и управления различными объектами. К таким объектам можно отнести системы телекоммуникаций, банковские системы, атомные станции, системы управления воздушным и наземным транспортом, а также системы обработки и хранения секретной и конфиденциальной информации. Для нормального и безопасного функционирования этих систем необходимо поддерживать их безопасность и целостность.

В настоящее время многие системы шифрования основываются на криптосистеме RSA. В своей работе я подробно описала алгоритм шифрования RSA, различные методы разложения чисел, привела примеры разложения.

Целью моей работы является рассмотрение алгоритма шифрования RSA, задачи разложения на множители, построение алгоритма шифрования на примере.

Задачи моей магистерской работы:

1. Анализ основных свойств криптосистемы RSA.

2. Анализ некоторых методов разложения.

3. Построение алгоритма шифрования RSA на примере.

Схема RSA представляет собой блочный шифр, в котором и открытый текст, и шифрованный текст представляются целыми числами из диапазона от 0 до $n-1$, для некоторого n.

Данная магистерская работа состоит из введения, трех разделов, заключения и списка литературы. Первый раздел состоит из двух подразделов. В ней показаны основные системы шифрования информации с открытым ключом, подробно рассмотрена система шифрования RSA. Второй раздел состоит так-

же из двух подразделов. Он включает в себя различные методы разложения и криптографические применения разложения целых чисел на множители. В третьем разделе подробно рассмотрены некоторые виды атак на криптосистему RSA.

В данной работе решены некоторые примеры с помощью алгоритмов разложения чисел на множители.

Основное содержание работы При подборе функции с замком f для системы с открытым ключом желательно, чтобы идея шифрования была проста концептуально и несложна в реализации. Но, с другой стороны, нужно иметь достаточное эмпирическое обоснование, основанное на многолетних попытках построения алгоритма для f^{-1}, свидетельствующее о том, что дешифрование не осуществимо без знания секретного ключа дешифрования. По этой причине естественно обратить внимание на старинную проблему теории чисел — задачу полной факторизации большого составного целого числа при неизвестных заранее простых множителях. Успех так называемой криптосистемы «RSA» (названной по именам ее создателей: Rivest, Shamir и Adleman), являющейся одной из самых старых и самых популярных криптосистем с открытым ключом, определен чрезвычайной трудностью этой задачи.

Как «работает» криптосистема RSA? Сначала каждый пользователь выбирает два очень больших простых числа p и q и вычисляет $n = pq$. Зная факторизацию числа n, несложно вычислить $\phi(n) = (p-1)(q-1) = n+1-p-q$. Затем пользователь выбирает случайно целое число e между 1 и $\phi(n)$, которое взаимно просто с $\phi(n)$.

Когда говорится «случайно», то всегда подразумевается, что число выбрано с помощью датчика случайных (или псевдослучайных) чисел, то есть компьютерной программы, генерирующей последовательность цифр, которую никто не может повторить или предугадать, и которая, по-видимому, имеет те же статистические свойства, что и истинно случайная последовательность. В системе RSA генератор случайных чисел используется не только для выбора e, но и для выбора больших простых чисел p и q. Что понимается под «случайно выработанным» простым числом? Сначала вырабатывается большое случайное целое число m. Если m четное, то оно заменяется на $m + 1$. Потом, чтобы проверить, является ли это нечетное число простым,

используется тест на простоту числа. Если число m не является простым, то проверяются $m+2$, потом $m+4$ и т. д., пока не будет получено первое простое число, превосходящее m, которое и берется в качестве «случайного» простого. Поскольку по теореме о простых числах доля простых среди целых вблизи m составляет примерно $1/\log m$, можно ожидать, что для нахождения первого простого, большего или равного m, потребуется проверить $O(\log m)$ чисел.

Подобным образом вырабатывается «случайное» число e, взаимно простое с $\phi(n)$. Сначала вырабатывается случайное (нечетное) целое число подходящего размера, которое последовательно увеличивается, пока не будет найдено e с НОД $(e(\phi(n))) = 1$.

Каждый пользователь A выбирает два простых числа p_A и q_A, а вслед за этим — случайное число e_A, которое не имеет общих множителей с $(p_A - 1)(q_A - 1)$. Далее, A вычисляет $n_A = p_A q_A$, $(\phi(n_A)) = n_A + 1 - p_A - q_A$ и число, обратное относительно умножения к e_A по модулю $(\phi(n_A))$: $d_A = e_A^{-1}\ (mod\ \phi(n_A))$. Ключ шифрования $K_{E,A} = (n_A, e_A)$ делается открытым, а ключ дешифрования $K_{D,A} = (n_A, d_A)$ — секретным. Шифрующее преобразование — это отображение $Z/n_A Z$ в себя по формуле $f(P) \equiv P^{e_A}\ (mod\ n_A)$. Дешифрующее преобразование — это отображение $Z/n_A Z$ в себя по формуле $f^{-1}(C) \equiv C^{d_A}\ (mod\ n_A)$. Нетрудно заметить, что согласно выбору d_A эти два отображения взаимно обратны. А именно, последовательное применение в любом порядке f и f^{-1} приводит к возведению в степень $d_A e_A$. Поскольку $d_A e_A$ дает при делении на $\phi(n_A)$ остаток 1, это эквивалентно возведению в первую степень.

Можно определить задачу разложения на множители числа n (над кольцом $\mathbb{Z}$) как нахождение всех простых делителей числа n. Аналоги задачи разложения существуют для любого факториального кольца.

Лемма 2.1. Задача разложения числа $n = pq$ и задача вычисления функции Эйлера $\phi(n)$ полиномиально эквивалентны.

Теорема 2.1. Задача вычисления секретного показателя d сводится с полиномиальной сложностью к задаче вычисления функции $\phi(n)$.

В общем случае задача дешифрования системы RSA эквивалентна задаче извлечения корня степени e в кольце $\mathbb{Z}/n\mathbb{Z}$.

Если составное число заданной длины имеет более двух различных простых делителей, то задача разложения облегчается. Это обусловлено тем, что число различных представлений числа n в виде произведения двух (не обязательно простых) сомножителей растет экспоненциально с ростом числа различных простых делителей. Если число различных простых делителей составного числа n равно k, то таких представлений ровно $2^{k-1}-1$. Например, в случае четырех простых делителей возможны следующие семь представлений:

$$n = a(bcd) = b(acd) = c(abd) = d(abc) = (ab)(cd) = (ac)(bd) = (ad)(bc).$$

Задача разложения в этом случае решается рекурсивно: число n раскладывается на два множителя и для каждого из них решается задача разложения. Поскольку размер множителей меньше, чем размер первоначального числа, то для них задача разложения является более простой, чем первоначальная.

Сложность задачи разложения по отношению к ряду методов определяется размером минимального простого делителя. Чем больше число простых делителей, тем меньше размер минимального из них, а значит, тем проще выполняется разложение. Поэтому сложность разложения составного числа обычно оценивается для случая, когда это число содержит два различных простых делителя.

Если в системе RSA у двух пользователей общее составное число n, но различные открытые e_1, e_2 и секретные d_1, d_2 ключи, причем $e_1 d_1 \equiv e_2 d_2 \equiv 1\ (mod\ \phi(n))$, то каждый пользователь может найти разложение числа n и тем самым узнать секретный ключ другого пользователя. Для этого достаточно найти значение квадратного корня из единицы, отличное от ± 1. Если $t^2 \equiv 1\ (mod\ n), t \neq \pm 1$, то $t^2 - 1 = (t+1)(t-1) \equiv 0\ (mod\ n)$, и $(t+1), (t-1)$ имеют нетривиальные общие делители с n. Разложение можно выполнить по аналогии с псевдопростым тестом Миллера-Рабина следующим алгоритмом.

Алгоритм 2.1. Разложение составного числа на множители по известным показателям RSA.

Вход. Число n, показатели e, d такие, что — $ed \equiv 1\ (mod\ \phi(n))$.

Выход. Делители p и q числа n.

Метод.

1. Положить $N \leftarrow ed - 1$. Представить N в виде $N = (2^f)s$, где s — нечетное число.

2. Выбрать случайное a и вычислить $b \leftarrow a^s \ (mod\ n)$.

3. Вычислять $b^{2^0} \equiv b \ (mod\ n), b^{2^1} \equiv (b^{2^0})^2 \ (mod\ n), b^{2^2} \equiv (b^{2^1})^2 \ (mod\ n), ...$ до получения m такого, что $b^{2^m} \equiv 1 \ (mod\ n)$. Если $b^{2^{m-1}} \equiv -1 \ (mod\ n)$, то вернуться на шаг 2, иначе положить $t \leftarrow b^{2^{m-1}} \ (mod\ n)$.

Метод Ферма. Для любого положительного нечетного числа n существует взаимно однозначное соответствие между множеством делителей числа n, больших чем $\sqrt{n}$, и множеством пар неотрицательных чисел $\{s, t\}$ таких, что $n = s^2 - t^2$, а именно, если $n = pq$, где $p \geq q$, то $s = (p+q)/2, t = (p-q)/2$ (обратно, если $n = s^2 - t^2 = (s+t)(s-t)$, то $p = s + t, q = s - t$).

Если числа p и q близки друг к другу (что характерно для системы RSA), то число t мало, а значит, s немного больше, чем $\sqrt{n}$. В этом случае можно найти p и q, последовательно перебирая числа $s = \lfloor\sqrt{n}\rfloor + 1, \lfloor\sqrt{n}\rfloor + 2, ...$ до тех пор, пока не получится s такое, что разность $s^2 - n$ является полным квадратом, то есть равна t^2. Метод работает, если число n не является полным квадратом.

Метод диофантовой аппроксимации. Этот метод разложения, основанный на нахождении кратчайших векторов решетки, был предложен К. Шнорром. [22]

Назовем u гладким (также называют D-гладким) относительно базы $D = \{p_1, ..., p_r\}^2$, где p_i - малые простые числа, если u раскладывается на простые множители из базы D. Если допускаются отрицательные числа u, то в базу D включают еще $p_0 = -1$.

Для получения разности $s^2 - t^2 \equiv 0 \ (mod\ n)$ используется произведение чисел u и $u - vn$, где u, v — взаимно простые гладкие числа. Параметрами метода являются числа α и $c > 1$. Число $p_0 = -1$ и наименьшие простые числа $p_1, ..., p_r$, где $p_r \approx (\ln n)^\alpha$, образуют базу разложения. Вещественные логарифмы чисел p_i по произвольному (например, натуральному) основанию образуют базис решетки. В ходе работы алгоритма ищутся векторы, близкие

к вектору

$$n = (\underbrace{0, 0, ...0}_{\text{r-раз}}, n^c, \ln n).$$

Алгоритм 2.1.1. Разложение методом диофантовой аппроксимации.

Вход. Число n, числа α и $c > 1$.

Выход. Нетривиальный делитель q числа n.

Метод.

1. Выбрать базу разложения $p_1, ..., p_r$, где $p_r = (\ln n)^\alpha$.

2. Положить $d_0 \leftarrow n$, положить $d_1, ..., d_r$ равными строками матрицы

$$\begin{pmatrix} \ln 2 & 0 & \dots & 0 & n^c \ln 2 \\ 0 & \ln 3 & \dots & 0 & n^c \ln 3 \\ 0 & 0 & \dots & 0 & n^c \ln 5 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \ln p_r & n^c \ln p_r \end{pmatrix}$$

Векторы d_i образуют решетку L.

3. Найти на решетке L не менее $r+2$ нетривиальных векторов $z_i \leftarrow \sum\limits_{j=0}^{r} a_{ij} d_j$ таких, что $a_{ij} \in \mathbb{Z}$ и для $u_i = \prod\limits_{a_{ij}>0} p_j^{a_i j}$, $v_i = \prod\limits_{a_{ij}<0} p_j^{|a_i j|}$ абсолютная величина $|u_i - v_i n|$ раскладывается на множители из базы разложения. Найти представление

$$u_i - v_i n = \prod_{j=0}^{r} p_j^{b_i j}.$$

4. Положить $a_i \leftarrow (a_{i0}, ..., a_{ir})$, $b_i \leftarrow (b_{i0}, ..., b_{ir})$ для $a_i j$, $b_i j$, найденных на предыдущем шаге.

5. Найти непустое множество линейно зависимых над $\mathbb{F}_2$ векторов вида $\sum\limits_{i=1}^{r+2} c_i(a_i + b_i)$, где $c_i \in \mathbb{F}_2$.

6. Положить

$$s \leftarrow \prod_{j=0}^{r} p_j^{\sum\limits_{i=1}^{r+2} c_i (a_{ij}+b_{ij})/2} \ (mod\ n),$$

$$t \leftarrow \prod_{j=0}^{r} p_j^{\sum_{i=1}^{r+2} c_i a_{ij}} \ (mod\ n).$$

7. При $s \neq \pm t\ (mod\ n)$ положить $q \leftarrow$НОД$(s \pm t, n)$.

8. Результат: q.

Метод непрерывных дробей. Непрерывной, или цепной, дробью называется выражение вида

$$a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \ddots}}} = \left[a_0; \frac{1}{a_1}, \frac{1}{a_2}, ..., \frac{1}{a_n}, ...\right].$$

Элементами $a_0, a_i, (i = 1, 2, ...)$ непрерывной дроби могут быть вещественные или комплексные числа, а также функции одной или нескольких переменных. Дроби $a_0 = \dfrac{a_0}{1}, \dfrac{1}{a_i}, (i = 1, 2, ...)$ называются звеньями непрерывной дроби. Пусть $a_i \neq 0$.

Алгоритм 2.1.2. Разложение методом непрерывных дробей.

Вход. Число n.

Выход. Нетривиальный делитель q числа n.

Метод.

1. Положить $P_{-1} \leftarrow 1, P_0 \leftarrow \lfloor\sqrt{n}\rfloor, a_0 \leftarrow \lfloor\sqrt{n}\rfloor, x_0 \leftarrow \sqrt{n} - a_0$.

2. Вычислить $P_0^2\ (mod\ n)$.

3. Для $k = 1, 2, ...$ выполнить следующие действия.

а) Положить $a_k \leftarrow \lfloor 1/x_{k-1}\rfloor, x_k \leftarrow 1/x_{k-1} - a_k$.

б) Положить $P_k \leftarrow a_k P_{k-1} + P_{k-2}\ (mod\ n)$.

в) Найти абсолютно наименьший вычет $P_k^2\ (mod\ n)$ и разложить его на множители.

4. Составить базу разложения $D = \{p_0, p_1, ..., p_r\}, p_0 = -1$, из тех простых чисел p_i, которые встречаются в разложении хотя бы в двух $P_k^2\ (mod\ n)$ или в четной степени хотя бы в одном $P_k^2\ (mod\ n)$ для $k = 0, 1, ...$

5. Для каждого числа $P_k^2 \ (mod\ n)$, являющегося D-гладким, то есть представимого в виде произведения степеней чисел p_i из базы D:

$$P_k^2 \ (mod\ n) = \prod_{i=0}^{r} p_i^{e_i^{(k)}},$$

составить вектор $e^{(k)} \leftarrow (e_0^{(k)}, ..., e_r^{(k)})$.

6. Если можно, найти множество $K = \{k | 0 \leq k \leq r, \bigoplus_{k \in K} e^{(k)} = \mathbf{0}\}$[3] линейно зависимых над $\mathbb{F}_2$ векторов $e^{(k)}$. В противном случае проделать шаг 3 еще для одного k, увеличивая по необходимости базу разложения.

7. Положить

$$s \longleftarrow \prod_{k \in K} P_k (\mod n); t \leftarrow \prod_{k \in K}^{r} P_j^{\gamma_j} (\mod n),$$

где $\gamma_j = \frac{1}{2} \sum\limits_{k \in K} e_j^{(k)}$.

8. Вычислить $q \leftarrow$НОД$(s \pm t, n)$.

9. При $q \neq n$ результат: q. В противном случае вернуться на шаг 6 и найти другое множество K. Если это невозможно, то проделать шаг 3 еще для одного k, увеличивая по необходимости базу разложения.

Метод квадратичного решета. Этот метод, как и предыдущий, относится к методам базы разложения. В ходе вычислений находятся вспомогательные числа, которые раскладываются на простые множители из базы разложения.

Алгоритм 2.1.3. Разложение методом квадратичного решета.

Вход. Число n.

Выход. Нетривиальный делитель q числа n.

Метод.

1. Построить базу разложения $D = \{-1, p_1, ..., p_r\}, p_0 = -1$, где каждое p_i, $1 \leq i \leq r$, — i-е простое число, для которого n является квадратичным вычетом, то есть выполняется равенство для символов Лежандра $\left(\frac{n}{p_i}\right) = 1$.

2. Найти $r + 2$ чисел s_k следующим образом.

а) Положить $k \leftarrow 1$.

б) Положить $y \leftarrow (\lfloor\sqrt{n}\rfloor + x)^2 - n$, — где значения x выбираются в порядке $0, \pm 1, \pm 2, ...$

в) Методом пробного деления на элементы базы D проверить, является ли число y D-гладким, то есть

$$y = \prod_{i=0}^{r} p_i^{e_i^{(k)}}.$$

Если это не так, то вернуться на шаг б) и выбрать новое.

г) Положить $s_k \leftarrow \lfloor\sqrt{n}\rfloor + x, e^{(k)} \leftarrow (e_0^{(k)}, ..., e_r^{(k)})$.

д) Положить $k \leftarrow k + 1$ и вернуться на шаг б).

3. Найти непустое множество $K = \{k | 0 \leq k \leq r + 1, \bigoplus_{k \in K} e^{(k)} = \mathbf{0}\}$ линейно зависимых над $\mathbb{F}_2$ векторов $e^{(k)}$.

4. Положить

$$s \leftarrow \prod_{k \in K} s_k (mod n); t \leftarrow \prod_{j=1}^{r} p_j^{\gamma_j} \ (mod\ n),$$

где $\gamma_j = \frac{1}{2} \sum_{k \in K} e_j^{(k)}$.

5. Вычислить $q \leftarrow$НОД$(s \pm t, n)$.

6. При $q \neq n$ результат: q. В противном случае вернуться на шаг 3 и найти другое множество K. Если это невозможно, то заменить некоторые числа s_k.

Разложение на эллиптической кривой. Метод разложения на эллиптической кривой был предложен Х. Ленстрой. В основе метода лежит устанавливаемый китайской теоремой об остатках изоморфизм групп

$$E(\mathbb{Z}/n\mathbb{Z}) \cong E(\mathbb{F}_p) \oplus E(\mathbb{F}_q)$$

где $n = pq$ и $p \neq q$.

Алгоритм 2.1.4. Разложение на эллиптической кривой.

Вход. Число n, размер m базы D.

Выход. Нетривиальный делитель d числа n.

Метод.

1. Выбрать случайную эллиптическую кривую $E(\mathbb{Z}/n\mathbb{Z})$ и точку Q на ней.

2. Положить $i \leftarrow 0, Q_i \leftarrow Q$.

3. При $i > m$ вернуться на шаг 1; в противном случае найти i-е простое число p_i.

4. Положить $i \leftarrow i+1, \alpha_i = \left\lfloor 0,5\dfrac{\log n}{\log p_i} \right\rfloor, j \leftarrow 0$.

5. При $j > \alpha_i$ перейти на шаг 3. В противном случае выполнить следующие действия.

а) Положить $Q_1 \leftarrow p_i Q_1$. При каждом сложении точек вычислять $d = \text{НОД}(n, \lambda_1)$. При $1 < d < n$ результат: d.

б) Положить $j \leftarrow j+1$ и вернуться на шаг 5.

6. Если нетривиальный делитель числа n не найден, то вернуться на шаг 1.

Теорема 3.2. Пусть n, e, d — натуральные числа, $ed \equiv 1 (mod \phi(n)), e < \phi(n), d < n^{1/4}/3$ и пусть $n = pq$, где p и q — нечетные простые числа, такие что $q < p < 2q$. Тогда существует полиномиальный алгоритм, при помощи которого, зная числа n и e, можно найти число d.

Доказательство. Идея состоит в том, чтобы использовать известное число n для приближения неизвестного $\phi(n)$. При этом можно считать $d > 1$.

Поскольку $ed \equiv 1 (mod \phi(n))$ существует такое целое число $k \geq 1$, что

$$ed - k\phi(n) = 1 \tag{1}$$

Согласно условию имеем $q < p < 2q$, так что $q^2 < pq = n$ и $p+q < 3q < 3\sqrt{n}$. Поскольку $\phi(n) = n - p - q + 1$, то находим

$$0 < n - \phi(n) = p + q - 1 < 3\sqrt{n} - 1. \tag{2}$$

Из (1) и (2) получаем, что

$$\left| \frac{e}{n} - \frac{k}{d} \right| = \left| \frac{ed - k\phi(n) - kn + k\phi(n)}{nd} \right| =$$

$$= \frac{k(n - \phi(n)) - 1}{nd} < \frac{3k\sqrt{n}}{nd} = \frac{3k}{d\sqrt{n}}.$$

При этом из соотношения $k\phi(n) = ed-1 < ed$ и неравенства $e < \phi(n)$ следует, что $k < d < n^{1/4}/3$. Значит,

$$\left|\frac{e}{n} - \frac{k}{d}\right| < \frac{3k}{d\sqrt{n}} < \frac{3}{\sqrt{n}} < \frac{1}{2d^2}.$$

Из теории цепных дробей известно, что в случае выполнения такого неравенства число k/d является некоторой подходящей дробью числа e/n. В силу неравенства (1) числа k и d взаимно просты, поэтому d в точности совпадает со знаменателем одной из подходящих дробей числа e/n. Количество же этих дробей есть величина порядка $O(\ln n)$, то есть число d восстанавливается за линейное время.

Заключение

В ходе выполнения магистерской работы были изучены материалы по теме «Система RSA и задача разложения» и решена намеченная цель – рассмотрение алгоритма шифрования RSA, задачи разложения на множители, построение алгоритма шифрования на примере. Для достижения этой цели, мной в ходе выполнения работы были решены следующие задачи:

1. Определены различные алгоритмы, в том числе и алгоритм шифрования RSA, максимально полно покрывающие тему «Система RSA и задача разложения».

2. Подробно разобраны примеры некоторых методов разложения.

3. Построен алгоритма шифрования RSA на примере.