Алгебра. Глава 11. Теория чисел и криптография

Д. В. Карпов

2024

Криптосистема RSA (Rivest–Shamir–Adleman, 1977)

- Пусть p, q — большие простые числа, $n = pq$.
- Тогда $\varphi(n) = (p-1)(q-1)$.
- Пусть $e \in \mathbb{N}$, $e < \varphi(n)$ и $(e, \varphi(n)) = 1$.
- Пусть $d \in \mathbb{N}$ — обратный вычет к e по модулю $\varphi(n)$ ($d < \varphi(n)$ и $ed \equiv 1 \pmod{\varphi(n)}$).
- Чаще всего числа e и d стараются выбирать так, чтобы число d было большим, а числа e — достаточно небольшим (но и не слишком маленьким).
- Пара (n, e) — открытый ключ. Он используется для шифрования сообщений и публикуется в открытом доступе.
- Пара (n, d) — секретный ключ. Он используется для дешифрования сообщений и должен храниться в секрете.
- Сообщение — число от 0 до $n-1$ (более длинные сообщения разбиваются на блоки, которые шифруются по отдельности).
- Шифрование — функция $P : [0..n-1] \to [0..n-1]$, где $P(m) \equiv m^e \pmod{n}$.
- Дешифрование — функция $S : [0..n-1] \to [0..n-1]$, где $S(m) \equiv m^d \pmod{n}$.

Криптосистема RSA. Доказательство корректности

Теорема 1

$S(P(m)) = P(S(m)) = m.$

Доказательство. • Нужно доказать, что $m^{ed} \equiv m \pmod{n}$.

• Для этого достаточно доказать, что $m^{ed} \equiv m \pmod{p}$ и $m^{ed} \equiv m \pmod{q}$.

• Заметим, что $ed \equiv 1 \pmod{\varphi(n)}$. То есть, $ed = (p-1)(q-1)k + 1$, где $k \in \mathbb{N}$.

• Пусть $m \not\vdots\, p$. Тогда $m^{p-1} \equiv 1 \pmod{p}$. Следовательно, $m^{ed} = m^{(p-1)(q-1)k+1} = (m^{p-1})^{(q-1)k} \cdot m \equiv 1^{(q-1)k} \cdot m \equiv m \pmod{n}$.

• Если $m \vdots\, p$, то $m^{ed} \equiv 0 \equiv m \pmod{n}$.

• Итак, во всех случаях получаем, что $m^{ed} \equiv m \pmod{p}$.

• То, что $m^{ed} \equiv m \pmod{q}$, доказывается аналогично. $\square$

Немного об истории вопроса

• В 1977 году авторы алгоритма (R. L. Rivest, A. Shamir, L. M. Adleman) опубликовали тестовый пример, в котором число n состояло из 129 десятичных (425 двоичных) знаков.

• Тестовый пример был расшифрован в 1994 году при помощи распределенных вычислений: для этого потребовалось полгода работы сети из 1600 компьютеров.

• В настоящее время надежными считаются системы, в которых n содержит порядка 2000 двоичных знаков.

Криптосистема RSA. О выборе p и q

- Выбирая простые числа p и q стоит придерживаться некоторых ограничений.
- Числа p и q не должны быть близки друг к другу. Обычно их выбирают так, чтобы длина их записи отличалась на несколько разрядов.
- Действительно, $pq = \left(\frac{p+q}{2}\right)^2 - \left(\frac{p-q}{2}\right)^2$.
- Если $|p-q|$ мал, то $\left(\frac{p+q}{2}\right)^2$ — точный квадрат, ненамного превосходящий n.
- Тогда перебирая точные квадраты, большие n, мы быстро найдем такое a, что $a^2 - n$ — точный квадрат.
- Далее, положив $a = \frac{p+q}{2}$ и $\sqrt{a^2 - n} = \frac{p-q}{2}$, мы легко найдем p и q.
- $(p-1, q-1)$ должен быть маленьким.
- Каждое из чисел $p-1, q-1$ должно иметь большой простой делитель.

Проверка простоты числа

• Как мы видим, в криптографии возникает вопрос: а как убедиться, что предъявленное нам число — простое

• Тривиальный алгоритм: перебираем все числа от 2 до $\sqrt{n}$ и проверяем, делится ли n на каждое из них.

• Этот алгоритм экспоненциален относительно длины входа ($\log n$) и для чисел интересующего нас размера неприменим.

• На данный момент известен единственный алгоритм проверки простоты с доказанным полиномиальным (относительно длины входа) временем работы — его придумали M. Agrawal, N. Kayal, N. Saxena в 2004 году.

• Однако, на практике и этот алгоритм работает очень долго и расходует много памяти.

• Чаще всего на практике используют вероятностные тесты, которые работают гораздо быстрее.

Вероятностные алгоритмы

- Вероятностный алгоритм — это алгоритм, ход и результаты работы которого, помимо входа, зависят от выбора некоторого случайного параметра a.
- Как правило, a — это натуральное число, которое случайным образом выбирается из некоторого диапазона.
- При определенных значениях a алгоритм может давать ошибочный результат, но вероятность этого должна быть не слишком велика (т. е. не превосходить некоторой заранее фиксированной константы).
- Например, бывают вероятностные алгоритмы, для которых вероятность ошибки меньше $\frac{1}{2}$.
- Для снижения вероятности ошибки можно многократно запустить вероятностный алгоритм. При каждом запуске алгоритма, параметр a выбирается заново, случайным и независимым от предыдущих запусков образом.
- Для проверки простоты числа нас будут интересовать вероятностные алгоритмы с односторонней ошибкой.
- Если такой алгоритм отвечает, что число составное, то это гарантировано так. А вот ответ простое может быть ошибочным.

Тест Ферма

Вход: нечётное натуральное число n.

- Выбираем случайным образом параметр $a \in [2..n-2]$;
- вычисляем a^{n-1} по модулю n;
- если $a^{n-1} \equiv 1 \pmod{n}$, то ответ "простое",
- если $a^{n-1} \not\equiv 1 \pmod{n}$, то ответ "составное".

- Из Теоремы Эйлера следует, что ответ "составное" не может быть ошибочным.
- В то же время, ответ "простое" ошибочным быть может.
- Отметим, что сравнение $a^{n-1} \equiv 1 \pmod{n}$ может быть выполнено только в случае $(a, n) = 1$.
- К сожалению, существуют такие нечетные составные числа n, для которых $a^{n-1} \equiv 1 \pmod{n}$ при всех a взаимно простых с n.
- Такие числа называются числами Кармайкла. Они проходят тест Ферма почти при любом выборе a. Пример: $n = 561$.
- В 1994 году доказано, что чисел Кармайкла бесконечно много. Встречаются они относительно редко.

Символ Якоби

Определение

Пусть $n = p_1^{k_1} \dots p_\ell^{k_\ell}$ — каноническое разложение нечетного числа, $a \in \mathbb{N}$. Тогда Символ Якоби — это $\left(\frac{a}{n}\right) := \prod_{i=1}^{\ell} \left(\frac{a}{p_i}\right)^{k_i}$.

• Из мультипликативности символа Лежандра следует, что $\left(\frac{ab}{n}\right) = \left(\frac{a}{n}\right) \cdot \left(\frac{b}{n}\right)$.

• Из $\left(\frac{a}{n}\right) = 1$ не следует, что существует квадрат, сравнимый с a по модулю n.

Лемма 1

Пусть $n \in \mathbb{N}$, $n \not\vdots\, 2$. Тогда $\left(\frac{2}{n}\right) = (-1)^{\frac{n^2-1}{8}}$.

Доказательство. • Пусть $f(k) := (-1)^{\frac{k^2-1}{8}}$.

• Рассмотрев все пары остатков по модулю 8, можно сделать вывод, что для любых нечетных k_1 и k_2 выполнено $f(k_1k_2) = f(k_1)f(k_2)$.

• По Лемме 4.7, $\left(\frac{2}{p_i}\right) = (-1)^{\frac{p^2-1}{8}} = f(p)$ для любого нечетного простого p.

• Теперь из определения символа Якоби следует утверждение леммы для нечетного $n = p_1^{k_1} \dots p_\ell^{k_\ell}$.

Теорема 2

(Закон взаимности для символа Якоби.)
Пусть $n, m \in \mathbb{N}$ *нечетны. Тогда* $\left(\frac{n}{m}\right) = (-1)^{\frac{n-1}{2}\cdot\frac{m-1}{2}} \cdot \left(\frac{m}{n}\right)$.

Доказательство. • Если $(m, n) > 1$, то $\left(\frac{n}{m}\right) = \left(\frac{m}{n}\right) = 0$ и теорема доказана.

• Далее $(m, n) = 1$, пусть $n = p_1 \dots p_k$ и $m = q_1 \dots q_s$ — их разложения на простые множители (не обязательно различные).

• Тогда $\left(\frac{n}{m}\right) = \prod_{i=1}^{k} \prod_{j=1}^{s} \left(\frac{p_i}{q_j}\right)$ и $\left(\frac{m}{n}\right) = \prod_{i=1}^{k} \prod_{j=1}^{s} \left(\frac{q_j}{p_i}\right)$.

• Значит, чтобы перейти от $\left(\frac{n}{m}\right)$ к $\left(\frac{m}{n}\right)$, нам нужно перевернуть ks символов Лежандра вида $\left(\frac{p_i}{q_j}\right)$, превратив их в $\left(\frac{q_j}{p_i}\right)$.

• Один такой переворот по закону взаимности Гуасса (Теореме 4.3) меняет знак символа Лежандра, если и только если оба простых числа $p_i, q_j \equiv 3 \pmod 4$.

• Пусть в разложении n ровно k' простых, сравнимых с 3 по модулю 4, а в разложении m ровно s' таких простых.

• Тогда $\left(\frac{n}{m}\right)$ и $\left(\frac{m}{n}\right)$ имеют разный знак, если и только если $k's'$ нечетно.

• Отметим, что $k' \not\vdots\, 2 \iff n \equiv 3 \pmod 4$ и $s' \not\vdots\, 2 \iff m \equiv 3 \pmod 4$.

• Остается отметить, что $m \equiv n \equiv 3 \pmod 4 \iff \frac{(m-1)(n-1)}{2} \not\vdots\, 2$. $\square$

• Благодаря Лемме 1 и Теореме 2 вычислить символ Якоби $\left(\frac{m}{n}\right)$ можно достаточно быстро, причем для этого не нужно знать разложение числа n на простые множители (а найти такое разложение для большого числа как раз — трудная задача).

Первообразные корни

• Пусть $n \in \mathbb{N}$. Через $\mathbb{Z}_n$ обозначается кольцо вычетов по модулю n, а через $\mathbb{Z}_n^*$ — множество всех обратимых элементов этого кольца (то есть, вычетов, взаимно простых с n — из Пр.СВ по модулю n).

• По Теореме Эйлера, для любого $a \in \mathbb{Z}_n^*$ мы знаем, что $a^{\varphi(n)} \equiv 1 \pmod{n}$.

Определение

Пусть $a \in \mathbb{Z}_n^*$, $d \in \mathbb{N}$. Будем говорить, что вычет a принадлежит к показателю d по модулю n, если $a^d = 1$, но $a^s \neq 1$ при $s \in \mathbb{N}$, $s < d$. Обозначение: $a \in_n d$.

• Аналогично Лемме 4.1 несложно доказать, что если $a \in_n d$, то $d \mid \varphi(n)$.

Определение

Пусть $n \in \mathbb{N}$. Вычет $a \in \mathbb{Z}_n^*$, первообразный корень по модулю n, если $a \in_n \varphi(n)$.

• По Теореме 4.1 существуют первообразные корни по модулю $p \in \mathbb{P}$. Кроме того, первообразные корни существуют по модулю p^n и $2p^n$, где $p \in \mathbb{P}$ нечетно, а также по модулю 4. По остальным модулям первообразных корней нет.

Теорема 3

Пусть $n \in \mathbb{N}$, a — первообразный корень по модулю n. Тогда a, a^2, ..., $a^{\varphi(n)} = 1$ — ПрСВ $\pmod{n}$, то есть, в точности все вычеты из $\mathbb{Z}_n^*$.

Доказательство. • Достаточно доказать, что $a^i \neq a^j$ при $1 \le j < i \le \varphi(n)$.

• Предположим противное, пусть $a^i = a^j \iff a^j(a^{i-j} - 1) = 0$.

• Однако, $a^j \neq 0$ и $a^{i-j} \neq 1$, так как $0 < i - j < \varphi(n)$. Противоречие. $\square$

• Если a — первообразный корень по модулю n, то любой вычет $b \in \mathbb{Z}_n^*$ представляется в виде $b = a^k$, где $1 \le k \le \varphi(n)$.

Теорема 4

Для простого $p \in \mathbb{P}$ существует первообразный корень по модулю p^2.

Доказательство. • Напомним, что $\varphi(p^2) = p(p-1)$.

• Достаточно найти такое $b \in \mathbb{N}$, что $b^{p(p-1)} \equiv 1 \pmod{p}^2$, но $b^s \not\equiv 1 \pmod{p}^2$ при $s < p(p-1)$.

• Так как существует первообразный корень по модулю p, существует и такое $a \in \mathbb{N}$, что $a^{p-1} \equiv 1 \pmod{p}$, но $a^s \not\equiv 1 \pmod{p}$ при $s < p-1$.

• Тогда $(a, p) = 1$, а значит, a и $a + p$ — разные вычеты вычеты из $\mathbb{Z}^*_{p^2}$.

• Если $a^s \equiv 1 \pmod{p^2}$, то $a^s \equiv 1 \pmod{p} \Rightarrow s \vdots p - 1 \Rightarrow s \in \{p-1, p(p-1)\}$.

• Аналогичное верно и для $a + p$.

• Предположим, что ни a, ни $a + p$ нам не подходит. Тогда $(a+p)^{p-1} \equiv a^{p-1} \equiv 1 \pmod{p^2}$.

• Но $(a+p)^{p-1} - a^{p-1} = \sum_{k=1}^{p-1} \mathrm{C}^k_{p-1} p^k a^{p-1-k} \equiv p(p-1)a^{p-2} \not\equiv 0 \pmod{p^2}$, противоречие.

Эйлеровы псевдопростые

Определение

Нечетное составное число n называется эйлеровым псевдопростым по основанию a, если $(a, n) = 1$ и $a^{\frac{n-1}{2}} \equiv \left(\frac{a}{n}\right) \pmod{n}$.

Теорема 5

Нечетное составное число n является эйлеровым псевдопростым по основанию не более чем $\frac{\varphi(n)}{2}$ чисел, взаимно простых с n и меньших n.

Доказательство. • Назовем число $b \in \mathbb{N}$, $(b, n) = 1$, хорошим, если $b^{\frac{n-1}{2}} \not\equiv \left(\frac{b}{n}\right) \pmod{n}$ и плохим, если это сравнение выполнено.

• Наша цель — доказать, что не более чем половина вычетов из $\mathbb{Z}_n^*$ — плохие.

Утверждение 1

Пусть $a, b \in \mathbb{Z}_n^*$, причем a — плохой вычет, а b — хороший. Тогда ab — хороший.

Доказательство. • Предположим, что ab — плохой вычет.

• Тогда $\left(\frac{a}{n}\right) \equiv a^{\frac{n-1}{2}} \pmod n$ и $\left(\frac{ab}{n}\right) \equiv (ab)^{\frac{n-1}{2}} \pmod n$.

• Так как $(a, n) = (b, n) = 1$, имеем $\left(\frac{a}{n}\right), \left(\frac{b}{n}\right), \left(\frac{ab}{n}\right) \in \{1, -1\}$ и

$$\left(\frac{b}{n}\right) = \left(\frac{b}{n}\right) \cdot \left(\frac{a}{n}\right)^2 = \left(\left(\frac{b}{n}\right) \cdot \left(\frac{a}{n}\right)\right) \cdot \left(\frac{a}{n}\right) = \left(\frac{ab}{n}\right) \cdot \left(\frac{a}{n}\right)$$

$$\equiv_p (ab)^{\frac{n-1}{2}} \cdot a^{\frac{n-1}{2}} \equiv_p b^{\frac{n-1}{2}} \cdot \left(a^{\frac{n-1}{2}}\right)^2 \equiv_p b^{\frac{n-1}{2}}.$$

• В последнем переходе мы использовали, что $a^{\frac{n-1}{2}} \equiv \pm 1 \pmod n$. □

Утверждение 2

Если b — хороший вычет, то плохих вычетов не более чем $\frac{\varphi(n)}{2}$.

Доказательство. • Пусть $a_1, \ldots, a_k$ — все плохие вычеты.

• По Утверждению 1 тогда $a_1 b, \ldots, a_k b$ — хорошие вычеты, и все они, очевидно, различны. □

• Осталось доказать существование хорошего вычета. Разберем два случая.

Случай 1: $n \vdots p^2$, где $p \in \mathbb{P}$.

• По Теореме 4 существует первообразный корень по модулю p^2 — пусть это $b \le p^2 - 1 \le n - 1$.

• Пусть $b \in_n d$. Тогда $b^d - 1 \vdots n \vdots p^2$, откуда следует $d \vdots \varphi(p^2) = p(p-1) \vdots p$.

• Очевидно, $n - 1 \not\vdots\, p$. Поэтому, $b^{n-1} \not\equiv 1 \pmod{n}$.

• Если $b^{\frac{n-1}{2}} \equiv \left(\frac{b}{n}\right) \pmod{n}$, то

$b^{\frac{n-1}{2}} \equiv \pm 1 \pmod{n} \Rightarrow b^{n-1} \equiv 1 \pmod{n}$, противоречие.

Случай 2: n свободно от квадратов.

• Пусть $n \vdots p$, где $p \in \mathbb{P}$.

• По КТО существует такое число $b \in [1..n-1]$, что $\left(\frac{b}{p}\right) = -1$ (то есть, b сравнимо по модулю p с любым квадратичным невычетом) и $b \equiv 1 \pmod{\frac{n}{p}}$. Ясно, что $(b, n) = 1$.

• Тогда $\left(\frac{b}{q}\right) = 1$ для любого простого $q|n$ отличного от p, откуда $\left(\frac{b}{n}\right) = -1$.

• Если $b^{\frac{n-1}{2}} \equiv \left(\frac{b}{n}\right) \equiv -1 \pmod{n}$, то $b^{\frac{n-1}{2}} \equiv -1 \pmod{\frac{n}{p}}$, что не так. Значит, b — хороший вычет.

Тест Соловея-Штрассена

Вход: нечётное натуральное число n.

- Выбираем случайным образом параметр $a \in [2..n-2]$;
- вычисляем a^{n-1} по модулю n;
- вычисляем $\left(\frac{a}{n}\right)$ по модулю n.
- Если оба эти числа сравнимы $a^{n-1} \equiv \left(\frac{a}{n}\right) \equiv \pm 1 \pmod{n}$, то ответ "простое",
- иначе ответ "составное".

• Из определений символа Лежандра и символа Якоби следует, что ответ "составное" не может быть ошибочным.

• В то же время, ответ "простое" ошибочным быть может.

• По Теореме 5, вероятность ошибки в тесте Соловея-Штрассена менее $\frac{1}{2}$.

• Повторив тест с числом n независимо k раз, получим вероятность ошибки менее $\frac{1}{2^k}$.

Тест Миллера-Рабина

Вход: нечётное натуральное число n.

- Пусть $n - 1 = 2^t \cdot u$, где $t, u \in \mathbb{N}$ и $u \not\vdots\, 2$.
- Выбираем случайным образом параметр $a \in [2..n-2]$.
- Вычисляем $a^u, a^{2u}, \dots, a^{2^{t-1}u}$ по модулю n (получившаяся последовательность называется последовательностью Миллера-Рабина).
- Ответ "простое" дается в следующих двух случаях:
  - если $a^u \equiv 1 \pmod{n}$;
  - если $a^{2^k u} \equiv -1 \pmod{n}$ при некотором $k \in [0..t-1]$.
- Во всех остальных случаях, дается ответ "составное".
- Ответ "простое" при выполнении теста Миллера-Рабина может быть ошибочным.

Определение

Нечетное составное число n называется сильно псевдопростым по основанию a, если тест Миллера-Рабина для числа n с параметром a дает ответ "простое".

Лемма 2

Если $n \in \mathbb{P}$, то тест Миллера-Рабина выдаст ответ "простое".

Доказательство. • По теореме Эйлера $a^{2^t \cdot u} \equiv 1 \pmod{n}$, так что в последовательности Миллера-Рабина есть хотя бы одна единица.

• Рассмотрим такое наименьшее k, что $a^{2^k \cdot u} \equiv 1 \pmod{n}$.

• Если $k = 0$, то $a^u \equiv 1 \pmod{n}$ и тогда дан ответ "простое ".

• Пусть $k > 0$. Тогда $a^{2^k \cdot u} \equiv 1 \pmod{n}$ и $a^{2^{k-\mathbf{1}} \cdot u} \not\equiv 1 \pmod{n}$.

• Следовательно, $(a^{2^{k-\mathbf{1}} \cdot u} - 1)(a^{2^{k-\mathbf{1}} \cdot u} + 1) = a^{2^k \cdot u} - 1 \,\vdots\, n$.

• Поскольку $n \in \mathbb{P}$ и $a^{2^{k-\mathbf{1}} \cdot u} \not\equiv 1 \pmod{n}$, получаем, что $a^{2^{k-\mathbf{1}} \cdot u} + 1 \,\vdots\, n$.

• В этом случае тоже дан ответ "простое". $\square$

• Итак, тест Миллера-Рабина — вероятностный тест с односторонней ошибкой.

• Можно доказать, что вероятность ошибки в тесте Миллера-Рабина не превосходит $\frac{1}{4}$, но доказательство весьма технически сложное.