Лекция №11
ρ-метод Полларда

Лектор: Н.Ю. Золотых Записал: А. Фель

22 ноября 2008

Предположим, что нам известно, что некоторое натуральное число n составное. Требуется разложить его на множители. Метод пробных делений на простые числа, меньшие $\sqrt{n}$ может потребовать порядка $O(\sqrt{n})$ операций.

Задача факторизации сводится к нахождению хотя бы одного делителя числа n. Остальные делители находятся рекурсивно.

1. Изначальная идея

Допустим, что мы знаем два числа $x, x' \in Z_n$, такие что

$$\text{НОД}(x - x', n) > 1$$

Ясно, что при этом делитель числа n найден.

Утверждение 1. *Пусть d является делителем n:*

$$d|n, 1 < d < n,$$

пусть $\exists x, x' \in Z_n, x \neq x'$, такие что

$$x \equiv x'(\mod d),$$

тогда

$$1 < \text{НОД}(x - x', n) < n$$

Доказательство. Так как $d|n$, то $n = c \cdot d$

$x \equiv x'(\mod d)$, то

$$x = a_1 \cdot d + b$$
$$x' = a_2 \cdot d + b$$
$$x - x' = (a_1 - a_2) \cdot d$$
$$\text{НОД}(x - x', n) = \text{НОД}((a_1 - a_2)d, cd) = d \cdot \text{НОД}(a_1 - a_2, c) > 1$$

□

Утверждение 2. *Пусть $n = pq$, p, q - простые, $x, x' \in Z_n, x \neq x'$. Тогда*

$$(x \equiv x'(\mod d) \Leftrightarrow \text{НОД}(x - x', n) = p)$$

Но как же найти такие x, x'? Возьмем случайные $x_1, x_2, \ldots, x_m$ из Z_n. Переберем всевозможные пары (x_i, x_j). Будем искать $\text{НОД}(x_i - x_j, n)$. Требуется найти такие (x_i, x_j), которые дают нетирвиальный НОД. Сколько же потребуется взять чисел, чтобы получить нетривиальный НОД?

Теорема 1 (Парадокс дней рождения)**.** *Для случайной выборки из n элементов объема m вероятность того, что все элементы попарно различны*

$$p_{nm} < e^{-\frac{(m-1)^2}{2n}}$$

Доказательство.

$$p_{nm} = \frac{n(n-1)\cdot\dots\cdot(n-m+1)}{n^m} = \prod_{i=1}^{m-1}\left(1-\frac{i}{n}\right)$$

Возьмем логарифм от обеих частей:

$$\ln p_{nm} = \sum_{i=1}^{m-1}\ln\left(1-\frac{i}{n}\right)$$

Разложим $\ln\left(1-\frac{i}{n}\right)$ в ряд Тейлора:

$$\ln\left(1-\frac{i}{n}\right) = -\frac{i}{n} - O\left(\left(\frac{i}{n}\right)^2\right) < -\frac{i}{n}$$

Тогда

$$\ln p_{nm} < -\sum_{i=1}^{m-1}\frac{i}{n} = -\frac{m(m-1)}{2n} < -\frac{(m-1)^2}{2n},$$

то есть

$$p_{nm} < e^{-\frac{(m-1)^2}{2n}}$$

□

Пример 1. Сколько человек необходимо опросить, чтобы с вероятностью $1/2$ утверждать, что по крайней мере у двух из них совпадут дни рождения?

$n = 365, p_{nm} = \frac{1}{2}, m = ?$

$$m < 1 + \sqrt{-2n\ln p_{nm}} = 1 + \sqrt{-2\cdot 365\cdot\ln\frac{1}{2}} \approx 22,49$$

то есть достаточно 23 человека.

Следствие 1. *Пусть* $\lambda > 0, m = \left\lfloor\sqrt{2\lambda n}+1\right\rfloor$, *тогда*

$$p_{nm} < e^{-\lambda}$$

Достаточным условием того, что НОД$(x_i - x_j, n)$ - нетривиальный является $x_i \equiv x_j(\mod d)$, где $d|n$ (по утверждению 1). Используя теорему 1 мы можем оценить вероятность того, что в выборке из n элементов объема m окажутся такие x_i, x_j, что $x_i \equiv x_j(\mod d)$. Для этого в качестве n в условиях теоремы 1 возьмем d. Таким образом мы делаем выборку из элементов $\{1, 2, \dots, d\}$. Вероятность того, что все элементы в выборке объема $m = \left\lfloor\sqrt{2\lambda d}+1\right\rfloor$ различны, меньше $e^{-\lambda}$. То есть с вероятностью, большей $(1-e^{-\lambda})$, в выборке присутствуют $x_i = x_j (i \neq j)$.Поскольку любой делитель d числа n меньше или равен $\sqrt{n}$, то

$$m = \left\lfloor\sqrt{2\lambda\sqrt{n}}+1\right\rfloor = \left\lfloor\sqrt{2\lambda}\sqrt[4]{n}+1\right\rfloor$$

$$m = O(\sqrt[4]{n})$$

Таким образом при заданном λ нам потребуется выбрать $m = \left\lfloor\sqrt{2\lambda}\sqrt[4]{n}+1\right\rfloor$ элементов, чтобы утверждать, что в выборке окажутся такие x_i, x_j, что НОД$(x_i - x_j, n)$ - нетривиальный, с вероятностью $(1-e^{-\lambda})$.

Но для проверки нетривиальности НОД$(x_i - x_j, n)$ для всех элементов в выборке потребуется порядка $\sqrt{n}$ сравнений (сравнение каждого с каждым). Но при этом мы не получаем выигрыша относительно обычного метода пробных делений на простые числа. Требуется как-то обойти необходимость сравнения каждого с каждым. Идею предложил Поллард в 1978 году.

2. ρ-метод Полларда

Возьмем некоторое случайное отображение

$$f : Z_n \to Z_n$$

которое сгенерирует некоторую случайную последовательность

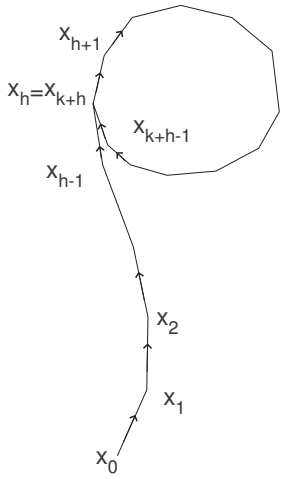
$$x_0, x_1, x_2, \ldots$$

где $x_i = f(x_{i-1})$.

Функция f имеет не более, чем n значений, поэтому последовательность зациклится.

$$x_{k+h} = x_h$$

h называется индексом вхождения, k - длиной цикла



Изображение похоже на греческий символ ρ, поэтому метод так и назван.

Отметим, что если x_{k+h} совпадет с x_h , то $\forall i > 0$ $x_{k+h+i} = x_{h+i}$. Поэтому для того, чтобы не сравнивать на очередной итерации x_h со всеми предыдущими $x_1, \ldots, x_{h-1}$, будем проводить сравнения по следующей схеме:

$$\left| \begin{array}{c} x_0 \\ x_1 \end{array} \right| \begin{array}{c} x_1 \\ x_2, x_3 \end{array} \left| \begin{array}{c} x_3 \\ x_4, x_5, x_6, x_7 \end{array} \right| \begin{array}{c} x_7 \\ x_8, \ldots, x_{15} \end{array} \left| \ldots \right.$$

Таким образом мы сможем определить факт зацикливания возможно не с первого вхождения, но при этом существенно уменьшим количество сравнений: нам потребуется не более, чем h + 2k сравнений.

Итак, собственно алгоритм Полларда:

```
Pollard(n)
    j = 1
    x = random(0, n − 1)
    while true
        x' = x
        for i = 1, . . . , j
            x = f(x)
            d = НОД(x − x', n)
            if d > 1
                return d
            end
        end
```

$j = 2j$
end
end

Для любого заданного λ в алгоритме Полларда количество итераций, которое потребуется для нахождения делителя n не превосходит $\left\lfloor \sqrt{2\lambda}\sqrt[4]{n} + 1 \right\rfloor$ с вероятностью $(1 - e^{-\lambda})$.

Остается еще один вопрос: как на практике выбирать функцию f? Функция должна быть не слишком сложной для вычисления, но в то же время не должна быть линейным многочленом, а также не должна порождать взаимнооднозначное отображение. Обычно в качестве f берут функцию

$$f(x) = x^2 \pm 1 \pmod{n}$$

или

$$f(x) = x^2 \pm a \pmod{n}$$