Сложность вычислений

Вероятностная проверка на простоту без ошибок

Чернис Константин, группа 694

Содержание

В данном проекте доказываются избранные факты вероятностной проверки чисел на простоту, а также проводятся некоторые эксперименты.

1. Введение в сложностные классы

Для начала опишем сложностные классы, затрагиваемые данной задачей:

Определение 1.1. Вероятностной машиной Тьюринга называется детерминированная машина Тьюринга M с двумя аргументами x (аргумент вероятностной машины) и r (случайные биты), где длина r есть некоторая функция от длины x. Результатом работы M на входе x будет вероятностое распределение, индуцированное данным x и равномерным на всех значениях r. Временем работы M на данном x будем считать максимальное время работы $M(x, r)$ для всех r указанной длины. Так же определяется и использованная память.

Определение 1.2. Классом **RP** называется класс языков A, для которых существует полиномиальный в худшем случае вероятностный алгоритм V, такой что:

- если $x \in A$, то $P_r[V(x, r) = 1] \geqslant \frac{1}{2}$;
- если $x \notin A$, то $P_r[V(x, r) = 1] = 0$.

Определение 1.3. Классом **coRP** называется класс языков A, для которых существует полиномиальный в худшем случае вероятностный алгоритм V, такой что:

- если $x \in A$, то $P_r[V(x, r) = 1] = 1$;
- если $x \notin A$, то $P_r[V(x, r) = 1] \leqslant \frac{1}{2}$.

Определение 1.4. Классом **ZPP** называется класс языков A, для которых существует вероятностный алгоритм A, такой что

$$x \in A \iff \forall r\ V(x, r) = 1,$$

а для каждого x ожидаемое по r время работы полиномиально.

Обозначение **ZPP** расшифровывается как "zero-error probabistic polynomial".

Утверждение 1.1. $\mathbf{ZPP} = \mathbf{RP} \cap \mathbf{coPR}$.

Таким образом, для вероятностной проверки чисел на простоту достаточно предоставить алгоритмы проверки чисел на простоту из **RP** и **coRP**, после чего запускать их по очереди до тех пор, пока один из алгоритмов не выдаст ответ, в котором он уверен. Вероятность отсутствия ответа будет уменьшаться минимум в 4 раза после каждой итерации цикла проверки, так что за полиномиальное число шагов вероятность станет экспоненциально малой и можно будет применить детерминированный экспоненциальный алгоритм.

В следующей секции будет описан алгоритм из **coRP**, а в секции 3 — из **RP**.

2. Алгоритм Миллера-Рабина

Большинство алгоритмов вероятностной проверки на простоту из **coRP** опираются на какое-либо свойство простых чисел, то есть проверяют необходимое условие. Наиболее популярным среди них является алгоритм Миллера-Рабина, который гарантирует, что для нечётного составного минимум 75% чисел от 1 до $n-1$ позволяют определить его непростоту.

Говоря в терминах Определения 1.3, A — множество простых чисел, и для $x \notin A$ $P_r[V(x,r) = 1] \leqslant \frac{1}{4}$, где $r \in \overline{1, n-1}$. Кроме того, как будет показано ниже, проверяемое условие действительно является необходимым, то есть для $x \in A$ $P_r[V(x,r) = 1] = 1$, то есть алгоритм Миллера-Рабина лежит в **coNP**.

2.1 Описание алгоритма

Заданное нечётное целое число $n > 1$ можно представить в виде $n - 1 = 2^e k$, где $e \geqslant 1$ (т.к. n нечётно) и k нечётное. Применяя к $x^{n-1} - 1 = x^{2^e k} - 1$ формулу разности квадратов, получаем:

$$\begin{aligned} x^{2^e k} - 1 &= \left(x^{2^{e-1}k}\right)^2 - 1 \\ &= \left(x^{2^{e-1}k} - 1\right)\left(x^{2^{e-1}k} + 1\right) \\ &= \left(x^{2^{e-2}k} - 1\right)\left(x^{2^{e-2}k} + 1\right)\left(x^{2^{e-1}k} + 1\right) \\ &\vdots \\ &= \left(x^k - 1\right)\left(x^k + 1\right)\left(x^{2k} + 1\right)\left(x^{4k} + 1\right)\ldots\left(x^{2^{e-1}k} + 1\right) \end{aligned}$$

Если n простое и $a \in \overline{1, n-1}$, то по малой теореме Ферма $a^{n-1} - 1 \equiv 0 \bmod n$. Используя разложение, полученное выше, имеем

$$\left(x^k - 1\right)\left(x^k + 1\right)\left(x^{2k} + 1\right)\left(x^{4k} + 1\right)\ldots\left(x^{2^{e-1}k} + 1\right) \equiv 0 \bmod n$$

Таким образом, для простого n один из множителей должен делиться на n, то есть необходимым условием, нарушение которого означает, что число составное, является

$$a^k \equiv 1 \bmod n \text{ или } a^{2^i k} \equiv -1 \bmod n \text{ для некоторого } i \in \overline{0, e-1}.$$

Определение 2.1. Представим нечётное $n > 1$ в виде $n - 1 = 2^e k$, где e нечётно и выберем $a \in \overline{1, n-1}$. Тогда a называется свидетелем для числа n, если не выполнено необходимое условие, то есть

$$a^k \not\equiv 1 \bmod n \text{ и } a^{2^i k} \not\equiv -1 \bmod n \ \forall i \in \overline{0, e-1}.$$

Если же необходимое условие выполнено, то есть

$$a^k \equiv 1 \bmod n \text{ или } a^{2^i k} \equiv -1 \bmod n \text{ для некоторого } i \in \overline{0, e-1},$$

то a не является свидетелем для n.

Отметим, что уже сейчас можно построить вероятностный алгоритм проверки на простоту со сколь угодно малой вероятностью ошибки:

Data: проверяемое число n, количество итераций t
Result: является ли число n простым
for $i \in \overline{1,t}$ **do**
  выбрать случайное a из $\overline{1, n-1}$;
  if a *является свидетелем для* n **then**
    return "n составное";
  end
end
return "n простое с вероятностью минимум $1 - 1/4^t$";

2.2 Доказательство оценки на число свидетелей

Для начала покажем, что оценка 75% неулучшаема:

Утверждение 2.1. Доля свидетелей для $n = 9$ составляет 3/4.

Доказательство. $n - 1 = 8 = 2^3$, так что $e = 3$ и $k = 1$, и для проверки необходимого условия надо перебрать (a, a^2, a^3). Из приведённой ниже таблицы видно, что свидетелями среди $\overline{1,8}$ являются $2, 3, 4, 5, 6, 7$, что составляет $6/8 = 1/4$, что и требовалось.

a mod 9	1	2	3	4	5	6	7	8
a^2 mod 9	1	4	0	7	7	0	4	1
a^3 mod 9	1	7	0	4	4	0	7	1

■

Существует также доказательство неулучшаемости оценки при $n \to \infty$, оно приведено в [3].

Теорема 2.1. Пусть $n > 1$ нечётное составное.

Доля целых чисел среди $\overline{1, n-1}$, являющихся свидетелями числа n, превышает 75%, за исключением $n = 9$, для которого доля составляет 75%.

Другими словами, доля целых чисел среди $\overline{1, n-1}$, не являющихся свидетелями числа n, меньше 25%, за исключением $n = 9$, для которого доля составляет 25%.

Докажем более слабое утверждение:

Теорема 2.2. Если $n > 1$ нечётное и составное, то доля свидетелей числа n превышает 50%. Другими словами, больше 50% из $a \in \overline{1, n-1}$ удовлетворяют $a^k \not\equiv 1 \bmod n$ и $a^{2^i k} \not\equiv -1 \bmod n \; \forall i \in \overline{0, e-1}$.

Доказательство. Докажем, что доля не свидетелей для n меньше 50%, показав, что они образуют собственную подгруппу группы обратимых чисел mod n. В силу того, что порядок собственной подгруппы составляет максимум половину от порядка группы, множество свидетелей числа n содержит минимум половину обратимых чисел mod n и все необратимые числа mod n среди $\overline{1, n-1}$ (множество необратимых непусто в силу того, что n составное). Таким образом, доля свидетелей для числа n первышает 50%.

Случай 1: n является степенью простого числа, то есть $n = p^\alpha$, где p — нечётное простое и $\alpha \geqslant 2$.

Утверждение 2.2. Если $n = p^\alpha$ для простого p и $\alpha \geqslant 1$, то не свидетели для n являются корнями уравнения $a^{p-1} \equiv 1 \bmod p^\alpha$, которые образуют группу по умножению $\bmod\ n$.

Доказательство. Обоснование приведено в [2]. ■

Согласно Утверждению 2.2 свидетели непростоты образуют группу по умножению $\bmod\ n$. Порядок числа a, являющегося решением уравнения $a^{p-1} \equiv 1 \bmod n$, делит $p-1$, так что он не делится на p. В то же время существуют обратимые $\bmod\ n$ числа, порядок которых делится на p: примером такого числа является $1+p$, чей порядок $\bmod\ p^\alpha$ составляет $p^{\alpha-1}$ (этот факт можно показать индукцией по r: база — $1 + kp \equiv 1 \bmod p$, переход — $(1 + kp^r)^p \equiv 1 \bmod p^{r+1}$). Таким образом, не свидетели $\bmod\ n$ образуют собственную подгруппу в группе обратимых чисел $\bmod\ n$, что заканчивает доказательство этого случая.

Случай 2: n не является степенью простого. Пусть $i_0 \in \overline{0, e-1}$ — максимальное число, такое что $\exists\, a_0 \in \mathbb{Z}$ такой что $a^{2^{i_0}} \equiv -1 \bmod n$. (В силу того, что $(-1)^{2^0} = -1$, требуемый i_0 существует, причём a_0 взаимно прост с n).

Множество
$$G_n = \{\, a \in \overline{1, n-1} \mid a^{2^{i_0}k} \equiv \pm 1 \bmod n \,\}$$
является группой по умножению $\bmod\ n$ и содержит все a, удовлетворяющие одному из двух условий:

(1) $a^k \equiv 1 \bmod n$,

(2) $a^{2^i k} \equiv 1 \bmod n$ для одного из $i \in \overline{0, e-1}$.

Если $a^k \equiv 1 \bmod n$, то $a^{2^{i_0}k} \equiv 1 \bmod n$. Если же $a^{2^i k} \equiv 1 \bmod n$ для некоторого $i \in \overline{0, e-1}$, то $\left(2^k\right)^{2^i} \equiv -1 \bmod n$, причём $i \leqslant i_0$ в силу максимальности i_0. Таким образом, $a^{2^{i_0}} \equiv -1 \bmod n$, если $i = i_0$, и $a^{2^{i_0}} = 1 \bmod n$, если $i < i_0$. Отсюда все $a \in \overline{1, n-1}$, удовлетворяющие (1) или (2), лежат в G_n.

Покажем, что G_n является собственной подгруппой обратимых чисел $\bmod\ n$, для чего найдём обратимое число, не лежащее в G_n. Пусть p — простой делитель n, тогда представим n в виде $n = p^\alpha n'$, где $\alpha \geqslant 1$ и $p \nmid n'$. p^α и n' нечётные и не равны 1 (в силу того, что n не является степенью простого) $\implies p^\alpha, n' \geqslant 3$.

Согласно китайской теореме об остатках, $\exists\, a \in \overline{1, n-1}$, удовлетворяющий следующим двум уравнениям:

$$a \equiv a_0 \bmod p^\alpha, \quad a \equiv 1 \bmod n'.$$

Выше показали, что $(a_0, n) = 1 \implies (a, n) = 1$ (т.к. $(a, n') = 1$), то есть a является обратимым $\bmod\ n$. Тогда для доказательства того, что подгруппа G_n не является собственной, остаётся показать, что $a \notin G_n$.

$$a^{2^{i_0}k} \equiv a_0^{2^{i_0}k} \equiv (-1)^k \equiv -1 \bmod p^\alpha \implies a^{2^{i_0}k} \not\equiv 1 \bmod n$$

в силу того, что $-1 \not\equiv 1 \bmod p^\alpha$ (т.к. $p^\alpha \geqslant 3$). Кроме того,

$$a^{2^{i_0}k} \equiv 1 \bmod n' \implies a^{2^{i_0}k} \not\equiv -1 \bmod n$$

в силу того, что $-1 \not\equiv 1 \bmod n'$ (т.к. $n' \geqslant 3$). Таким образом, $a \notin G_n$, что завершает доказательство данного случая, а с ним и всей теоремы. ■

Теорема 2.1 доказазывается аналогичным образом, оценка $1/4$ на число не свидетелей достигается за счёт двукратного применения приёма с собственной подгруппой. Полное доказательство описано в [2].

3. Алгоритм Эдельмана-Хуана

4. Алгоритм ECPP

5. Список литературы

[1] Д.В. Мусатов. “Сложность вычислений.”

[2] Conrad, Keith. (2017). “The Miller – Rabin Test.”

[3] Monier, Louis. (1980). "Evaluation and comparison of two efficient probabilistic primality testing algorithms". Theoretical Computer Science. 12. 97–108.