

PRIMALITY

RANDOM VS FIXED INPUT

② PRIMALITY OF RANDOM INPUTS

In RSA we must generate large random primes ($p, q : |p|, |q| \approx 1024$ or larger).

How many primes are there?

PRIME NUMBER THEOREM (no proof)

Let $\pi(n) = |\{1 \leq p \leq n : p \text{ prime}\}|$ (# primes in $\{1, \dots, n\}$)

Then $\lim_{n \rightarrow \infty} \frac{\pi(n)}{n/\ln n} = 1$

That is, $\pi(n) = \frac{n}{\ln n} + O\left(\frac{n}{\ln n}\right)$

In fact, it can be proved that for $n \geq 17$

$$\pi(n) \geq \frac{n}{6\ln n}$$

Therefore, for $n \geq 17$:

$$\Pr(X = \text{RANDOM}\{1, \dots, n\} \text{ is prime}) \geq \frac{1}{6\ln n}$$

The PNT states that primes are quite dense:

there is a prime approximately every $\ln n$ numbers:

$$n = 2^{1024} \quad (\text{extraction of numbers of } 1024 \text{ bits})$$

On average, $\ln(n) = \ln(2^{1024}) \approx 710$ extractions to obtain a prime.

NOTA BENE: If I have a good primality test for random numbers, I can obtain a large prime $\leq n$ by applying the test $\sim \ln n$ times to $\text{RANDOM}(1, 2, \dots, n)$

OBSERVATION: We do not have to worry that we extract a prime $p \ll n$! e.g.

the probability that $P < \frac{n}{\ln n}$ (thus, represented on $\sim 10^4 - 30$ bits) is $< 10^{-9}$ (all 30 most significant bits set to 0)



A simple test for checking primality of random numbers

Recall that Fermat's Little Theorem states that if n is prime:

$$\forall a \in \mathbb{Z}_n^* (\neq \mathbb{Z}_n^+): a^{n-1} \bmod n = 1$$

This is only a necessary condition! There are composite numbers satisfying:

$$\forall a \in \mathbb{Z}_n^* (\neq \mathbb{Z}_n^+): a^{n-1} \bmod n = 1$$

These are called Carmichael's numbers and "behave" like primes on \mathbb{Z}_n^* .

Fortunately these numbers are rather rare therefore unlikely to be selected at random.

DEF Given $n \in \mathbb{N}^+$ composite and $a \in \mathbb{Z}_n^*$, we say that n is a a base- a pseudoprime if $a^{n-1} \pmod{n} = 1$

In other words, n "behaves like a prime" w.r.t. a .

OBSERVATION: If $b \in \mathbb{Z}_n^+ - \mathbb{Z}_n^*$ we surely have

$$b^{n-1} \pmod{n} \neq 1 \quad \text{or otherwise}$$

$$b \cdot b^{n-2} = b^{n-2} \equiv 1 \pmod{n} \Rightarrow b^{n-2} = b^{-1}$$

When n is composite: implausible $b \in \mathbb{Z}_n^*$

- $a \in \mathbb{Z}_n^*$: $a^{n-1} \equiv 1 \pmod{n}$ is a "cheater" (makes n "look" like a prime)
- $b \in \mathbb{Z}_n^+ - \mathbb{Z}_n^*$: $b^{n-1} \not\equiv 1 \pmod{n}$ is a certificate of nonprimality ($b \in \mathbb{Z}_n^*$ or $b \in \mathbb{Z}_n^+ - \mathbb{Z}_n^*$)

For some numbers n , values b : $b^{n-1} \not\equiv 1 \pmod{n}$ may be difficult to find.

e.g. for Carmichael numbers:

$b^{n-1} \not\equiv 1 \pmod{n}$ only for $b \in \mathbb{Z}_n^+ - \mathbb{Z}_n^*$ (very few)

However, the vast majority of composite numbers admits 2 as a certificate of nonprimality!

Pomerance Theorem (1981)

Let X be a random number in $\{1, \dots, n\}$.
Then

$\text{Prob}(X \text{ is base-2 pseudoprime}) \leq ce^{-\frac{1}{2} \ln \frac{\ln \ln n}{\ln \ln n}}$

$2^{x-1} \equiv 1 \pmod{x}$ $\rightarrow 0$
 $n \rightarrow \infty$

Based on Pomerance Theorem we can devise this simple test:

PSEUDOPRIME(x)

if $\text{HOD-EXP}(2, x-1, x) \neq 1$
then return COMPOSITE
else return PRIME

PSEUDOPRIME is always correct when it returns COMPOSITE but may be wrong when it returns PRIME if X is a base-2 pseudoprime. (one-sided error)

However, by Pomerance Theorem, if X is randomly chosen in $\{1, \dots, n\}$:

$$\Pr(\text{PSEUDOPRIME}(x) \text{ is incorrect} \mid \text{on random } X \in \{1, \dots, n\}) \leq ce^{-\frac{1}{2} \ln \frac{\ln \ln n}{\ln \ln n}}$$

We can use PSEUDOPRIME(x) for the generation of random primes as follows:

```

RANDOM_PRIME( $n, s$ ) → parametro di precisione
repeat  $s$ . times
     $x \leftarrow \text{RANDOM}(\{1, \dots, n\})$ 
    if PSEUDOPRIME( $x$ ) = PRIME
        then return  $x$ 
return FAILURE

```

CORRECTNESS ($n \geq 17$) event A \Rightarrow event B
 $\Pr(A) \leq \Pr(B)$

$$\begin{aligned}
 1) \Pr(\text{RP}(n, s) = \text{FAILURE}) &\leq \Pr(\text{extract } s \text{ non composite numbers}) \leq \\
 &\leq \left(1 - \frac{\pi(n)}{n}\right)^{s \cdot \pi(n)} \stackrel{\pi(n) \geq \frac{n}{2}}{\leq} \left(1 - \frac{n}{2n} \cdot \frac{n}{s}\right)^{s \cdot \frac{n}{2}} \\
 &= \left(1 - \frac{1}{2s}\right)^{s \cdot \frac{n}{2}} \leq e^{-s} \quad \left(1 - \frac{1}{t}\right)^t \leq e^{-1}
 \end{aligned}$$

We can make this probability very small by choosing large enough s . E.g., $s=21 \Rightarrow < 10^{-9}$

$$\begin{aligned}
 2) \Pr(\text{RP}(n, s) \text{ is composite}) &= \\
 &= \Pr(\text{RP}(n) \text{ returns a base-2 pseudoprime}) \\
 &\leq \Pr\left(\bigcup_{i=1}^{s \cdot \pi(n)} \text{"at iteration } i, \text{ RANDOM}(\{1, \dots, n\}) \text{ is a base-2 pseudoprime"}\right) = \Pr\left(\bigcup_{i=1}^{s \cdot \pi(n)} E_i\right) \\
 &\leq s \cdot \pi(n) \cdot \Pr(\text{RANDOM}(\{1, \dots, n\}) \text{ is a base-2 pseudop.})
 \end{aligned}$$

\uparrow union of $s \cdot \pi(n)$ events

$$\leq \text{Shu } n e^{-\frac{1}{2} \ln n \frac{\ln \ln n}{\ln n}} \rightarrow 0 \text{ for } n \rightarrow \infty$$

(for 1024-bit numbers, $s=21$:

$$\Pr(\text{RP}(n) \text{ is composite}) < 10^{-43}!$$

clearly

$$\Pr(\text{RP}(n,s) \text{ is incorrect}) \leq 1 + 2 \\ (< 10^{-9} + 10^{-43} \text{ for 1024-bit numbers and } s=21)$$

RUNNING TIME

Each iteration of $\text{RP}(n)$ executes

$\text{MOD-EXP}(2, x-1, x)$, $x \leq n$, running in time $\mathcal{O}(|\log n|^3)$. Since $\text{Shu } n = \mathcal{O}(s|\log n|)$:

$$T_{\text{RP}}(|\log n|, s) = \mathcal{O}(s |\log n|^4)$$

Worst-case analysis very conservative. In practice much faster (use speed-up tricks, e.g. test for divisibility by small factors)

MILLER-RABIN'S (MR) PRIMALITY TEST

The test will check whether a specific number $n > 1$ is prime or composite.

OBSERVATION : Very different from the PSEUDOPRALITY test for random inputs that uses a deterministic algorithm:

- always correct on a large fraction of the inputs
- always wrong on a small set of inputs (base-2 pseudoprimes)

The MR primality test works in the worst case (e.g., even if the input is a base-2 pseudoprime or a Carmichael number).

The MR test is a randomized algorithm: it makes use of RANDOM(.) to check for primality/compositeness.

The correctness analysis is performed w.r.t the probability space induced by the calls to RANDOM(.)

MAIN IDEA : Given a fixed $n > 1$:

- randomized search for a "certificate" of nonprimality (compositeness).
- The search for such a certificate is repeated s times. The probability of "missing" such a certificate when n is composite is $\leq \frac{1}{2^s}$ (becomes quickly very small)
- If no certificate is found in the s iterations, MR(n) declares that n is prime
- MR is a randomized, one-sided error algorithm:
 1. always correct if n is prime
 2. may be incorrect if n is composite (but with probability that can be made very small)

CERTIFICATES OF NONPRIMALITY USED IN MR(n)

1. We look for random values $a \in \mathbb{F}_n^\times$:

$$a^{n-1} \not\equiv 1 \pmod{n}$$

(in other words, look for a : n is not a base- a pseudoprime)

REMARK: This extends the test for random numbers based on base-2 pseudoprimes to arbitrary values of $a \in \mathbb{Z}_n^+$. Clearly this is not sufficient for "hard" inputs!

Carmichael's numbers n

$$\forall a \in \mathbb{Z}_n^*: a^{n-1} \equiv 1 \pmod{n}.$$

For these numbers, the only valid certificates may come from

$$a \in \mathbb{Z}_n^+ - \mathbb{Z}_n^* \quad (\text{since } a^{n-1} \not\equiv 1 \pmod{n})$$

but these values may be very few w.r.t n

(in fact, Carmichael numbers tend to have very few prime factors $\Rightarrow |Z_n^+ - Z_n^*| \ll n$)

2. Look for nontrivial square roots of unity:

$$d \in \mathbb{Z}_n^+: (d \not\equiv \pm 1 \pmod{n}) \wedge (d^2 \equiv 1 \pmod{n})$$

(these are numbers $d \in \mathbb{Z}_n^*: d \neq \pm 1 \pmod{n}$)

Why?

THEOREM:

If $\exists d \in \mathbb{Z}_n^+: (d \not\equiv \pm 1 \pmod{n}) \wedge (d^2 \equiv 1 \pmod{n})$
 $\Rightarrow n$ is composite

EXAMPLE: 4 in \mathbb{Z}_{15} : $(4 \not\equiv \pm 1 \pmod{15}) \wedge (4^2 \equiv 1 \pmod{15})$

PROOF : We prove the contrapositive:

n prime \Rightarrow the only values $d \in \mathbb{Z}_n$:
 $d^2 \equiv 1 \pmod{n}$ are $d \equiv \pm 1 \pmod{n}$

Consider the congruent equation:

$$\begin{aligned}x^2 &\equiv 1 \pmod{n} \Leftrightarrow \\x^2 - 1 &\equiv 0 \pmod{n} \Leftrightarrow \\(x+1)(x-1) &\equiv 0 \pmod{n} \Leftrightarrow \\(x+1)(x-1) &= kn, \quad k \in \mathbb{N}\end{aligned}$$

Thus

$$n \mid (x+1)(x-1).$$

Since n is prime:

$$n \mid ab \Rightarrow (n \mid a) \vee (n \mid b)$$

or otherwise

$$\gcd(n, a) = 1 \text{ and}$$

$$\gcd(n, b) = 1 \text{ therefore}$$

$$\gcd(n, ab) = 1 \quad (\text{Bezout})$$

which contradicts $n \mid ab \quad (\Rightarrow \gcd(n, ab) = n)$

Therefore

$$n \mid (x+1) \quad \text{or} \quad n \mid (x-1)$$

$$n \mid (x+1) \Rightarrow x+1 = k_1 n \Rightarrow x = -1 + k_1 n$$

$$n \mid (x-1) \Rightarrow x-1 = k_2 n \Rightarrow x = 1 + k_2 n$$

It must be

$$x \equiv \pm 1 \pmod{n}$$

Q.E.D.