Public Key Cryptography

Lecture 3

Primality Tests

Index

Record Primes

http://primes.utm.edu/

The largest known prime:

$$2^{2^{136279841}} - 1$$

announced in October 2024.

- It has 41024320 digits.
- It is a Mersenne prime (probably Mersenne 52).
  Mersenne number: a number of the form $2^n - 1$.
  In order to have $2^n - 1$ prime, n must be prime.
- Discovered by GIMPS (The Great Internet Mersenne Prime Search) project (http://www.mersenne.org/).
- Coordination of clusters of PCs.

Record Primes (cont.)

The “Top Ten” Record Primes (as of October 21, 2024):

prime	digits	year	Reference
$2^{136279841}$	41024320	2024	Mersenne 52?
$2^{82589933} - 1$	24862048	2018	Mersenne 51?
$2^{77232917} - 1$	23249425	2018	Mersenne 50?
$2^{74207281} - 1$	22338618	2016	Mersenne 49?
$2^{57885161} - 1$	17425170	2013	Mersenne 48
$2^{43112609} - 1$	12978189	2008	Mersenne 47
$2^{42643801} - 1$	12837064	2009	Mersenne 46
$516693^{2097152} - 516693^{1048576} + 1$	11981518	2023	
$465859^{2097152} - 465859^{1048576} + 1$	11887192	2023	
$2^{37156667} - 1$	11185272	2008	Mersenne 45

Primality: the Problem

Theorem (Euclid)

There are infinitely many primes.

Proof. Suppose that

$$p_1 = 2 < p_2 = 3 < ... < p_r$$

are all of the primes. Denote

$$P = p_1 p_2 ... p_r + 1$$

and let p be a prime dividing P. Then p can not be any of $p_1, p_2, \ldots, p_r$, otherwise p would divide $P - p_1 p_2 \ldots p_r = 1$, impossible. So P is still another prime, contradiction.

The problem

Decide if a given (large) number is prime.

Primality Tests

Definition

- *primality test: a criterion to decide if a number is prime.*
- *compositeness test: a criterion to decide if a number is composite.*

Even if they are not the same, we will call them generically *primality tests*.

- If n passes a primality test, then it *may* be prime.
- If n passes a whole lot of primality tests, then it is more likely to be prime.
- If n fails a single primality test, then it is surely composite.

A Bit of History of Primality Tests

- Elementary primality tests:
  trial division, sieve of Eratosthenes
- Fermat test
  It is based on Fermat's Little Theorem. It became the basis for many efficient primality tests.
- Randomized polynomial-time algorithms (1970's and 1980's): e.g. Solovay-Strassen and Miller-Rabin tests.
- True primality tests: e.g. Lucas-Lehmer test for Mersenne primes.
- Unconditional deterministic polynomial-time algorithm:
  Agrawal, Kayal, Saxena (2003), Lenstra

Elementary Primality Tests

- **Trial Division**
  Take an odd $m \neq 1$ and check if $m|n$.
  If n passes the trial division tests for more and more values of m, it becomes more and more likely that n is prime.
  We know that if n passes the trial division tests for every $m \leq \sqrt{n}$, then n is prime.
  If n fails a single trial division test for some m, then n is surely composite.
  Weak point: complexity $O(n^{\frac{1}{2}})$.
- **The Sieve of Eratosthenes**
  This generates all primes less then n.
  The best method for small primes (up to 1000000).
  Weak point: a lot of memory for storage.

Fermat Test

In what follows let n be a large odd natural number.
By Fermat's Little Theorem, if n is prime, then $\forall b \in \mathbb{Z}$ (enough $b < n$) with $(b, n) = 1$ we have

$$b^{n-1} = 1 \pmod{n} \tag{1}$$

If n is not prime, it is still possible (but probably not very likely) that (1) holds.

Definition

An odd composite natural number n is called *pseudoprime to the base b* if $(b, n) = 1$ and (1) holds.

Remarks. (a) A pseudoprime is a number that "pretends" to be prime by passing the test (1).
(b) Every odd natural number is pseudoprime to the bases $b = \pm 1$.
(c) $\forall b \in \mathbb{Z}$ with $|b| \geq 2$, there are infinitely-many pseudoprimes to the base b.

Fermat Test (cont.)

Example. $n = 91$ is pseudoprime to the base $b = 3$, because $3^{90} = 1 \pmod{91}$. But 91 is not pseudoprime to the base 2, because $2^{90} = 64 \pmod{91}$.
If we did not already know that 91 is composite, the fact that $2^{90} \neq 1 \pmod{91}$ would tell us that it is.

Theorem

Let $n \in \mathbb{N}$ be an odd composite.
(i) n pseudoprime to $b \Rightarrow n$ pseudoprime to $-b$ and b^{-1}, where b^{-1} is the inverse modulo n of b.
(ii) n pseudoprime to b_1 and $b_2 \Rightarrow n$ pseudoprime to $b_1 b_2$.
(iii) If n fails (1) for a single base $b < n$, then n fails (1) for at least half of the possible bases $b < n$.

Fermat Test (cont.)

- Unless n happens to pass the test (1) for every b with $(b, n) = 1$, there is at least a 50% chance that n will fail (1) for a randomly chosen b.
- If n is composite, then Fermat's test reveals this fact with a 100% probability and if n is prime, then Fermat's test reveals this fact with a high probability. If (1) does not hold for any b, then n is surely composite.
- Suppose that we have considered k different values for b and n is pseudoprime to all these bases. Then the probability that n is still composite despite passing the k tests is at most $\frac{1}{2^k}$, unless n happens to have the very special property that (1) holds for every $b \in \mathbb{Z}$. Hence if k is large, we can say with a high probability that n is prime.
- Such a method is called a *probabilistic* method. A *deterministic* method would tell us with a 100% certainty whether n is either composite or prime.

Fermat Test (cont.)

Fermat Primality Test

- Fermat(n,k)
- Input: $n \in \mathbb{N}$, $n \geq 3$ odd and $k \in \mathbb{N}^*$.
- Output: n is either composite or, with a high probability $(1 - \frac{1}{2^k})$, prime.
- Algorithm:

For $i = 1$ to k do
    Randomly choose $1 < b < n - 1$;
    Compute $r := b^{n-1} \pmod{n}$;
    If $r \neq 1$ then output COMPOSITE;
Output PRIME.

Fermat Test (cont.)

Remarks

- If the algorithm gives the answer COMPOSITE, then this is for sure.
- If the algorithm gives the answer PRIME, then the probability that n is composite is less than $\frac{1}{2^k}$.

Weak point: Carmichael numbers.

Definition

A composite natural number n is called a *Carmichael number* if (1) holds $\forall b \in \mathbb{Z}$ with $(b, n) = 1$.

Fermat Test (cont.)

Theorem

Let $n \in \mathbb{N}$ be odd composite.
(i) If n is divisible by a perfect square different of 1, then n is not a Carmichael number.
(ii) If n is square free (that is, it is not divisible by the square of any prime), then n is a Carmichael number $\Leftrightarrow p-1|n-1$ for every prime $p|n$.

Example. $n = 561 = 3 \cdot 11 \cdot 17$ is a Carmichael number, because 560 is divisible by 2,10 and 16. This is the least Carmichael number.
It has been proved that there are infinitely-many Carmichael numbers, they being relatively rare.
For instance, there are only 105212 Carmichael numbers less than 10^{15}.

Miller-Rabin Test

- widely used in practice for RSA
- relies on the notion of strong pseudoprime

Let $n \in \mathbb{N}$ be odd and $b \in \mathbb{Z}$ with $(b, n) = 1$.
If n is pseudoprime to b, then $b^{n-1} = 1 \pmod{n}$.

Idea of the Miller-Rabin test:
Successively extract the square roots from the previous congruence, that is, raise b to $\frac{n-1}{2}, \frac{n-1}{4}, \ldots, \frac{n-1}{2^s}$, where $t = \frac{n-1}{2^s}$ is odd. Then the first result different of 1 has to be -1 if n is prime, because ± 1 are the only square roots modulo a prime of 1.

In practice, we write $n - 1 = 2^s t$ for some odd t. Then compute $b^t \pmod{n}$. If it is not 1, then we compute its successive squares $b^{2t} \pmod{n}$, $b^{2^2 t} \pmod{n}$ etc. until we get 1 and the algorithm stops because in the step immediately before getting 1, we should have obtained -1, otherwise n being composite.

Miller-Rabin Test (cont.)

Miller-Rabin Test relies on the following result:

Theorem

Let p be a prime. Then the equation

$$a^2 = 1 \pmod{p}$$

has only the solutions $a = 1 \pmod{p}$ and $a = -1 \pmod{p}$.

Proof. We may assume that $a \in \{0, \ldots, p-1\}$.
We have

$$a^2 = 1 \pmod{p} \Leftrightarrow p|(a-1)(a+1).$$

It follows that $p|a-1$ or $p|a+1$.
If $p|a-1$, then $a-1=0$, because $a-1<p$. Hence $a=1$.
If $p|a+1$, then $a+1=0$ or $a+1=p$, because $a+1<p+1$.
Hence $a=p-1=-1$.

Miller-Rabin Test (cont.)

Definition

Let $n \in \mathbb{N}$ be odd composite and write $n-1 = 2^s t$ for some odd t. Let $b \in \mathbb{Z}$ with $(b,n) = 1$. If n and b satisfy the condition

$$b^t = 1 \pmod{n} \text{ or } \exists 0 \le j < s : b^{2^j t} = -1 \pmod{n} \quad (2)$$

then n is called *strong pseudoprime to the base b.*

One can show that (2) holds for n prime and $(b,n) = 1$.

Theorem

Strong pseudoprime to the base b $\Rightarrow$ pseudoprime to the base b.

Example. Let $n = 65$ and $b = 14$. We have $65 - 1 = 2^6 \cdot 1$. Then $14 \neq \pm 1 \pmod{65}$, $14^2 = 1 \pmod{65}$, hence $14^{2^j} = 1 \neq -1 \pmod{65}$ for $1 \le j < s = 6$. Thus 65 is not strong pseudoprime to the base 14. But $b^{n-1} = 14^{64} = 1 \pmod{65}$, hence 65 is pseudoprime to the base 14.

Miller-Rabin Test (cont.)

Theorem

Let $n \in \mathbb{N}$ be an odd composite.
(i) If n is a strong pseudoprime to b, then n is a strong pseudoprime to b^k for every $k \in \mathbb{Z}$.
(ii) n is a strong pseudoprime to b for at most 25% of the values $0 < b < n$.

In general, if n is a strong pseudoprime to a base b_1 and to a base b_2, then it does not follow that n is a strong pseudoprime to the base $b_1 b_2$.

Example. Consider $n = 65$. The number of possible bases is $N = \varphi(n) = 4 \cdot 12 = 48$. Then n is:
(i) pseudoprime to the bases
$\pm1, \pm8, \pm12, \pm14, \pm18, \pm21, \pm27, \pm31.$ $(N/3)$
(ii) strong pseudoprime to the bases
$\pm1, \pm8, \pm18.$ $(N/8)$

Miller-Rabin Test (cont.)

- Let p be a prime. Write $p - 1 = 2^s \cdot t$, where t is odd.
- Choose $1 < a < p$.
- Consider the following sequence (computed by the repeated squaring modular exponentiation):

$$a^t, a^{2t}, a^{2^2 t}, \ldots, a^{2^s t}$$

  where each number is reduced modulo p.
- Characteristics of the sequence:
  (1) Eventually it gets to the value 1 (and remains 1).
  [It follows by Fermat's Little Theorem: $a^{2^s t} = a^{p-1} = 1 \pmod{p}$, because p is prime.]
  (2) The previous number in the sequence (if it does exist) to the first value 1 must be $-1 \pmod{p}$.
  [It follows by the fact that ± 1 are the only square roots modulo p of 1.]

Miller-Rabin Test (cont.)

Miller-Rabin Test

- Miller-Rabin(n,k)
- Input: $n \in \mathbb{N}$, $n \geq 3$ odd, and $k \in \mathbb{N}^*$.
- Output: n is composite or, with probability $1 - \frac{1}{4^k}$, n is prime.
- Algorithm:

Step 0. Write $n - 1 = 2^s t$, where t is odd.

Step 1. Choose (randomly) $1 < a < n$.

Step 2. Compute (by the repeated squaring modular exponentiation) the following sequence (modulo n):

$$a^t, a^{2t}, a^{2^2 t}, \ldots, a^{2^s t}$$

Step 3. If either the first number in the sequence is 1 or if one gets the value 1 and its previous number -1, then n is possible to be prime and one repeats Steps 1-3 at most k times.

If one does not get to Step 4, then the algorithm stops and n is probable prime.

Step 4. The algorithm stops and n is composite.

Miller-Rabin Test (cont.)

Remarks

- If the algorithm gives the answer COMPOSITE, then this is for sure.
- If the algorithm gives the answer PRIME, then the probability of correct answer is $1 - \frac{1}{4^k}$, where k is the number of repetitions.
- For $k = 50$, the probability that the Miller-Rabin Test gives a wrong PRIME answer is at most

$$\frac{1}{4^{50}} = \frac{1}{1267650600228229401496703205376}.$$

This is much less than the probability to obtain incorrect results because of a hardware error.

Miller-Rabin Test (cont.)

Example. Let us check with the Miller-Rabin test if $n = 409$ is prime (with 3 repetitions if necessary).

Step 0. Write $n - 1 = 408 = 2^3 \cdot 51$, hence $s = 3$ and $t = 51$.

$$\boxed{\text{k=1}}$$

Step 1. Choose $a = 2$.
Step 2. Compute the following sequence (modulo $n = 409$):

$$2^{51}, 2^{2 \cdot 51}, 2^{2^2 \cdot 51}, 2^{2^3 \cdot 51}.$$

Step 3. We have:

- $2^{51} = 143 \pmod{409}$ (repeated squaring modular exp.),
- $2^{2 \cdot 51} = (2^{51})^2 = 143^2 = 408 = -1 \pmod{409}$,
- $2^{2^2 \cdot 51} = (2^{2 \cdot 51})^2 = (-1)^2 = 1 \pmod{409}$,
- $2^{2^3 \cdot 51} = (2^{2^2 \cdot 51})^2 = 1 \pmod{409}$.

Hence $n = 409$ is possible to be prime [the sequence is: 143,-1,1,1].

$$\boxed{\text{k=2}}$$

Step 1. Choose $a = 3$.
Step 2. Compute the following sequence (modulo $n = 409$):

$$3^{51}, 3^{2\cdot 51}, 3^{2^2\cdot 51}, 3^{2^3\cdot 51}.$$

Step 3. We have:

- $3^{51} = 266 \pmod{409}$ (repeated squaring modular exp.),
- $3^{2\cdot 51} = (3^{51})^2 = 266^2 = 408 = -1 \pmod{409}$,
- $3^{2^2\cdot 51} = (3^{2\cdot 51})^2 = (-1)^2 = 1 \pmod{409}$,
- $3^{2^3\cdot 51} = (3^{2^2\cdot 51})^2 = 1 \pmod{409}$.

Hence $n = 409$ is possible to be prime [the sequence is: 266,-1,1,1].

Miller-Rabin Test (cont.)

$\boxed{\text{k=3}}$

Step 1. Choose $a = 5$.
Step 2. Compute the following sequence (modulo $n = 409$):

$$5^{51}, 5^{2\cdot 51}, 5^{2^2\cdot 51}, 5^{2^3\cdot 51}.$$

Step 3. We have: $5^{51} = 1 \pmod{409}$ (repeated squaring modular exp.).
Hence n is possible to be prime [the sequence is: 1,1,1,1].

According to the algorithm, $n = 409$ is probable prime. The probability of error is less than $1/4^3$.

Miller-Rabin Test (cont.)

Example. Let us check with the Miller-Rabin test if $n = 413$ is prime (with 3 repetitions if necessary).

Step 0. Write $n - 1 = 412 = 2^2 \cdot 103$, hence $s = 2$ and $t = 103$.

$$\boxed{\text{k=1}}$$

Step 1. Choose $a = 2$.
Step 2. Compute the following sequence (modulo $n = 413$):

$$2^{103}, 2^{2 \cdot 103}, 2^{2^2 \cdot 103}.$$

Step 3. We have:

- $2^{103} = 72 \pmod{413}$ (repeated squaring modular exp.),
- $2^{2 \cdot 103} = (2^{103})^2 = 72^2 = 228 \pmod{413}$,
- $2^{2^2 \cdot 103} = (2^{2 \cdot 103})^2 = 228^2 = 359 \pmod{413}$.

Hence $n = 413$ is surely composite [the sequence is: 72,228,359].

Miller-Rabin Test (cont.)

- In practice, we check just for few bases. For instance, there is only one composite number $< 2,5 \cdot 10^{10}$ that is strong pseudoprime to all the bases $b = 2, 3, 5, 7$.
- Let $p_1, p_2, \ldots, p_l$ be the first l primes and ψ_l the smallest positive composite integer which is a strong pseudoprime to all the bases $p_1, p_2, \ldots, p_l$. In order to determine the primality of any integer $n < \psi_t$, it is enough to apply Miller-Rabin to n with $b = p_1, \ldots, p_l$. In this way, the answer returned by Miller-Rabin is always correct.

l	ψ_l
1	2047
2	1373653
3	25326001
4	3215031751
5	2152302898747

AKS Test

- Agrawal, Kayal, Saxena (2002)
- the first deterministic general polynomial-time algorithm for testing primality
- nevertheless, Miller-Rabin Test is used in practice, because AKS Test has a rather high (even if polynomial) complexity

Based on the following generalization of Fermat's Little Theorem:

Theorem

Let $n \in \mathbb{N}$, $n \geq 2$ and $a \in \mathbb{Z}$ such that $(a, n) = 1$. Then n is prime $\Leftrightarrow$ the following polynomial congruence holds

$$(X + a)^n = X^n + a \pmod{n}.$$

AKS Test (cont.)

A simple test for primality would be: given an input n, choose an a with $(a, n) = 1$ and test whether the congruence is satisfied.

However, this takes time $O(n)$ because we need to evaluate n coefficients in the worst case.

A simple way to reduce the number of coefficients is to evaluate both sides of the congruence modulo a polynomial of the form $X^r - 1$ for an appropriately chosen small r. In other words, test if the following equation is satisfied:

$$(X + a)^n = X^n + a \pmod{X^r - 1, n} \tag{1}$$

where for $f, g, h \in \mathbb{Z}_n[X]$ we use the notation $f = g \pmod{h, n}$ to represent the equation $f = g$ in the ring $\mathbb{Z}_n[X]/(h)$ (see a subsequent chapter on polynomials and finite fields).

AKS Test (cont.)

All primes n satisfy the equation (1) for all values of a and r.

But some composites n may also satisfy the equation for a few values of a and r (and indeed they do).

However, we can almost restore the characterization: one shows that, for appropriately chosen r, if the equation (1) is satisfied for several a's then n must be a prime power. The number of a's and the appropriate r are both bounded by a polynomial in $\log n$ and therefore, we get a deterministic polynomial-time algorithm for testing primality.

Given $r \in \mathbb{N}$ and $a \in \mathbb{Z}$ with $(a, r) = 1$, we denote by $o_r(a)$ the order of a modulo r (that is, the smallest non-zero power k of a such that $a^k \bmod r = 1$). We have $o_r(a) | \varphi(r)$ (Euler's function) for any a with $(a, r) = 1$.

AKS Test (cont.)

AKS Test

- *Input*: $n \in \mathbb{N}$, $n \geq 2$.
- *Output*: n is prime or composite.
- *Algorithm*:
  1. If ($n = a^b$ for $a \in \mathbb{N}$ and $b > 1$), output COMPOSITE.
  2. Find the smallest r such that $o_r(n) > 4\log^2 n$.
  3. If $1 < (a, n) < n$ for some $a \leq r$, output COMPOSITE.
  4. If $n \leq r$, output PRIME.
  5. For $a = 1$ to $[2\sqrt{\varphi(r)}\log n]$ do
     If $(X + a)^n \neq X^n + a \pmod{X^r - 1, n}$,
     then output COMPOSITE.
  6. Output PRIME.

Theorem

AKS Test returns PRIME if and only if n is prime.

Selective Bibliography

- M. Agrawal, N. Kayal, N. Saxena, *PRIMES is in P*, Annals of Mathematics, 160 (2004), 781–793.
- M. Cozzens, S.J. Miller, *The Mathematics of Encryption: An Elementary Introduction*, American Mathematical Society, 2013.
- N. Koblitz, *A Course in Number Theory and Cryptography*, Springer, 1994.
- A.J. Menezes, P.C. van Oorschot, S.A. Vanstone, *Handbook of Applied Cryptography*, CRC Press, 1997. [http://www.cacr.math.uwaterloo.ca/hac]
- http://primes.utm.edu