The Miller Rabin Test

CS 719 Course Report

Aryaman Maithani

November 4, 2021

Contents

§1. Introduction

In this report, we are concerned with finding an algorithm for the following problem.

Input: An integer $n > 1$.
Output: isPrime(n).

The simplest way to do this is by trial division. Indeed, we simply divide n by $2, 3, 4$, and so on, and see if the remainder is 0 in any case. As we know, we only need to divide by numbers up to $\sqrt{n}$. The issue with this algorithm is that it is extremely inefficient, requiring $\Theta(\sqrt{n})$ operations, which is *exponential* in the *bit length* $\text{len}(n)$. For example, if n has 100-decimal digits, it would take more than 10^{33} years to perform $\sqrt{n}$ divisions.

However, note that the above algorithm does *more* than what we expected from our algorithm. Namely, it not only tells us that the number is prime but also produces a nontrivial factor in the case that n is composite. Naïvely, one might think that it is necessary for us to produce a prime factor to claim that a number is composite. However, that is not the case.

In this report, we describe a much faster primality testing. This is a polynomial time algorithm. It allows for 100-decimal digits numbers to be tested in less than a second. Unlike the earlier algorithm, it does *not* give us a prime factor in the case that n is composite.

However, this algorithm is *probabilistic*. This means that the algorithm can make a mistake. Fortunately, one has control over this probability, and can make it arbitrarily small (but not zero).

For the rest of the talk, we shall assume that $n > 1$ is an *odd* integer. Let $n = p_1^{e_1} \cdots p_r^{e_r}$ be its prime factorisation.
By $\mathbb{Z}_n$, we shall denote the ring of integers modulo n. We have a ring homomorphism

$$\begin{aligned} \theta : \mathbb{Z}_n &\to \mathbb{Z}_{p_1^{e_1}} \times \cdots \times \mathbb{Z}_{p_r^{e_r}} \\ [a]_n &\mapsto ([a]_{p_1^{e_1}}, \cdots, [a]_{p_r^{e_r}}). \end{aligned}$$

In fact, the Chinese Remainder Theorems tells us that the above is an isomorphism. This gives us a group isomorphism between the group of invertible elements of the two rings as

$$(\mathbb{Z}_n)^* \xrightarrow{\cong} (\mathbb{Z}_{p_1^{e_1}})^* \times \cdots \times (\mathbb{Z}_{p_r^{e_r}})^*. \tag{1}$$

Several probabilistic primality tests, including the Miller–Rabin test, have the following general structure.

> Define $\mathbb{Z}_n^+$ to be the set of nonzero elements of $\mathbb{Z}_n$. Note that $|\mathbb{Z}_n^+| = n - 1$. Moreover, $\mathbb{Z}_n^+ = \mathbb{Z}_n^*$ iff n is prime. Suppose also that we define a set $L_n \subseteq \mathbb{Z}_n^+$ such that:
>
> 1. there is an efficient algorithm that on input n and $\alpha \in \mathbb{Z}_n^+$, determines if $\alpha \in L_n$;
> 2. if n is prime, then $L_n = \mathbb{Z}_n^*$; and
> 3. if n is composite, $|L_n| \leqslant c(n-1)$ for some universal constant $c < 1$.

Algorithm. To test for primality, we set a "repetition parameter" k, and choose random elements $\alpha_1, \ldots, \alpha_k \in \mathbb{Z}_n^+$. If $\alpha_i \in L_n$ for all $i \in \{1, \ldots, k\}$, then we output `true`; otherwise, we output `false`.

Observation 1. Let us note some properties of the above algorithm.

1. The algorithm is efficient since we can check $\alpha \in L_n$ efficiently.
2. If n is prime, then the algorithm outputs `true`, and it does so *correctly*.
3. If n is composite, then the algorithm *may* output `true`, with probability at most c^k.

In particular, note that there is a *one-sided error*. In fancy language, this is a *Monte Carlo algorithm*.

§2. First attempt

We now try to define a suitable candidate for L_n.

Definition 2.

$$L_n := \{\alpha \in \mathbb{Z}_n^+ : \alpha^{n-1} = 1\}. \tag{2}$$

Note that we can test $\alpha \in L_n$ efficiently, using a repeated-squaring algorithm.

Observation 3. It is easy to see that $L_n \subseteq \mathbb{Z}_n^*$. Indeed, α^{n-2} acts as the inverse of $\alpha \in L_n$.

However, one can even note that L_n is a *subgroup* of $\mathbb{Z}_n^*$. Indeed, defining $\varphi : \mathbb{Z}_n^* \to \mathbb{Z}_n^*$ to be the $(n-1)$-power map $x \mapsto x^{n-1}$, one sees that $L_n = \ker(\varphi)$.

Theorem 4. If n is prime, then $L_n = \mathbb{Z}_n^*$.
If n is composite *and* $L_n \subsetneq \mathbb{Z}_n^*$, *then* $|L_n| \leqslant \frac{1}{2}(n-1)$.

Proof. The first statement is clear. For the second, one recalls that L_n is a subgroup of Z_n^*. Thus, $\frac{|\mathbb{Z}_n^*|}{|L_n|}$ is a positive integer. Therefore, if the integer is not 1, it is at least 2. Hence, we see

$$|L_n| \leqslant \frac{1}{2}|\mathbb{Z}_n^*| \leqslant \frac{1}{2}(n-1).$$ □

However, there *are* infinitely many odd composite n for which $L_n = \mathbb{Z}_n^*$ and thus, they cannot be ignored.

Definition 5. An odd composite number n such that $L_n = \mathbb{Z}_n^*$ is called a *Carmichael number*.

Example 6. The smallest Carmichael number is $561 = 3 \cdot 11 \cdot 17$.

Theorem 7. n is a Carmichael number iff n is of the following form:

1. $n = p_1 \cdots p_r$ for distinct primes p_i,
2. $r \geqslant 3$,
3. $(p_i - 1) \mid (n-1)$ for all $i \in \{1, \dots, r\}$.

Proof. Let $n = p_1^{e_1} \cdots p_r^{e_r}$ be a Carmichael number. Recalling (1), we have

$$\mathbb{Z}_n^* \cong \mathbb{Z}_{p_1^{e_1}}^* \times \cdots \times \mathbb{Z}_{p_r^{e_r}}^*.$$

Since $n-1$ annihilates the left group, it annihilates the right group. But this happens iff

$$p_i^{e_i - 1}(p_i - 1) \mid (n-1)$$

for all $i \in \{1, \ldots, r\}$ (since each factor on the right is a cycle group). In particular, $(p_i - 1) \mid (n-1)$. Moreover, if $e_i > 1$ for some i, then $p_i \mid n-1$, a contradiction. Thus, $e_i = 1$ for all i.

Now, we must show that $r \geqslant 3$. For the sake of contradiction, assume that $r = 2$. In this case, we have $n = p_1 p_2$ for some $p_1 > p_2$. We note that

$$n - 1 = p_1 p_2 - 1 = (p_1 - 1)p_2 + (p_2 - 1).$$

The above shows that $p_1 - 1 \mid p_2 - 1$, a contradiction since $p_1 > p_2$.

Conversely, suppose n has the given form. Let a be coprime to n and hence, to each p_i. Then, by Fermat's Little Theorem, we have $a^{p_i - 1} \equiv 1 \mod p_i$. Since $n-1$ is a multiple of $p_i - 1$, we get

$$a^{n-1} \equiv 1 \mod p_i$$

for all $i \in \{1, \ldots, r\}$. By the Chinese Remainder Theorem, we are now done. □

§3. The Miller-Rabin test

We now define a new set L_n' as follows.

Definition 8. Let $n - 1 = t2^h$ where t is odd, and $h \geqslant 1$.

$$\begin{aligned} L_n' := \{\alpha \in \mathbb{Z}_n^+ : \alpha^{t2^h} &= 1 \text{ and} \\ \alpha^{t2^{j+1}} &= 1 \Rightarrow \alpha^{t2^j} = \pm 1 \\ &\text{for } j = 0, \ldots, h-1\}. \end{aligned} \tag{3}$$

The Miller-Rabin test uses this set L_n'. By definition, it is clear that $L_n' \subseteq L_n$, since we have the condition (2) from earlier.
In fact, L_n' is precisely the set of those elements of L_n which also satisfy (3).

Testing whether a given $\alpha \in \mathbb{Z}_n^+$ belongs to L_n' can be done using the following algorithm:

Algorithm (Testing membership).

1. $\beta \leftarrow \alpha^t$
2. if $\beta = 1$ then return `true`
3. for $j \leftarrow 0$ to $h-1$ do
   - if $\beta = -1$ then return `true`
   - if $\beta = 1$ then return `false`
   - $\beta \leftarrow \beta^2$
4. return `false`

This algorithm runs in time $O(\text{len}(n)^3)$ and thus, satisfies the first criteria.

Theorem 9. If n is prime, then $L'_n = \mathbb{Z}_n^*$. If n is composite, then $|L'_n| \leqslant \frac{1}{4}(n-1)$.

Proof. **Case 1.** n is prime.

Note that we have $L'_n \subseteq L_n = \mathbb{Z}_n^*$. Thus, it suffices to prove that $L_n \subseteq L'_n$. But this follows because $x^2 = 1 \Rightarrow x = \pm 1$ in a field.

Case 2. $n = p^e$ for a prime $p \geqslant 3$ and $e \geqslant 2$.

Recall that L_n is the kernel of the $(n-1)$-power map. Since $\mathbb{Z}_n^*$ is cyclic, it follows that $|L_n| = \gcd(\varphi(n), n-1)$. We can explicitly calculate it to get

$$\left|L'_n\right| \leqslant |L_n| = \gcd(p^{e-1}(p-1), p^e - 1) = p - 1 = \frac{p^e - 1}{p^{e-1} + \cdots + 1} \leqslant \frac{n-1}{4}.$$

Case 3. $n = p_1^{e_1} \cdots p_r^{e_r}$ is the standard prime factorisation of n, with $r > 1$.

Let $\theta : \mathbb{Z}_n \to \mathbb{Z}_{p_1^{e_1}} \times \cdots \times \mathbb{Z}_{p_r^{e_r}}$ be the ring isomorphism from earlier.
Write $n - 1 = t2^h$ and $\varphi(p_i^{e_i}) = t_i 2^{h_i}$ in the usual way, and let $g := \min\{h, h_1, \ldots, h_r\}$.
Note that $g \geqslant 1$, and that each $\mathbb{Z}_{p_i^{e_i}}^*$ is a cyclic group of order $t_i 2^{h_i}$.

We first show that $\alpha^{t2^g} = 1$. By definition of L'_n, we may assume $g < h$. Now, suppose $\alpha^{t2^g} \neq 1$, and let j be the smallest index in $g, \ldots, h-1$ such that $\alpha^{t2^{j+1}} = 1$. By definition of L'_n, we have $\alpha^{t2^j} = -1$. Let i be such that $g = h_i$. Writing $\theta(\alpha) = (\alpha_1, \ldots, \alpha_r)$, we have $\alpha_i^{t2^j} = -1$. Thus, the order of α_i^t (in $\mathbb{Z}_{p_i^{e_i}}^*$) is equal to 2^{j+1}. But this is a contradiction since it does not divide $\left|\mathbb{Z}_{p_i^{e_i}}^*\right| = t_i 2^{h_i}$. $\qquad (\because j \geqslant g = h_i)$

For $j = 0, \ldots, h$, define ρ_j to be the $(t2^j)$-power map on $\mathbb{Z}_n^*$. From the previous claim, and the definition of L_n', it follows that $\alpha^{t2^{g-1}} = \pm 1 \ \forall \alpha \in L_n'$. Thus, $L_n' \subseteq \rho_{g-1}^{-1}(\{\pm 1\})$ and hence,

$$|L_n|' \leqslant 2|\ker(\rho_{g-1})|. \tag{4}$$

Also,

$$|\ker(\rho_j)| = \prod_{i=1}^{r} \gcd(t_i 2^{h_i}, t2^j) \qquad \forall j \in \{0, \ldots, h\}.$$

Since $g \leqslant h$ and $g \leqslant h_i$ for all i, we get

$$2^r|\ker(\rho_{g-1})| = |\ker(\rho_g)| \leqslant |\ker(\rho_h)|. \tag{5}$$

Combining (4)-(5), we get

$$|L_n'| \leqslant 2^{-r+1}|\ker(\rho_h)| = \frac{|L_n|}{2^{r-1}}.$$

If $r \geqslant 3$, then we are done since $|L_n| \leqslant |Z_n^*| \leqslant n-1$, and $2^{r-1} \geqslant 4$.

If $r = 2$, then n is not a Carmichael number and thus,

$$\frac{|L_n|}{2^{r-1}} = \frac{|L_n|}{2} \leqslant \frac{1}{4}(n-1),$$

and we are again done. □