Acta Univ. Sapientiae, Mathematica, 1, 2 (2009) 161–168

Experimental results on probable primality

Omar Khadir
University of Hassan II-Mohammedia
BP. 146, Mohammedia,
Morocco
email: khadir@hotmail.com

László Szalay
Institute of Mathematics and Statistics
University of West Hungary
H-9400 Sopron, Erzsébet 9, Hungary
email: laszalay@ktk.nyme.hu

Abstract. In this paper we present experimental results on probable primality. More than four billion of randomly chosen integers having 256, 512 and 1024 bits were tested. We realized more experiments than Rivest did in 1991, and can confirm his observation: Miller–Rabin test does not ameliorate the small prime divisors test followed by Fermat test with the only base 2.

1 Introduction

Prime integers play a fundamental role in mathematics. They have always been a source of interest and fascination. Since the appearance of public key cryptography at the end of 1970's (see, for instance [1,9]), they have become more and more useful. RSA, Rabin cryptosystem, elliptic curve method, discrete logarithm problem and many digital signature protocols are completely based on large prime integers. By large, we mean that the considered numbers have at least 256 binary digits, or around 77 decimal digits.

It is well-known (see, for instance, [5]) that the running time of algorithms for constructing cryptosystem keys is dominated by the running time for generating prime integers. Finding rapid procedures for this latter task has, therefore, great importance.

AMS 2000 subject classifications: 11Y11, 11A41, 94A60
Key words and phrases: Primality, Fermat test, Miller–Rabin test, cryptography

It is also well known that there is no efficient and practical deterministic algorithm for quickly producing prime integers. That is why we only look for non deterministic algorithms which give us integers that are primes with a strong probability. There is a new look at the primality concept: an integer is taken as prime, not because it is really prime in an exact mathematical sense, but instead of that, it is prime because one thinks that nobody can factorize it. Recently, an integer is called industrial-grade prime (the term is due to H. Cohen) if its primality has not been proven, but it has undergone probable prime test(s).

The purpose of this work is to confirm what was concluded by Rivest in [10], as we made more experiments than Rivest. By analyzing experimental results on 4.13 billion randomly selected large integers, we show that a particular probabilistic algorithm for generating large prime integers based on three tests is likely equivalent to a similar algorithm, but based on only two tests. More precisely, our experimental results tend to indicate that using only two tests, division by small prime divisors followed by the Fermat test (see, for example [3,12]) produces the same results as using three tests: division by small primes, then the application of Fermat test, followed by Rabin-Miller test (see, for example [7,8]) with eight random bases. The Miller–Rabin test seems to be a waste of time when added as the third one to the first two aforesaid tests.

The paper is organized as follows. In section 2 we review the three tests composing the main algorithm and specify their formal parameters. In section 3 we briefly recall Rivest experimental results, and then we describe our own experiments, present and analyze the computing results. Section 4 contains conclusion and suggestion on possible forthcoming work.

In the sequel, we will adopt classical notation. In particular, $\mathbb{N}$ is the set of non-negative integers. Let $a, b, c \in \mathbb{N}$. Then $\gcd(a, b)$ denotes the great common divisor of a and b, while the remainder of a, when divided by b is denoted by a mod b. We write $a = b\ [c]$ if c divides the difference $a - b$. As usual, let $\pi(x)$ denote the number of primes less than or equal to the real number x. Finally, the bit length l_b of a positive integer $n = \sum_{i=0}^{k-1} 2^i a_i$ is $l_b = k$, where $a_{k-1} = 1$, and $a_i \in \{0, 1\}$ if $i = 0, \dots k-2$.

2 Three known tests

In this section we review three known tests and specify their formal parameters. Let $n > 1$ be an odd integer for which we want to test primality.

2.1 Small division test T_1

This test is the trial division by small divisors, namely by primes that are less than a fixed bound B. We divide n by all primes less than B. If we find one divisor, then n is composite, otherwise n is a candidate to be prime. Eratostenes sieve is applied to generate all primes between 2 and the bound B.

2.2 Fermat test T_2

Here we use a test based on the little Fermat theorem (see, for example [3,12]). If an integer a satisfies $\mathsf{gcd}(a, n) = 1$, we calculate $a^{n-1} \bmod n$ and compare it to 1. If $a^{n-1} \neq 1$ $[n]$, then n is composite, otherwise n is a candidate to be prime.

2.3 Miller–Rabin test T_3

Miller–Rabin test (see, for example [7,8]) is more efficient than Solovay and Strassen probabilistic test (see, for instance [11,6]). Since n is odd, we can uniquely find two positive integers r and s such that $n - 1 = 2^r s$. Let a be any integer such that $\mathsf{gcd}(a, n) = 1$. If $a^s \neq 1$ $[n]$ and $\forall\, j \in \{0, 1, \dots, r-1\}$: $a^{2^j s} \neq -1$ $[n]$, then n is composite.

If n passes all three tests, then it is probably a prime integer. In other words, we believe in its primality.

3 Results of our experiments

In this section, first we recall Rivest experimental tests [10], and then describe our own experiments providing the main results.

3.1 Rivest experiences

In 1991, Rivest examined 718 million randomly chosen 256-bit integers. Firstly he tested them by small divisors with the upper bound $B = 10^4$. $43,741,404$ passed this first test. Of those, $4,058,000$ passed Fermat test with the base 2. Of those, no one was eliminated by Miller–Rabin test with 8 random bases.

3.2 Our own experiments

Three kinds of experiments were realized with Maple software, versions 9.5 and 10, depending on the bit length: 256, 512 or 1024. We used ordinary personal computers working with Pentium IV 3.4 GHz processor and 248 MB of RAM. The main parameters were taken as in Rivest experiments:

- the upper limit of small primes is $B = 10^4$,
- the Fermat base is $b = 2$,
- the eight bases in the Miller–Rabin test are randomly chosen from the set $\{3, 4, \dots, n-2\}$.

In our case, we used blocks of integers and the number of randomly selected integers in each block was mainly between 5 and 10 million. Sometimes we used smaller or larger blocks as well.

We summarize the results in the next table where numbers N, N_1, N_2 and N_3 are defined as follows.

- N is the number of the randomly selected integers,
- N_1 denotes the number of integers which passed the first test T_1,
- N_2 is the number of integers which passed both T_1 and T_2,
- N_3 shows the number of integers which passed all the three tests.

Moreover, for $i = 1,\ 2$, and 3 let $R_i = 100\dfrac{N_i}{N}$.

We began to test more than one billion of integers whose bit lenght is 256, more than what was tested by Rivest. We found that the time required by PCs to run every range of 5 million of integers is around 40 minutes and around 80 minutes for every range of 10 million. For data see Table 1.

bit length	N	N_1	N_2	N_3
256	1.13×10^9	68 781 054	6 381 145	6 381 145
512	10^9	60 875 654	2 820 804	2 820 804
512	10^9	60 893 522	2 822 109	2 822 109
1024	10^9	60 876 414	1 408 923	1 408 923

bit length	N	$R_1(\%)$	$R_2(\%)$	$R_3(\%)$
256	1.13×10^9	6.0868189	0.5647031	0.5647031
512	10^9	6.0875654	0.2820804	0.2820804
512	10^9	6.0893522	0.2822109	0.2822109
1024	10^9	6.0876414	0.1408923	0.1408923

Table 1.

Then we tested two times one billion integers with bit lenght 512. And, finally, we tested one billion of integers with $l_b = 1024$. For comparison, see again Table 1.

We emphasized that, in the three kinds of experiment, we found $N_3 = N_2$ implying $R_3 = R_2$.

4 Conclusions

I. In this work, we realized new experiments on large integers in order to determine their primality. We tested more than four billion integers having 256, 512 and 1024 bits. They were all selected randomly. The main fact is that, from those which passed the small divisor test and the Fermat test, no one was blocked by the Miller–Rabin test. This result, based on more experiments, confirms what was already observed by Rivest. With the parameters mentioned above, the Miller–Rabin test does not improve the probabilistic algorithm based on the two first tests. Hence it seems that the Miller–Rabin test is unnecessary as the third stage of the three tests.

On the other hand, for future work, we suggest to replace the Miller–Rabin test by an alternative one and to verify experimentally if this modification brings any amelioration or not.

II. It seems that the upper bound on small primes is unnecessarily high. Both Rivest and us first used $B = 10^4$, but now we suggest $B = 300$ or $B = 3000$ instead. Why? Because with $B = 10^4$ we filtered 93.91% of the attendants independently from the bit length (supposing that l_b is large enough). If we have all the primes $p_1 = 2,\ p_2 = 3,\ ...,\ p_m \le B$, then in the first step of the 3 tests they exclude expectedly

$$1 - \left(1 - \frac{1}{p_1}\right)\left(1 - \frac{1}{p_2}\right)\cdots\left(1 - \frac{1}{p_m}\right) \tag{1}$$

part of the attendants. This formula gives 50% for $B = 2$, approximately 66.667% for $B = 3$, and so on, and provides $93,911\%$ for $B = 10^4$ (this was the preferred case). But for $B = 300$ we already have $90,245\%$, and going further, for $B = 3000$ we obtain $93,003\%$, which almost coincides with what we had for $B = 10^4$ before.

The following table shows the comparison of running experiences of different values B if $l_b = 256$. One can observe, that if we decrease B, then the number of random integers which failed the small prime divisor test also decreases, but the final ratio of the integers survived all the three tests is approximately constant. Furthermore, the values in the third column of Table 2 coincide with the values forecasted by (1).

B	Size of sample	Failed T_1	Passed $T_1 \wedge T_2 \wedge T_3$
10000	1.13 billion	93.913181	0.564703
3000	100 million	93.005612	0.564496
300	100 million	90.251749	0.564111

Table 2.

III. In the experiment we randomly chose a huge number of integers to classify them by three consecutive primality tests. Therefore, it is natural to compare the number of integers passing through all three tests (the number of industrial-grade primes) and the expected value of primes. Now we recall the thesis of Dusart [2], providing good approximations of the function $\pi(x)$.

Theorem 1 *(Dusart, [2], p.36.) If* $x \geq 1.332 \cdot 10^{10}$, *then*

$$\frac{x}{\ln x}\left(1+\frac{1}{\ln x}+\frac{1.8}{\ln^2 x}\right) \leq \pi(x) \leq \frac{x}{\ln x}\left(1+\frac{1.0992}{\ln x}\right).$$

Let π_n and $d_n = \frac{\pi_n}{2^{n-1}}$ denote the number of primes and the density of the primes in the interval $I_n = [2^{n-1}; 2^n - 1]$, respectively. By Theorem 1, we obtain

$$\begin{aligned} 0.0056\,424 \leq \quad & d_{256} = \frac{\pi_{256}}{2^{255}} && \leq 0.0056\,509\ , \\ 0.0028\,194 \leq \quad & d_{512} = \frac{\pi_{512}}{2^{511}} && \leq 0.0028\,217\ , \\ 0.001409\,299 \leq \quad & d_{1024} = \frac{\pi_{1024}}{2^{1023}} && \leq 0.001409\,875\ . \end{aligned}$$

Note, that in the experiment we investigated 1.13 and 2 and 1 billion random integers from the interval I_{256}, I_{512} and I_{1024}, respectively. Hence, with the given cardinality of the samples, the expected values E_{256}, E_{512} and E_{1024} of primes, by Dusart's theorem, satisfy the inequalities

$$\begin{aligned} 6375919 \ \leq \ & E_{256} \ \leq \ 6385530\ , \\ 5638897 \ \leq \ & E_{512} \ \leq \ 5643385\ , \\ 1409299 \ \leq \ & E_{1024} \ \leq \ 1409874\ . \end{aligned}$$

The following table recalls the statistics about the candidates for primality (the integers survived the three tests).

	256	512	1024
cardinality of sample	1.13 billion	2 billion	1 billion
number of candidates	6381145	5642913	1408923

Table 3.

When the bit length is 256, then we gained 6381145 industrial-grade primes and this number is in the interval [6375919; 6385530] bounding E_{256}. Similarly it is true when we choose random integers from I_{512}. In the case of longest bit length, 1408923 is not in the interval around E_{1024}, but less then its lower limit 1409299 (better case).

These data also reinforce the primality of integers passing through the three tests.

Acknowledgement

The research is supported by János Bolyai Scholarship of HAS, by Hungarian National Foundation for Scientific Research Grant No. T 61800 FT, and by Mexican-Hungarian bilateral TeT.

References

[1] W. Diffie and M. E. Hellman, New directions in cryptography, *IEEE Transactions on Information Theory*, **22** (1976), 644–654, (1976).

[2] P. Dusart, *Autour de la fonction qui compte le nombre de nombres premiers*, Thesis, University of Limoges, France, 1998.

[3] N. Koblitz, *A Course in Number Theory and Cryptography*, Graduate Texts in Mathematics, 2nd ed., Vol. 114, Springer-Verlag, 1994.

[4] A. K. Lenstra and E. R. Verheul, Selecting Cryptographic Key Sizes, *Journal of Cryptology*, **14** (2001), 255–293.

[5] C. Lu, A. L. M. dos Santos, F. R. Pimentel, Implementation of Fast RSA Key Generation on Smart Cards, *Proceedings of the 2002 ACM Symposium on Applied Computing*, (2002), 214–220.

[6] A. J. Menezes, P. C. van Oorschot, S. A. Vanstone, *Handbook of Applied Cryptography*, CRC Press, Boca Raton, Florida, 1997.

[7] G. L. Miller, Reimann's Hypothesis and a Test for Primality, *J. Comp. and System Sci.,* **13** (1976), 300–317.

[8] M. O. Rabin, Probabilistic Algorithm for Testing Primality, *J. Number Theory*, **12** (1980), 128–138.

[9] R. Rivest, A. Shamir, L. Adleman, A method for obtaining digital signatures and public key cryptosystems, *Comm. ACM*, **21** (1978), 120–126.

[10] R. L. Rivest, Finding four Million Large Random Primes, *Proceeding of the 10th Conference on Advanced Cryptology*, LNCS, **537** (1991), 625–626.

[11] R. Solovay and V. Strassen, A Fast Monte Carlo Test for Primality, *SIAM, Journal on Computation*, **6** (1978), 84–85.

[12] D. R. Stinson, *Cryptography, Theory and Practice*, Third Edition, Chapman & Hall/CRC, 2006.

Received: February 14, 2009