THE ROLE OF NUMBER THEORY IN MODERN CRYPTOGRAPHY

Robert R. Marceau*
M.S. Program in Computer Science, Rivier College

***Keywords**: bully children*

Abstract

The basic premise of modern cryptography is that the encryption and decryption processes should be computationally easy when the secret key is known, but very difficult without the key. Very difficult may well mean that with the current state of computation, it would take longer than the estimated age of the universe. One typical problem of this class is the factoring of a large composite integer that was generated by the multiplication of two large prime numbers. The multiplication of integers consisting of hundreds of digits is trivial for a computer. The factoring of a similar size number is a very difficult problem.

1 Historical Notes

Number theory may be traced back to the Greeks. Diophantus (c. 250) was interested in integer solutions to various equations. Several other great mathematicians have made contributions to number theory including Fermat, Gauss, and Euler. [1, page 120] It is only recently that great interest in number theory has developed due to its application to cryptography.

2 Preliminaries

2.1 Prime

The number $n \in \mathbb{Z}$, with $n > 1$ is called *prime* if and only if the only positive factors of n are 1 and n. If $n > 1$ is not prime, it is called composite. [2, page 210] [3, page 109]

2.2 Greatest Common Divisor

Let $a, b \in \mathbb{Z}$, not both 0. The largest $d \in \mathbb{Z}$ such that $d|a \wedge d|b$ is called the *greatest common divisor* of a and b. [2, page 215] [1, page 80

2.3 Fundamental Theorem of Arithmetic

Every positive integer greater than 1 can be written uniquely as a prime or the product of two or more primes where the prime factors are written in order of nondecreasing size. [2, page 211] [4, page 210] [3, page 110] [1, page 97]

2.4 Infinite number of primes

A proof of this theorem was provided by Euclid. [2, page 212] [4, page 11] This proof illustrates the classic “proof by contradiction” method.

ISSN 1559-9388 (online version), ISSN 1559-9396 (CD-ROM version).

> Assume the contrary, that there is only a finite number of primes, *n*. These prime numbers may be listed as $p_1, p_2, \ldots, p_n$. Consider the number $p_1 \cdot p_2 \cdots p_n + 1$. This number yields a remainder of 1 when it is divided by any prime. We have just produced a number that is not divisible by any prime on our list. This contradiction leads us to the conclusion that our initial assumption is false and that there are in fact an infinite number of primes.

An alternative direct proof is found in [1, pages 66–67]:

> Consider the value $Q_n = n! + 1$. This value is not divisible by any integer from 1 to n. By the Fundamental Theorem of Arithmetic, it must be either prime or divisible by a prime. Either way, it has a prime factor q_n, which must be greater than n. Since we have found a prime greater than *n*, for every positive integer n, there must be infinitely many primes.

2.5 Pseudo-prime to the base *b*

Let $b, n \in \mathbb{Z}^+$, n composite. If $b^{n-1} \equiv 1 \pmod{n}$ then n is called a *pseudo-prime to the base* b. [2, page 240] [5, pages 95–97] [4, page 36]

2.6 Carmichael Number

If a composite integer n is a pseudo-prime $\forall b \in \mathbb{Z}^+$ such that $gcd(b, n) = 1$ then n is called a *Carmichael number*. [2, page 240] [4, page 37] [1, page 207–208]

2.7 Inverse modulo *m*

If $a \in \mathbb{Z}, m \in \mathbb{Z}^+$ are relatively prime, then $\exists!\ \bar{a}$ such that $a\bar{a} \equiv 1 \pmod{m}$. (If two integers a, and $m > 0$ have no common divisors, then there is a unique integer $\bar{a}$ such that a times $\bar{a}$ has a remainder of 1 when divided by m). [2, page 234][4, page 24] [1, pages 140, 200]

2.8 Euler's Totient ϕ function

The Totient function $\phi(n)$ is defined as the number of positive integers less than or equal to *n* that are relatively prime to n. Two numbers p and q are said to be relatively prime to each other if and only if the Greatest Common Divisor of p and q is 1. [3, page 104] [4, page 30] [5, page 129–131]

2.9 Multiplicative function

If a function $f(n)$ defined for all positive integers n is called multiplicative if $f(nm) = f(n)f(m)$ whenever n and m are relatively prime. [1, page 222]

2.10 Mersenne Numbers

A *Mersenne Number*, M_n is defined as $2^n - 1$ where n is a non-negative integer. [5, pages 51–53]

3 Totient Function $\phi(n)$

If p is a prime, then $\phi(p) = p - 1$. All of the numbers from 1 through $p - 1$ are relatively prime to p. There are $p - 1$ of these numbers. It should be noted that the converse is also true: If there are $n - 1$ numbers relatively prime to n that are less than n, then n must be prime. Otherwise, n would have a factor greater than 1 and less than n. [3, page 104] [4, page 30] [5, page 129–131]

The Totient function is a Multiplicative function. If *n* and *m* are relatively prime, then $\phi(nm) = \phi(n)\phi(m)$. [1, page 224] It is these two properties of the Totient function that make the calculation of the Totient function computationally equivalent to factoring a number. If n and m are two large prime numbers, factoring nm is prohibitly expensive.

4 Mersenne Numbers

If M_n is a Mersenne Number, and n is composite, then M_n is also composite. [5, page 52] Consider $n = rs$, where r and s are positive integers. Then we have:

$$\begin{aligned} M_n &= 2^n - 1 \\ &= 2^{rs} - 1 \\ &= (2^r - 1)(2^{r(s-1)} + 2^{r(s-2)} + \cdots + 2^r + 1) \\ &= M_r\left(2^{r(s-1)} + 2^{r(s-2)} + \cdots + 2^r + 1\right) \end{aligned}$$

If r divides n, then M_r divides M_n. The converse, however, is not true. Consider the Mersenne number M_{11}. The number 11 is prime, however, $M_{11} = 2^{11} - 1 = 2047 = 23 \cdot 89$.

5 Prime Testing

5.1 Fermat's Little Theorem

If *p* is a prime and a is a positive integer where *p* does not divide a (since *p* is a prime, this is the same as the condition that *a* and *p* have no common factors greater than 1) then $a^{p-1} \equiv 1 \pmod p$. [3, pages 128, 135, 190] [1, pages 199–201]

Fermat's Little Theorem may be used to prove a given number is composite. Unfortunately, the converse is not true. A counter example is provided by the number 341. We have $2^{340} \equiv 1 \pmod{341}$, but $341 = 11 \cdot 31$ is clearly composite.

For any chosen base, there are Pseudo-primes to that base. It has been proved that there exists an infinite number of Pseudo-primes. [1, page 206]

5.2 Euler's Theorem

If *m* is a positive integer and *a* is an integer that is relatively prime to *m*, then $a^{\phi(m)} \equiv 1 \pmod m$. It should be noted that this is a generalization of Fermat's Little theorem, replacing the restriction that the modulus be prime with the condition that *a* and *m* have no common factors greater than 1. This is true since $\phi(p) = p - 1$ for all prime numbers *p*. [5, pages 25–26] [1, page 217]

5.3 Lucas-Lehmer Test

The Lucas-Lehmer Test may be used to determine if a Mersenne number is prime or composite. [5, pages 146-149] [3, page 193] [1, page 243] A sequence of positive integers is defined recursively as:

$$S_0 = 4 \quad and \quad S_{k+1} = S_k^2 - 2$$

Let *p* be a prime number. The Mersenne number M_p is prime if and only if $S_{p-2} \equiv 0 \, (mod \, M_p)$.

5.4 Miller Test

As seen by Fermat's Little theorem, $a^{p-1} \not\equiv 1 \, (mod \, p)$ is sufficient to demonstrate that the number *p* is not prime. The pseudo-prime numbers (and the Carmichael numbers in particular) demonstrate that it does not lead to a test for primeness. The Miller Test will eliminate some additional composite number from consideration, but, like Fermat's Little Theorem, it can only prove some numbers are composite. The test proceeds as follows (given an odd number $n > 0$ and a base b, where $1 < b < n - 1$): [5, pages 100–102] [1, pages 209–210]

1. Divide *n*−1 by 2 until an odd factor, *q* is found. We now have $n - 1 = 2^k q$.
2. Set *i* = 0 and $r = b^q \, mod \, n$.
3. If *i* = 0 and *r* = 1 terminate the test with an inconclusive result.
4. If *i* ≥ 0 and *r* = *n* − 1 terminate the test with an inconclusive result.
5. Increase *i* by 1 and set $r = r^2 \, mod \, n$
6. If *i* < *k* proceed to step 4, otherwise, *n* is composite.

6 Diffie–Hellman

As strange as it may seem, it is possible for two individuals to decide on a secret number by only exchanging public messages. Using the typical Bob and Alice cryptography character names, the procedure works as follows: [4, pages 50–51] [3, pages 270–271] [1, pages 299–300] [6, pages 188–190]

> Alice and Bob first agree on a large prime p and an integer g with $2 \leq g \leq p - 2$ such that the *order* of $g \, mod \, p$ is sufficiently high. The order of $g \, mod \, p$ is defined as the smallest integer x such that $g^x \equiv 1 \, (mod \,)p$. It should be noted that for all $a, 0 \leq a < x, g^a$ has a unique inverse, g^{a-x}. The values p and g are publicly known. Alice also chooses a random positive integer $a, a \leq p - 2$ and calculates $A = g^a \, mod \, p$. The value A is transmitted to Bob (over the insecure channel), while the exponent a is kept secret. Bob chooses a random positive integer $b, b \leq p - 2$ and calculates $B = g^b \, mod \, p$, which is then transmitted to Alice.
>
> Both Alice and Bob are able to now calculate the common key K. Alice, knowing a in addition to p and B calculates $b^a \, mod \, p$, which is equal to $g^{ab} \, mod \, p$. In a similar manner, Bob calculates $A^b \, mod \, p$, which is also equal to $g^{ab} \, mod \, p$. (The multiplication of the exponents a and b is commutative).

The selection of g deserves further discussion. If g is a *primitive* root mod *p*, then it has an order of $p - 1$. For a given prime, p, there are $\phi(p - 1)$ primitive roots mod p [6, page 66]. Finding one of these

primitive roots is computationally equivalent to factoring the integer $p-1$. Give the size of p, factoring $p-1$ is in general intractable. However, if p is chosen such that $(p-1)/2$ is also prime, say q, we can use an efficient method to test whether a randomly chosen g is a primitive root. (We have chosen p such that we know the prime factorization of $p-1=2\cdot q$).

The number of primitive roots mod p is $\phi(p-1)$, but since $p-1=2\cdot q$, we can calculate $\phi(p-1)=\phi(2\cdot q)=\phi(2)\phi(q)=q-1$. Roughly half of the values between 1 and $p-2$ are primitive roots. A randomly chosen g can be quickly tested by calculating $g^2 \bmod p$ and $g^p \bmod p$. If both of these values are not equal to 1, then g is a primitive root mod p.

7 RSA

Two large prime numbers p and q are chosen and kept secret. The product $n=pq$ is calculated and another small integer e which is relatively prime to $\phi(n)$ is chosen. $\phi(n)$ is easy to calculate if p and q are known, since $\phi(n)=\phi(p)\phi(q)=(p-1)(q-1)$. The pair of numbers (n,e) is the public key. Two functions are defined:

$$E(b)=b^e \bmod n$$
$$D(a)=a^d \bmod n$$

The function $E(b)$ is used to encrypt the message (which has been broken into a sequence of numbers b, with $0\le b\le n-1$). The number d is the inverse of $e \bmod \phi(n)$. We know such an inverse exists since e and ϕ are relatively prime. The extended Euclidean algorithm may be used to easily calculate d, provided $\phi(n)$ and e are known. The public key is the pair (n,e) which may be used by anyone to encrypt a message. The private key is the pair (n,d), which is required to decrypt a message. Combining results, we have:

$$D\big(E(b)\big)\equiv (b^e)^d \bmod n$$

However, d is the inverse of e modulo $\phi(n)$. It can further be shown that $b^{ed}\equiv b\,(mod\,p)$ and $b^{ed}\equiv b\,(mod\,q)$. At this point, we have pq divides $b^{ed}-b$. Since $n=pq$, we finally have $b^{ed}\equiv b\,(mod\,n)$. [5, pages 163–171]

Calculating b knowing only $b^e \bmod n$ and (n,e) is computationally equivalent to factoring n. It may also be viewed as the discrete logarithm problem. That is determine the value x such that $a\equiv g^x\,(mod\,p)$. [3, page 205] [1, pages 332–333] [6, pages 186–187, 213–214]

8 Exponentiation mod *n*

There is a very efficient method to calculate $a^m \bmod n$. It is reminiscent of Horner's rule to evaluate a polynomial for a particular value of x. [6, page 46] [4, pages 34–35] The procedure is as follows:

The exponent m is expressed in binary, where each $e_i\in\{0,1\}$ is chosen such that $m=\sum_{i=0}^{r} e_i 2^i$. (Each e_i is a binary digit for the binary expansion of m, e_0 being the least significant bit. Here $r+1$ is the number of bits required to represent m in binary). The following equations illustrate how g^m may be calculated:

$$m = \sum_{i=0}^{r} e_i 2^i$$

$$g^m = g^{\sum_{i=0} e_i r^i}$$

$$= \prod_{i=0}^{r} (g^{2^i})^{e_i}$$

$$g^{2^{i+1}} = (g^{2^i})^2$$

The following pseudo-code demonstrates how this may be done:

```
int power(int base, int exponent)
{
        int result = 1;
        while (e > 0) {
                if (exponent & 1)
                result = result * base;
                base = base * base;
                exponent = exponent / 2;
        }
        return result;
}
```

The process only requires r multiplications and divisions, plus another s multiplications, where s is the number of one bits in the binary expansion of the exponent. For cryptographic purposes, the result and base would be reduced modulus n after each iteration. The running time for the procedure above to calculate b^n is $\theta(\log n)$. This is a substantial improvement over the naïve approach which has a running time of $\theta(n)$.

9 Conclusion

As seen in the sections describing Diffie–Hellman key exchanges and the RSA encryption algorithm, the ability to find large (200 decimal digits) prime numbers is of critical importance. Methods to perform modular arithmetic, in particular raising a number to a very large power are also necessary. The security relies upon the fact that calculating the Totient function for a large number is equivalent to finding the prime factorization of the number. In the applications discussed above, two large primes are multiplied to give a number n for which calculating the Totient function is very difficult.

References

[1] Kenneth H. Rosen. *Elementary Number Theory and Its Applications (4th Edition)*. Addison Wesley, 2000.
[2] Kenneth Rosen. *Discrete Mathematics and Its Applications*. McGraw-Hill Science/Engineering/Math, 2006.
[3] James A. Anderson and James M. Bell. *Number Theory with Applications*. Prentice Hall, 1997.

[4] William Stein. *Elementary Number Theory: Primes, Congruences, and Secrets: A Computational Approach (Undergraduate Texts in Mathematics)*. Springer, 2008.
[5] S.C. Coutinho. *The Mathematics of Ciphers: Number Theory and RSA Cryptography*. A K Peters/CRC Press, 1999.
[6] Johannes Buchmann. *Introduction to Cryptography (Undergraduate Texts in Mathematics)*. Springer, 2004.
[7] John Stillwell. *Elements of Number Theory*. Springer, 2002.
[8] Underwood Dudley. *Elementary Number Theory: Second Edition*. Dover Publications, 2008.
[9] William J. LeVeque. *Fundamentals of Number Theory*. Dover Publications, 1996.
[10] G. Everest and Thomas Ward. *An Introduction to Number Theory (Graduate Texts in Mathematics)*. Springer, 2005.
[11] Gareth A. Jones and Josephine M. Jones. *Elementary Number Theory*. Springer, 1998.
[12] Ronald S. Irving. *Integers, Polynomials, and Rings: A Course in Algebra (Undergraduate Texts in Mathematics)*. Springer, 2003.

* **ROBERT MARCEAU** wrote his first computer program in October, 1969 on a DEC PDP-8/I running TSS/8. Since that time, he has earned a B.S. and M.S. in Mathematics from the University of Massachusetts - Lowell and is currently enrolled in the Computer Science Ph.D. program there. He is also expecting to complete his M.S. in Computer Science at Rivier University in spring 2013. He has spent over thirty years in the software industry and is currently an Adjunct Faculty member at Nashua Community College teaching a variety of Mathematics and Computer Science courses.