# A Review of the Applications of Number Theory in RSA Cryptosystem 

Argha Sengupta<br>Chandigarh University,<br>Punjab, India.<br>21MSM3210@cuchd.in

Nikita Madaan<br>Chandigarh University,<br>Punjab, India.

nikita.e12455@cumail.in

Shikha Tuteja<br>Chandigarh University<br>Punjab, India<br>shikhatuteja85@gmail.com


#### Abstract

This paper is a review of prime numbers and their applications in cryptography, discussed in special reference to RSA (Rivest-Shamir-Adleman) cryptosystem. Firstly, we discuss prime numbers and important related theorems, viz. Unique Factorization Theorem and Chinese Remainder Theorem. Then we discuss modular arithmetic and introduce Euler's totient function and discuss Euler's theorem, which forms the backbone of RSA cryptosystem. Next, Cryptography and public key cryptography are introduced and the implementation of RSA cryptosystem is discussed. RSA is widely used because of the difficulty of finding the prime factorization of large composite numbers. Implementation of RSA cryptosystem in real-world applications is also discussed along with the conclusions.

Keywords-Prime Numbers, Modular Arithmetic, Public-Key Cryptography, RSA Cryptosystem


## I. Introduction

The rise of human intelligence gave way to many unique phenomena that were unobserved before, one of them being the art of hiding information. The need to hide information has been around for as long as complex civilizations have existed. The mechanism of conveying information from one entity to the other is of the utmost importance in times of conflict, with each party manoeuvring to outclass and surprise the other. In the ancient world, the Romans and the Greeks were the inventors of simple cryptographic techniques which involved the transposition of letters of the alphabet according to some simple rules, such as shifting the letters forward by three (Caesar Cipher) or transposing the letters in accordance with a more complex set of rules (Scytale).

Number theory is a branch of mathematics that studies integers. It has developed over thousands of years, with ancient Babylonians laying the ground-work for Greek mathematicians (Pythagoras, Euclid, Diophantus, et al.) to bring about the dawn of mathematics. Concurrently, in ancient India, many prominent mathematicians (Aryabhatta, Brahmagupta, Bhaskaracharya, et al.) worked independently in the areas of trigonometry, number theory and quadratic equations. Renaissance Europe saw a revival and ultimately the maturity of number theory with contributions by Fermat, Euler, Lagrange, Legendre, Gauss, et al into elementary number theory, analytic theory, algebraic number theory and Diophantine geometry.

Professor Alan Turing and his team [1] was instrumental in the Allied code-breaking efforts against the German Enigma machine and the widespread use of mathematics for code-breaking efforts. The post-war era saw an exponential rise in the scope and applications of mathematics and cryptography. For the better part of the twentieth century, symmetric key cryptography (i.e., using the same key for encryption and decryption) was the only known cryptographic technique in implementation until the introduction of asymmetric key cryptography (i.e., using separate keys for encryption and decryption) in 1976, with the idea being that it would be computationally infeasible to compute the private key, knowing the public key. It is this concept that is discussed in this review paper in detail.
Shortly later, in 1978, Ron Rivest, Adi Shamir, and Leonard Adleman presented a practical implementation of public key cryptography, which is named RSA cryptosystem.

This paper is subdivided into two parts. The first one discusses the basics of number theory the all the prerequisite theorems that are necessary for a comprehensive understanding of the modern-day cryptography. The second part discusses public key cryptography and RSA (Rivest-Shamir-Adleman) algorithm in its mathematical form. Proofs and examples are presented to assist the reader to grasp the concepts and get a feel for the real-life implementation of the system.

## II. Number Theory

## II a. Prime Numbers

Prime numbers [2] are introduced quite early in the mathematics curriculum of schools. The basic notion of prime numbers lies in the realization that there exist some natural numbers ( $\mathbb{N}$ ) that cannot be obtained by multiplying smaller natural numbers. The most obvious examples are 2. 3. 5. 7. .... etc.. Thus, a prime number has only a single factor, i.e., 1 and itself.

Definition: Formally, a natural number $n \in \mathbb{N}$ is prime if $n>$ 1 and $\operatorname{gcd}(\mathrm{n}, \mathrm{k})=1, \forall k \in \mathbb{N}$ and $\mathrm{k}<\mathrm{n}$.

Prime numbers are the building blocks of natural numbers. A number which is not prime is called a composite number. Thus, $4,6,8,9,10, \ldots$, etc. are composite numbers. Each of these composite numbers can be written as a product of
prime numbers. For example, $4=2 \times 2,6=2 \times 3,8=2 \times 2 \times$ $2,9=3 \times 3$ and so on. Notice that each composite number is represented by a unique product of primes.

## II b. Unique Factorization Theorem

Any positive integer greater than 1 can be represented as a unique product of prime numbers. Formally, for some $n \in \mathbb{N}$ and $n>1, n=p_{x} \times q y \times r_{z} \ldots$ where $p . q . r \ldots$ are primes and the exponents $x . y . z \ldots$ are integers. This says that a number, say 1000 , for example, can be represented only as the product $2 \times 2 \times 2 \times 5 \times 5 \times 5$ and nothing else, i.e., no other prime number can appear in the prime factorization of 1000.

II c. Chinese Remainder Theorem

Given a system of equations

$$
\begin{align*}
& x=a(\bmod (m))  \tag{1}\\
& x=b(\bmod (n)) \tag{2}
\end{align*}
$$

where $\operatorname{gcd}(m, n)=1$, if $x 1$ and $x 2$ are solutions to (1) and (2) respectively, then

$$
x l \equiv x 2 \bmod (m, n)
$$

is the unique solution [3]. This is easily extended to a system of multiple equations.

## II d. Modular Arithmetic

Assume that there exists a set S , where $\mathrm{S}=$ $\{1,2,3,4,5,6,7,8,9,10,11,12\}$. Also, consider any two arbitrary numbers $\mathrm{a}, \mathrm{b} \in \mathrm{S}$. The objective is to carry out the usual operations of addition and multiplication $(a+b)$ and $(a \times b)$ such that the resultant, c , of the operation is also contained in the set, without any restrictions on a and $b$. This seems impossible at first,

Consider this: $2 \times 11=22 \notin S$, or $10+9=19$, which is again not in $S$. To be able to carry out the objective, addition and multiplication operations need to be modified.

Let $S$ be a set where $S=\{1,2,3,4, \ldots . n\}$. Modular addition is defined as follows:

$$
\begin{cases}a+b ; & a+b \leq n \\ a+b-12 ; & a+b>n\end{cases}
$$

Setting $n=12$, we see that $S$ gets reduced to $\{1,2, \ldots, 12\}$. For any two numbers, say a and $b$, the resultant of modular addition defined above lies always in $S$. For example, if $a=$ 5 and $b=6, a+b=5+6=11 \in \mathrm{~S}$, but if $a=5$ and $b=9$, then, according to the definition above, $a+b=5+9=14$ > 12 , thus, $a+b-12=5+9-12=2 \epsilon \mathrm{~S}$. Thus, the resultant lies always in S .

Let $a$ and $b$ be two numbers in $S$. Then, modular multiplication is defined as,

$$
a \cdot b=a b(\bmod (12)) \equiv \mathrm{r}
$$

where $r$ is the remainder when $(a b)$ is divided by 12 . For example, take
$a=2$ and $b=11$. Then, $a . b=2.11=2.11(\bmod (12))=$ $22(\bmod (12))=10$, which is in $S$. Similarly, it can be verified that for any choice of $a$ and $b$, that the resultant $a b(\bmod (12))$ always lies in $S$.

Observe that

$$
5.5(\bmod (12)) \equiv 25(\bmod (12)) \equiv 1
$$

Similarly,

$$
11.6(\bmod (13)) \equiv 66(\bmod (13)) \equiv 1,
$$

and

$$
8.2(\bmod (15)) \equiv 16(\bmod (15)) \equiv 1
$$

An integer $p$ is said to be the multiplicative inverse of another integer $q$ modulo $n$ if

$$
p \cdot q(\bmod (n)) \equiv 1
$$

## II e. Congruences

The idea of modular multiplication leads us to the concept of congruences. Geometrically, two shapes are said to be congruent to each other if they are identical, i.e. can be superimposed on one another, fitting perfectly. Similarly, two numbers $a$ and $b$ are said to be congruent modulo $m$ if the remainder of $a$ when divided by $m$ is equal to the remainder of $b$ when divided by $m$. If $a(\bmod (m)) \equiv b$, then a is congruent to $b(\bmod (m))$.

Properties of Properties of Congruences [4]:
$-a \equiv a(\bmod (n)) \forall a \in \mathbb{Z}+$, i.e., congruence is reflexive.

- If $a \equiv b(\bmod (n))$, then $b \equiv a(\bmod (n))$, i.e., congruence is symmetric.
- If $a \equiv b(\bmod (n))$, and if $a \equiv b(\bmod (n))$, then $a \equiv$ $\mathrm{c}(\bmod (n))$, i.e., congruence is transitive.
- If $a \equiv b(\bmod (n))$, and $\mathrm{c} \equiv \mathrm{d}(\bmod (n))$, then $(\mathrm{a}+\mathrm{b}) \equiv(\mathrm{c}+$ d) $(\bmod (n))$, i.e., congruence is additive.
- If $a \equiv b(\bmod (n))$, and $\mathrm{c} \equiv \mathrm{d}(\bmod (n))$, then $(\mathrm{a} . \mathrm{b}) \equiv$ $(c . d)(\bmod (n))$, i.e., congruence is multiplicative.


## II f. Fermat's Little Theorem

Let $m$ be a prime number and let $n$ be some other integer such that $m \nmid n$ ( $m$ does not divide $n$, or $m$ is not a factor of $n$ ), then

$$
n^{m-1} \equiv 1(\bmod (m))
$$

## II g. Coprime Numbers

The basic idea behind coprime numbers is similar to the one behind prime numbers (i.e. prime numbers have only two factors, viz. 1 and themselves).

Two positive integers a and b are relatively prime (or coprime) if their Greatest Common Divisor (alternatively called Highest Common Factor) is 1, i.e., $\operatorname{gcd}(\mathrm{a}, \mathrm{b})=1$. The integers a and b need not be prime to be relatively prime. For example, if $a=9$ and $b=25$, then $\operatorname{gcd}(9,25)=1$, even though 9 and 25 are composite numbers. It is seen that larger the integer, greater the number of co-primes that it has. For example,

- Consider the number 14 . It is relatively prime with $1,3,5$, 9,11 , and 13 .
- Similarly, consider the number 30. It is relatively prime with $1,7,11,13,17,19,23$, and 29.

Thus, given any positive integer $n$, there are at max $k$ integers relatively prime to $n$, where $k \leq n$. The integers k are called totatives of $n$ and $\operatorname{gcd}(n, k)=1$. This introduces the idea of distribution of prime numbers. Given a positive integer $n$, the objective is to find the number of totatives of $n$ (i.e. how many integers $k$ less than $n$ exist such that $\operatorname{gcd}(n, k)=1)$. This is given by Euler's Totient Function, $\phi$.

## II h. Euler's Totient Function

The totient $\phi(n)$ of a positive integer $n$ is defined as the number of positive integers less than $n$ such that they are relatively prime to $n$. Here, $\phi(n)$ is the number of positive integers less than $n$ and relatively prime to $n$. For example, $\phi(14)=6$ (six integers, $1,3,5,9,11$, and 13 are relatively prime to 14). Similarly, $\phi(30)=8$. An important consequence of this is that for any prime number $n$, $\phi(n)=n-1$. For example, when $n=13, \phi(13)=12$ (twelve integers, $1,2,3,4, . ., 12$ are relatively prime to 13 ).


Fig: Plot [5] of Euler's Totient Function for $n \in[0,100]$.

## II i. Euler's Totient Function

For a given $n \in \mathbb{N}$ and $\forall a \in \mathbb{N}$ such that $\operatorname{gcd}(a, n)=1$,

$$
a^{\phi(n)} \equiv 1(\bmod (n)),
$$

where $\phi(n)$ is Totient Function. Alternatively, it can be stated as follows,

$$
a^{\phi(n)}(\bmod (n))=1
$$

where the conditions for coprimality remain the same. Raising a to the power $\phi(\mathrm{n})$ and dividing by n , the remainder obtained is 1 . This forms the backbone of RSA cryptosystem.

## III. CRyptography

III a. Background

The science of securing communication between multiple parties is called cryptography [2]. It encompasses all the techniques involved in ensuring the security of a communication channel. Cryptographic techniques have been in use since the advent of civilizations. Ancient Greeks and Romans used rudimentary cryptographic techniques which involved linear shifting by fixed intervals of the letters so as to make the resultant text gibberish. This is an example of a substitution cipher. Substitution ciphers use a set of rules to substitute each character of the message with some other character using a set of rules. The famous German Enigma machine is a poly-substitution
cipher, a more complex type of a substitution cipher. Stream ciphers have also been used to encrypt information using a string of random characters whose length is exactly equal to the length of the message. One-Time-Pad (OTP) is an implementation of a stream cipher.
Two entities, Bob and Alice (could be people, bank and a person, two banks, etc.) want to communicate in the presence of an eavesdropper John (intrusive third party). The objective is to send a message between Bob and Alice in the presence of John, without John being able to understand what the message is. The original message, $m$, called plaintext, is to be encrypted to ciphertext, $c$, and sent from Bob to Alice (or vice-versa). The encrypted message $c$ is then to be decrypted by the receiver to get the original plaintext message, $m$. This can be done as follows.

- Consider a plaintext $m$ originating from Bob. Using an encryption key $k, \mathrm{~m}$ is transformed (encrypted) to $c$. Mathematically, this operation is
represented as $k(m)=c$, where $k$ is the encryption function.
- The ciphertext $c$, received is by Alice. It is then decrypted using a decryption key $d$ back to $m$. Mathematically, this operation is represented as $d(c)=m$.

Thus, the functions $k$ and $d$ are inverses of each other, i.e., $d$ $=\mathrm{k}^{\wedge}(-1)$.

## III b. Public-Key Cryptography

As discussed in the previous section, to encrypt a message $m$, an encryption key, $k$ and a decryption key, $d$ are required. If the same keys are used for encryption and decryption, i.e., $k=d$, then the system is called symmetric key cryptography. If different keys are used for encryption and decryption, i.e., $k \neq d$, then the system is called asymmetric key cryptography.

A subset of the latter is public key cryptography. A publickey cryptosystem uses a unique, publicly distributed encryption key $k$ to encrypt a private message $m \in M$, where $M$ is the set of all messages, giving the ciphertext $c$. Mathematically, this operation is represented as

$$
\mathrm{k}(\mathrm{~m})=\mathrm{c}
$$

The decryption key, $d$ is private, and unique for each receiver. When $d$ operates on $c$, the original message, $m$ is obtained, i.e.

$$
d(c)=m
$$

Diffie and Hellman suggested in 1976 [6] that cryptosystems resistant to known plaintext attacks (i.e. the eavesdropper knows a few pairs of plaintexts and
their corresponding ciphertexts for key recovery) could be used to construct one-way functions, called trapdoor functions. These functions would be easy to calculate only in the forward direction (i.e., during encryption) and extremely difficult to invert (i.e., during decryption). They suggested the use of modular exponentiation as the trapdoor. This trapdoor requires the inversion of modular exponentiation which relies on the difficulty of integer factorization. Unless the key, $k$, is known, it is extremely difficult to invert the encryption.

The trapdoor satisfies the following properties [7]:

- The encryption should be easy, i.e., $k(m)=c$ should be easy to compute given the key, k , and the message, m . Also, given the ciphertext c and a decryption key d, it should be easy to recover m.
- The decryption should be difficult given only c.


## III c. RSA Cryptosystem

RSA cryptosystem is a public key cryptosystem which was developed by Ron Rivest, Adi Shamir and Leonard Adleman in 1977 [8]. Assume that Bob wants to send a message to Alice. The following stages are involved in RSA:

1. Key Generation: Alice generates two equally sized (same number of digits) prime numbers, $p$ and $q$ such that the product, $p \times q=n$ is of the required bit size ( 1024 bits, 2048 bits, and so on). Next, Alice generates a number $e$ such that $1<e<$ $\phi(n) . p$ and $q$ are private while $n$ and $e$ are public.
2. Key Sharing: Alice shares her public key ( $n, e$ ) over an insecure channel and bob collects it.
3. Encryption: Bob has a message $m$ (a combination of texts, digits, images, etc.) which he wishes to share. He converts $m$ to binary and adds padding. Then, to encrypt $m$, Bob carries out the computation

$$
m^{e}(\bmod (n)) \equiv c,
$$

where $c$ is Bob's encrypted message, i.e. the ciphertext.
3.a Rabin suggested in 1979 [9] that $\mathrm{m}^{\wedge} 2(\bmod (n))$ $\equiv c$ be the encryption scheme to make decryption convenient for Alice. But Goldwasser and Micali [10] identified some problems that they remedied in their paper.
4. Ciphertext Sharing: Bob shares $c$ over an insecure channel where Alice collects it.
5. Decryption: Alice has the ciphertext $c$, her public key ( $n, \mathrm{e}$ ) and the private, prime integers ( $\mathrm{p}, \mathrm{q}$ ). Recall that $1<\mathrm{e}<\phi(n)$ and $\operatorname{gcd}(\mathrm{e}, \mathrm{n})=1$, i.e., e and n share no common factors. To decrypt [11], Alice must devise a scheme that would reverse the encryption. Thus, given c , e, and n , she needs to find an integer $d$ such that

$$
c^{d} \bmod (n) \equiv m
$$

i.e.

$$
\begin{gathered}
\left(m^{e}\right)^{d} \bmod (n) \equiv m \\
\Longrightarrow m^{e d} \bmod (n) \equiv m
\end{gathered}
$$

6. Decryption Scheme: Since $\operatorname{gcd}(m, n)=1$, using Euler's Theorem:

$$
\begin{align*}
& \Longrightarrow m^{\phi(n)} \equiv 1(\bmod (n))  \tag{1}\\
& \Longrightarrow m^{k} \phi(n) \equiv 1^{k}(\bmod (n))  \tag{2}\\
& \Longrightarrow m \cdot m^{k \phi(n)} \equiv \operatorname{m(\operatorname {mod}(n))}  \tag{3}\\
& \Longrightarrow m^{k \phi(n)+1} \equiv m(\bmod (n)) \tag{4}
\end{align*}
$$

Comparing $\mathrm{m}^{\wedge} \mathrm{k}(\phi(\mathrm{n})+1) \equiv \mathrm{m}(\bmod (\mathrm{n}))$ and $\mathrm{m}^{\wedge}(\mathrm{ed}) \equiv$ $m(\bmod (n))$, we have

$$
\begin{align*}
& \Longrightarrow e d=k \phi(n)+1  \tag{5}\\
& \Longrightarrow d=\frac{k \phi(n)+1}{e}  \tag{6}\\
& \left.\Longrightarrow c^{\frac{k \phi(n)+1}{e}} \equiv \operatorname{m(mod}(n)\right) \tag{7}
\end{align*}
$$

## IV. SECURITY of RSA

Consider that an eavesdropper, Eve, has access to the public key, i.e., $1 n$. eo and the ciphertext c. The problem of breaching RSA encryption, also called RSA Problem, is to find the message $m$ given only the public key $(n, e)$ and $c$. Eve does not know the private key $d$. Knowledge of $d$ would enable Eve to perform the operation

$$
\mathrm{c}^{\wedge} \mathrm{d}(\bmod (\mathrm{n})) \equiv \mathrm{m}
$$

Recall form (6) that

$$
d=\frac{k \phi(n)+1}{e}
$$

Since, $\mathrm{n}=\mathrm{p} . \mathrm{q}$,

$$
=\mathrm{M}^{\wedge} \mathrm{e}
$$

$$
\begin{equation*}
\Longrightarrow d=\frac{k \phi(p \cdot q)+1}{e} \tag{8}
\end{equation*}
$$

## IV a. Strong RSA Assumption

It is evident from (8) that breaching the encryption is dependent on finding the factorization of $n$ [12], since other variables, viz. ( $\mathrm{n}, \mathrm{e}$ ) are known and the calculation of k is trivial. Thus, the security of ESA encryption depends on the hardness of prime factorization of n . To construct RSA cryptosystems, it is assumed that Eve knows, or is allowed to choose the public exponent e according to her wish. This is called Strong RSA Assumption.

## IV b. Possible Avenues of Attack on RSA

RSA algorithm has a few shortcomings that can exploited as follows [14]:
IV.b.1. Lack of Forward Secrecy: A communication channel possesses forward security if the cryptographic keys used to encrypt information during a communication session cannot
be derived even if the channel is compromised in the future. RSA lacks forward secrecy. Let's see how.
Assume that a client (Bob) makes contact with an ecommerce website like Flipkart or Myntra (Alice). The client first downloads Flipkart server's certificates to authenticate and exchange the receive the public key of the server to encrypt future messages. Imagine that an eavesdopper, Eve, employs a packet sniffer in the client's LAN. The eavesdropper gains access to all the encrypted information owing through the client's network, including the public keys of the website (Flipkart). The only protection
against a breach is that the eavesdropper lacks the private key $d$. If, somehow d is leaked, then the encryption is broken. The private key can be (and has been in the past) leaked [15] by cheating employees.

This security threat is overcome by not sharing public keys over the channel, and instead using Diffie-Hellman algorithm to establish a session key without transmitting over the internet.
IV.b.1. Attacks on Some Chosen Ciphertexts: Messages that are shorter than the modulus can be compromised quite easily due to the inherent vulnerability of modular exponentiation, explained by the example below. Assume that $n=209(19 \times 11)$ and $e=3$. Let $M=5$. Then,

$$
\begin{aligned}
\mathrm{M}^{\wedge} \mathrm{e}(\bmod (209)) & =5^{\wedge} 3(\bmod (209)) \\
& =125(\bmod (209)) \\
& =125 \\
& =5^{\wedge} 3
\end{aligned}
$$

Here, $\mathrm{M}^{\wedge} \mathrm{e}<\mathrm{n}$, and thus, knowing the public exponent e , the attacker could easily know break the encryption by finding the cube root of the message. Therefore, for all messages $M$ such that $\mathrm{M}^{\wedge} \mathrm{e}<\mathrm{n}$, only the $\mathrm{e}^{\wedge}$ th root of the
ciphertext $c$ needs to be calculated, thereby partially breaking the encryption. There's no need to derive the decryption key $d$. To avoid this, it must be ensured that $M e i$ $n$. This is done by adding redundant bits according to defined standards. Physically, padding refers to excess area added to an object to increase cushioning. Similarly, padding in cryptography refers to the addition of redundant bits to a message to make the bit-length match the required standards. This is accomplished according to the standards defined in PKCS\#1v2.2 [16].
IV.b.1. Integer Factorization of Modulus N : Mathematical attack refers to finding integer factorization algorithms [17], [18], [19] for the modulus. Recall the expression for the private key,

$$
d=\frac{k \phi(n)+1}{e}
$$

Finding d essentially boils down to finding $\phi(n)$. Since the eavesdropper knows that $\phi(n)=(p-1) \times(\mathrm{q}-1)$, and also that $n=p \times q$, finding $\phi(n)$ would allow the eavesdopper to find $p$ and $q$. There are no existing algorithms that can solve the problem in polynomial time. [20] Trial division is the most trivial method which is useless to actually crack RSA encryption.
(a) Pollard's rho method [21], [22] is a clever technique that uses the same principle as the birthday paradox [23] to calculate all random pairs $\left(x_{i}, x_{j}\right)$ such that

$$
g c d|\mathrm{xj}-\mathrm{xi}|, \mathrm{N})>1
$$

Define $f(\mathrm{x})=x^{\wedge} 2+1$. Assume that $x_{0}=2$. Then, $x_{1}=f(\mathrm{x} 0) \bmod (N)$
and

$$
f(\mathrm{x} 2)=f\left(x_{1}\right) \bmod (N)=f(f(\mathrm{x} 0) \bmod (N)
$$

As soon as $\operatorname{gcd}(|x j-x i|, N)>1$ is satisfied, a factor is obtained. The longer the algorithm runs, the higher the probability of finding a solution.
(b) Wiener [24] showed that the modulus $N$ could be factorized if the decryption key, $d$, is small enough such that $d<(1 / 3) \mathrm{N}^{\wedge}(1 / 4)$. Blomer and May [25] generalized this method for every public key pair $(N, e)$ satisfying ex+y $=$ $0(\bmod (\mathrm{~N}))$ where $\mathrm{x}<(1 / 3) \mathrm{N}^{\wedge}(1 / 4)$ and $\bmod (\mathrm{y})=\mathrm{O}\left(\mathrm{N}^{\wedge}(-\right.$ 3ex/4).
(c) Overmars and Venkatraman [20] proved that all semiprimes $N=p \times q$, where $\mathrm{p}, \mathrm{q}$ are Pythagorean primes that can be represented as a sum of four square integers,
$\left(a c^{\wedge} 2\right)+\left(b c^{\wedge} 2\right)+\left(a d^{\wedge} 2\right)+\left(b d^{\wedge} 2\right)$. This sum could be expressed as the sum of two different squares $r^{\wedge} 2+s^{\wedge} 2$, where $r^{\wedge} 2=\left(a c^{\wedge} 2\right)+\left(b c^{\wedge} 2\right)$ and $s^{\wedge} 2=\left(a^{\wedge} 2\right)+\left(b d^{\wedge} 2\right)$ by applying Euler's factorization method.
(d) Number field sieve method is used to factor integers of the form $r^{\wedge} e \pm s$ where $r, s \in \mathbb{Z}+$ explained on-rigorously here [26].

## V. Conclusions and Challenges Ahead

It's been 45 years since Rivest, Shamir, and Adleman presented their seminal work on public-key cryptography. RSA is the most widely used cryptographic scheme even today, with 2048 -bit keys recommended by NIST until 2030. Even the moderate strength 1024 -bit key hasn't been factored yet using classical algorithms. Needless to say, RSA is still the preferred cryptosystem for commercial applications. The rise of quantum computing has exponentially increased computing power and raised concerns over the security of public-key cryptography [27], and RSA in particular due to a polynomial-time integer factoring algorithm by Shor (1995) [28]. The largest semi prime factored till date using quantum algorithms
is $\quad 1, \quad 099, \quad 551, \quad 473, \quad 989$ (10000000000000010011000000000000101000101) in
binary, which is 41 -bits long, still a far-cry from the moderate strength 1024-bit key. Part of the problem lies in the real-world realization and economic costs of implementing quantum computers capable of factoring large integers ( $\sim 1024$ bits). As of 2022, RSA-250 (250 decimal digits or 829 binary digits) has been factored successfully. Also, with NIST regularly updating key
sizes, it remains to be seen how RSA fares against postquantum algorithms.

## VI. References

[1] Alan Mathison Turing. The essential turing. Oxford University Press, 2004.
[2] Douglas R Stinson. Cryptography: theory and practice. Chapman and Hall/CRC, 2005.
[3] Kenneth Ireland, Michael Ira Rosen, and Michael Rosen. A classical introduction to modern number theory, volume
84. Springer Science \& Business Media, 1990.
[4] William P Wardlaw. The rsa public key cryptosystem. In Coding theory and cryptography, pages 101-123. Springer, 2000.
[5] Florentin Smarandache and Octavian Cira. Solving Diophantine Equations. 102014.
[6] Whitfield Diffie and Martin E Hellman. New directions in cryptography. In Secure communications and asymmetric cryptosystems, pages 143-180. Routledge, 2019.
[7] Andrew C Yao. Theory and application of trapdoor functions. In $23^{\text {rd }}$ Annual Symposium on Foundations of Computer Science (SFCS 1982), pages 80-91. IEEE, 1982.
[8] Ronald L Rivest, Adi Shamir, and Leonard Adleman. A method for obtaining digital signatures and public-key cryptosystems. Communications of the ACM, 21(2):120126, 1978.
[9] Michael O Rabin. Digitalized signatures and public-key functions as intractable as factorization. Technical report, Massachusetts Inst of Tech Cambridge Lab for Computer Science, 1979.
[10] Shai Goldwasser and Silvio Micali. Probabilistic encryption \& how to play mental poker keeping secret all partial information. In Providing sound foundations for cryptography: on the work of Shafi Goldwasser and Silvio Micali, pages 173-201. 2019.
[11] Megan Maxey. A modern day application of euler's theorem: The rsa cryptosystem. 2012.
[12] Divesh Aggarwal and Ueli Maurer. Breaking rsa generically is equivalent to factoring. In Annual International Conference on the Theory and Applications of Cryptographic Techniques, pages 36-53. Springer, 2009.
[13] Dan Boneh. Strong RSA Assumption, pages 597-597. Springer US, Boston, MA, 2005.
[14] Avi Kak. Lecture 12: Public-key cryptography and the rsa algorithm. Lecture Notes on "Computer and Network Security", Purdue University, pages 3-7, 2015.
[15] Zakir Durumeric, Frank Li, James Kasten, Johanna Amann, Jethro Beekman, Mathias Payer, Nicolas Weaver, David Adrian, Vern Paxson, Michael Bailey, et al. The matter of heartbleed. In Proceedings of the 2014
conference on internet measurement conference, pages 475488, 2014.
[16] Kathleen Moriarty, Burt Kaliski, Jakob Jonsson, and Andreas Rusch. PKCS \#1: RSA Cryptography Specifications Version 2.2. RFC 8017, November 2016.
[17] Joseph Zalaket and Joseph Hajj-Boutros. Prime factorization using square root approximation. Computers \& Mathematics with Applications, 61(9):2463-2467, 2011.
[18] Melissa e O'neill. The genuine sieve of eratosthenes. Journal of Functional Programming, 19(1):95-106, 2009.
[19] Burt Kaliski. Rsa factoring challenge., 2005.
[21] JM Pollard. Monte carlo methods for index computation $(\bmod \mathrm{p})$. MATHEMATICS OF COMPUTATION, 32(143):918-924, 1978.
[22] Nagaratna Hegde and P Deepthi. Pollard rho algorithm for integer factorization and discrete logarithm problem. International Journal of Computer Applications, 121(18):14\{17, 2015.
[23] Kazuo Nishimura and Masaaki Sibuya. Probability to meet in the middle. Journal of Cryptology, 2(1):13-22, 1990.
[24] Michael J Wiener. Cryptanalysis of short rsa secret exponents. IEEE Transactions on Information theory, 36(3):553-558, 1990.
[25] Johannes Blomer and Alexander May. A generalized wiener attack on rsa. In International Workshop on Public Key Cryptography, pages 1-13. Springer, 2004.
[26] Arjen K Lenstra, Hendrik W Lenstra Jr, Mark S Manasse, and John M Pollard. The number field sieve. In Proceedings of the twenty-second annual ACM symposium on Theory of computing, pages 564-572, 1990.
[27] Daniel J Bernstein, Nadia Heninger, Paul Lou, and Luke Valenta. Post-quantum rsa. In International Workshop on Post-Quantum Cryptography, pages 311-329. Springer, 2017.
[28] Peter W Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. SIAM review, 41(2):303-332, 1999.

