Topics in Discrete Mathematics

George Voutsadakis[1]

[1]Mathematics and Computer Science
Lake Superior State University

LSSU Math 216

Subsection 1

Introduction

Problem and Solution Idea

- Alice wants to tell Bob a secret. But Eve is eavesdropping. Can Alice tell Bob the secret? Can they hold a private conversation?
- Perhaps they can create a secret code and converse only in this code. Eve can overhear everything, including the details of their secret code!
- Alice and Bob could make up their code in private (where Eve can't hear). But, this option could be impractical, slow, and expensive.
- It seems impossible for Alice and Bob to hold a private conversation while Eve is listening to everything they say. It is therefore an amazing fact that private communication in a public forum is possible!
- The key is to develop a secret code with the following property:

  Revealing the encryption procedure does not undermine the secrecy of the decryption procedure.

- The idea is to find a procedure that is relatively easy to do, but extraordinarily difficult to undo. For example, it is not hard to multiply two enormous prime numbers, but factoring the resulting product is extremely hard.

Conjecture on the Inefficiency of Factoring

- Suppose p and q are large prime numbers, e.g., 500 digits each.
- It is not difficult to multiply these numbers. The result, $n = pq$, is a 1000-digit composite number. On a computer, this computation takes less than a second.
- Suppose we are given their product $n = pq$ and want to factor n to recover the prime factors p and q. If we use trial division, we need to do about 10^{500} divisions. This would take a very long period of time even on an ultra fast computer.
- If instead of using 500-digit primes p and q, we use 1000-digit primes, $n = pq$ increases from 1000 to 2000 digits, the time to multiply quadruples, but the time to factor gets 10^{1000} times bigger!

Conjecture (Inefficiency of Factoring)

There is no efficient procedure for factoring positive integers.

- The security of public-key cryptosystems relies on not having an efficient factoring algorithm!

Encoding the Message in ASCII

- Alice's message to Bob will be a large integer.
- A system is needed for converting a message into a number.
- Suppose her message is

  `Dear Bob, Do you want to go to the movies tonight? Alice`

  First, Alice converts this message into a positive integer. A standard way to convert the Roman alphabet into numbers is the ASCII code. In ASCII, Alice's message, rendered as numbers, is

D	e	a	r	spc	B	o	b	,	...
068	101	097	114	032	066	111	098	044	...

  Next, Alice combines these three-digit numbers into one large integer

$$M = 68, 101, 097, 114, 032, 066, 111, 098, \ldots, 099, 101.$$

  Since Alice's original message is about 50 characters long, this message is about 150 digits long.
- Now Alice is ready to send her message to Bob.

The Communication Protocol

- Bob creates a pair of functions, D and E, which are inverses of one another, i.e., $D(E(M)) = M$.
- Bob tells Alice the function E. At this point, Eve gets to see the function E.
- The function E is fairly easy to compute, but it is very hard to figure out D knowing only E.
- Alice uses Bob's public encryption function E. She computes $N = E(M)$ and sends the integer N to Bob. Eve gets to see this integer as well.
- Bob now uses his private decryption function D to compute $D(N)$. The result is $D(N) = D(E(M)) = M$, so Bob gets the message M.
- Since Eve does not know D, she cannot figure out what M is.
- The challenge is to create functions E and D that work for this protocol.

Summary of the Exchange



		Bob creates a public encryption function E and a secret decryption function D.
	Bob sends his public encryption function E to Alice.	
Alice writes her message M in ASCII. Uses Bob's function E to calculate $N = E(M)$.		
	Alice sends N to Bob.	
		Bob uses his function D to calculate $M = D(N)$.

Subsection 2

Rabin's Method

Square Roots in $\mathbb{Z}_n$

- The challenge is to create good encryption and decryption functions.
  - They should be relatively easy to compute.
  - Revealing E should not provide enough information to figure out D.
- In Rabin's Cryptosystem,
  - the encryption function is especially simple: For n be a large integer, the encryption function is

$$E(M) = M^2 \mod n.$$

  - Decryption involves taking a square root (in $\mathbb{Z}_n$).
- The integer n needs to be chosen in a special manner (to be described below).
- To understand how to decrypt messages and why Rabin's method is secure, we need to understand how to take square roots in $\mathbb{Z}_n$.

Square Roots in $\mathbb{Z}_n$: Examples

- Example: In $\mathbb{Z}_{59}$, when we ask for the square roots of 17, we seek those elements $x \in \mathbb{Z}_{59}$ for which $x^2 = x \otimes x = 17$.
  The calculator's value of $\sqrt{17} = 4.1231056\ldots$ is not of any help.
  There are only 59 different elements in $\mathbb{Z}_{59}$. We can simply square all of them and see which (if any) gives 17 as a result. This is painful to do by hand but fast on a computer. We find that 17 has two square roots in $\mathbb{Z}_{59}$: 28 and 31.
- What is $\sqrt{18}$ in $\mathbb{Z}_{59}$?
  We find that 18 does not have a square root in $\mathbb{Z}_{59}$.
- When we search for square roots of 17 in $\mathbb{Z}_{1121}$, we find four answers: $146, 500, 621$ and 975.

Quadratic Residues in $\mathbb{Z}_p$

Definition (Quadratic Residue)

Let n be a positive integer and let $a \in \mathbb{Z}_n$. If there is an element $b \in \mathbb{Z}_n$ such that $a = b \otimes b = b^2$, we call a a **quadratic residue modulo** n. Otherwise, i.e., if there is no such b, we call a a **quadratic nonresidue**.

Proposition (At Most Two Square Roots in $\mathbb{Z}_p$)

Let p be a prime and $a \in \mathbb{Z}_p$. Then a has at most two square roots in $\mathbb{Z}_p$.

- Suppose that a has three (or more) square roots in $\mathbb{Z}_p$. Notice that if x is a square root of a, then so is $-x \equiv p - x$ because $(p-x)^2 = p^2 - 2px + x^2 \equiv x^2 \equiv a \pmod{p}$. Since a has three (or more) square roots, we can choose two square roots, $x, y \in \mathbb{Z}_p$, such that $x \neq \pm y$. Now $(x-y)(x+y) = x^2 - y^2 \equiv a - a = 0 \pmod{p}$. The condition $x \neq \pm y$ implies that $x + y \not\equiv 0 \pmod{p}$ and $x - y \not\equiv 0 \pmod{p}$. This means that p is not a factor of either $x + y$ or $x - y$. Yet p is factor of $(x+y)(x-y)$, a contradiction!

Square Roots in $\mathbb{Z}_p$ (p prime, $p \equiv 3 \pmod 4$)

Proposition

Let p be a prime with $p \equiv 3 \pmod 4$. Let $a \in \mathbb{Z}_p$ be a quadratic residue. Then the square roots of a in $\mathbb{Z}_p$ are

$$[\pm a^{(p+1)/4}] \mod p.$$

- Let $b = a^{(p+1)/4} \mod p$. We need to prove that $b^2 = a$. By hypothesis, a is a quadratic residue in $\mathbb{Z}_p$, so there is an $x \in \mathbb{Z}_p$ such that $a = x \otimes x = x^2$. Calculate

$$\begin{aligned} b^2 &\equiv [a^{(p+1)/4}]^2 \equiv [(x^2)^{(p+1)/4}]^2 \\ &\equiv [x^{(p+1)/2}]^2 \equiv x^{p+1} \\ &\equiv x^p x^1 \equiv x^2 \equiv a \pmod p. \end{aligned}$$

Of course, if $b^2 \equiv a \pmod p$, then also $(-b)^2 \equiv a \pmod p$. By the preceding proposition, there can be no other square roots in $\mathbb{Z}_p$.

Examples

- Example: $p = 59$ is prime and $59 = 3 \pmod 4$. In $\mathbb{Z}_{59}$ we have $17^{(p+1)/4} = 17^{15} = 28$ and $28^2 = 31 \otimes 31 = 17$.
- Example: 17 has four square roots in $\mathbb{Z}_{1121}$. This is not a contradiction to the proposition, because $1121 = 19 \cdot 59$ is not prime. We now describe how to find the four square roots of 17.
  Suppose x is a square root of 17 in $\mathbb{Z}_{1121}$. Then $x \otimes x = 17$, whence $x^2 = 17 \pmod{1121}$, i.e., $x^2 = 17 + 1121k$, for some integer k. This can be rewritten as $x^2 = 17 + 19 \cdot (59k)$ and $x^2 = 17 + 59 \cdot (19k)$. Therefore, $x^2 = 17 \pmod{19}$ and $x^2 = 17 \pmod{59}$. Thus, to solve $x^2 = 17 \pmod{1121}$, we should first solve the two equations

$$x^2 = 17 \pmod{19} \quad \text{and} \quad x^2 = 17 \pmod{59}.$$

  - We have already found in $\mathbb{Z}_{59}$ the square roots of 17 are 28 and 31.
  - Since $19 \equiv 3 \pmod 4$, we can use the formula of the proposition $17^{(19+1)/4} = 17^5 \equiv 6 \pmod{19}$. The other square root is $-6 \equiv 13$.

The Subproblems for Finding $\sqrt{17} \mod 1121$

- Summarizing
  - We want to find $\sqrt{17}$ in $\mathbb{Z}_{1121}$;
  - We have $1121 = 19 \cdot 59$.
  - In $\mathbb{Z}_{19}$ the square roots of 17 are 6 and 13.
  - In $\mathbb{Z}_{59}$ the square roots of 17 are 28 and 31.
- Furthermore, if x is square root of 17 in $\mathbb{Z}_{1121}$, then (after we reduce x modulo 59) it is also a square root of 17 in $\mathbb{Z}_{59}$, and (after we reduce x modulo 19) it is also a square root of 17 in $\mathbb{Z}_{19}$.
- Thus x satisfies: $x = 6$ or $13 \pmod{19}$ and $x = 28$ or $31 \pmod{59}$.
- Now, we need to solve each of the four congruence systems:

$$\begin{array}{cccc} x \equiv 6 \pmod{19} & x \equiv 6 \pmod{19} & x \equiv 13 \pmod{19} & x \equiv 13 \pmod{19} \\ x \equiv 28 \pmod{59} & x \equiv 31 \pmod{59} & x \equiv 28 \pmod{59} & x \equiv 31 \pmod{59} \end{array}$$

We use the Chinese Remainder Theorem.

Solving a Subproblem Using Chinese Remainder Theorem

- We do only the calculations for $\begin{cases} x \equiv 13 \pmod{19} \\ x \equiv 28 \pmod{59} \end{cases}$.

  Since $x \equiv 13 \pmod{19}$, we write $x = 13 + 19k$, for some integer k.
  Substituting into the other equation $13 + 19k = 28 \pmod{59}$ or $19k \equiv 15 \pmod{59}$.
  Multiply both sides by $19^{-1} = 28$ in $\mathbb{Z}_{59}$: $28 \cdot 19k \equiv 28 \cdot 15 \pmod{59}$, whence $k \equiv 7 \pmod{59}$. Therefore, we get $k = 7 + 59j$, for some integer j.
  Substituting back in $x = 13 + 19k$, we get

$$x = 13 + 19(7 + 59j) = 146 + 1121j.$$

  So $x = 146$ is one of the four square roots of 17 in $\mathbb{Z}_{1121}$.
  The other three are 500, 621 and 975.

Efficiency of Factoring Subject to Availability of Roots

Theorem

Let $n = pq$ where p and q are primes. Suppose $x \in \mathbb{Z}_n$ has four distinct square roots a, b, c, d. If these four square roots are known, then there is an efficient computational procedure to factor n.

- Suppose $x \in \mathbb{Z}_n$, where $n = pq$ with p, q prime. Assume $x = a^2 = b^2 = c^2 = d^2$ in $\mathbb{Z}_n$, with a, b, c, d distinct. Since a is a square root of x, so is $-a$. We may assume that $b = -a$, but $c \neq \pm a$. Note $(a-c)(a+c) = a^2 - c^2 \equiv x - x = 0 \pmod{n}$, whence $(a-c)(a+c) = kpq = kn$, where k is some integer. Furthermore, since $c \neq \pm a$ (in $\mathbb{Z}_n$), $a - c \not\equiv 0$ and $a + c \not\equiv 0 \pmod{n}$. Therefore $\gcd(a-c, n) \neq n$. If $\gcd(a-c, n) = 1$, then neither p nor q is a divisor of $a - c$, and since $(a-c)(a+c) = kpq$, p and q must be factors of $a + c$, which contradicts that $a + c \not\equiv 0 \pmod{n}$. Since the only other divisors of n are p and q, we must have $\gcd(a-c, n) = p$ or $\gcd(a-c, n) = q$. Thus, we can factor n by finding gcd's.

Illustration of the Theorem

- Let $n = 38989$. The four square roots of 25 in $\mathbb{Z}_n$ are $a = 5$, $b = -5 = 38984$, $c = 2154$, and $d = -2154 = 36835$.
  Now we calculate

$$\gcd(a - c, n) = \gcd(-2149, 38989) = 307,$$

$$\gcd(a + c, n) = \gcd(2159, 38989) = 127$$

  and, indeed, $127 \times 307 = 38989$.
  Although there may be other procedures to find square roots in $\mathbb{Z}_{pq}$, an efficient procedure would be a contradiction to the non-existence conjecture. So, it is believed there is no computationally efficient procedure to find square roots in $\mathbb{Z}_{pq}$.

Rabin's Communication Protocol

- Alice wants to send a message to Bob.
- To prepare for this, Bob finds two large prime numbers p and q with $p \equiv q \equiv 3 \pmod 4$. He calculates $n = pq$.
- Bob then sends the integer n to Alice. Of course, Eve now knows n as well, but because factoring is difficult, neither Alice nor Eve knows the factors p and q.
- Next, Alice forms M by converting into ASCII and using the ASCII codes as the digits of M. Then, she calculates $N = M^2 \mod n$.
- Now Alice sends N to Bob. Eve receives the number N as well.
- To decrypt, Bob computes the four square roots of N (in $\mathbb{Z}_n$). Because Bob knows the factors of n (namely, p and q), he can compute the square roots. This gives four possible square roots, only one of which is the message M that Alice sent. The other three square roots give nonsense.
- Eve cannot decrypt (she does not know how to find square roots).

Subsection 3

RSA (**R**ivest, **S**hamir and **A**dleman)

Reminder of Euler's Theorem and Euler's Totient

- The RSA cryptosystem is named after its inventors, Rivest, Shamir and Adleman.
- It is based on Euler's extension to Fermat's Little Theorem:

Euler's Theorem

Let n be a positive integer and let a be an integer relatively prime to n. Then $a^{\varphi(n)} \equiv 1 \pmod{n}$, where φ is Euler's totient, i.e., $\varphi(n)$ is the number of integers from 1 to n that are relatively prime to n.

- For use with the RSA system, we are especially interested in $\varphi(n)$ with $n = pq$ where p and q are distinct prime numbers.
- In this case, recall that

$$\varphi(n) = \varphi(pq) = pq - p - q + 1 = (p-1)(q-1).$$

The Communication Protocol

- Bob finds two large prime numbers p and q and calculates their product $n = pq$. He also finds two integers e and d, satisfying some special properties explained below.
- The encryption and decryption functions are

$$E(M) = M^e \mod n \quad \text{and} \quad D(N) = N^d \mod n.$$

  These calculations can be done efficiently on a computer.
- Bob tells Alice his encryption function E. In so doing, he reveals the numbers n and e not only to Alice but also to Eve. He keeps the function D secret; that is, he does not reveal the number d.
- Alice forms her message M and calculates $N = E(M)$.
- Alice, then, sends the result to Bob. Eve gets to see N, but not M.
- Bob calculates $D(N) = D(E(M)) \stackrel{?}{=} M$. For Bob to be able to decrypt the message, it is important that we have $D(E(M)) = M$.

  Working in $\mathbb{Z}_n$, we want $D(E(M)) = D(M^e) = (M^e)^d = M^{ed} \stackrel{?}{=} M$.

Ensuring Decryption Inverts Encryption

- To ensure $D(E(M)) = M$, we use Euler's theorem, i.e., that if $M \in \mathbb{Z}_n^*$, then $M^{\varphi(n)} = 1$ in $\mathbb{Z}_n^*$.
- Raising both sides to a positive integer k gives $M^{k\varphi(n)} = 1$. If we multiply both sides by M, we get $M^{k\varphi(n)+1} = M$, so if $ed = k\varphi(n) + 1$, then we have $D(E(M)) = M^{ed} = M$. In other words, we want $ed \equiv 1 \pmod{\varphi(n)}$.
- Now we explain how to choose e and d:
  - Bob selects e to be a random value in $\mathbb{Z}_{\varphi(n)}^*$, i.e., e is an integer between 1 and $\varphi(n)$ that is relatively prime to $\varphi(n)$.
  - Bob knows the prime factors of n, so he can calculate $\varphi(n)$.
  - Next he computes $d = e^{-1}$ in $\mathbb{Z}_{\varphi(n)}^*$ (by finding x, y, with $\varphi(n)x + ey = 1$).
  - This ensures that in $\mathbb{Z}_n^*$, $D(E(M)) = M^{ed} = M^{k\varphi(n)+1} = (M^{\varphi(n)})^k \otimes M = 1^k \otimes M = M$.
  - With this choice of e and d, Bob can decrypt Alice's message.

Review of the Method

- Bob picks primes $p = 1231$, $q = 337$; computes $n = pq = 414847$.
- He can also compute $\varphi(n) = (p-1)(q-1) = 1230 \cdot 336 = 413280$.
- He chooses e at random in $\mathbb{Z}^*_{413280}$ - say, $e = 211243$.
- Finally, he calculates (in $\mathbb{Z}^*_{413280}$) $d = e^{-1} = 166147$.

The RSA Communication Procedure

- Bob finds two very large prime numbers p and q. He calculates $n = pq$ and $\varphi(n) = (p-1)(q-1)$.
- Bob chooses a random number $e \in \mathbb{Z}^*_{\varphi(n)}$ and calculates, using Euclid's Algorithm, $d = e^{-1}$, where the inverse is in the group $\mathbb{Z}^*_{\varphi(n)}$.
- Bob tells Alice the numbers n and e (but keeps the number d secret). Eve gets to see n and e.
- Alice forms her message M and calculates $N = E(M) = M^e \mod n$.
- Alice sends the number N to Bob. Eve gets to see this number as well.
- Bob calculates $D(N) = N^d = (M^e)^d = M$ and reads Alice's message.

Eve's Troubles

- Eve knows Bob's public encryption function $E(M) = M^e \mod n$, but she does not know the two prime factors of n.
- She also knows $E(M)$ (the encrypted form of Alice's message), but she does not know M.
- If Eve can guess the message M, then she can check her guess because she too can compute $E(M)$.
- Otherwise, Eve can try to break Bob's code.
  - One way she can do this is to factor n. Once she has n, she can compute $\varphi(n)$ and then get $d = e^{-1}$ (in $\mathbb{Z}^*_{\varphi(n)}$). Our supposition is that factoring is too hard for this to be feasible.
  - But Eve does not really need to know the prime factors of n. She would be happy just knowing $\varphi(n)$, so she can calculate d.

Proposition

Let p and q be primes and let $n = pq$. Suppose we are given n, but we do not know p or q. If we are also given $\varphi(n)$, then we can efficiently calculate the prime factors of n.

Proof and Example

- We know that $n = pq$, and $\varphi(n) = (p-1)(q-1)$. This is a system of two equations in two unknowns (p and q) that we can simply solve. We write $q = \frac{n}{p}$ and substitute this into the second equation, which we solve via the quadratic formula.
- Example: If $n = 414847$, then $\varphi(n) = 413280$. We want to solve $\left\{ \begin{array}{rcl} pq & = & 414847 \\ (p-1)(q-1) & = & 413280 \end{array} \right\}$. Substitute $q = \frac{414847}{p}$ into $(p-1)(q-1) = 413280$ to get

$$\begin{array}{rcl} (p-1)(\frac{414847}{p} - 1) = 413280 & & \\ \Longrightarrow & 414848 - \frac{414847}{p} - p = 413280 & \\ \Longrightarrow & p^2 - 1568p + 414847 = 0. & \end{array}$$

Using the quadratic formula, we find roots $p = 337$ and 1231. The prime factors of 414847 are, indeed, 337 and 1231.

Reasons for Eve's Troubles

- Eve does not need $\varphi(n)$; she just needs to know d. This is unlikely:

Proposition

Let p, q be large primes and $n = pq$. Suppose there is an efficient procedure that, given e with $\gcd(e, \varphi(n)) = 1$, produces d with $ed \equiv 1 \pmod{\varphi(n)}$. Then there is an efficient procedure to factor n.

- Thus, if factoring is intractable, then there is no efficient way for Eve to recover the exponent d just from knowing e and n.
- To break Bob's code, Eve needs to solve the equation $M^e \equiv N \pmod{n}$, where she knows e, N and n.
- We considered the possibility that Eve would recover the decryption function (especially the integer d) and compute M from N the same way Bob might. However, there may be other ways to solve this equation that we have not considered.

  Open Problem: Prove that breaking RSA is as hard as factoring.