Number Theory: Cryptography Supplement

MAT 421

Spring 2019

The goal of this assignment is to perform some cryptography that may be *basic*, but not *trivial*.

Background

Encoding and encryption are not the same.

- *Encryption* is the process of modifying a text so as to hide its meaning.
  - Decrypting an encrypted message should be fast and easy *only* for the intended recipient. For anyone else, decryption should be impossible, at least in principle. At least one part of the key should be secret.
- *Encoding* is the process of turning a message into numbers.
  - Decoding an encoded message should be fast and easy for anyone who receives the message. There are no secrets.
  - The simplest way to encode a message is by turning A into 0, B into 1, C into 2, … Z into 25. One then ignores spaces and punctuation, which is usually pretty easy to do anyway:

    ILOVEMATHEMATICSMORETHANLIFEITSELF

  - It's quite possible to encode punctuation, as well: just add more numbers. For instance, one might add the following values:

character	(space)	,	.	;	!	?	:	'	—
value	26	27	28	29	30	31	32	33	34

    We can now encode

    I LOVE MATHEMATICS, MORE THAN LIFE ITSELF!

  - Old computers and the internet encoded text using ASCII.[1] These days, Unicode is increasingly common.

Why does this matter? Modern cryptography relies on number theory, which means we need numbers, which means we have to encode text as numbers before we can encrypt it.

[1] American Standard Code for Information Interchange

Multi-letter encryption: our approach

To encode a message, we combine 2 or more letters as one number. The technique is similar to base-n arithmetic.

1. First convert *each* letter to *one* number using the scheme described above:

   I LOVE MATHEMATICS, MORE THAN LIFE ITSELF!

   turns into

   8, 26, 11, 14, 21, 4, 26, 12, 0, 19, 7, 4, 12, 0, 19, 8, 2, 18, 27, 12,
   14, 17, 4, 26, 19, 7, 0, 13, 26, 11, 8, 5, 4, 26, 8, 19, 18, 4, 11, 5, 30

2. Now we combine two numbers into one. Our encoding has 35 characters total, so we'll encode as "base-35" numbers. For each pair of numbers, we multiply the second by 35, then add it to the first, like so:

$$\begin{aligned}(8,26) &\rightarrow 8+35\times 26=918\\(11,14) &\rightarrow 11+35\times 14=501\\(21,4) &\rightarrow 21+35\times 4=161\\&\vdots\\(5,30) &\rightarrow 5+35\times 30=446\ .\end{aligned}$$

   We now have the sequence

   918, 501, 161, 446, 665, 147, 12, 299, 632, 937,
   502, 157, 691, 7, 923, 291, 145, 306, 649, 389, 1055.

   Note: If the message lacks enough letters to form pairs, we add a random letter at the end.

3. Now we encrypt. For this example, we'll use RSA: first, choose p and q such that $N = pq$ is larger than the largest integer we encrypt. Suppose we choose $p = 29$ and $q = 37$; we have $N = 1073$. Choose $e = 13$. The sequence of integers above encrypts as

   670, 76, 426, 311, 73, 739, 534, 622, 807, 941,
   361, 12, 835, 1069, 429, 320, 377, 861, 224, 708, 79.[2]

[2] If someone tries to decode these encrypted numbers using the decoding method described below, the result won't make sense:

$$670 \rightarrow (5,19) \quad 76 \rightarrow (6,2) \quad 426 \rightarrow (6,12)\ldots.$$

This corresponds to F,T,G,C,G,M,... which is illegible. Notice that the encryption is not one-to-one: the message's third letter, L, and its fifth letter, V, have both encrypted as G. This is another advantage of combining letters.

4. To decrypt the message, we first need to determine the decryption exponent. Since $\phi(N) = 1008$, the Extended Euclidean Algorithm gives us $d = 853$. The decryption takes place as

$$670^{853} \equiv 918 \pmod{1073} \quad 76^{853} \equiv 501 \pmod{1073} \quad \ldots .$$

Altogether, the sequence above decrypts as

$$918, 501, 161, 446, 665, 147, 12, 299, 632, 937, 502, 157, 691, 7, 923, 291, 145, 306, 649, 389, 1055.$$

Notice that we have received the sequence of numbers at the end of step 2.

5. To decode each number, reverse the encoding: divide the number by 35, and the remainder gives us the first letter, while the quotient gives us the second. In our case,

$$918 \rightarrow (8, 26) \quad 501 \rightarrow (11, 14) \quad \ldots .$$

Notice that we have received the sequence of numbers at the end of step 1, which corresponds to the original message!

Assignment

1. Verify every step of the encryption above *by hand.* You may use a calculator or a computer algebra system, but only for making the exponentiation and multiplication easier. If you're particularly inclined *and able to do so in reasonable time,* you may write a computer program to do it for you; just include the source code with some *reasonable* documentation.[3]

2. Consider the message, "MISSISSIPPI".

   (a) Encode it using our 2-letter encoding.

   (b) Encrypt the result of part (a) using RSA with the parameters $N = 1073$, $e = 13$.

   (c) Decrypt the result of part (b).
   Hint: The decryption parameters are given in the discussion above.

   (d) Decode the result of part (c). If you don't get "MISSISSIPPI" then you have made a mistake.

3. Repeat #2, but this time:

   (a) Verify that 2 is a primitive number modulo 1087.
   Hint: Don't compute all 1087 powers of 2. Here's a trick based on facts that come from Modern Algebra: $2^a \equiv 1 \pmod{1087}$ only if a is a divisor of 1086 (*not a typo).* Find the prime factorization of 1086, which isn't too hard, then list all its possible divisors, which

[3] I "speak" all of Ada, C, C++, Eiffel, Java, Kotlin, Modula-2, Modula-3, Oberon, Pascal, Python, Sage, and SPARK almost natively.

"Reasonable time": I am able to write such a program in less than an hour in Sage. Of course, I've been writing programs regularly for about 25 years. Only you can judge how long it might take.

"Reasonable documentation": The reader should understand the general idea. Not every line needs a comment.

isn't too hard (there are 8). Cross out 1086 and $543 = {}^{1086}/_{2}$. That leaves you with 6 divisors, all smaller than 543. If none of them gives you $2^a \equiv 1086 \equiv -1 \pmod{1087}$, then you know that 2 is primitive.
Bonus: What fact from Modern Algebra is the trick based on?
Hint for the bonus: $\{1, 2, \ldots, 1087\}$ is a group under multiplication, with 1 as its identity.

(b) Encrypt the result of 2(a) using Elgamal encryption, with the parameters $p = 1087$, $\alpha = 2$, $\alpha^a = 829$, and $b = 17$.
(Notice that I haven't told you a, because you don't need it.)

4. The following message has been encrypted using Elgamal with 2-letter encoding, $p = 1087$, $\alpha = 2$, $\alpha^{ab} = 348$, and $b = 17$. Again, you don't know a, and you don't need it! Decrypt the message.

$$946, 243, 187, 633.$$