8 Cryptography

8.1 Character Ciphers

---

1. **Introduction:** The goal of this entire chapter (and the rest of the course) is to talk about encryption and cryptography.

2. **Terminology:** We have the following:

   (a) *Cryptology*: The study of encryption/decryption.

   (b) *Cryptography*: The study of methods of encryption/decryption.

   (c) *Cipher*: A particular method of encryption.

   (d) *Cryptanalysis*: Breaking of systems of encryption.

   (e) *Plaintext*: The human-readable text we wish to encryp.

   (f) *Encryption*: The process of applying a cipher to plaintext.

   (g) *Ciphertext*: The human-non-readable result.

   (h) *Decryption*: The process of getting the plaintext back.

   (i) *Some Names*:

      i. Alice: encrypts and sends

      ii. Bob: receives and decrypts

      iii. Eve: eavesdropper

3. **Basic Methods:**

   (a) **Character Assignment:** To begin, we will assign a number to each letter of the alphabet:

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25

   Note: For now we will exclude lower-case, punctuation and spaces, but we could include those and use a different modulus.
   Note: This can be confusing since `A` is the first leter of the alphabet and so we would naturally want to assign it to 1. We use this for purposes of making our modular arithmetic easier.

   (b) **Shift Cipher:** For each plaintext letter P we assign ciphertext

$$C \equiv P + b \bmod 26$$

Ex. Encrypt `LEIBNIZ` with $b = 3$.

$$\begin{aligned}
L: &\quad P = 11, 11+3 \equiv 14 = C : O \\
E: &\quad P = 4, 4+3 \equiv 7 = C : H \\
I: &\quad P = 8, 8+3 \equiv 11 = C : L \\
B: &\quad P = 1, 1+3 \equiv 4 = C : E \\
N: &\quad P = 13, 13+3 \equiv 16 = C : Q \\
I: &\quad P = 8, 8+3 \equiv 11 = C : L \\
Z: &\quad P = 25, 25+3 \equiv 2 = C : C
\end{aligned}$$

Which then results in `OHLEQLC`. To decrypt we simply reverse: $C \equiv P + b \bmod 26$, $P \equiv C - b \bmod 26$.

(c) **Affine Cipher:** Choose a and b and encrypt via $C = aP + b \bmod 26$. How will decryption work? $C \equiv aP + b \bmod 26$, $aP \equiv C - b \bmod 26$ there needs to be a unique P. To have this we need $\gcd(a, 26) = 1$ so that a has a multiplicative inverse. Then $P \equiv a^{-1}(C - b) \bmod 26$. How many choices? $\phi(26) = 12$ for a and 26 choices for b.

Ex. If we choose $a = 5$ and $b = 7$ then encryption is $C \equiv 5P + 7 \bmod 26$ and decryption is $5P \equiv C - 7 \bmod 26 \implies P \equiv 21(C - 7) \bmod 26$ (calculated from 21 being the multiplicative inverse of 5).

4. **Breaking Shift Ciphers:** To break a shift cipher, we only need b. For example, if we manage to find a specific C_0 for a specifice P_0, then we know that $C_0 \equiv P_0 + b \bmod 26$ so $b \equiv C_0 - P_0 \bmod 26$. How might we do this? With frequency analysis.

   Frequency Analysis: In english, the most frequent letter is `E`, note this is $P_0 = 4$. Find the most frequent ciphertext letter. If that is C_0 we guess at that.

5. **Breaking Aphine Ciphers:** One C_0 and P_0 pair is not sufficient! Since knowing $C_0 \equiv aP_0 + b \bmod 26$ is not enough to find a and b. However, having another pair is good enough because:

$$\begin{aligned}
&C_0 \equiv aP_0 + b \bmod 26 \\
&C_1 \equiv aP_1 + b \bmod 26 \\
&\hline
&C_0 - C_1 \equiv a(P_0 - P_1) \bmod 26
\end{aligned}$$

   This will have solutions if and only if $\gcd(P_0 - P_1, 26) \mid C_0 - C_1$, and if so there will be $\gcd(P_0 - P_1, 26)$ solutions.

   Note: Keep in mind this is valid cipher text. There is an a (which Alice chose). So there will be solutions. There may be more than 1. If multiple

possible a, for each, find b, simply try all of those a, b combinations until we get proper plaintext.

8.2 Exponentiation Ciphers

---

1. **Introduction:** Can we find a process which is harder to invert?
   First we will modify the table of letters slightly:

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25

   Now, we can put letters together unambigiously. For example `JU` can be assigned to 0920 or just 920. Without the leading 0 it is unclear what something like 111 means. It could be $111 \implies 0111$ or $111 \implies 1101$.

   Fermat's Little Theorem: Recall, if p is prime and $a \in \mathbb{Z}$ with $p \nmid a$ then $a^{p-1} \equiv 1 \bmod p$.

2. **Exponentiation Cipher**

   (a) **Encryption:** Let p be an odd prime (typically very large) and let e be a positive integer with $\gcd(e, p-1) = 1$ (use Euclidean Algorithm for this). We then take the plaintext and group the letters into blocks so no block is larger than p.
   For example,

   - If $p = 29$ then blocksize is 1 since $\texttt{z} \leftrightarrow 25 < p$.
   - If $p = 3001$ then blocksize is 2 since $\texttt{zz} \leftrightarrow 2525 < p$.
   - If $p = 377173$ then blocksize is 3 since $\texttt{zzz} \leftrightarrow 252525 < p$.

   We then pad the plaintext with junk letters at the end if needed so that the plaintext length is a multiple of the blocksize. Traditionally `X` is used but any letter can be used. To encrypt, Alice needs to divide full plaintext into blocks. For each block P we do

$$C \equiv P^e \bmod p$$

   Ex. Alice wants to encrypt `LOVENOTE` with $(e, p) = (479, 3001)$ and $\gcd(479, 3000) = 1$.

	`LO`	`VE`	`NO`	`TE`
	1114	2104	1314	1904
	1114^{479}	2104^{479}	1314^{479}	1904^{479}
$\equiv$	0169	0317	0017	1697

   So we get `0169 0317 0017 1697` as the ciphertext that Alice would send to Bob.

(b) **Decryption:** This process is invertible since the fact that $\gcd(e, p-1)$ guarantees that there exists some d with $de \equiv 1 \bmod p$. Then for a ciphertext block raised to d:

$$C^d \equiv (P^e)^d \equiv P^{ed} \equiv P^{1+k(p-1)} \equiv P(P^{p-1})^k \equiv P(1)^k \equiv P \bmod p$$

Here the fact that $P^{p-1} \equiv 1 \bmod p$ is guaranteed by FLiT. Note that $p \nmid P$ since $P < p$.

Thus, to decrypt ciphertext, Bob simply takes C and raises it to d, $C^d \bmod p$.

Ex. Alice sends cipher text to `2672 0317 1665 2110 0246 1749 0017 2112` Bob. She encrypted using Bob's choice of $(e, p) = (479, 3001)$. To decrypt Bob would have to use $e^{-1} = d = 119$.

	2672	0317	1665	2110	0246	1749	0017	2112
	2672^{119}	0317^{119}	1665^{119}	2110^{119}	0246^{119}	1749^{119}	0017^{119}	2112^{119}
$\equiv$	1800	2104	2414	2017	1804	1105	1314	2223
	`SA`	`VE`	`YO`	`UR`	`SE`	`LF`	`NO`	`WX`

Then, Bob can obviously see the message `SAVE YOURSELF NOW` which is padded with an `X` to make the length a multiple of two characters.

Note: Bob chose $p = 3001$ and $e = 479$ and provided them to Alice so she can send him messages. Two things

- If the two of them keep the p and e secret this is fairly secure.
- Since Alice knows e, she can calculate d. So Alice can decrypt anything else sent to Bob if that specific p and e are used. So Bob would have to use a different p, e with each person.

This is symmetric: knowing $p, e \equiv$ knowing p, d. Is there a way to provide an encryption method so that even if you know how to encrypt you cannot figure out how to decrypt?

8.4 Public Key Encryption and RSA

---

1. **Introduction:** The primary problem with a technique like an exponentiation cipher is that given (p, e) it is easy to find (p, d).
2. **RSA:**
   (a) **Encryption:** Bob picks two distinct large primes p and q and calculated $n = pq$. This will be his modulus. He then chooses e with $\gcd(\phi(n), e) = 1$. Note that $\phi(n) = \phi(pq) = (p-1)(q-1)$ so he can

choose an e pretty easily via the Euclidean Algorithm. Bob makes the pair (n, e) publicly avaliable.

Alice takes her message and breaks it up just like the exponentiation cipher, breaking it into blocks with numerical value not possible more than n. For each plaintext block P, she calculates the ciphertext block as the least nonnegative residue:

$$C = P^e \bmod n$$

She then sends all the ciphertext blocks to Bob.

Ex. Suppose Bob chooses $p = 59$ and $q = 73$. Then $n = (59)(73) = 4307$ and $\phi(n) = (58)(72) = 4176$. He then chooses $e = 7$ with $\gcd(e, \phi(n)) = 1$. Alice wishes to encrypt and send `WORD`. She divides it into blocks of length 2 and does:

	`WO`	`RD`	
	2214	1703	
	2214^7	1703^7	
$\equiv$	3918	1655	mod 4307

(b) **Decryption:** Since Bob knows $\phi(n) = (p-1)(q-1)$ he can easily find d with $ed \equiv 1 \bmod \phi(n)$. Then for each ciphertext block$C$ he can decrypt by calculating the least nonnegative residue $C^d \bmod n$.

Proof. Claim that $C^d \equiv P \bmod n$ because p, q are coprime it suffices to show that $C^d \equiv P \bmod p$ and $C^d \equiv P \bmod q$. Let's show $C^d \equiv P \bmod p$. $C^d \equiv (P^e)^d \equiv P^{ed}$ and $ed = 1 + k\phi(n)$, then

$$P^{ed} \equiv P^1 P^{k\phi(n)} = P\big(P^{\phi(n)}\big)^k \equiv P\big(P^{(p-1)(q-1)}\big)^k \equiv P\big(P^{(p-1)}\big)^{(q-1)k}$$

$$\equiv P(1)^{(q-1)k} \bmod p$$

But we can't guarantee $\gcd(P, p) = 1$ then certainly we get $C^d \equiv P \bmod p$. However, if $\gcd(P, p) \neq 1$ then $p \mid P$ and then, $C^d \equiv (P^e)^d \equiv 0 \equiv P \bmod p$. Together, we see that $C^d \equiv P \bmod p$ always and a similar argument for q gives us $C^d \equiv P \bmod q$ always. Thus,

$$C^d \equiv P \bmod n$$

□

Ex. If Bob receives `1611` from Alice, he has computed $d = 2983$ so $ed = 1 \bmod \phi(n)$. He then does $1611^{2983} \equiv 0704 \bmod 4307$ to receieve `HE` from Alice.

(c) **Security:** If Alice (or Eve) knows (n, e) only, how hard is it to find d? The short answer, no, it's extremely hard. Eve wants d with $ed \equiv 1 \bmod \phi(n)$, to do this she needs to know $\phi(n)$, which means she has

to factor n to get $n = pq$ to get $\phi(n) = (p-1)(q-1)$. The issue with this is that factoring seems to be hard. Is it possible that Eve can find $\phi(n)$ without factoring n? Well, if she can factor, then she knows $\phi(n)$. Suppose she knows $\phi(n)$. She also knows n. Observe:

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

$$p - q = \sqrt{(p-q)^2} = \sqrt{(p+q)^2 - 4pq} = \sqrt{(n - \phi(n) + 1)^2 - 4n}$$

Then notice,

$$p = \frac{1}{2}((p+q) + (p-q)) \text{ and } q = \frac{1}{2}((p+q) - (p-q))$$

What all this shows is if we have $n, \phi(n)$ we can get p, q. So, factoring n is equivalently difficult to finding $\phi(n)$.

(d) **Digital Signatures:** Suppose Alice has public key (n_A, e_A) and private key (n_A, d_A) while Bob has public key (n_B, e_B) and private key (n_B, d_B). If Alice wants to send a message to Bob, is there a way to "sign" her message such that Bob knows that she sent it and no one else could have.

Alice takes the plaintext block P and she signs it by doing $S \equiv P^{d_A} \mod n_A$ (only Alice can do this!). So S is the signed plaintext. Note: Since (n_A, e_A) is public, anyone can do S^{e_A} and get $(P^{d_A})^{e_A} \equiv P \mod n_A$ so anyone/everyone can verify that it is "signed" by Alice and only Alice. Then she does S^{e_B} to get $C \equiv S^{e_B} \mod n_B$ this is the encrypted signed plaintext. So in summary,

$$C = \left(P^{d_A} \mod n_A\right)^{e_B} \mod n_B$$

To decrypt and unsign it, Bob does

$$P = \left(C^{d_B} \mod n_B\right)^{e_A} \mod n_A$$

Note: Alice may need to re-block the text here. This is beause the result of signing a block might result in a signed block with a numerical value larger than Bob's encryption modulus.

8.5 RSA Attacks

---

1. **Introduction:** RSA is (currently) incredibly hard to break. Most methods for breaking the encrpytion are well beyond the scope of this course (technical and physical attacks alike). So for this section the "attacks" we cover are less attacks and more so warnings for users of RSA encryptions to be wary of.

2. **Common Modulus Attack:** Suppose Bob1 and Bob2 use the same n. Suppose for security they use coprime e. Bob1 has (n, e_1) and Bob2 has (n, e_2). Suppose then that Alice wants to send P to both of them (scandalous). Then suppose that Eve intercepts both ciphertexts C_1 and C_2. Remeber that (n, e) are public. Eve knows C_1 and C_2 as well as $C_1 = P^{e_1} \bmod n$ and $C_2 = P^{e_2} \bmod n$. However, she does not know P. Since she can discover that $\gcd(e_1, e_2) = 1$ so $\exists \alpha, \beta$ such that $\alpha e_1 + \beta e_2 = 1$. Then she does:

$$C_1^\alpha C_2^\beta = (P^{e_1})^\alpha (P^{e_2})^\beta = P^{\alpha e_1 + \beta e_2} = P^1 \equiv P \bmod n$$

3. **Hastad Broadcast Attack:** This generalizes but the simple version is with three Bobs. Suppose we have Bob1, Bob2, and Bob3 each use $e = 3$ for their encryption exponent, but they all choose pairwise coprime moduli n_1, n_2, and n_3. Suppose then that Alice sends P to all of them;

$$C_1 \equiv P^3 \bmod n_1$$
$$C_2 \equiv P^3 \bmod n_2$$
$$C_3 \equiv P^3 \bmod n_3$$

Suppose then that Eve intercepts all of them, and then creates the following system of linear congruences.

$$x \equiv C_1 \equiv P^3 \bmod n_1$$
$$x \equiv C_2 \equiv P^3 \bmod n_2$$
$$x \equiv C_3 \equiv P^3 \bmod n_3$$

Then by the Chinese Remainder Theorem she can find a unique solution mod $n_1 n_2 n_3$. So she has $x \equiv P^3 \bmod n_1 n_2 n_3$. However, $P < n_1$, $P < n_2$, and $P < n_3$ so in fact $P^3 < n_1 n_2 n_3$ so we then have $P^3 = x \implies P = \sqrt[3]{x}$. (We know that $P^3 = x$ since x is congruent to P^3 and they have the same bounds of $0 \leq x < n_1 n_2 n_3$.)

4. **Interception/Resend Attack:** (Burn Your Trash!) Suppose Bob uses public key (n, e) and private key (n, d). Alice wants to send P to Bob so of course she does $C \equiv P^e \bmod n$. Suppose then that Eve intercepts this C, she will then choose r such that $\gcd(r, n) = 1$ and then sends Bob $\bar{C} \equiv Cr^e \bmod n$. Bob then (not knowing that his message has been tampered with) receieves $\bar{C}$ and attempts to decrypt it, finding:

$$(\bar{C})^d \equiv (Cr^e)^d \equiv (P^e r^e)^d \equiv P^{ed} r^{ed} \equiv Pr \bmod n$$

Which is incomprehensible to him so he bins the message, at which point Eve retrieves it and multiplies by r^{-1} to get P.

8.6 Problems

1. Given the plaintext `LISTENTOITTWICE`. Encrypt using an affine cipher with $a = 11$ and $b = 8$.

2. Suppose Eve intercepts the message `USWNRSCHISPWRCVSHGKCNSBINMRCNPSDN` sent from Alice to Bob using an affine cipher.

   (a) Use frequency analysis to find the values of a and b. Make your steps clear with explanations.

   (b) Decrypt the message.

3. Use the exponentiation cipher with $p = 3637$ and $e = 71$ to encrypt the message:

   `NEEDBACKUPNOWX`

4. Suppose Bob receives the following ciphertext:

   `1333 0513 0452 0767 2130 1395 1097 3597`

   which he knows Alice encrypted using an exponentiation cipher with $p = 3637$ and $e = 71$.

   (a) Find the least nonnegative residue of the decryption exponent d and make sure it's clear what the modulus is.

   (b) Decode the message.

5. Eve intercepts the following ciphertext from Alice to Bob

   `11,17,00,12,10,24,14,00,13,10,11`

   which she knows Alice encrypted using an exponentiation cipher with $p = 29$ and (obviously) using single-character chunks. Eve does not know e or d but she discovered that the first character of the plaintext is `S`.

   (a) Write down the discrete logarithm problem that corresponds to the encrypti on of the first character.

   (b) It is a fact that the integer $r = 2$ is a primitive root modulo $p = 29$. Use this fact along with index arithmetic to solve for e.
   Note: You don't need to write down the entire table of indices for $r = 2$ since you only need two specific values. You can find these by trial-and-error on Wolfram Alpha if you like.

   (c) Use e to solve for d.

(d) Use d to decrypt the message.

6. Given:

   - Alice: Public key $(e_a, n_a) = (103, 3551)$ and private key $d_a = 2599$.
   - Bob: Public key $(e_b, n_b) = (27, 4189)$ and private key $d_b = 1203$.

   (a) Suppose Alice wishes to send the following message to Bob, signed and encrypted:

   `EVEISLISTENING`

   i. Break the plaintext up into two-character strings, padded with an `X` if necessary and assign numerical values.
   ii. Sign the text.
   iii. Encrypt the signed text.

   (b) Suppose Bob receives the following from Alice:

   `0502 0684 2713 1962 3755 1695`

   i. First decrypt the message. The result is signed and hence still unreadable.
   ii. Un-sign the message to reveal the message.

7. Suppose you intercept the following ciphertext from Alice to Bob:

   ```
   160574 069934 062359 171345 116991 061338 246034 232780
                  197240 238665 264414
   062793 213172 090175 151722 269709 259093 194899 145138
                  280675 059999 147437
   ```

   You know that Bob's public key is $(e, n) = (5201, 288319)$. Bob thinks this is secure because he doesn't believe that his n can be factored easily. Factor n, find $\phi(n)$, find d and then decrypt the message. Be clear about the steps you take.

8. Suppose you intercept the two ciphertext blocks $C_1 = 4280$ and $C_2 = 0330$ sent to Bob1 and Bob2. You know that the Bobs' public keys are $(e_1, n) = (100, 4757)$ and $(e_2, n) = (49, 4757)$. Use a common modulus attack to find P.

9. Suppose you intercept the three ciphertext blocks $C_1 = 1533$, $C_2 = 3561$, and $C_3 = 0835$ sent to Bob1, Bob2, and Bob3. You know that the Bobs' public keys are $(e_1, n) = (3, 5353)$, $(e_2, n) = (3, 5251)$, and $(e_3, n) = (3, 5893)$. Use a Hastad Broadcast Attack to find P.

10. Suppose you intercept the ciphertext block $C = 0156$ sent to Bob. You know that Bob's public key is $(e_B, n_B) = (27, 4189)$ so you choose $r = 888$ with $\gcd(888, 4189) = 1$ and perform an intercept and resend attack as follows:

(a) Find the $\overline{C}$ you would resend to Bob.

(b) Bob attempts to decrypt it and gets trash. You retrieve the trash and find it to be 0662. Find the multiplicative inverse of r and use it to find P.