Cryptography

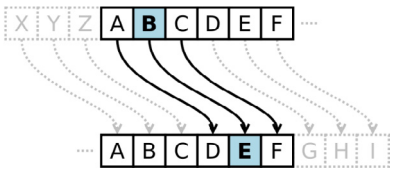
When people need to secretly store or communicate messages, they turn to cryptography. Cryptography involves using techniques to obscure a message so outsiders cannot read the message. It is typically split into two steps: encryption, in which the message is obscured, and decryption, in which the original message is recovered from the obscured form.

Substitution Ciphers

One simple encryption method is called a **substitution cipher**.

> **Substitution Cipher**
> A substitution cipher replaces each letter in the message with a different letter, following some established mapping.

A simple example of a substitution cipher is called the **Caesar cipher**, sometimes called a shift cipher. In this approach, each letter is replaced with a letter some fixed number of positions later in the alphabet. For example, if we use a shift of 3, then the letter A would be replaced with D, the letter 3 positions later in the alphabet. The entire mapping would look like: [1]



Original: ABCDEFGHIJKLMNOPQRSTUVWXYZ
Maps to: DEFGHIJKLMNOPQRSTUVWXYZABC

Example 1

Use the Caesar cipher with shift of 3 to encrypt the message: "We ride at noon"

We use the mapping above to replace each letter. W gets replaced with Z, and so forth, giving the encrypted message: ZH ULGH DW QRRQ.

Notice that the length of the words could give an important clue to the cipher shift used. If we saw a single letter in the encrypted message, we would assume it must be an encrypted A or I, since those are the only single letters than form valid English words.

To obscure the message, the letters are often rearranged into equal sized blocks. The message ZH ULGH DW QRRQ could be written in blocks of three characters as ZHU LGH DWQ RRQ.

[1] http://en.wikipedia.org/w/index.php?title=File:Caesar3.svg&page=1. PD

Example 2

Decrypt the message GZD KNK YDX MFW JXA if it was encrypted using a shift cipher with shift of 5.

We start by writing out the character mapping by shifting the alphabet, with A mapping to F, five characters later in the alphabet.

Original: `ABCDEFGHIJKLMNOPQRSTUVWXYZ`
Maps to: `FGHIJKLMNOPQRSTUVWXYZABCDE`

We now work backwards to decrypt the message. The first letter G is mapped to by B, so B is the first character of the original message. Continuing, our decrypted message is BUY FIF TYS HAR ESA.

Removing spaces we get BUYFIFTYSHARESA. In this case, it appears an extra character was added to the end to make the groups of three come out even, and that the original message was "Buy fifty shares."

Try it Now 1

Decrypt the message BNW MVX WNH if it was encrypted using a shift cipher with shift 9 (mapping A to J).

Notice that in both the ciphers above, the extra part of the alphabet wraps around to the beginning. Because of this, a handy version of the shift cipher is a cipher disc, such as the Alberti cipher disk shown here[2] from the 1400s. In a cipher disc, the inner wheel could be turned to change the cipher shift. This same approach is used for "secret decoder rings."

The security of a cryptographic method is very important to the person relying on their message being kept secret. The security depends on two factors:

1. The security of the method being used
2. The security of the encryption key used

In the case of a shift cipher, the method is "a shift cipher is used." The encryption key is the specific amount of shift used.

Suppose an army is using a shift cipher to send their messages, and one of their officers is captured by their enemy. It is likely the method and encryption key could become compromised. It is relatively hard to change encryption methods, but relatively easy to change encryption keys.

[2] http://en.wikipedia.org/wiki/File:Alberti_cipher_disk.JPG

During World War II, the Germans' Enigma encryption machines were captured, but having details on the encryption method only slightly helped the Allies, since the encryption keys were still unknown and hard to discover. Ultimately, the security of a message cannot rely on the method being kept secret; it needs to rely on the key being kept secret.

Encryption Security
The security of any encryption method should depend only on the encryption key being difficult to discover. It is not safe to rely on the encryption method (algorithm) being kept secret.

With that in mind, let's analyze the security of the Caesar cipher.

Example 3.

Suppose you intercept a message, and you know the sender is using a Caesar cipher, but do not know the shift being used. The message begins EQZP. How hard would it be to decrypt this message?

Since there are only 25 possible shifts, we would only have to try 25 different possibilities to see which one produces results that make sense. While that would be tedious, one person could easily do this by hand in a few minutes. A modern computer could try all possibilities in under a second.

Shift	Message	Shift	Message	Shift	Message	Shift	Message
1	DPYO	7	XJSI	13	RDMC	19	LXGW
2	COXN	8	WIRH	14	QCLB	20	KWFV
3	BNWM	9	VHQG	15	PBKA	21	JVEU
4	AMVL	10	UGPF	16	OAJZ	22	IUDT
5	ZLUK	11	TFOE	17	NZIY	23	HTCS
6	YKTJ	**12**	**SEND**	18	MYHX	24	GSBR
						25	FRAQ

In this case, a shift of 12 (A mapping to M) decrypts EQZP to SEND. Because of this ease of trying all possible encryption keys, the Caesar cipher is not a very secure encryption method.

Brute Force Attack
A brute force attack is a method for breaking encryption by trying all possible encryption keys.

To make a brute force attack harder, we could make a more complex substitution cipher by using something other than a shift of the alphabet. By choosing a random mapping, we could get a more secure cipher, with the tradeoff that the encryption key is harder to describe; the key would now be the entire mapping, rather than just the shift amount.

Example 4

Use the substitution mapping below to encrypt the message "March 12 0300"

Original: `ABCDEFGHIJKLMNOPQRSTUVWXYZ0123456789`
Maps to: `2BQF5WRTD8IJ6HLCOSUVK3A0X9YZN1G4ME7P`

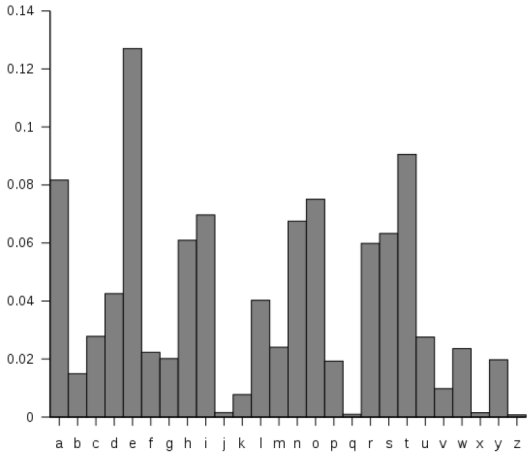
Using the mapping, the message would encrypt to 62SQT ZN Y1YY

Try it Now 2

Use the substitution mapping from Example 4 to decrypt the message C2SVX2VP

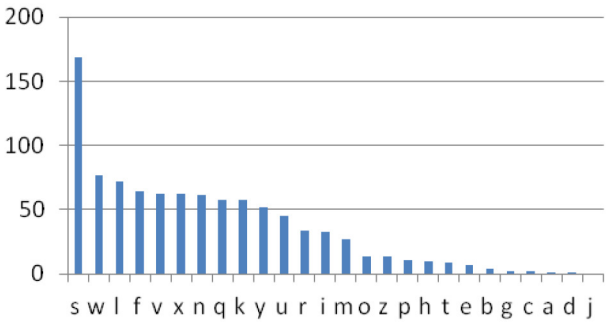
While there were only 25 possible shift cipher keys (35 if we had included numbers), there are about 10^{40} possible substitution ciphers[3]. That's much more than a trillion trillions. It would be essentially impossible, even with supercomputers, to try every possible combination. Having a huge number of possible encryption keys is one important part of key security.

Unfortunately, this cipher is still not secure, because of a technique called frequency analysis, discovered by Arab mathematician Al-Kindi in the 9th century. English and other languages have certain letters than show up more often in writing than others.[4] For example, the letter E shows up the most frequently in English. The chart to the right shows the typical distribution of characters.



Example 5

The chart to the right shows the frequency of different characters in some encrypted text. What can you deduce about the mapping?



Because of the high frequency of the letter S in the encrypted text, it is very likely that the substitution maps E to S. Since W is the second most frequent character, it likely that T or A maps to W. Because C, A, D, and J show up rarely in the encrypted text, it is likely they are mapped to from J, Q, X, and Z.

---

[3] There are 35 choices for what *A* maps to, then 34 choices for what *B* maps to, and so on, so the total number of possibilities is 35*34*33*…*2*1 = 35! = about 10^{40}

[4] http://en.wikipedia.org/w/index.php?title=File:English_letter_frequency_(alphabetic).svg&page=1 PD

In addition to looking at individual letters, certain pairs of letters show up more frequently, such as the pair "th." By analyzing how often different letters and letter pairs show up an encrypted message, the substitution mapping used can be deduced[5].

Transposition Ciphers

Another approach to cryptography is **transposition cipher.**

> **Transposition Ciphers**
> A transposition cipher is one in which the order of characters is changed to obscure the message.

An early version of a transposition cipher was a Scytale[6], in which paper was wrapped around a stick and the message was written. Once unwrapped, the message would be unreadable until the message was wrapped around a same-sized stick again.

One modern transposition cipher is done by writing the message in rows, then forming the encrypted message from the text in the columns.

Example 6

Encrypt the message "Meet at First and Pine at midnight" using rows 8 characters long.

We write the message in rows of 8 characters each. Nonsense characters are added to the end to complete the last row.

```
MEETATFI
RSTANDPI
NEATMIDN
IGHTPXNR
```

We could then encode the message by recording down the columns. The first column, reading down, would be MRNI. All together, the encoded message would be MRNI ESEG ETAH TATT ANMP TDIX FPDN IINR. The spaces would be removed or repositioned to hide the size of table used, since that is the encryption key in this message.

Example 7

Decrypt the message CEE IAI MNL NOG LTR VMH NW using the method above with a table with rows of 5 characters.

Since there are total of 20 characters and each row should have 5 characters, then there will be 20/5 = 4 rows.

---

[5] For an example of how this is done, see http://en.wikipedia.org/wiki/Frequency_analysis

[6] http://en.wikipedia.org/wiki/File:Skytala%26EmptyStrip-Shaded.png

We start writing, putting the first 4 letters, CEEI, down the first column.

```
CALLM
EINTH
EMORN
INGVW
```

We can now read the message: CALL ME IN THE MORNING VW. The VW is likely nonsense characters used to fill out the message.

More complex versions of this rows-and-column based transposition cipher can be created by specifying an order in which the columns should be recorded. For example, the method could specify that after writing the message out in rows that you should record the third column, then the fourth, then the first, then the fifth, then the second. This adds additional complexity that would make it harder to make a brute-force attack.

To make the encryption key easier to remember, a word could be used. For example, if the key word was "MONEY", it would specify that rows should have 5 characters each. The order of the letters in the alphabet would dictate which order to read the columns in. Since E, the 4th letter in the word, is the earliest letter in the alphabet from the word MONEY, the 4th column would be used first, followed by the 1st column (M), the 3rd column (N), the 2nd column (O), and the 5th column (Y).

Example 8

Encrypt the message BUY SOME MILK AND EGGS using a transposition cipher with key word MONEY.

Writing out the message in rows of 5 characters:

```
BUYSO
MEMIL
KANDE
GGSPK
```

We now record the columns in order 4 1 3 2 5:
SIDP BMKG YMNS UEAG OLEK
As before, we'd then remove or reposition the spaces to conceal evidence of the encryption key.

Try it Now 3

Encrypt the message "Fortify the embassy" using a transposition cipher with key word HELP

To decrypt a keyword-based transposition cipher, we'd reverse the process. In the example above, the keyword MONEY tells us to begin with the 4th column, so we'd start by writing SIDP down the 4th column, then continue to the 1st column, 3rd column, etc.

Example 9

Decrypt the message RHA VTN USR EDE AIE RIK ATS OQR using a row-and-column transposition cipher with keyword PRIZED.

The keyword PRIZED tells us to use rows with 6 characters. Since D comes first in the alphabet, we start with 6th column. Since E is next in the alphabet, we'd follow with the 5th column. Continuing, the word PRIZED tells us the message was recorded with the columns in order 4 5 3 6 2 1.

For the decryption, we set up a table with 6 characters in each row. Since the beginning of the encrypted message came from the last column, we start writing the encrypted message down the last column.

					R
					H
					A
					V

The 5th column was the second one the encrypted message was read from, so is the next one we write to.

				T	R
				N	H
				U	A
				S	V

Continuing, we can fill out the rest of the message.

A	I	R	S	T	R
I	K	E	O	N	H
E	A	D	Q	U	A
R	T	E	R	S	V

Reading across the rows gives our decrypted message: AIRSTRIKEONHEADQUARTERSV

Unfortunately, since the transposition cipher does not change the frequency of individual letters, it is still susceptible to frequency analysis, though the transposition does eliminate information from letter pairs.

Advanced shared symmetric-key methods

Both the substitution and transposition methods discussed so far are shared **symmetric-key** methods, meaning that both sender and receiver would have to have agreed upon the same secret encryption key before any methods could be sent.

All of the methods so far have been susceptible to frequency analysis since each letter is always mapped to the same encrypted character. More advanced methods get around this weakness. For example, the Enigma machines used in World War II had wheels that rotated. Each wheel was a substitution cipher, but the rotation would cause the substitution used to shift after each character.

For a simplified example, in the initial setup, the wheel might provide the mapping

Original: `ABCDEFGHIJKLMNOPQRSTUVWXYZ0123456789`

Maps to: `2BQF5WRTD8IJ6HLCOSUVK3A0X9YZN1G4ME7P`

After the first character is encrypted, the wheel rotates, shifting the mapping one space, resulting in a new shifted mapping:

Original: `ABCDEFGHIJKLMNOPQRSTUVWXYZ0123456789`
Maps to: `P2BQF5WRTD8IJ6HLCOSUVK3A0X9YZN1G4ME7`

Using this approach, no letter gets encrypted as the same character over and over.

Example 10

Encrypt the message "See me". Use a basic Caesar cipher with shift 3 as the initial substitution, but shift the substitution one place after each character.

The initial mapping is

Original: `ABCDEFGHIJKLMNOPQRSTUVWXYZ`
Maps to: `DEFGHIJKLMNOPQRSTUVWXYZABC`

This would map the first letter, S to V. We would then shift the mapping by one.

Original: `ABCDEFGHIJKLMNOPQRSTUVWXYZ`
Now maps to: `EFGHIJKLMNOPQRSTUVWXYZABCD`

Now the next letter, E, will map to I. Again we shift the cipher

Original: `ABCDEFGHIJKLMNOPQRSTUVWXYZ`
Now maps to: `FGHIJKLMNOPQRSTUVWXYZABCDE`

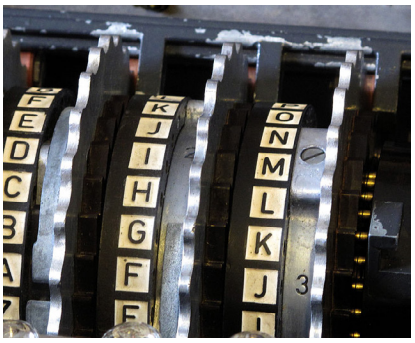
The next letter, E, now maps to J. Continuing this process, the final message would be VIJSL.

Notice that frequency analysis is much less useful now, since the character E has been mapped to three different characters due to the shifting of the substitution mapping.

Try it Now 4

Decrypt the message KIQRV if it was encrypted using a basic Caesar cipher with shift 3 as the initial substitution, but shifting the substitution one place after each character.

The actual Engima machines used in WWII were more complex. Each wheel consisted of a complex substitution cipher, and multiple wheels were used in a chain[7]. The specific wheels used, order of the wheels, and starting position of the wheels formed the encryption key. While captured Engima devices provided the Allied forces details on the encryption method, the keys still had to be broken to decrypt messages.



[7] http://en.wikipedia.org/wiki/File:Enigma_rotors_with_alphabet_rings.jpg

These code breaking efforts led to the development of some of the first electronic computers by Alan Turing at Bletchley Park in the United Kingdom. This is generally considered the beginnings of modern computing[8].

In the 1970s, the U.S. government had a competition and ultimately approved an algorithm deemed DES (Data Encryption Standard) to be used for encrypting government data. It became the standard encryption algorithm used. This method used a combination of multiple substitution and transposition steps, along with other steps in which the encryption key is mixed with the message. This method uses an encryption key with length 56 bits, meaning there are 2^{56} possible keys.

This number of keys make a brute force attack extremely difficult and costly, but not impossible. In 1998, a team was able to find the decryption key for a message in 2 days, using about $250,000 worth of hardware. However, the price and time will go down as computer power increases.

From 1997 to 2001 the government held another competition, ultimately adopting a new method, deemed AES (Advanced Encryption Standard). This method uses encryption keys with 128, 192, or 256 bits, providing up to 2^{256} possible keys, making brute force attacks essentially impossible.

Public Key Cryptography

Suppose that you are connecting to your bank's website. It is possible that someone could intercept any communication between you and your bank, so you'll want to encrypt the communication. The problem is that all the encryption methods we've discussed require than both parties have already agreed on a shared secret encryption key. How can you and your bank agree on a key if you haven't already?

This becomes the goal of public key cryptography – to provide a way for two parties to agree on a key without a snooping third party being able to determine the key. The method relies on a one-way function; something that is easy to do one way, but hard to reverse. We will explore the Diffie-Hellman-Merkle key exchange method.

As an example, let's consider mixing paint. It's easy to mix paint to make a new color, but much harder to separate a mixed paint into the two original colors used.[9][10]

[8] For a good overview, see http://www.youtube.com/watch?v=5nK_ft0Lf1s

[9] http://en.wikipedia.org/w/index.php?title=File:Diffie-Hellman_Key_Exchange.svg&page=1

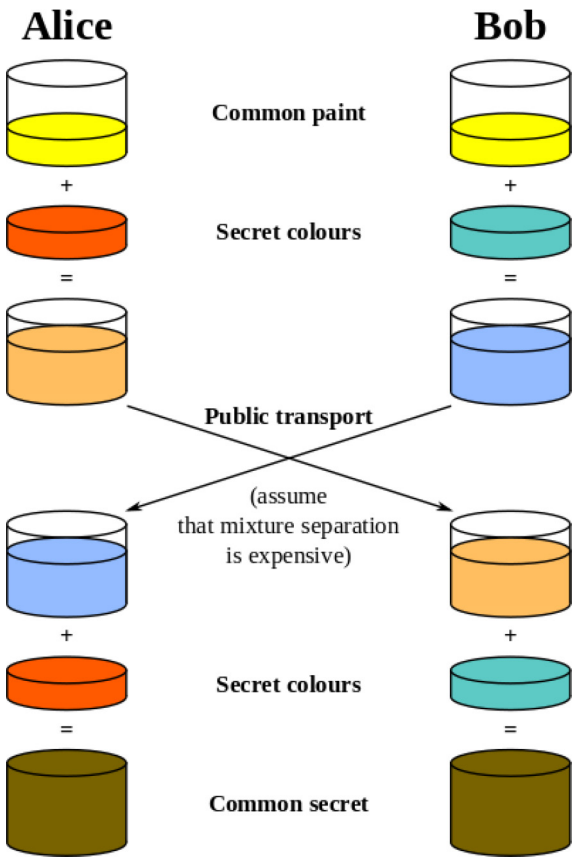
[10] For a video overview of this process, see http://www.youtube.com/watch?v=YEBfamv-_do

Using this analogy, Alice and Bob publically agree on a common starter color. Each then mixes in some of their own secret color. They then exchange their mixed colors.

Since separating colors is hard, even if a snooper were to obtain these mixed colors, it would be hard to obtain the original secret colors.

Once they have exchanged their mixed colors, Alice and Bob both add their secret color to the mix they obtained from the other person. In doing so, both Alice and Bob now have the same common secret color, since it contains a mix of the original common color, Alice's secret color, and Bob's secret color.

They now have a common secret color they can use as their encryption key, even though neither Alice nor Bob knows the other's secret color.
Likewise, there is no way for a snooper to obtain the common secret color without separating one of the mixed colors.



To get this process to work for computer communication, we need to have the process result in a share common number to act as the common secret encryption key. For this, we need a numerical one-way function.

Modular arithmetic

If you think back to doing division with whole numbers, you may remember finding the whole number result and the remainder after division.

Modulus[11]
The **modulus** is another name for the remainder after division.
For example, 17 mod 5 = 2, since if we divide 17 by 5, we get 3 with remainder 2.

Modular arithmetic is sometimes called clock arithmetic, since analog clocks wrap around times past 12, meaning they work on a modulus of 12. If the hour hand of a clock currently points to 8, then in 5 hours it will point to 1. While 8+5 = 13, the clock wraps around after 12, so all times can be thought of as modulus 12. Mathematically, 13 mod 12 = 1.

[11] Sometime, instead of seeing 17 mod 5 = 2, you'll see 17 ≡ 2 (mod 5). The ≡ symbol means "congruent to" and means that 17 and 2 are equivalent, after you consider the modulus 5.

Example 11

Compute: a) 10 mod 3 b) 15 mod 5 c) 2^7 mod 5

a) Since 10 divided by 3 is 3 with remainder 1, 10 mod 3 = 1
b) Since 15 divided by 5 is 3 with no remainder, 15 mod 5 = 0
c) $2^7 = 128$. 128 divide by 5 is 25 with remainder 3, so 2^7 mod 5 = 3

Try it Now 5

Compute: a) 23 mod 7 b) 15 mod 7 c) 2034 mod 7

Recall that when we divide 17 by 5, we could represent the result as 3 remainder 2, as the mixed number $3\frac{2}{5}$, or as the decimal 3.4. Notice that the modulus, 2, is the same as the numerator of the fractional part of the mixed number, and that the decimal part 0.4 is equivalent to the fraction $\frac{2}{5}$. We can use these conversions to calculate the modulus of not-too-huge numbers on a standard calculator.

Modulus on a Standard Calculator

To calculate *a* mod *n* on a standard calculator
1) Divide *a* by *n*
2) Subtract the whole part of the resulting quantity
3) Multiply by *n* to obtain the modulus

Example 12

Calculate 31345 mod 419

31345 / 419 = 74.8090692	Now subtract 74 to get just the decimal remainder
74.8090692 – 74 = 0.8090692	Multiply this by 419 to get the modulus
0.8090692 * 419 = 339	This tells us 0.8090692 was equivalent to $\frac{339}{419}$

In the text above, only a portion of the decimal value was written down. In practice, you should try to avoid writing down the intermediary steps, and instead allow your calculator to retain as many decimal values as it can.

The one-way function

When you use a prime number *p* as a modulus, you can find a special number called a generator, *g*, so that g^n mod *p* will result in all the values from 1 to *p* – 1.

In the table to the right, notice that when we give values of n from 1 to 6, we get out all values from 1 to 6. This means 3 is a generator when 7 is the modulus.

n	3^n	$3^n \bmod 7$
1	3	3
2	9	2
3	27	6
4	81	4
5	243	5
6	729	1

This gives us our one-way function. While it is easy to compute the value of $g^n \bmod p$ when we know n, it is difficult to find the exponent n to obtain a specific value.

For example, suppose we use $p = 23$ and $g = 5$. If I pick n to be 6, I can fairly easily calculate $5^6 \bmod 23 = 15625 \bmod 23 = 8$.

If someone else were to tell you $5^n \bmod 23 = 7$, it is much harder to find n. In this particular case, we'd have to try 22 different values for n until we found one that worked – there is no known easier way to find n other than brute-force guessing.

While trying 22 values would not take too long, when used in practice much larger values for p are used, typically with well over 500 digits. Trying all possibilities would be essentially impossible.

The key exchange

Before we can begin the key exchange process, we need a couple more important facts about modular arithmetic.

Modular Exponentiation Rule
$(a^b \bmod n) = (a \bmod n)^b \bmod n$

Example 13

Compute $12^5 \bmod 7$ using the exponentiation rule.

Evaluated directly: $12^5 = 248{,}832$, so $12^5 \bmod 7 = 248{,}823 \bmod 7 = 3$.

Using the rule above, $12^5 \bmod 7 = (12 \bmod 7)^5 \bmod 7 = 5^5 \bmod 7 = 3125 \bmod 7 = 3$.

You may remember a basic exponent rule from algebra: $(a^b)^c = a^{bc} = a^{cb} = (a^c)^b$
For example: $64^2 = (4^3)^2 = 4^6 = (4^2)^3 = 16^3$

We can combine the modular exponentiation rule with the algebra exponent rule to define the modular exponent power rule.

Modular Exponent Power Rule
$(a^b \bmod n)^c \bmod n = (a^{bc} \bmod n) = (a^c \bmod n)^b \bmod n$

Example 14

Verify the rule above if $a = 3$, $b = 4$, $c = 5$, and $n = 7$

$(3^4 \bmod 7)^5 \bmod 7 = (81 \bmod 7)^5 \bmod 7 = 4^5 \bmod 7 = 1024 \bmod 7 = 2$

$(3^5 \bmod 7)^4 \bmod 7 = (243 \bmod 7)^4 \bmod 7 = 5^4 \bmod 7 = 625 \bmod 7 = 2$, the same result.

Try it Now 6

Use the modular exponent rule to calculate 10000 mod 7, by noting $10000 = 10^4$.

This provides us the basis for our key exchange. While it will be easier to understand in the following example, here's the process:

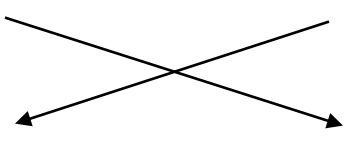
1. Alice and Bob agree publically on values a prime p and generator g.
2. Alice picks some secret number a, while Bob picks some secret number b.
3. Alice computes $A = g^a \bmod p$ and sends it to Bob.
4. Bob computes $B = g^b \bmod p$ and sends it to Alice.
5. Alice computes $B^a \bmod p$, which is $(g^b \bmod p)^a \bmod p$
6. Bob computes $A^b \bmod p$, which is $(g^a \bmod p)^b \bmod p$

The modular exponent power rule tells us $(g^a \bmod p)^b \bmod p = (g^b \bmod p)^a \bmod p$, so Alice and Bob will arrive at the same shared value to use as a key, even though neither knows the other's secret number, and no eavesdropper can determine this value knowing only g, p, A, and B.

Example 15

Alice and Bob publically share a generator and prime modulus. In this case, we'll use 3 as the generator and 17 as the prime.

(This example is continued on the next page)

	Alice		Bob
Alice and Bob publically share a generator and prime modulus.	$g = 3, p = 17$	Common info	$g = 3, p = 17$
Each then secretly picks a number n of their own.	$n = 8$	secret number	$n = 6$
Each calculates $g^n \bmod p$	$3^8 \bmod 17 = 16$		$3^6 \bmod 17 = 15$
They then exchange these resulting values.	$A = 16$  $B = 15$		$B = 15$ $A = 16$
Each then raises the value they received to the power of their secret $n \bmod p$.	$B^n \bmod p =$ $15^8 \bmod 17 = 1$	mix in secret number	$A^n \bmod p =$ $16^6 \bmod 17 = 1$
The result is the shared secret key.	1	shared secret key	1

The shared secrets come out the same because of the modular exponent power rule $(a^b \bmod n)^c \bmod n = (a^c \bmod n)^b \bmod n$. Alice computed $(3^6 \bmod 17)^8 \bmod 17$ while Bob computed $(3^8 \bmod 17)^6 \bmod 17$, which the rule says will give the same results.

Notice that even if a snooper were to obtain both exchanged values $A = 16$ and $B = 15$, there is no way they could obtain the shared secret key from these without having at least one of Alice or Bob's secret numbers. There is no easy way to obtain the secret numbers from the shared values, since the function was a one-way function.

Using this approach, Alice and Bob can now use the shared secret key obtained as the key for a standard encryption algorithm like DES or AES.

Try it Now 7

Suppose you are doing a key exchange with Kylie using generator 5 and prime 23. Your secret number is 2. What number do you send to Kylie? If Kylie sends you the value 8, determine the shared secret key.

RSA

There are several other public-key methods used, including RSA, which is very commonly used. RSA involves distributing a public encryption key, which anyone can use to encrypt messages to you, but which can only be decrypted using a separate private key. You can think of this as sending an open padlock to someone – they can lock up information, but no one can unlock it without the key you kept secret.

RSA's security relies on the difficulty of factoring large numbers. For example, it's easy to calculate that 53 times 59 is 3127, but given the number 12,317 that is a product of two primes, it's much harder to find the numbers that multiply to give that number. It's exponentially harder when the primes each have 100 or more digits. Suppose we find two primes p and q and multiply them to get $n = pq$. This number will be very hard to factor. If we also know p and q, there are shortcuts to find two numbers e and d so that $m^{ed} \bmod n = m \bmod n$ for all numbers m. Without knowing the factorization of n, finding these values is very hard.

To use RSA, we generate two primes p and q and multiply them to get $n = pq$. Since we know the factorization, we can easily find e and d so $m^{ed} = m \bmod n$. Now, we lock away p, q, and d. We then send the values e and n out publically. To encrypt a message m, the sender computes $S = m^e \bmod n$. As we saw earlier, the modulus is a one-way function which makes the original message very hard to recover from S. However, we have our private key d we can use to decrypt the message. When we receive the secret message S, we compute $S^d \bmod n = \left(m^e\right)^d \bmod n = m^{ed} \bmod n = m \bmod n$, recovering the original message[12].

Example 16

Suppose that Alice has computed $n = 3127$, $e = 3$, and $d = 2011$. Show how Bob would encrypt the message 50 and how Alice would then decrypt it.

Bob would only know his message, $m = 50$ and Alice's public key: $n = 3127$ and $e = 3$. He would encrypt the message by computing $m^e \bmod n$: $50^3 \bmod 3127 = 3047$.

Alice can then decrypt this message using her private key d by computing $S^d \bmod n$: $3047^{2011} \bmod 3127 = 50$.

This method differs from Diffie-Hellman-Merkle because no exchange process is needed; Bob could send Alice an encrypted message using Bob's public key without having to communicate with Alice beforehand to determine a shared secret key. This is especially handy for applications like encrypting email, where both parties might not be online at the same time to perform a Diffie-Hellman-Merkle style key exchange.

[12] Many details have been left out, including how e and d are determined, and why this all works. For a bit more detail, see http://www.youtube.com/watch?v=wXB-V_Keiu8, or http://doctrina.org/How-RSA-Works-With-Examples.html

Other secret keeping methods

While this chapter has focused on cryptography methods for keeping a message secure, there is another area of secret keeping called **steganography** which focuses on hiding the existence of a message altogether. Historically, steganography included techniques like invisible ink, watermarks, or embedding a secret message inside a longer document that appeared unimportant. The digital age has provided many new ways to hide messages in digital photos or in the background noise of a music file.

Resources

There are several calculation tools available based on the material in this chapter available at: http://www.opentextbookstore.com/mathinsociety/apps/

Try it Now Answers

1. SEND MONEY

2. PARTY AT 9

3. HELP gives column order 2 1 3 4.

```
FORT
IFYT
HEEM
BASS
YPLR
```

Encrypted text: OFE APF IHB YRY ESL TTM SR

4. The initial mapping was:

```
Original:     ABCDEFGHIJKLMNOPQRSTUVWXYZ
Now maps to:  DEFGHIJKLMNOPQRSTUVWXYZABC
```

Using this, we can see the first character of the encrypted message, K, can be decrypted to the letter H. We now shift the mapping by one character.

```
Original:     ABCDEFGHIJKLMNOPQRSTUVWXYZ
Now maps to:  EFGHIJKLMNOPQRSTUVWXYZABCD
```

The second character in the message, I, can be decrypted to the letter E. Continuing this process of shifting and decrypting, KIQRV decrypts to HELLO.

5. a) 2 b) 1 c) 4

6. $10000 \bmod 7 = 10^4 \bmod 7 = (10 \bmod 7)^4 \bmod 7 = 3^4 \bmod 7 = 81 \bmod 7 = 4$

7. To compute the number we'd send to Kylie, we raise the generator to the power of our secret number modulus the prime: $5^2 \bmod 23 = 25 \bmod 23 = 2$.

If Kylie sends us the value 8, we determine the shared secret by raising her number to the power of our secret number modulus the prime: $8^2 \bmod 23 = 64 \bmod 23 = 18$. 18 would be the shared secret.

Exercises

Substitution ciphers
In the questions below, if it specifies an alphabetic cipher, then the original map used letters only: ABCDEFGHIJKLMNOPQRSTUVWXYZ. If it specifies an alphanumeric cipher, then the original map used letters and numbers:
ABCDEFGHIJKLMNOPQRSTUVWXYZ0123456789

1. Encrypt the message "SEND SUPPLIES" using an alphabetic Caesar cipher with shift 7 (mapping A to H).

2. Encrypt the message "CANCEL CONTRACT" using an alphanumeric Caesar cipher with shift 16 (mapping A to Q).

3. Decrypt the message "2R1 ONO 5SN OXM O" if it was encrypted using an alphanumeric Ceasar cipher with shift 10 (mapping A to K).

4. Decrypt the message "RJJY NSAJ SNHJ" if it was encrypted using an alphabetic Ceasar cipher with shift 5 (mapping A to F).

For questions 5-8 use this substitution mapping:

Original:	`ABCDEFGHIJKLMNOPQRSTUVWXYZ0123456789`
Maps to:	`HLCO2BQF5WRTZN1G4D8IJ6SUVK3A0X9YME7P`

5. Use the substitution mapping to encrypt the message "DEAR DIARY"

6. Use the substitution mapping to encrypt the message "ATTACK AT SUNRISE"

7. Use the substitution mapping to decrypt the message "Z2DQ 2D1N"

8. Use the substitution mapping to decrypt the message "Z22 IHI3 YX3"

Transposition ciphers

9. Encrypt the message "Meet in the library at ten" using a tabular transposition cipher with rows of length 5 characters.

10. Encrypt the message "Fly surveillance over the northern county" using a tabular transposition cipher with rows of length 8 characters.

11. Decrypt the message "THE VHI NIE SAN SHT STI MQA DAN SDR S" if it was encrypted using a tabular transposition cipher with rows of length 7 characters.

12. Decrypt the message "DOLR UTIR INON KVEY AZ" if it was encrypted using a tabular transposition cipher with rows of length 6 characters.

13. Encrypt the message "Buy twenty million" using a tabular transposition cipher with the encryption keyword "RENT".

14. Encrypt the message “Attack from the northeast” using a tabular transposition cipher with the encryption keyword “POWER”.

15. Decrypt the message “RYL OEN ONI TPM IEE YTE YDH WEA HRM S” if it was encrypted using a tabular transposition cipher with the encryption keyword “READING”.

16. Decrypt the message “UYH SRT ABV HLN SEE L” if it was encrypted using a tabular transposition cipher with the encryption keyword “MAIL”.

Shifting substitution ciphers

17. Encrypt the message “SEND SUPPLIES” using an alphabetic Caesar cipher that starts with shift 7 (mapping A to H), and shifts one additional space after each character is encoded.

18. Encrypt the message “CANCEL CONTRACT” using an alphabetic Caesar cipher that starts with shift 5 (mapping A to F), and shifts one additional space after each character is encoded.

Modular arithmetic

19. Compute
    a. 15 mod 4
    b. 10 mod 5
    c. 257 mod 11

20. Compute
    a. 20 mod 4
    b. 14 mod 3
    c. 86 mod 13

21. Determine if 4 is a generator modulus 11

22. Determine if 2 is a generator modulus 13

23. Use the modular exponent rule to calculate 157^{10} mod 5

24. Use the modular exponent rule to calculate 133^{8} mod 6

Diffie-Hellman-Merkle key exchange

25. Suppose you are doing a key exchange with Marc using generator 5 and prime 23. Your secret number is 7. Marc sends you the value 3. Determine the shared secret key.

26. Suppose you are doing a key exchange with Jen using generator 5 and prime 23. Your secret number is 4. Jen sends you the value 8. Determine the shared secret key.

RSA

27. Suppose that Alice has computed $n = 33$, $e = 7$, and $d = 3$. Show how Bob would encrypt the message 5 and how Alice would then decrypt it.

28. Suppose that Alice has computed $n = 55$, $e = 7$, and $d = 13$. Show how Bob would encrypt the message 8 and how Alice would then decrypt it.

Extensions

29. To further obscure a message, sometimes the usual alphabet characters are replaced with other symbols. Design a new set of symbols, and use it to encode a message. Exchange with a friend and see if they can decode your message.

30. To make an encryption harder to break, sometimes multiple substitution and transposition ciphers are used in sequence. For example, a method might specify that the first letter of the encryption keyword be used to determine the initial shift for a Caesar cipher (perhaps with a rotating cipher), and also be used for a transposition cipher. Design your own sequence of encryption steps and encrypt a message. Exchange with a friend and see if they can follow your process to decrypt the message.

31. When using large primes, computing values like $67^{24} \bmod 83$ can be difficult on a calculator without using additional tricks, since 67^{24} is a huge number. We will explore an approach used.
    a. Notice that $67^2 \bmod 83$ is fairly easy to calculate: $67^2 \bmod 83 = 4489 \bmod 83 = 7$.

       Since $67^4 \bmod 83 = (67^2)^2 \bmod 83$ can be rewritten using the modular exponent rule as $(67^2 \bmod 83)^2 \bmod 83$, this is also easy to evaluate:
       $67^4 \bmod 83 = (67^2 \bmod 83)^2 \bmod 83 = 7^2 \bmod 83 = 49$.

       This process can be continued to find $67^8 \bmod 83$ as $(6^4)^2 \bmod 83$. Find this value, then find $67^{16} \bmod 83$ and $67^{32} \bmod 83$.

    b. There is a rule that $(ab) \bmod n = (a \bmod n)(b \bmod n) \bmod n$
       Noting that 17000 = 170*100, calculate 17000 mod 83 using the rule above.

    c. Note that $67^5 = 67^4 67$. Use this, along with the rule from above and the results from part *a* to compute $67^5 \bmod 83$.

    d. Note that $67^7 = 67^{4+2+1} = 67^4\ 67^2\ 67^1$. Compute $67^7 \bmod 83$.

    e. Write 67^{24} as a product of powers of 67, and use this to compute $67^{24} \bmod 83$

32. Use the process from the previous question to evaluate $23^{34} \bmod 37$

33. To encrypt text messages with RSA, the words are first converted into a string of numbers, and then encrypted. Several characters are usually combined together to produce a message number smaller than the modulus, but approximately the same size. Look up an ASCII table to convert the message "SCALE THE WALLS" to numbers, then encrypt it using the RSA public key $n = 10823$, $e = 5$. Since ASCII characters are two digits, pair up characters to form four-digit numbers before encoding. For example A is 65 and B is 66, so the character pair AB could be treated as the number 6566 and encrypted as 10148

34. Explore approaches to steganography that don't require specialized software. Attempt to hide a message using one of these techniques, and see if a fellow student can detect the message.

35. When you visit a secure website, your web browser will report that the site's identity has been verified by a third party, called a certificate authority. This is meant to assure you that you are visiting the actual company's website. Research how these certificates work.