Shippensburg Area Math Circle

Cryptography Toolkit:
The Science and Art of
Encryption, Decryption, and Code-Breaking

Facilitated by Dr. Lance Bryant & Dr. Sarah Bryant

Contents

I. Introduction to Criminal Codes and Ciphers

Introduction

For as long as man has had the ability to communicate, secrecy has been sought. Over the centuries various methods of secret writing, or cryptography, have been developed for numerous purposes. The two major categories of cryptographic systems are ciphers and codes, both of which are used extensively by criminals to conceal clandestine records, conversations, and writings.

Cryptology is the scientific study of cryptography and includes cryptanalytics, which deals with methods of solving cryptographic systems. This article is an introduction to the variety of secret writing encountered in law enforcement and describes the role of FBI cryptanalysts in examining and deciphering these criminal codes and ciphers.

Cipher Systems

Ciphers involve the replacement of true letters or numbers (plain text) with different characters (cipher text) or the systematic rearrangement of the true letters without changing their identities to form an enciphered message. Cipher systems have been common since ancient times and vary in degree of complexity and sophistication. The Enigma Cipher Machine used by the Germans during World War II, for example, was thought to be unbreakable. Only after the fighting had concluded did it become known that the Allies had broken the cipher and had been reading secret German communications throughout the war.

Criminals have a long history of using cipher systems. During the Prohibition Era, rum runners in ships off the East and West Coasts of the United States used a variety of cipher systems, including advanced cipher machines, to communicate with their confederates on shore. The United States Coast Guard and the Department of Commerce pooled their resources to intercept and decipher the rum runners' messages. In 1969 the Zodiac Killer, who terrorized California's Bay Area during the 1960s and 1970s, sent a three-part cipher message to area newspapers explaining his motive for killing. This complex cipher used more than fifty shapes and symbols to represent the 26 letters of the alphabet but was broken in hours by a high school history teacher and his wife.

Criminals typically use homemade, simple substitution cipher systems that use a single cipher text character to replace a plain text character. Those most likely to use such ciphers include criminals involved in clandestine activities that require incriminating records, such as drug trafficking, loansharking, and illegal bookmaking. Incarcerated criminals also use cipher systems to communicate with cohorts inside and outside of prison.

II. Ciphers: Encrypt and Decrypt

In this section we will go through examples of the following ciphers:

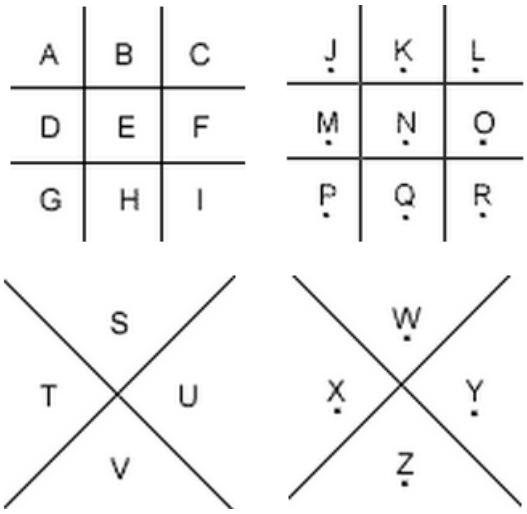
1. Masonic Cipher
2. Caesar Cipher
3. Affine Cipher
4. Vigenère Cipher
5. Autokey Cipher

Be sure that you understand how to encrypt and decrypt each cipher above. You should also have a working knowledge of the uses of each cipher and what their weaknesses are for code breakers.

Section III provides descriptions of even more ciphers.

Cipher 1: Masonic Cipher (aka PigPen Cipher)

Letters are placed inside grids, one possible arrangement of letters is given below:



Example:

Encrypt:

Decrypt:

Cipher 2: Additive Cipher (aka Shift Cipher aka Caesar Cipher)

A relatively basic form of substitution cipher is the Caesar Cipher, named for its Roman origins. The Caesar Cipher involves writing two alphabets, one above the other. The lower alphabet is shifted by one or more characters to the right or left and is used as the cipher text to represent the plain text letter in the alphabet above it.

Plain Text

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
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	A

Cipher Text

Example:
The phrase 'Lucky Dog' would be enciphered as follows:

Plain Text:	L	U	C	**K**	Y	D	O	G
Cipher Text:	M	V	D	**L**	Z	E	P	H

Encrypt (key = 7): MEET AT BRIDGE

Decrypt (key = 2): YCVEJ QWVHQ TUCJN K

Frequency Analysis: Hints and Tips

Solving Simple Substitution Ciphers

If the cryptanalyst knows which language the cipher was written in and has enough cipher text to work with, simple substitution ciphers can often be solved easily. Cryptanalysts use the following procedures when decrypting an unknown cipher:

- The cipher text message is identified from other cipher text or plain text on the document.
- The number of different cipher text characters or combinations are counted to determine if the characters or combinations represent plain text letters, numbers, or both letters and numbers.
- Each cipher text character is counted to determine the frequency of usage.
- The cipher text is examined for patterns, repeated series, and common combinations.

After these analyses have been completed, the cryptanalyst begins to replace cipher text characters with possible plain text equivalents using known language characteristics. For example:

- The English language is composed of 26 letters. However, the nine high-frequency letters E, T, A, O, N, I, R, S, and H constitute 70 percent of plain text.
- EN is the most common two-letter combination, followed by RE, ER, and NT.
- Vowels, which constitute 40 percent of plain text, are often separated by consonants.
- The letter A is often found in the beginning of a word or second from last. The letter I is often third from the end of a word.

Using these and many other known language characteristics, a cryptanalyst can decipher a simple substitution cipher.

Cipher 3: Affine Cipher

Modular Arithmetic is also known as clock arithmetic. This type of arithmetic (addition, subtraction, multiplication, and division) uses only the numbers 0,…., n-1. When have you ever seen that? Well, on a clock of course. 8 hours past 6:00 pm is not 14:00 it is 2:00 am. On the clock we use only the numbers 1,…,12. When working with the alphabet we will use numbers 0,…,25. When we use these numbers we say we are working "mod 26."

Now, the number 8 is 8 mod 26, the number 11 is 11 mod 26 but what is 28 mod 26?

We have gone "around the clock." So to get the simplified version we simply subtract 26 to simplify.
28 mod 26= 2 mod 26.

Try it: 55 mod 26= _____mod 26.

Reduce the following:
(24+8)mod 26 =
(17 * 3) mod 26=
(13*5 + 7)mod 26 =

When using the affine cipher, we first match the letters a to z to numbers 0 to 25. Then we multiply the plaintext number by *a* and next add *b*. The last step is to simplify the result mod 26.

For example, with the key numbers (9, 7) we do the following:
C = (9P+7) mod 26 where P = plaintext number and C = ciphertext number

Given plaintext letter G encrypt with the key (9, 7).

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
7	16	25	8	17	0		18	1	10	19	2	11	20	3	12	21	4	13	22	5	14	23	6	15	24

To decrypt, if someone knows the key (a,b) they must simply reverse by doing the inverse operations.

The affine cipher is susceptible to frequency analysis attacks and computer attacks where the computer tries every possible (a,b) combination. Why is this feasible?

Ciphers 4 and 5: Vigenère and Autokey Cipher

Assign numbers to letters: A = 0, B = 1, C = 2, …, Y = 24, Z = 25

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

Write your plaintext message as a single line. Below that write your keyword repeatedly.
"Add" the two lines to get your ciphertext using mod 26 (wrap around) arithmetic.

Plaintext: B U **R** Y T H E G O L D U N D E **R** T H E **R** O S E B U S H
Keyword: + B L A Z E B L A Z E B L A Z E B L A Z E B L A Z E B L
Ciphertext: C F **R** X X I P G N P E F N C I **S** E H D **V** P D E A Y T S

Notice, for example, that the plaintext R appears three times and it is encrypted differently each time. Also notice that X appears twice in the cryptext but is not from the same plaintext letter.

To decrypt, write the ciphertext message as a single line, below that write your keyword repeatedly, then *subtract* the keyword from the ciphertext.

III. SUMMARY OF CRYPTOGRAPHY METHODS

1. Substitution Ciphers

In these ciphers, each letter of plaintext is replaced by a different symbol (usually another letter).

a. <u>The Caesar cipher</u>

- Shift alphabetically by 3: plain A becomes cipher D. Plain HELLO is encoded as KHOOR.
- Note that letters "wrap around." Plaintext LAZY is encoded as the ciphertext ODCB.

b. <u>Shift ciphers</u>

- Pick a number from 0 to 25; this is your *key*. To encode plaintext, shift all letters by the key amount. The Caesar cipher is really a shift 3 cipher.
- The cipher wheel is useful for encoding/decoding shift ciphers. The final page of this handout contains a cipher wheel you can cut out.

c. <u>Substitution ciphers</u>

- Make a rule that substitutes plaintext letter with ciphertext one. For example, if the rule is

  Plain: 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

  Cipher: D T L Y W U R P K H B F O G J M Q S V X Z A E I N C

  then plain HELLO gets encoded as PWFFJ. The key is the bottom row. If the key is created by just randomly mixing up the alphabet, it can be hard to remember.

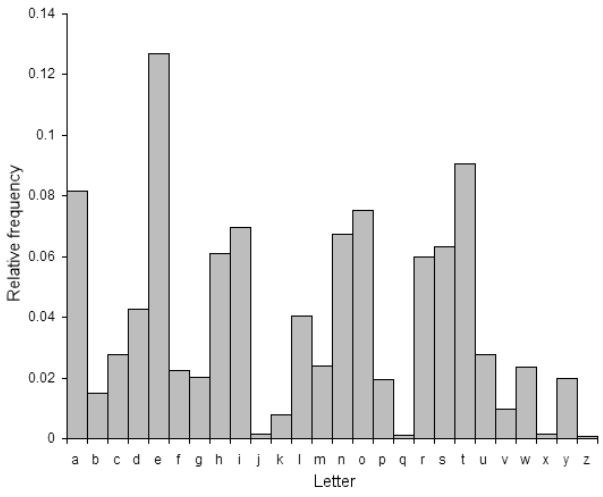
- "Keyed substitution with reversed leftovers" is a convenient way to create a substitution cipher.
  - Pick a keyword, eliminate repeated letters, write the remaining letters in the alphabet in reverse order. Use that as the bottom line for your substitution rule.
  - For example, if the keyword is CONFIDENTIAL, then our substitution rule is

    Plain: 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

    Cipher: C O N F I D E T A L Z Y X W V U S R Q P M K J H G B

d. <u>Frequency analysis</u>

- To crack a ciphertext message encrypted using substitution, compare the frequency of letters in the ciphertext to standard English letter frequencies.



2. Permutation Ciphers

In these ciphers, the letters of the plaintext are scrambled to form the ciphertext.

a. <u>The Rail Fence cipher</u>

- Write the plaintext in a zig-zag, starting from top left.
  To get the ciphertext, read the result off by rows.

  A T C A D W
- ATTACK AT DAWN -> T A K T A N -> ATCA DWTA KTAN

- For a variation, you can change the number of rows and the offset (how far down the first "zig" you start). Here is that same message again, but with three rows and an offset of 1:

  A C D
- ATTACK AT DAWN -> T A K T A N -> TAKT ANTA WACD

  T A W

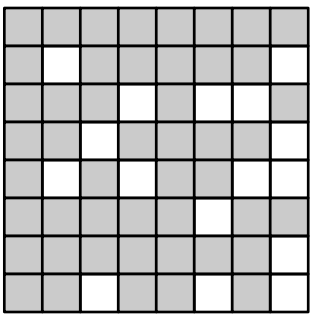
b. <u>Columnar Transposition</u>

- Write your keyword at the top. Below it, write your plaintext in a grid with one column for each letter in the keyword. Number your columns by the alphabetical order of the keyword letters, and read your plaintext off by columns in that order.

M	I	C	K	E	Y	M	O	U	S	E
6	4	1	5	2	11	7	8	10	9	3
B	U	R	Y	T	H	E	G	O	L	D
U	N	D	E	R	T	H	E	R	O	S
E	B	U	S	H						

  In the example here, the ciphertext is RDUT RHDS UNBY ESBU EEHG ELOO RHT.
- Decrypting can be a little tricky! Think carefully about where your letters go.

c. <u>The Turning Grille cipher</u>



- Cut an 8 x 8 grid out of paper and then cut 16 squares out of the grid so that when it is rotated 4 times all 64 of the underlying spaces are revealed.
- To encode a message, place your grille on a piece of paper, write the first 16 letters of your message, rotate the grille 90° clockwise, write the next 16 letters of your message, rotate the grille 90° clockwise, etc. If you have left-over spaces, fill them with dummy letters. At this point you have a square filled with 64 jumbled letters. Read off the rows of the square to create the ciphertext.
- You could also use grid sizes 6 x 6 or 10 x 10 or 12 x 12 or 14 x 14 or…
- How might you make and use a grille cipher using a rectangular 6 x 10 grid? (It's possible!)
- The final page of this handout contains a turning grille you can cut out and use.

3. Polyalphabetic Substitution

In these ciphers, a single letter might be encoded differently throughout the ciphertext.

a. The Vigenère cipher

- Assign numbers to letters: A = 0, B = 1, C = 2, …, Y = 24, Z = 25

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

- Write your plaintext message as a single line. Below that write your keyword repeatedly.
- "Add" the two lines to get your ciphertext using mod 26 (wrap around) arithmetic.
- Plaintext: B U R Y T H E G O L D U N D E R T H E R O S E B U S H
  Keyword: + B L A Z E B L A Z E B L A Z E B L A Z E B L A Z E B L
  Ciphertext: C F R X X I P G N P E F N C I S E H D V P D E A Y T S
- Notice, for example, that the plaintext R appears three times and it is encrypted differently each time.
- To decrypt, write the ciphertext message as a single line, below that write your keyword repeatedly, then *subtract* the keyword from the ciphertext.

b. The Autokey cipher

- A very clever variation on the Vigenère cipher! After writing the keyword **once**, start using the original plaintext.
- Plaintext: B U R Y T H E G O L D U N D E R T H E R O S E B U S H
  Keyword: + B L A Z E B U R Y T H E G O L D U N D E R T H E R O S
  Ciphertext: C F R X X I Y X M E K Y T R P U N U H V F L L F L G Z
- Remember how to decrypt the Autokey cipher?

4. Other methods

a. The Playfair cipher – uses digraph substitution
- This cipher is difficult to explain quickly, so we won't provide the details here. However, this cipher isn't hard to understand, and you might enjoy looking it up to learn how it works.
- This cipher is like substitution, but each *pair* of letters has a corresponding cipher pair. Since there are $26 \cdot 26 = 676$ possible pairs of letters, this makes frequency analysis much more difficult.

b. The Bifid cipher – uses fractioning

	1	2	3	4	5
1	C	O	N	F	I
2	D	E	T	A	L
3	B	G	H	K	M
4	P	Q	R	S	U
5	V	W	X	Y	Z

- Pick a keyword, eliminate repeated letters in it, and write it in a Polybius square followed by all the other letters in the alphabet. To the right, the keyword CONFIDENTIAL is used. *Note*: Since there are only 25 spaces, we **eliminate J** and replace it with I if necessary.
- Record the row and column of each letter. If the plaintext is RETREAT then…

Letter:	R	E	T	R	E	A	T
Row:	4	2	2	4	2	2	2
Column:	3	2	3	3	2	4	3

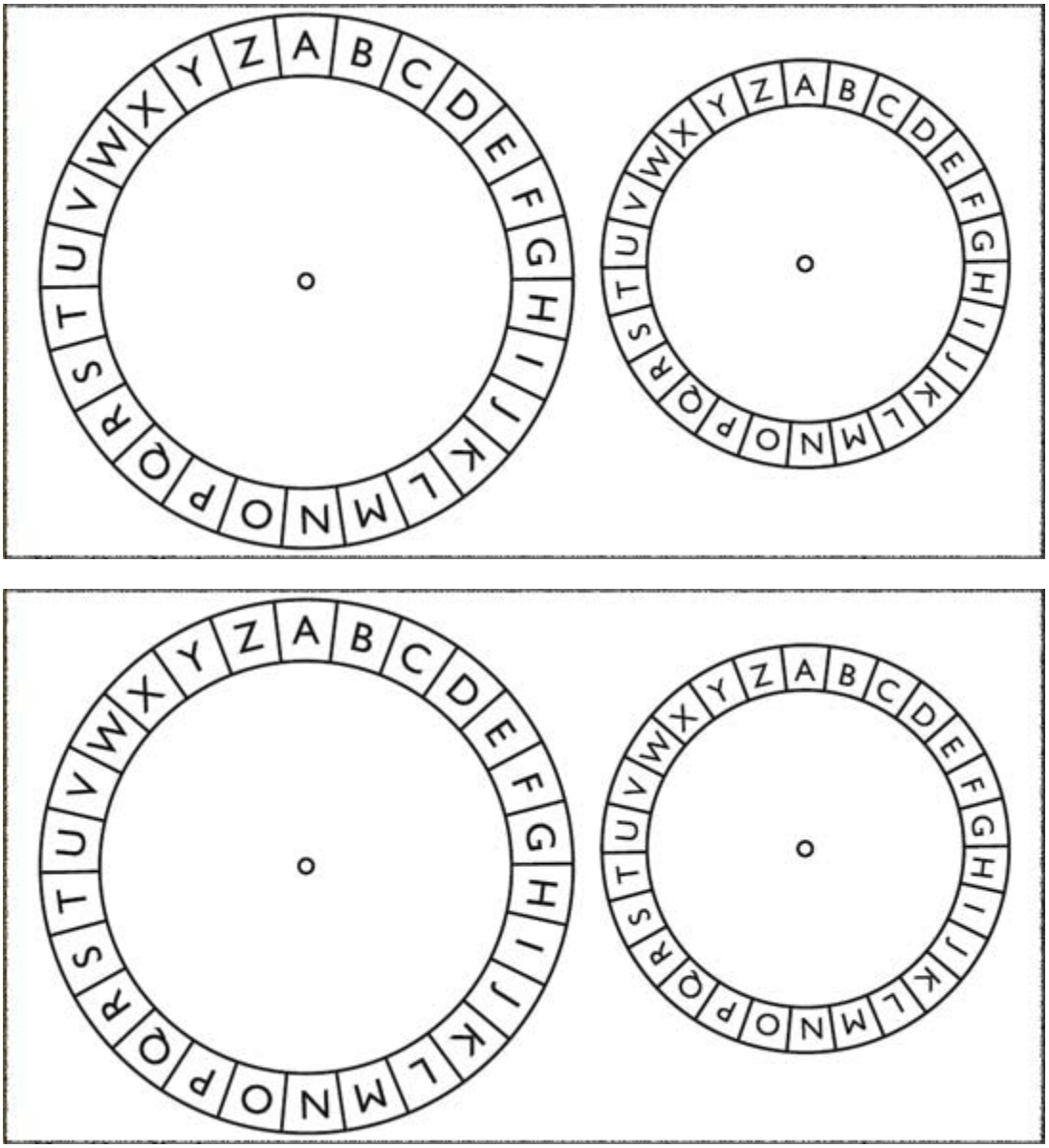
- Write all the row numbers followed by all the column numbers, and break them into pairs.
  42 24 22 23 23 32 43
- Write the letters that correspond to each of these pairs. This is your ciphertext.
  Q A E T T G A

Epilogue – The Key Agreement Problem

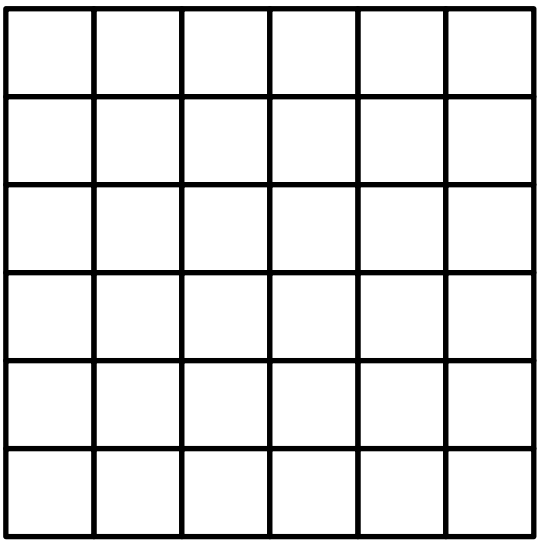
All of the ciphers described here rely on a shared *key* – a piece of information shared by the sender and receiver – to make them work. But how can two people agree on a key if they can only communicate over insecure channels? In the 1970's mathematical methods were devised that make this possible, and these methods form the basis of all modern cryptography today.

To see how this key agreement works, research "Diffie-Hellman key exchange." The mathematics involved is actually not all that complicated: it uses modular arithmetic (like with the Shift, Vigenère, and Autokey ciphers) and the rule of exponents that if $g, a,$ and b are positive integers, then $(g^a)^b = \left(g^b\right)^a = g^{ab}$. Curiously, while both parties end up with the same key at the end of this process, neither party actually *chooses* the key they both end up with. Each person provides part of the information that gets mixed together to form the key they share at the end.

The Cipher Wheel



The Turning Grille



IV. Resources

Online

www.rumkin.com/tools/cipher/

http://www.fbi.gov/about-us/lab/forensic-science-communications/fsc/jan2000/olson.htm

http://www.cryptoclub.org/

Book