Classical Ciphers

Dhananjoy Dey

Indian Institute of Information Technology, Lucknow
ddey@iiitl.ac.in

January 3, 2024

Disclaimers

1

All the pictures used in this presentation are taken from freely available websites.

2

If there is a reference on a slide all of the information on that slide is attributable to that source whether quotation marks are used or not.

Outline

Outline

History of Cryptography

History of Cryptography

- Classical cryptography
- Machine ciphers
- Modern cryptography, and
- post-quantum cryptogaphy

Transposition Ciphers

Transposition Ciphers

Transposition Ciphers

Transposition Ciphers

- **Columnar Transposition**
- **Row Transposition**

Columnar Transposition

- The columnar cipher is a type of transposition cipher.
- In this cipher, the text is written in rows usually of a specific length and read by columns.
- **Plaintext:** Cryptography grows ever more prominent in our lives

Columnar Transposition

- The columnar cipher is a type of transposition cipher.
- In this cipher, the text is written in rows usually of a specific length and read by columns.
- **Plaintext:** Cryptography grows ever more prominent in our lives

C	*r*	*y*	*p*	*t*	*o*
g	*r*	*a*	*p*	*h*	*y*
g	*r*	*o*	*w*	*s*	*e*
v	*e*	*r*	*m*	*o*	*r*
e	*p*	*r*	*o*	*m*	*i*
n	*e*	*n*	*t*	*i*	*n*
o	*u*	*r*	*l*	*i*	*v*
e	*s*	*z*	*z*	*z*	*z*

Columnar Transposition

- The columnar cipher is a type of transposition cipher.
- In this cipher, the text is written in rows usually of a specific length and read by columns.
- **Plaintext:** Cryptography grows ever more prominent in our lives

C	*r*	*y*	*p*	*t*	*o*
g	*r*	*a*	*p*	*h*	*y*
g	*r*	*o*	*w*	*s*	*e*
v	*e*	*r*	*m*	*o*	*r*
e	*p*	*r*	*o*	*m*	*i*
n	*e*	*n*	*t*	*i*	*n*
o	*u*	*r*	*l*	*i*	*v*
e	*s*	*z*	*z*	*z*	*z*

- CGGVENOE RRREPEUS YAORRNRZ PPWMOTLZ THSOMIIZ OYERINVZ

Row Transposition

- In row transposition cipher, first we fix the number of rows
- The plaintext is written in columns based on the number of rows.
- The ciphertext is read in row-wise
- **Plaintext:** Knowledge is the Deity in the worship of progress
  [No. of rows $= 4$]

Row Transposition

- In row transposition cipher, first we fix the number of rows
- The plaintext is written in columns based on the number of rows.
- The ciphertext is read in row-wise
- **Plaintext:** Knowledge is the Deity in the worship of progress [No. of rows $= 4$]

K	l	e	h	i	n	w	h	f	g	s
n	e	i	e	t	t	o	i	p	r	z
o	d	s	D	y	h	r	p	r	e	z
w	g	t	e	i	e	s	o	o	s	z

- KLEHINWHFGS NEIETTOIPRZ ODSDYHRPREZ WGTEIESOOSZ

Transposition Ciphers

Exercise

Find the plaintext for the following ciphertext

DKUWOECERKAOSIHDNSTALTHEROIASGCNHTE

Transposition Ciphers

Exercise

Find the plaintext for the following ciphertext

DKUWOECERKAOSIHDNSTALTHEROIASGCNHTE

Solution

- *Length of the column = 7*
- ***Plaintext:***

  Deadlock continues, Shah writes to Kharge

Keyword Columnar Transposition

- The columnar transposition cipher can be strengthened by using a keyword
- **Plaintext:** CRYPTOISFUN, **Keyword:** MATH

$$\begin{matrix} M & A & T & H \end{matrix}$$

$$\begin{bmatrix} C & R & Y & P \\ T & O & I & S \\ F & U & N & X \end{bmatrix}$$

- **Ciphertext:**

Keyword Columnar Transposition

- The columnar transposition cipher can be strengthened by using a keyword
- **Plaintext:** CRYPTOISFUN, **Keyword:** MATH

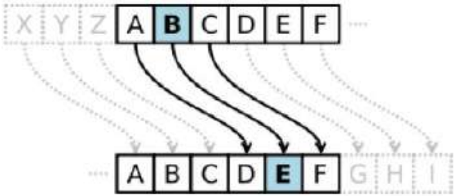
$$\begin{matrix} M & A & T & H \end{matrix}$$

$$\begin{bmatrix} C & R & Y & P \\ T & O & I & S \\ F & U & N & X \end{bmatrix}$$

- **Ciphertext:** ROUPSXCTFYIN

Caesar's Cipher(50 B.C.)

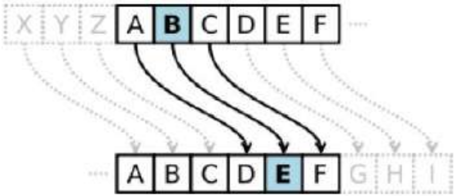
- Used by Caesar to communicate with his generals.
- Each letter is shifted by a constant ($= 3$) position in the alphabet.



E.g., LUCKNOW →

Caesar's Cipher(50 B.C.)

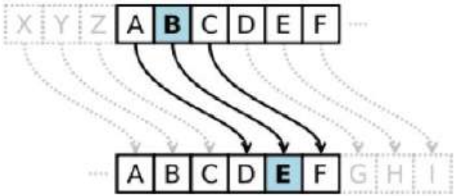
- Used by Caesar to communicate with his generals.
- Each letter is shifted by a constant ($= 3$) position in the alphabet.



E.g., LUCKNOW → OXFNQRZ

Caesar's Cipher(50 B.C.)

- Used by Caesar to communicate with his generals.
- Each letter is shifted by a constant ($= 3$) position in the alphabet.

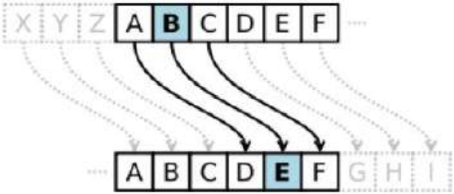


E.g., LUCKNOW → OXFNQRZ

- Shift cipher
- # of possibilities

Caesar's Cipher(50 B.C.)

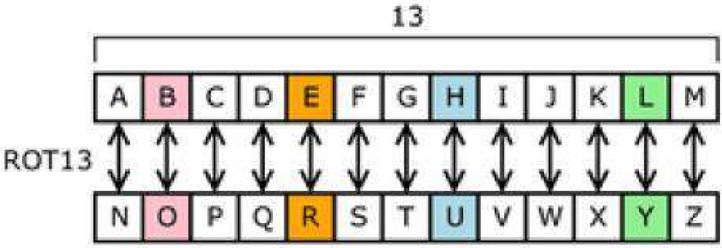
- Used by Caesar to communicate with his generals.
- Each letter is shifted by a constant ($= 3$) position in the alphabet.



E.g., LUCKNOW → OXFNQRZ

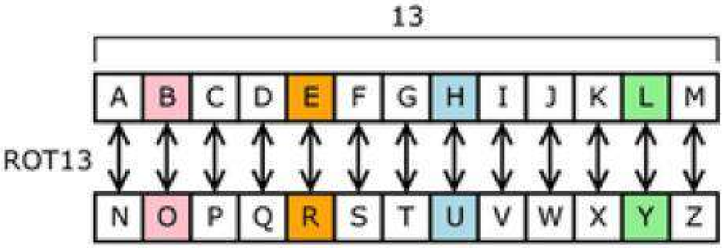
- Shift cipher
- # of possibilities $= 26$.
- On average, a plaintext will be computed after trying 13 decryption rules.

Shift Cipher



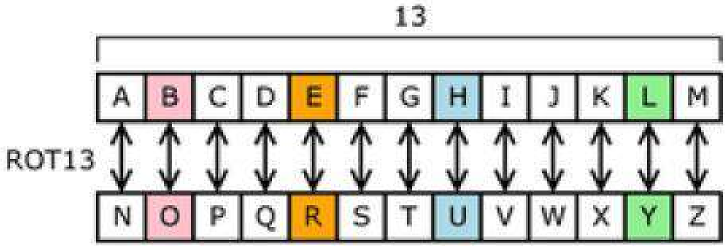
TECHNOLOGY →

Shift Cipher



TECHNOLOGY → GRPUABYBTL

Shift Cipher

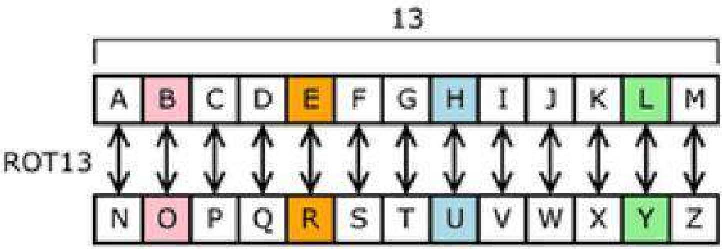


TECHNOLOGY → GRPUABYBTL

Exercise

1. Ciphertext : TBZR ZHFZ ADPAALY DPSS JOHUNL PAZ SVNV. Find the shift and the plaintext.

Shift Cipher

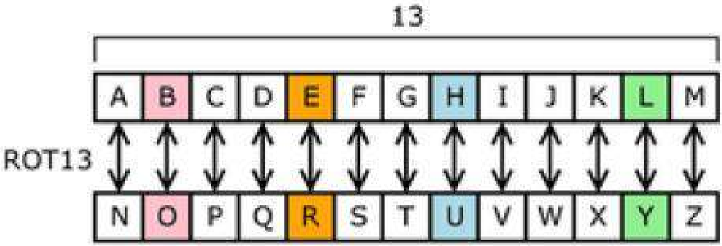


TECHNOLOGY → GRPUABYBTL

Exercise

1. *Ciphertext : TBZR ZHFZ ADPAALY DPSS JOHUNL PAZ SVNV. Find the shift and the plaintext.*
2. *Ciphertext : PWZY XFDV CPMCLYOD EHTEEPC LD I, CPAWLNTYR ESP MWFP MTCO DJXMZW. Find the shift and the plaintext.*

Shift Cipher



TECHNOLOGY → GRPUABYBTL

Exercise

1. *Ciphertext : TBZR ZHFZ ADPAALY DPSS JOHUNL PAZ SVNV. Find the shift and the plaintext.*
2. *Ciphertext : PWZY XFDV CPMCLYOD EHTEEPC LD I, CPAWLNTYR ESP MWFP MTCO DJXMZW. Find the shift and the plaintext.*
3. *Ciphertext : ZYCD-AEKXDEW MBIZDYQBKZRI: K NOMKNO YP BOFYVEDSYXSJSXQ SXDOBXOD COMEBSDI. Find the shift and the plaintext.*

Affine Cipher

- An affine cipher is a simple substitution where

$$c_i \equiv (ap_i + b)\ mod\ 26.$$

Affine Cipher

- An affine cipher is a simple substitution where

$$c_i \equiv (ap_i + b)\ mod\ 26.$$

- What is the key-space of this cipher?

Affine Cipher

- An affine cipher is a simple substitution where

$$c_i \equiv (ap_i + b)\ mod\ 26.$$

- What is the key-space of this cipher?

$$26\phi(26)$$

Affine Cipher

Example

Encrypt the following plaintext using (a, b) as $(3, 8)$

The Cryptographer Who Ensures We Can Trust Our Computers

Affine Cipher

Example

Encrypt the following plaintext using (a, b) as $(3, 8)$

The Cryptographer Who Ensures We Can Trust Our Computers

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

Table: Numerical values assigned to the English alphabet.

- To encrypt 'T', we calculate
  $3.19 + 8$

Affine Cipher

Example

Encrypt the following plaintext using (a, b) as $(3, 8)$

The Cryptographer Who Ensures We Can Trust Our Computers

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

Table: Numerical values assigned to the English alphabet.

- To encrypt 'T', we calculate
  $3.19 + 8 \equiv (3.(-7) + 8) \equiv -21 + 8 \equiv -13 \equiv 13 \mod 26 \rightarrow N$

Affine Cipher

Example

Encrypt the following plaintext using (a, b) as $(3, 8)$
The Cryptographer Who Ensures We Can Trust Our Computers

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

Table: Numerical values assigned to the English alphabet.

- To encrypt 'T', we calculate
  $3.19 + 8 \equiv (3.(-7) + 8) \equiv -21 + 8 \equiv -13 \equiv 13 \mod 26 \rightarrow N$
- **Ciphertext:** Ndu Ohcbnyahibduh Wdy Uvkqhuk Wu Oiv Nhqkn Yqh Oysbqnuhk

Exercises

Exercise

1. *Evaluate the following:*
   (a) $1234 \mod 87$
   (b) $-5678 \mod 91$

Exercises

Exercise

1. *Evaluate the following:*
   (a) $1234 \mod 87$
   (b) $-5678 \mod 91$
2. *If an encryption function* e_K *is identical to the decryption function* d_K*, then the key* K *is said to be an* ***involutory key****. Find all the involutory keys in the Shift Cipher over* $\mathbb{Z}_{26}$*.*

Exercises

Exercise

1. *Evaluate the following:*
   (a) $1234 \mod 87$
   (b) $-5678 \mod 91$
2. *If an encryption function e_K is identical to the decryption function d_K, then the key K is said to be an* ***involutory key****. Find all the involutory keys in the Shift Cipher over $\mathbb{Z}_{26}$.*
3. *Determine the* ***number of keys*** *in an Affine Cipher over $\mathbb{Z}_{500}$.*

Exercises

Exercise

1. *Evaluate the following:*
   (a) $1234 \mod 87$
   (b) $-5678 \mod 91$
2. *If an encryption function e_K is identical to the decryption function d_K, then the key K is said to be an* ***involutory key****. Find all the involutory keys in the Shift Cipher over $\mathbb{Z}_{26}$.*
3. *Determine the* ***number of keys*** *in an Affine Cipher over $\mathbb{Z}_{500}$.*
4. *List all the* ***invertible elements*** *in $\mathbb{Z}_{51}$.*

Mono-alphabetic Cipher

- Each letter is replaced with another letter, according to a fixed substitution

Mono-alphabetic Cipher

- Each letter is replaced with another letter, according to a fixed substitution

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

HELLO WORLD $\rightarrow$

Mono-alphabetic Cipher

- Each letter is replaced with another letter, according to a fixed substitution

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

HELLO WORLD → EZIIP MPWIU

Number of possible keys (Key space):

Mono-alphabetic Cipher

- Each letter is replaced with another letter, according to a fixed substitution

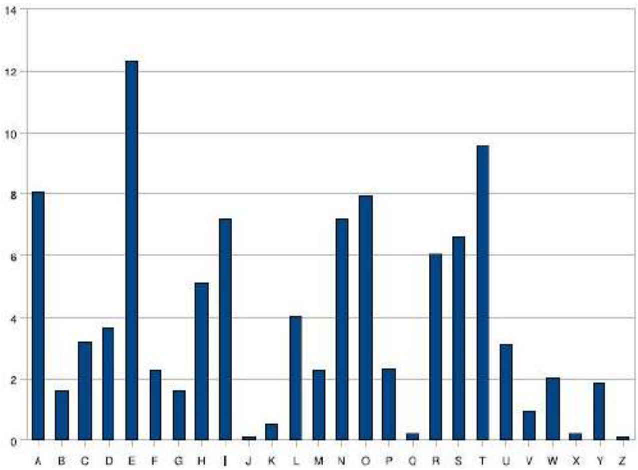
A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
C	G	H	U	Z	J	T	E	L	Y	X	I	F	O	P	K	J	W	V	A	B	D	M	S	N	Q

HELLO WORLD → EZIIP MPWIU

Number of possible keys (Key space): $26!$

Mono-alphabetic Cipher

Frequency Analysis



Mono-alphabetic Cipher

Frequency Analysis

E	12.7%	D	4.2%	P	1.9%
T	9.0%	L	4.0%	B	1.5%
A	8.2%	U	2.8%	V	1.0%
O	7.5%	C	2.8%	K	0.8%
I	7.0%	M	2.4%	Q	0.1%
N	6.7%	W	2.4%	X	0.1%
S	6.3%	F	2.2%	J	0.1%
H	6.1%	G	2.0%	Z	0.1%
R	6.0%	Y	2.0%		

Mono-alphabetic Cipher

Frequency Analysis

digram	*frequency*	*digram*	*frequency*	*digram*	*frequency*	*digram*	*frequency*
th	3.15	to	1.11	sa	0.75	ma	0.56
he	2.51	nt	1.10	hi	0.72	ta	0.56
an	1.72	ed	1.07	le	0.72	ce	0.55
in	1.69	is	1.06	so	0.71	ic	0.55
er	1.54	ar	1.01	as	0.67	ll	0.55
re	1.48	ou	0.96	no	0.65	na	0.54
es	1.45	te	0.94	ne	0.64	ro	0.54
on	1.45	of	0.94	ec	0.64	ot	0.53
ea	1.31	it	0.88	io	0.63	tt	0.53
ti	1.28	ha	0.84	rt	0.63	ve	0.53
at	1.24	se	0.84	co	0.59	ns	0.51
st	1.21	et	0.80	be	0.58	ur	0.49
en	1.20	al	0.77	di	0.57	me	0.48
nd	1.18	ri	0.77	li	0.57	wh	0.48
or	1.13	ng	0.75	ra	0.57	ly	0.47

Mono-alphabetic Cipher

Frequency Analysis

digram	frequency	digram	frequency	digram	frequency	digram	frequency
th	3.15	to	1.11	sa	0.75	ma	0.56
he	2.51	nt	1.10	hi	0.72	ta	0.56
an	1.72	ed	1.07	le	0.72	ce	0.55
in	1.69	is	1.06	so	0.71	ic	0.55
er	1.54	ar	1.01	as	0.67	ll	0.55
re	1.48	ou	0.96	no	0.65	na	0.54
es	1.45	te	0.94	ne	0.64	ro	0.54
on	1.45	of	0.94	ec	0.64	ot	0.53
ea	1.31	it	0.88	io	0.63	tt	0.53
ti	1.28	ha	0.84	rt	0.63	ve	0.53
at	1.24	se	0.84	co	0.59	ns	0.51
st	1.21	et	0.80	be	0.58	ur	0.49
en	1.20	al	0.77	di	0.57	me	0.48
nd	1.18	ri	0.77	li	0.57	wh	0.48
or	1.13	ng	0.75	ra	0.57	ly	0.47

Trigram: the, and, ent, ion, tio, for, nde, . . .

Extension of Mono-alphabetic Cipher

There are three ways to obfuscate the letter frequency:

- **homophone cipher**

Example

Beale cipher – The oldest known usage in 1401

Extension of Mono-alphabetic Cipher

There are three ways to obfuscate the letter frequency:

- **homophone cipher**

Example

Beale cipher – The oldest known usage in 1401

- **polyalphabetic cipher**

Example

Vigenére cipher, Enigma – The oldest known usage in 1568

Extension of Mono-alphabetic Cipher

There are three ways to obfuscate the letter frequency:

- **homophone cipher**

Example

Beale cipher – The oldest known usage in 1401

- **polyalphabetic cipher**

Example

Vigenére cipher, Enigma – The oldest known usage in 1568

- **polygraphic cipher**

Example

Playfair – The oldest known usage in 1854

Homophone Cipher

- The Homophonic Substitution Cipher involves replacing each letter with a variety of substitutes, the number of potential substitutes being proportional to the frequency of the letter.

Homophone Cipher

- The Homophonic Substitution Cipher involves replacing each letter with a variety of substitutes, the number of potential substitutes being proportional to the frequency of the letter.

A	09	12	33	47	53	67	78	92				
B	48	81										
C	13	41	62									
D	01	03	45	79								
E	14	16	24	44	46	55	57	64	74	82	87	98
F	10	31										
G	06	25										
H	23	39	50	56	65	68						
I	32	70	73	83	88	93						
J	15											
K	04											
L	26	37	51	84								
M	22	27										
N	18	58	59	66	71	91						
O	00	05	07	54	72	90	99					
P	38	95										
Q	94											
R	29	35	40	42	77	80						
S	11	19	36	76	86	96						
T	17	20	30	43	49	69	75	85	97			
U	08	61	63									
V	34											
W	60	89										
X	28											
Y	21	52										
Z	02											

Homophone Cipher

Exercise

Encrypt the plaintext: Information Systems Security

Homophone Cipher

Exercise

Encrypt the plaintext: Information Systems Security

Homophonic Cipher

Plaintext
Information System Security

Ciphertext
73 91 31 05 35 27 92 69 83 05 91 86 21 19 85 64 22 96 98 41 08 80 93 20 52

Polygraphic Cipher

- A polygraphic cipher is using substitution of a group of characters in the plaintext alphabet, known as "*poligraph*".

Playfair Cipher

- First choose an encryption key, say, POINTS.
- Enter the letters of the key in the cells of a 5×5 matrix in a left to right fashion starting with the first cell at the top-left corner.
- Fill the rest of the cells of the matrix with the remaining letters in alphabetic order.
- The letters I and J are assigned the same cell.

Polygraphic Cipher

Playfair Cipher

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

Polygraphic Cipher

Playfair Cipher

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

INSTITUTE →

Polygraphic Cipher

Playfair Cipher

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

INSTITUTE → INSTITUTEZ →

Polygraphic Cipher

Playfair Cipher

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

INSTITUTE → INSTITUTEZ → NTDPNPZDKV

Polygraphic Cipher

Playfair Cipher

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

INSTITUTE → INSTITUTEZ → NTDPNPZDKV

INDIAN →

Polygraphic Cipher

Playfair Cipher

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

INSTITUTE → INSTITUTEZ → NTDPNPZDKV

INDIAN → NTBTCO

Poly-alphabetic Cipher

Vigenére Cipher

- A key of the form $K = (k_o, k_1, \cdots, k_{n-1})$, where each $k_i \in \{0, 1, \cdots, 25\}$, is used to encipher the plaintext.
- Each k_i represents a particular shift of the alphabet.
- To encrypt a message

$$C_i \equiv (P_i + k_{i \; mod \; n}) \; mod \; 26$$

- To decrypt

$$P_i \equiv (C_i - k_{i \; mod \; n}) \; mod \; 26$$

Exercise

Find the key space of Vigenére Cipher when the length of keyword n

Poly-alphabetic Cipher

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

Poly-alphabetic Cipher

Vigenére Cipher

- **Plaintext:**

ATTA CKAT DAWN

- **Keyword:** TECH
- **Ciphertext:**

Poly-alphabetic Cipher

Vigenére Cipher

- **Plaintext:**

  ATTA CKAT DAWN

- **Keyword:** TECH
- **Ciphertext:**

  TXVH VOCA WEYU

Poly-alphabetic Cipher

Vigenére Cipher

- **Plaintext:**
  Quantum computers will become important tools as the next generation of problems comes to light
- **Keyword:** Tech
- **Ciphertext:**

Poly-alphabetic Cipher

Vigenére Cipher

- **Plaintext:**
  Quantum computers will become important tools as the next generation of problems comes to light
- **Keyword:** Tech
- **Ciphertext:**
  Jycumyo jhqrbmitz pmns uievfi ktistatrv ahsnz tw vox rgem kguxvcabsp vy ttvupgtl gqtxw vv emiom

Analysis

- A poly-alphabetic substitution cipher uses multiple simple substitutions to encrypt a message
- A polyalphabetic substitution does not preserve plaintext letter frequencies to the same degree as a mono-alphabetic substitution.
- However, if the length keyword is known and the message is long enough, we can transform this into class of simple substitution.

Analysis

How to determine the length of an unknown keyword

Analysis

How to determine the length of an unknown keyword

- **Kasiski Test**
  - It relies on the occasional coincidental alignment of letter groups in plaintext with the keyword.
  - It was described by Friedrich Kasiski in 1863; however, it was apparently discovered earlier, around 1854, by Charles Babbage.
  - It is based on the observation that 2 identical segments of plaintext will be encrypted to the same ciphertext whenever their occurrence in the plaintext is δ positions apart.
  - We find repeated letter groups in the ciphertext and tabulate the separations between them.
  - The gcd of these separations gives a possible length for the keyword.

Analysis

Example

CHREEVOAHMAERATBIAXXWTNXBEEOPHBSBQMQEQERBW
RVXUOAKXAOSXXWEAHBWGJMMQMNKGRFVGXWTRZXWIAK
LXFPSKAUTEMNDCMGTSXMXBTUIADNGMGPSRELXNJELX
VRVPRTULHDNQWTWDTYGBPHXTFALJHASVBFXNGLLCHR
ZBWELEKMSJIKNBHWRJGNMGJSGLXFEYPHAGNRBIEQJT
AMRVLCRREMNDGLXRRIMGNSNRWCHRQHAEYEVTAQEBBI
PEEWEVKAKOEWADREMXMTBHHCHRTKDNVRZCHRCLQOHP

WQAIIWXNRMGWOIIFKEE

Analysis

Example

CHREEVOAHMAERATBIAXXWTNXBEEOPHBSBQMQEQERBW
RVXUOAKXAOSXXWEAHBWGJMMQMNKGRFVGXWTRZXWIAK
LXFPSKAUTEMNDCMGTSXMXBTUIADNGMGPSRELXNJELX
VRVPRTULHDNQWTWDTYGBPHXTFALJHASVBFXNGLLCHR
ZBWELEKMSJIKNBHWRJGNMGJSGLXFEYPHAGNRBIEQJT
AMRVLCRREMNDGLXRRIMGNSNRWCHRQHAEYEVTAQEBBI
PEEWEVKAKOEWADREMXMTBHHCHRTKDNVRZCHRCLQOHP

WQAIIWXNRMGWOIIFKEE

Analysis

Example

CHREEVOAHMAERATBIAXXWTNXBEEOPHBSBQMQEQERBW
RVXUOAKXAOSXXWEAHBWGJMMQMNKGRFVGXWTRZXWIAK
LXFPSKAUTEMNDCMGTSXMXBTUIADNGMGPSRELXNJELX
VRVPRTULHDNQWTWDTYGBPHXTFALJHASVBFXNGLLCHR
ZBWELEKMSJIKNBHWRJGNMGJSGLXFEYPHAGNRBIEQJT
AMRVLCRREMNDGLXRRIMGNSNRWCHRQHAEYEVTAQEBBI
PEEWEVKAKOEWADREMXMTBHHCHRTKDNVRZCHRCLQOHP

WQAIIWXNRMGWOIIFKEE

- The string **CHR** appears at positions 1, 166, 236, 276, and 286.
- The distances from the 1st occurrence to the other 4 occurrences are 165, 235, 275, and 285 rsp.
- The $\gcd$ of these 4 integers is 5 , so that is very likely the keyword length.

Analysis

How to determine the length of an unknown keyword

Analysis

How to determine the length of an unknown keyword

- **Index of Coincidence**
  - The index of coincidence I is defined to be the probability that two randomly selected letters in the ciphertext represent the same plaintext symbol.
  - This concept was defined by William Friedman in 1920.

Analysis

How to determine the length of an unknown keyword

- **Index of Coincidence**
  - The index of coincidence I is defined to be the probability that two randomly selected letters in the ciphertext represent the same plaintext symbol.
  - This concept was defined by William Friedman in 1920.
  - Suppose $\mathbf{x} = x_1 x_2 \cdots x_n$ is a string of n alphabetic characters. The index of coincidence of $\mathbf{x}$ is defined as

$$I_c(\mathbf{x}) = \frac{\sum_{i=0}^{25} f_i(f_i - 1)}{n(n-1)}$$

Analysis

How to determine the length of an unknown keyword

- **Index of Coincidence**
  - The index of coincidence I is defined to be the probability that two randomly selected letters in the ciphertext represent the same plaintext symbol.
  - This concept was defined by William Friedman in 1920.
  - Suppose $\mathbf{x} = x_1 x_2 \cdots x_n$ is a string of n alphabetic characters. The index of coincidence of $\mathbf{x}$ is defined as

$$I_c(\mathbf{x}) = \frac{\sum_{i=0}^{25} f_i(f_i - 1)}{n(n-1)} \approx \sum_{i=0}^{25} p_i^2$$

  - The index of coincidence of English text ≈ 0.065.
  - I for a random text ≈ 0.03846.
  - For any English ciphertext the index of coincidence I must satisfy $0.03846 \leq I \leq 0.065$.

Poly-alphabetic Cipher

Hill Cipher[1]

- Encryption key,

$$K = \begin{pmatrix} k_{11} & k_{12} & k_{13} \\ k_{21} & k_{22} & k_{23} \\ k_{31} & k_{32} & k_{33} \end{pmatrix}$$

[1] Hill cipher was developed by **Lester S. Hill**, an American mathematician in 1929.

Poly-alphabetic Cipher

Hill Cipher[1]

- Encryption key,

$$K = \begin{pmatrix} k_{11} & k_{12} & k_{13} \\ k_{21} & k_{22} & k_{23} \\ k_{31} & k_{32} & k_{33} \end{pmatrix}$$

- The plaintext letters $p_1, p_2 \ \& \ p_3$ encrypted into ciphertext letters $c_1, c_2 \ \& \ c_3$ by

$$\begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} k_{11} & k_{12} & k_{13} \\ k_{21} & k_{22} & k_{23} \\ k_{31} & k_{32} & k_{33} \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix}$$

[1] Hill cipher was developed by **Lester S. Hill**, an American mathematician in 1929.

Hill Cipher

Example

- Use keyword “TECH"

Hill Cipher

Example

- Use keyword “TECH"

$$\mathbf{K} = \begin{pmatrix} 19 & 4 \\ 2 & 7 \end{pmatrix}$$

- **Encrypt:** Offering e-Learning support to its learners

Hill Cipher

Example

- Use keyword "TECH"

$$\mathbf{K} = \begin{pmatrix} 19 & 4 \\ 2 & 7 \end{pmatrix}$$

- **Encrypt:** Offering e-Learning support to its learners

$$\mathbf{K} = \begin{pmatrix} 19 & 4 \\ 2 & 7 \end{pmatrix} \begin{pmatrix} 14 \\ 5 \end{pmatrix} =$$

Hill Cipher

Example

- Use keyword “TECH"

$$\mathbf{K} = \begin{pmatrix} 19 & 4 \\ 2 & 7 \end{pmatrix}$$

- **Encrypt:** Offering e-Learning support to its learners

$$\mathbf{K} = \begin{pmatrix} 19 & 4 \\ 2 & 7 \end{pmatrix} \begin{pmatrix} 14 \\ 5 \end{pmatrix} = \mathbf{K} = \begin{pmatrix} -7 & 4 \\ 2 & 7 \end{pmatrix} \begin{pmatrix} -12 \\ 5 \end{pmatrix} = \begin{pmatrix} 104 \\ 11 \end{pmatrix} = \begin{pmatrix} 0 \\ 11 \end{pmatrix}$$

- **Ciphertext:**

Alhmrmlq q-Hyilvwde iypdyjl bg utw jyilvoxaD

Exercises

Exercise

1. *Let p be prime. Find the number of 3×3 invertible matrices over $\mathbb{Z}_p$.*
2. *Find the number of $n \times n$ invertible matrices over $\mathbb{Z}_p$.*
3. *Find the number of $n \times n$ invertible matrices over $\mathbb{Z}_{p^\alpha}$*
4. *Find the number of $n \times n$ invertible matrices over $\mathbb{Z}_m$*

Cryptography During The French and American Wars in Vietnam

CRYPTOGRAPHY DURING THE FRENCH AND AMERICAN WARS IN VIETNAM

PHAN DƯƠNG HIỆU AND NEAL KOBLITZ

ABSTRACT. After Vietnam's Declaration of Independence on 2 September 1945, the country had to suffer through two long, brutal wars, first against the French and then against the Americans, before finally in 1975 becoming a unified country free of colonial domination. Our purpose is to examine the role of cryptography in those two wars. Despite the far greater technological resources of their opponents, the communications intelligence specialists of the Việt Minh, the National Liberation Front, and the Democratic Republic of Vietnam had considerable success in both protecting Vietnamese communications and acquiring tactical and strategic secrets from the enemy. Perhaps surprisingly, in both wars there was a balance between the sides. Generally speaking, cryptographic knowledge and protocol design were at a high level at the central commands, but deployment for tactical communications in the field was difficult, and there were many failures on all sides.

http://eprint.iacr.org/2016/1136.pdf

Classical Ciphers

- These ciphers are too weak nowadays, too easy to break, especially with computers.
- However, these simple ciphers give a good illustration of several of the important ideas of the cryptography and cryptanalysis.
- Moreover, most of them can be very useful in combination with more modern cipher – to add a new level of security.

Outline

Codebook Cipher

- Literally, a book filled with “codes"
  - More precisely, **2 codebooks**, 1 for *encryption* and other for *decryption*
- Key itself is the codebook
- Security of cipher requires physical security for codebook
- Codebooks widely used through WWII

Codebook Cipher

- Literally, a book filled with “codewords"
- Zimmerman Telegram encrypted via codebook

Februar	13605
fest	13732
finanzielle	13850
folgender	13918
Frieden	17142
Friedenschluss	17149
⋮	⋮

Codebook Cipher

- Literally, a book filled with “codewords"
- Zimmerman Telegram encrypted via codebook

Februar	13605
fest	13732
finanzielle	13850
folgender	13918
Frieden	17142
Friedenschluss	17149
⋮	⋮

- Modern block ciphers are based on codebooks cipher

Permutation on Block of Characters

Example

AAAA	AAAB	AAAC	$\cdots$	ZZZZ
QAQZ	WIJT	ENTO	$\cdots$	MIHB

Permutation on Block of Characters

Example

AAAA	AAAB	AAAC	⋯	ZZZZ
QAQZ	WIJT	ENTO	⋯	MIHB

- ‘*code book*’

Permutation on Block of Characters

Example

AAAA	AAAB	AAAC	···	ZZZZ
QAQZ	WIJT	ENTO	···	MIHB

- ‘*code book*’
- If blocks are large enough, then frequency analysis becomes impossible (infeasible).

Block Cipher

- Avoid transport & storage of huge table
- Introduce computation rule to compute table elements:

$$T[X] = f_{key}(X)$$

- Design “good" rule f:

Block Cipher

- Avoid transport & storage of huge table
- Introduce computation rule to compute table elements:

$$T[X] = f_{key}(X)$$

- Design “good" rule f:
  - Secure
  - Efficient

Surprise Test (10 min)

Surprise Test (10 min)

Problem

(i) *First write your registration number of the form Ics2021abc/Iit2021xyz. Consider the decimal value of abc/xyz and then add 143/137. Denote the result of the addition as n. Consider the following functions:*

$$f : \mathbb{Z}_n \to \mathbb{Z}_n \quad \text{and} \quad g : \mathbb{Z}_n \to \mathbb{Z}_{26}, \quad \text{where}$$

$$f(x) = (49x + 7)/(51x + 13) \mod n \quad \text{and} \quad g(x) = f(x) \mod 26.$$

Convert TECH in numerical value using the following encoding method and write the output in alphabetic form after applying g on it.

$$C \to 2, \quad E \to 4, \quad H \to 7, \quad T \to 19$$

(ii) *How many one-to-one mapping f you can define of the form* $x \mapsto (\alpha x + \beta) \mod n$

4 +1

Outline

One-Time Pad

It was first described by Gilbert Vernam in 1917 for use in automatic encryption and decryption of telegraph messages.

One-Time Pad

It was first described by Gilbert Vernam in 1917 for use in automatic encryption and decryption of telegraph messages.

Encryption

e=000 h=001 i=010 k=011 l=100 r=101 s=110 t=111

Encryption: Plaintext $\oplus$ Key = Ciphertext

One-Time Pad

It was first described by Gilbert Vernam in 1917 for use in automatic encryption and decryption of telegraph messages.

Encryption

e=000 h=001 i=010 k=011 l=100 r=101 s=110 t=111

Encryption: Plaintext $\oplus$ Key = Ciphertext

	h	e	i	l	h	i	t	l	e	r
Plaintext:	001	000	010	100	001	010	111	100	000	101
Key:	111	101	110	101	111	100	000	101	110	000

One-Time Pad

It was first described by Gilbert Vernam in 1917 for use in automatic encryption and decryption of telegraph messages.

Encryption

e=000 h=001 i=010 k=011 l=100 r=101 s=110 t=111

Encryption: Plaintext ⊕ Key = Ciphertext

	h	e	i	l	h	i	t	l	e	r
Plaintext:	001	000	010	100	001	010	111	100	000	101
Key:	111	101	110	101	111	100	000	101	110	000
Ciphertext:	110	101	100	001	110	110	111	001	110	101
	s	r	l	h	s	s	t	h	s	r

One-Time Pad

Decryption

e=000 h=001 i=010 k=011 l=100 r=101 s=110 t=111

Decryption: Ciphertext $\oplus$ Key = Plaintext

One-Time Pad

Decryption

e=000 h=001 i=010 k=011 l=100 r=101 s=110 t=111

Decryption: Ciphertext $\oplus$ Key = Plaintext

	s	r	l	h	s	s	t	h	s	r
Ciphertext:	110	101	100	001	110	110	111	001	110	101
Key:	111	101	110	101	111	100	000	101	110	000
Plaintext:	001	000	010	100	001	010	111	100	000	101
	h	e	i	l	h	i	t	l	e	r

One-Time Pad

- **Provably secure** $\cdots$
  - Ciphertext provides no info about plaintext
  - All plaintexts are equally likely

One-Time Pad

- **Provably secure** $\cdots$
  - Ciphertext provides no info about plaintext
  - All plaintexts are equally likely
- $\cdots$ but, only when be used correctly
  - Key must be random, used only once
  - Key is known only to sender and receiver

One-Time Pad

- **Provably secure** · · ·
  - Ciphertext provides no info about plaintext
  - All plaintexts are equally likely
- · · · but, only when be used correctly
  - Key must be random, used only once
  - Key is known only to sender and receiver
- Note: Key is same size as message

One-Time Pad

- **Provably secure** · · ·
  - Ciphertext provides no info about plaintext
  - All plaintexts are equally likely
- · · · but, only when be used correctly
  - Key must be random, used only once
  - Key is known only to sender and receiver
- Note: Key is same size as message
- So, why not distribute message instead of pad?

Real-World One-Time Pad

- Project **VENONA**
  - Encrypted spy messages from U.S. to Moscow in 30’s, 40’s & 50’s
  - Nuclear espionage, etc.
  - Thousands of messages
- Spy carried one-time pad into U.S.
- Spy used key to encrypt secret messages
- Repeats within the “one-time" key made cryptanalysis possible

VENONA Decrypt (1944)

[C% Ruth] learned that her husband [v] was called up by the army but he was not sent to the front. He is a mechanical engineer and is now working at the ENORMOUS [ENORMOZ] [vi] plant in SANTA FE, New Mexico. [45 groups unrecoverable] detain VOLOK [vii] who is working in a plant on ENORMOUS. He is a FELLOWCOUNTRYMAN [ZEMLYaK] [viii]. Yesterday he learned that they had dismissed him from his work. His active work in progressive organizations in the past was cause of his dismissal. In the FELLOWCOUNTRYMAN line LIBERAL is in touch with CHESTER [ix]. They meet once a month for the payment of dues. CHESTER is interested in whether we are satisfied with the collaboration and whether there are not any misunderstandings. He does not inquire about specific items of work [KONKRETNAYa RABOTA]. In as much as CHESTER knows about the role of LIBERAL's group we beg consent to ask C. through LIBERAL about leads from among people who are working on ENOURMOUS and in other technical fields.

- “Ruth" == Ruth Greenglass, “Liberal" == Julius Rosenberg, “Enormous" == the atomic bomb

Outline

Self Study

Machine Ciphers

- **Enigma** – German
- **Purple** – Japan
- **Sigaba** (US Army)/**ECM** (Electric Cipher Machine (US Navy))

Enigma

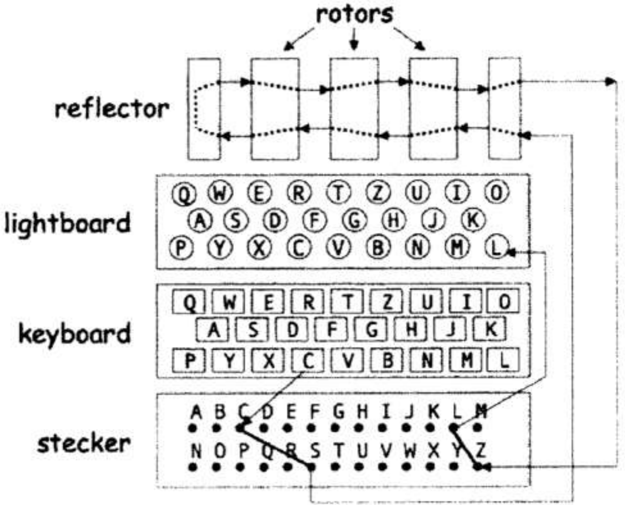
- Developed and patented in 1918 by Arthur Scherbius – Electo-mechanical
- Many variations on basic design
- Eventually adopted by Germany
  - For both military and diplomatic use
  - Many variations used
- Broken by Polish cryptanalysts Henryk Zygalski, Jerzy Rozycki, and Marian Rejewski, late 1930s
- Exploited throughout WWII
  - By Poles, British, Americans

Enigma

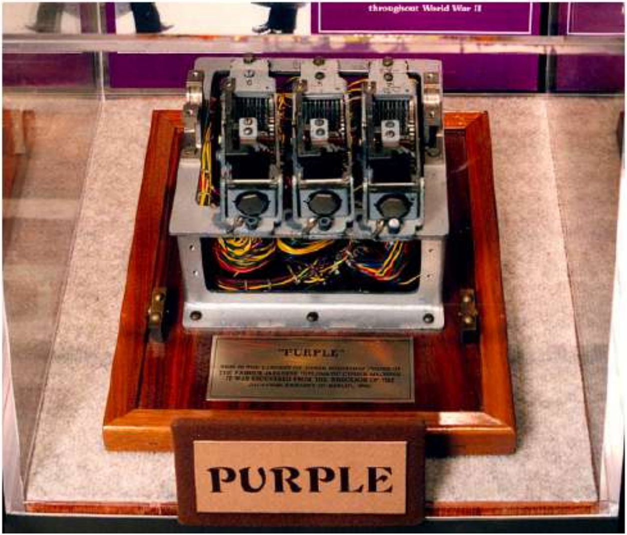
Enigma

- Enigma is a substitution cipher
- But not a simple substitution
  - Permutation changes with each letter typed
- Enigma is an example of a poly-alphabetic substitution

Enigma – Wiring Diagram



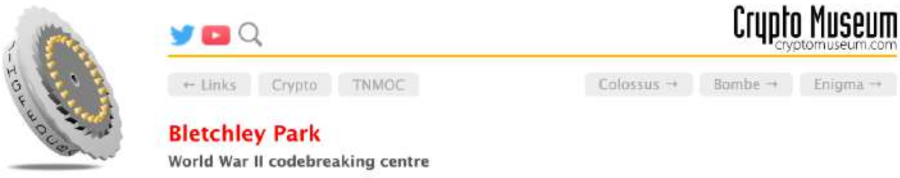
Purple



Sigaba

Crypto Museum

Crypto Museum



https://www.cryptomuseum.com/bp/

Classifying World War II Era Ciphers with ML

Classifying World War II Era Ciphers with Machine Learning

Brooke Dalton* Mark Stamp*†

July 4, 2023

- **Classified:** Enigma, M-209, Sigaba, Purple, and Typex

Classifying World War II Era Ciphers with ML

Classifying World War II Era Ciphers with Machine Learning

Brooke Dalton* Mark Stamp*†

July 4, 2023

- **Classified:** Enigma, M-209, Sigaba, Purple, and Typex

https://arxiv.org/pdf/2307.00501.pdf

The End

Thanks a lot for your attention!