University of Basel

Center for Innovative Finance

Bitcoin, Blockchain and Cryptoassets

Symmetric Cryptography

Prof. Dr. Fabian Schär
University of Basel

Release Ver.: (Local Release)
Version Hash: (None)
Version Date: (None)

Types of Secret Writing

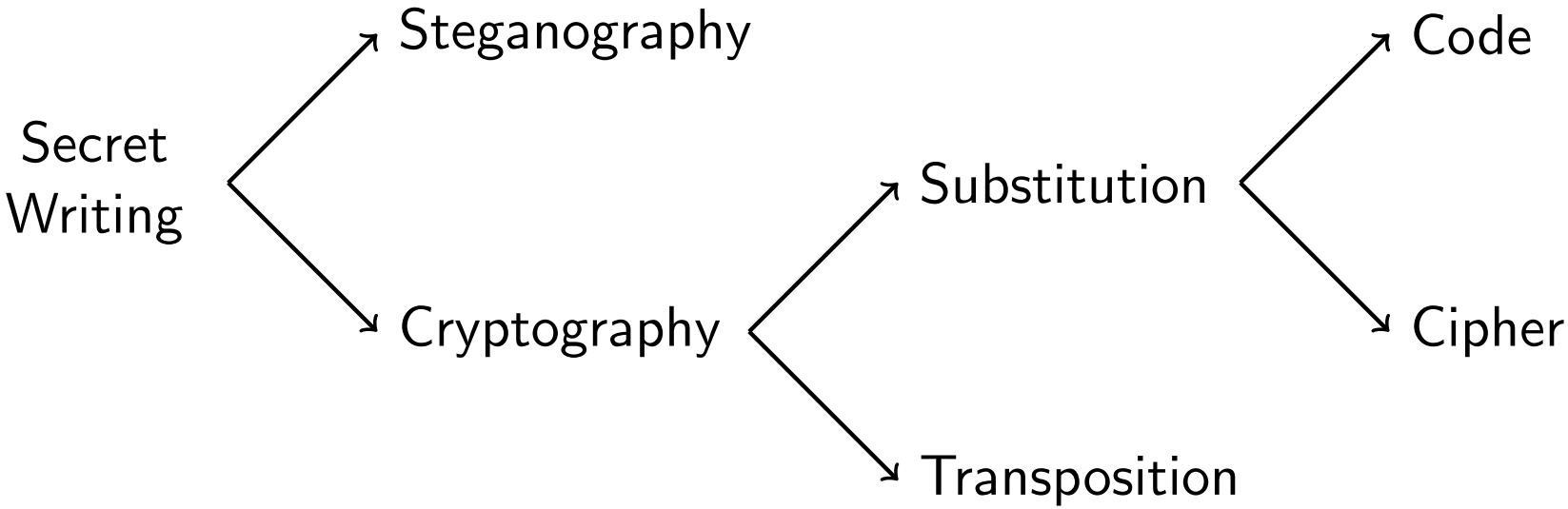

Figure: Types of secret writing. Based on [1]

Monoalphabetic Substitution

- Simple case: shift alphabet by $x \in \{1, 25\}$ positions

Plain alphabet	a	b	c	d	e	f	[...]	t	u	v	w	x	y	z
Cipher alphabet	D	E	F	G	H	I	[...]	W	X	Y	Z	A	B	C
Plain text	v	e	n	i,		v	i	d	i,		v	i	c	i
Cipher text	Y	H	Q	L,		Y	L	G	L,		Y	L	F	L

Table: Principle of the Caesar Cipher. Source [1]

- Results in 25 distinct cipher alphabets

Monoalphabetic Substitution

- More advanced: arbitrary letter mapping

Plain alphabet	a	b	c	d	e	f	[...]	t	u	v	w	x	y	z
Cipher alphabet	Z	F	L	V	A	R	[...]	Q	M	E	Y	P	C	W

- Available cipher alphabets:
  $26! = 403{,}291{,}461{,}126{,}605{,}635{,}584{,}000{,}000$

Breaking Monoalphabetic Substitution

YOEE UHFO, THM PWO PVPJBFS. YO GHFSW-PXMEPXO THM LHW ZOBFS P GMWBHMR PFU XAHWHMSA RXMUOFX. XABR RPVDEO XOIX YPR GAHROF BF P YPT XAPX XAO LBWRX XAWOO VHRX LWOKMOFX EOXXOWR YHMEU GHBFGBUO YBXA XAO XAWOO VHRX LWOKMOFX EOXXOWR BF XAO OFSEBRA EPFSMPSO. BF RAHWX XOIXR LWOKMOFGBOR GPF NPWT P EHX. RONOWPE PXXOVDXR PFU UONBPXBHFR LWHV RXWBGXET DPBWBFS XAO EOXXOWR PGGHWUBFS XH XAOBW LWOKMOFGT GPF ZO FOGORRPWT.

Letter	#	Frequency (%)	Letter	#	Frequency (%)
A	15	4.5	N	3	0.9
B	20	6.0	O	45	13.6
C	0	0.0	P	25	7.6
D	3	0.9	Q	0	0.0
E	13	3.9	R	22	6.6
F	24	7.3	S	9	2.7
G	13	3.9	T	7	2.1
H	19	5.7	U	8	2.4
I	2	0.6	V	6	1.8
J	1	0.3	W	24	7.3
K	4	1.2	X	40	12.1
L	7	2.1	Y	6	1.8
M	13	3.9	Z	2	0.6

Table: Frequency analysis of the encrypted message.

Breaking Monoalphabetic Substitution

Letter	Frequency (%)	Letter	Frequency (%)
a	8.04	n	7.23
b	1.48	o	7.64
c	3.34	p	2.14
d	3.82	q	0.12
e	12.49	r	6.28
f	2.40	s	6.51
g	1.87	t	9.28
h	5.05	u	2.73
i	7.57	v	1.05
j	0.16	w	1.68
k	0.54	x	0.23
l	4.07	y	1.66
m	2.51	z	0.09

Table: Relative frequencies in the English language. Source: norvig.com/mayzner.html

Letter	Frequency (%)	Letter	Frequency (%)
A	4.5	N	0.9
B	6.0	O	13.6
C	0.0	P	7.6
D	0.9	Q	0.0
E	3.9	R	6.6
F	7.3	S	2.7
G	3.9	T	2.1
H	5.7	U	2.4
I	0.6	V	1.8
J	0.3	W	7.3
K	1.2	X	12.1
L	2.1	Y	1.8
M	3.9	Z	0.6

Table: Frequency analysis of the encrypted text.

- Try to find common words like “the” in text
- Exploit relations amongst letters (e.g. “qu” or “th”)
- Insert deciphered letters $\rightarrow$ repeat

Improved Monoalphabetic Substitution

- Use symbols that delete preceding symbol
- Homophone encryption:
  - Use multiple symbols to encrypt one letter, according to its frequency
- Intentionally misspell words
- Replace single words with one symbol (= code)

Polyalphabetic 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
1	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
2	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
3	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
4	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
5	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
6	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
7	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
8	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
9	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
10	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
11	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
12	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
13	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
14	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
15	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
16	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
17	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
18	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
19	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
20	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
21	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
22	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
23	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
24	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
25	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
26	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

Table: A Vigenère Square; Based on [1]

Polyalphabetic Substitution

- An example with the code word "*CIF*"

Code word	C	I	F	C	I	F	C	I	F	C
Plain text	b	l	o	c	k	c	h	a	i	n
Cipher text	D	T	T	E	S	H	J	I	N	P

Table: Encryption with Vigenère square

Polyalphabetic Substitution

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
1	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
2	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
3	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
4	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
5	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
6	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
7	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
8	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
9	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
10	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
11	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
12	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
13	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
14	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
15	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
16	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
17	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
18	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
19	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
20	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
21	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
22	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
23	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
24	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
25	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
26	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

Table: Encryption of *b*, *l* and *o* with code word "*CIF*"

Monoalphabetic vs. Polyalphabetic

Method	Plain		Cipher
Monoalphabetic	1	$\leftrightarrow$	1
Homophone	1	$\leftrightarrow$	N
Polyalphabetic	N	$\leftrightarrow$	N

- Monoalphabetic: every cipher letter stands for exactly one plain text letter
- Homophone: multiple cipher symbols can stand for one plain text letter
- Polyalphabetic: the same cipher letter can stand for multiple plain text letters and a plain text letter can be encrypted by multiple cipher letters

Breaking Polyalphabetic Substitution

- Simple frequency analysis not possible anymore
- Mid 19th century: vulnerability discovered
- Use repetition of code word as starting point
- Find length of code word and use frequency anaylsis on every x-th cipher letter

Breaking Polyalphabetic Substitution

Code word	K	I	N	G	K	I	N	G	K	I	N	G	K	I	N	G	K	I	N	G	K	I	N	G
Plain text	t	h	e	s	u	n	a	n	d	t	h	e	m	a	n	i	n	t	h	e	m	o	o	n
Cipher text	D	P	R	Y	E	V	N	T	N	B	U	K	W	I	A	O	X	B	U	K	W	W	B	T

Table: Encryption with Vigenère Square. Based on [1]

- "the" can be encrypted in four ways: DPR, BUK, GNO, ZRM
- This depends on its relative position to the code word
- "the" is encrypted twice with BUK
- Thanks to this repetition the length of the code word can be guessed

Unbreakable Cipher

- Weakness of Vigenère-Cipher:
  Repetition of code word (= key)
- Solution:
  - Length of key = length of text
  - Random key (don't use words or lists)
  - Use each key only once
- "*onetime pad cipher*" $\rightarrow$ theoretically unbreakable!

Encryption in the Age of the Computer

- Electronics are much faster than mechanical parts
- Possibility to imagine hypothetical cipher machines
- Key difference:
  Numbers vs. Letters (ASCII)

A:	0	1	0	0	0	0	0	1	**N:**	0	1	0	0	1	1	1	0
B:	0	1	0	0	0	0	1	0	**O:**	0	1	0	0	1	1	1	1
C:	0	1	0	0	0	0	1	1	**P:**	0	1	0	1	0	0	0	0
D:	0	1	0	0	0	1	0	0	**Q:**	0	1	0	1	0	0	0	1
E:	0	1	0	0	0	1	0	1	**R:**	0	1	0	1	0	0	1	0
F:	0	1	0	0	0	1	1	0	**S:**	0	1	0	1	0	0	1	1
G:	0	1	0	0	0	1	1	1	**T:**	0	1	0	1	0	1	0	0
H:	0	1	0	0	1	0	0	0	**U:**	0	1	0	1	0	1	0	1
I:	0	1	0	0	1	0	0	1	**V:**	0	1	0	1	0	1	1	0
J:	0	1	0	0	1	0	1	0	**W:**	0	1	0	1	0	1	1	1
K:	0	1	0	0	1	0	1	1	**X:**	0	1	0	1	1	0	0	0
L:	0	1	0	0	1	1	0	0	**Y:**	0	1	0	1	1	0	0	1
M:	0	1	0	0	1	1	0	1	**Z**	0	1	0	1	1	0	1	0

Table: ASCII binary letters. Based on [1]

DES: Data Encryption Standard

- Companies need a standardized approach
- Encryption Method “Lucifer” by Horst Feistel in early 1970s
- Becomes “Data Encryption Standard (DES)”
- Advantage: Standardization and security
- Disadvantage: Key distribution problem persists

References

[1] Simon Singh, *The Code Book: The Secret History of Codes and Codebreaking*, Fourth Estate London, 1999.