

4. Polyalphabetische Substitution

Im 15. Jahrhundert kam der Mathematiker Leon Battista auf die Idee, statt nur einem Geheimtextalphabet mehrere zu verwenden und während der Verschlüsselung zwischen ihnen hin und her zu springen. Im 16. Jahrhundert verbesserte der Franzose Blaise de Vigenère diese Art der Verschlüsselung noch, indem er ganze 26 Geheimtextalphabete einführte, zwischen denen gewechselt wird. Dazu verwendete er das nach ihm benannte **Vigenère-Quadrat**:

	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

Für die Ver- und Entschlüsselung braucht man ein Schlüsselwort, als Beispiel verwenden wir hier das Wort **CODE**. Dieses Schlüsselwort gibt an, welche Geheimtextalphabete verwendet werden und in welcher Reihenfolge diese benutzt werden. Zuerst wird das Alphabet in Zeile **C** verwendet, dann das in Zeile **O**, dann **D**, dann **E**. Danach geht es wieder von vorne los.

Polyalphabetische Substitution

Als Beispiel verschlüsseln wir den folgenden Ausspruch: „lieber alternativ als alt und naiv“. Wir schreiben unser Schlüsselwort so oft über den zu verschlüsselnden Text, bis jedem Buchstaben des Textes ein Buchstabe des Schlüsselworts entspricht:

```

C O D E C O D E C O D E C O D E C O D E C O D E C O D E C
l i e b e r a l t e r n a t i v a l s a l t u n d n a i v

```

Den ersten Buchstaben des Textes (l) verschlüsseln wir mit dem Alphabet C. Dort finden wir für das l den Buchstaben N. Für das i verwenden wir das Alphabet O und finden dort den Buchstaben W etc. Damit ergibt sich insgesamt:

```

Schlüssel:  C O D E C O D E C O D E C O D E C O D E C O D E C
Klartext:   l i e b e r a l t e r n a t i v a l s a l t u n d n a i v
Geheimtext: N W H F G F D P V S U R C H L Z C Z V E N H X R F B D M X

```

Diese Art der Verschlüsselung heißt **polyalphabetisch**, da mehrere Geheimtextalphabete verwendet werden. Versucht man, diese Verschlüsselung mit der Häufigkeitsanalyse zu knacken, so stößt man auf gleich mehrere Probleme:

Erstens wird in unserem Beispielsatz z. B. das e einmal als H, beim zweiten Mal als G und beim dritten Mal als S verschlüsselt, d. h., ein und derselbe Buchstabe wird mit verschiedenen Buchstaben verschlüsselt, sodass sich nicht mehr ein einzelner Buchstabe des verschlüsselten Textes besonders häuft.

Zweitens werden z. B. sowohl das b als auch das r von „lieber“ mit F verschlüsselt, sodass die Häufung eines einzelnen Buchstabens im Geheimtext nicht einmal mehr auf die Häufung eines Buchstabens im Klartext schließen lässt.

16

Du überwachst seit Tagen eine Frau, die unter Umständen eine gegnerische Agentin sein könnte. Nachdem die ersten sechs Tage nichts Ungewöhnliches passiert ist, hat sie sich heute Morgen zum ersten Mal verdächtig verhalten. Du bist dir nicht ganz sicher, ob es etwas zu bedeuten hat, aber als Agent kann man nicht vorsichtig genug sein, darum informierst du deinen Kontaktmann lieber mittels einer verschlüsselten Nachricht. Wie lautet die unten stehende Botschaft, nachdem du sie mit der Vigenère-Verschlüsselung und dem Codewort „BIGMAC“ verschlüsselt hast?

ZIELPERSON HAT HEUTE MORGEN IM MCDONALDS DIE HERRENTTOILETTE BENUTZT

17

Entschlüssele den folgenden (mit der polyalphabetischen Vigenère-Verschlüsselung, Codewort „INFO“, verschlüsselten) Geheimtext:

Tvjpm Qollx tqexh iyx Xizjg Tnxh!

18

Als Beppo, der pensionierte Clown, vorige Woche in seinem Bett an Altersschwäche starb, hatte er einen Zettel in der Hand, auf dem das Wort „CLOWN“ und darunter die unten stehende Buchstabenkette stand. Was wollte er der Nachwelt als Letztes mitteilen?

Ypf znu wwafv tgp qqzt!

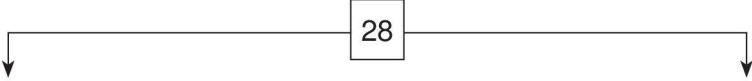
Polyalphabetische Substitution entschlüsseln

Über Jahrhunderte blieb die Vigenère-Verschlüsselung die Standardmethode, um geheime Nachrichten unleserlich zu machen. Erst im 19. Jahrhundert gelang es Charles Babbage (der übrigens vor allem für den ersten Entwurf eines modernen Computers bekannt ist), die Vigenère-Verschlüsselung zu knacken. Er hatte die Idee, nach Buchstabenfolgen zu suchen, die mehr als einmal im Geheimtext vorkommen.

Solche Wiederholungen können auf zweierlei Arten zustande kommen: Entweder wurden zwei identische Buchstabenfolgen im Klartext mit demselben Teil des Schlüssels verschlüsselt, oder aber zwei verschiedene Buchstabenfolgen im Klartext wurden mit verschiedenen Teilen des Schlüsselwortes chiffriert, wobei zufällig im Geheimtext dieselbe Buchstabenkombination entstanden ist.

Je länger die Buchstabenkombination, desto unwahrscheinlicher ist es, dass diese zufällig entstanden ist. Babbage betrachtete deshalb nur Buchstabenfolgen aus vier oder mehr Buchstaben, die mehrmals im Geheimtext auftauchten, und ging davon aus, dass diese jeweils für dieselben Klartextbuchstaben stehen mussten und mit demselben Teil des Schlüssels verschlüsselt wurden. Das tat er, weil er herausfinden wollte, wie lang das Schlüsselwort ist.

Die mehrfach vorkommenden Buchstabenkombinationen schrieb er, zusammen mit den Abständen zwischen ihnen, in eine Tabelle. Da beide Buchstabenfolgen ja mit dem gleichen Teil des Schlüssels verschlüsselt wurden, muss der Abstand zwischen den beiden identischen Abschnitten ein Vielfaches der Schlüssellänge sein. Die verschiedenen ermittelten Abstände müssen also allesamt durch die Schlüssellänge teilbar sein, deshalb bestimmte er den größten gemeinsamen Teiler der Abstände und hatte damit einen guten Kandidaten für die Schlüssellänge gefunden. Dazu als Beispiel der folgende Geheimtext:



OCUGMNCSOEWWRVNNGGCOURROLGAEH**OCUGMNC**SOEWWNVZBFXRFXJIR
 KMNRSOHLHVIQHRKBVWJFQTSHJJFTXWNRSSEAAGYSGTXRRXYSGXRGVFS
 TMDVIPFBGEEWYOEUFH**IWIA**LUAWWSYBEOIXSVGLBLSXBANZEDBNKDQMJG
 PAWNPGSSEIRKYIRUEEHJBRKIRWJSTBSPLYGPAARYRHHFDRRGITPIRJQCP
 DEAZTBFVHAIJJBGDRXWCVMFYMJUGLIR**RFQU**UUSJFZBWIRLJFMXNNFJFFB
NQJWSVNN**NQJW**CUNNQHNSTSFELWRKEZMYV**GDRV**SIAWFEEZBVFDNIRA
 RKLVGMMHFVHBRIOFNF**RVXQUTU**AYSRCFAHHJFAWAANTVAFALRFFQAEEESH
 EBTGEQZRLWVIBSVM**NBGM**GXXURVROAG**GDRVXQUTUGRFQU**OOERZBQLCU
 EZHVGDV**IWIA**WN**NBGM**REXIFWNUZBNHXJBUTLOIXHHGD

Es gibt sieben Buchstabenkombinationen mit mindestens vier Buchstaben, die mehrmals vorkommen:

OCUGMNC SOEWW (Abstand: 28)
 RVXQUTU (Abstand: 70)
 IWIA (Abstand: 322)
 RFQU (Abstand: 175)

NQJW (Abstand: 7)
 GDRV (Abstand: 105)
 NBGM (Abstand: 49)

Der größte gemeinsame Teiler aller Abstände ist die 7, deshalb besteht das Schlüsselwort mit an Sicherheit grenzender Wahrscheinlichkeit aus sieben Buchstaben. Doch was bringt uns diese Erkenntnis?

Polyalphabetische Substitution

Babbage machte sich klar, dass z. B. ein aus sieben Buchstaben bestehendes Schlüsselwort dazu führt, dass jeder siebte Buchstabe der Klartextnachricht mit demselben Geheimtextalphabet verschlüsselt wird. Er schrieb sich deshalb nur jeden siebten Buchstaben auf, und zwar einmal beginnend mit dem ersten, einmal beginnend mit dem zweiten etc., sodass er am Schluss sieben Botschaften $B_1, B_2, \dots, B_7$ hatte, die zwar entschlüsselt keinerlei Sinn mehr ergaben (da sie ja jeweils nur jeden siebten Buchstaben der Originalnachricht enthielten), dafür aber jeweils mit ein und demselben Geheimtextalphabet verschlüsselt waren. Er machte also quasi aus einer polyalphabetischen Verschlüsselung sieben monoalphabetische Verschlüsselungen. Und von diesen wusste er ja bereits, wie sie zu knacken waren: mit der Häufigkeitsanalyse.

Da die verschiedenen Geheimtextalphabete aus dem Vigenère-Quadrat allesamt nur Verschiebungen des normalen Alphabets (und nicht Neuaneordnungen der Buchstaben) sind, hatte er hier sogar noch den zusätzlichen Vorteil, dass er pro Alphabet nur einen einzigen Buchstaben richtig erraten musste und sich die anderen Buchstaben daraus zwangsläufig ergaben.

Bei längeren deutschen Texten lässt sich der häufigste Buchstabe mit einiger Sicherheit als E identifizieren. Da jedoch jeder der nun sieben Geheimtexte natürlich nur noch aus einem Siebtel der Buchstaben der Originalnachricht besteht, kann es bei kürzeren Texten durchaus passieren, dass das E nur der zweit- oder dritthäufigste Buchstabe ist. Im Beispiel oben ergibt sich die folgende Häufigkeitsverteilung für B_1 :

A: 0	B: 1	C: 0	D: 1	E: 0	F: 5	G: 2	H: 0	I: 1	J: 12	K: 0	L: 1	M: 3
N: 3	O: 2	P: 1	Q: 3	R: 4	S: 9	T: 2	U: 0	V: 0	W: 6	X: 4	Y: 5	Z: 4

Der häufigste Buchstabe ist also das J, das vermutlich für den Klartextbuchstaben E steht. Wenn aber E mit J verschlüsselt wird, so wird D mit I, C mit H, B mit G und A mit F verschlüsselt etc. Das verrät uns zusätzlich noch den ersten Buchstaben des Schlüsselworts, nämlich F, da das Geheimtextalphabet von B_1 mit F beginnt.

Obwohl die Anzahl der Buchstaben eigentlich nicht ausreicht, stimmt bei unserem Beispiel in vier der nun sieben Nachrichten der häufigste Geheimtextbuchstabe mit dem Klartextbuchstaben E überein. Bei der Häufigkeitsanalyse von B_3 hat man allerdings nicht ganz so viel Glück:

A: 6	B: 4	C: 2	D: 0	E: 6	F: 5	G: 4	H: 3	I: 0	J: 0	K: 0	L: 1	M: 1
N: 1	O: 0	P: 3	Q: 3	R: 7	S: 1	T: 5	U: 8	V: 7	W: 0	X: 0	Y: 1	Z: 1

Hier gibt es keinen deutlich häufigsten Buchstaben. U ist mit 8 Vorkommen am häufigsten vertreten, gefolgt von R und V mit jeweils 7 Vorkommen. Würde U dem Klartextbuchstaben E entsprechen, so ergäbe sich die folgende Häufigkeit für die Klartextbuchstaben:

K: 6	L: 4	M: 2	N: 0	O: 6	P: 5	Q: 4	R: 3	S: 0	T: 0	U: 0	V: 1	W: 1
X: 1	Y: 0	Z: 3	A: 3	B: 7	C: 1	D: 5	E: 8	F: 7	G: 0	H: 0	I: 1	J: 1

Man erkennt an mehreren Werten, dass das so vermutlich nicht stimmt: Im Text kommen anscheinend vier Qs vor, dafür kein einziges N, S, T und U (allesamt normalerweise recht häufige Buchstaben). Der Versuch, das E dem Geheimtextbuchstaben V zuzuordnen, scheitert ebenfalls. Für die Zuordnung des E zum R ergibt sich die – wesentlich plausiblere – folgende Häufigkeitstabelle:

N: 6	O: 4	P: 2	Q: 0	R: 6	S: 5	T: 4	U: 3	V: 0	W: 0	X: 0	Y: 1	Z: 1
A: 1	B: 0	C: 3	D: 3	E: 7	F: 1	G: 5	H: 8	I: 7	J: 0	K: 0	L: 1	M: 1

Damit lautet der dritte Buchstabe des Schlüsselworts N. Insgesamt ergibt sich als Schlüssel FONTANE. Hat man erst den Schlüssel, ist das Entschlüsseln ein Kinderspiel. Der entschlüsselte Text lautet:

JOHN MAYNARD

WER IST JOHN MAYNARD

JOHN MAYNARD WAR UNSER STEUERMANN

AUS HIELT ER BIS ER DAS UFER GEWANN

ER HAT UNS GERETTET ER TRAEGT DIE KRON

ER STARB FUER UNS UNSRE LIEBE SEIN LOHN

JOHN MAYNARD

DIE SCHWALBE FLIEGT UEBER DEN ERIESEE

GISCHT SCHAEUMT UM DEN BUG WIE FLOCKEN VON SCHNEE

VON DETROIT FLIEGT SIE NACH BUFFALO

DIE HERZEN ABER SIND FREI UND FROH

UND DIE PASSAGIERE MIT KINDERN UND FRAUN

IM DAEMMERLICHT SCHON DAS UFER SCHAUN

UND PLAUDERND AN JOHN MAYNARD HERAN

TRITT ALLES WIE WEIT NOCH STEUERMANN

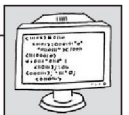
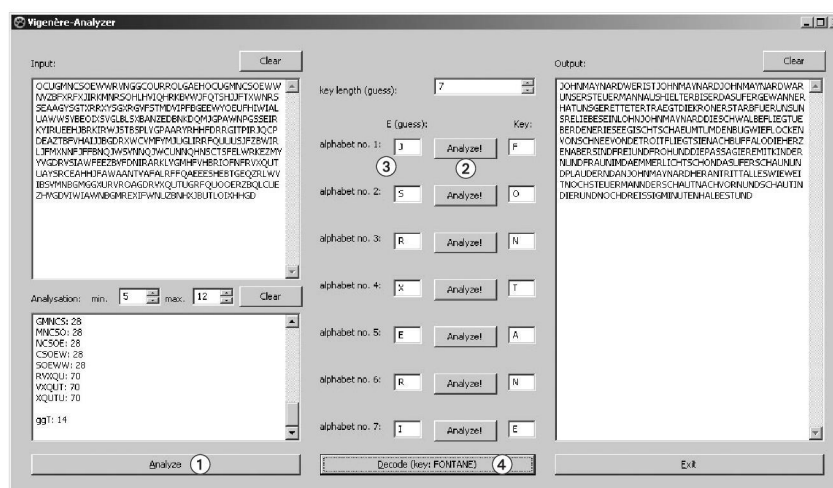
DER SCHAUT NACH VORN UND SCHAUT IN DIE RUND

NOCH DREISSIG MINUTEN HALBE STUND

19

Entschlüsse den (mit der Vigenère-Verschlüsselung verschlüsselten) Geheimtext in der Datei Aufgabe 19.txt mithilfe des Programms „Vigenère-Analyzer“ (Textdatei und Programm werden vom Lehrer bereitgestellt).

Das Programm führt genau die oben beschriebenen Schritte durch: ① Es sucht nach mehrfach vorkommenden Ketten im Geheimtext und rät die Schlüssellänge entsprechend. ② Anschließend führt es auf Knopfdruck für jedes Alphabet eine Häufigkeitsanalyse durch. ③ Wenn du dann noch für jedes Alphabet das E rätst, erscheint ein Button ④, mit dem du den Text entschlüsseln kannst. Außerdem wird das entsprechende Schlüsselwort angezeigt, sodass man einzelne falsche Buchstaben gut erkennen kann.



20

Entschlüsse den Geheimtext in der Datei Aufgabe 20.txt mithilfe des Programms „Vigenère-Analyzer“.

