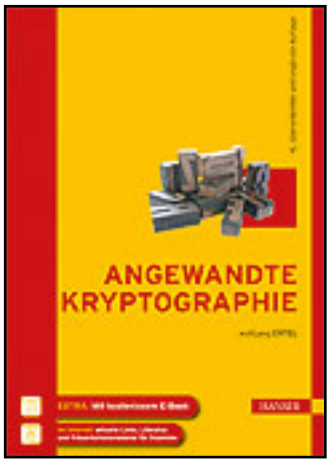
Angewandte Kryptographie

Bearbeitet von
Prof. Dr. Wolfgang Ertel

4., aktualisierte Auflage 2012. Taschenbuch. 224 S. Paperback
ISBN 978 3 446 42756 3
Format (B x L): 17,6 x 240,4 cm
Gewicht: 515 g

HANSER



Leseprobe

Wolfgang Ertel

Angewandte Kryptographie

ISBN (Buch): 978-3-446-42756-3

ISBN (E-Book): 978-3-446-43196-6

Weitere Informationen oder Bestellungen unter

http://www.hanser-fachbuch.de/978-3-446-42756-3

sowie im Buchhandel.

3 Klassische Chiffren

Als klassische Chiffren bezeichnet man alle bis etwa 1950 entwickelten und benutzten Verfahren. Von diesen werden hier nur diejenigen vorgestellt, die für das Verständnis der modernen Verfahren notwendig oder hilfreich sind. Etwas genauer gehen wir auf Technik und die Geschichte der legendären Chiffriermaschine Enigma ein. Für weitere Details sei der historisch interessierte Leser verwiesen auf [Bau00, Dew89, Kah67, Har95, Sin00].

Alle hier vorgestellten Chiffren arbeiten mit mathematischen Methoden auf einem endlichen Alphabet, denn jede Chiffriermaschine sowie auch jeder Computer kann nur eine feste endliche Menge von Zeichen darstellen. Nun ist aber zum Beispiel die Addition natürlicher Zahlen auf keiner endlichen Teilmenge abgeschlossen. Eine Chiffre, die Addition benutzt, würde manchmal undefinierte Werte liefern. Um dies zu verhindern wird, wie beim Teilen von ganzen Zahlen mit Rest, die modulare Arithmetik benutzt. Der damit nicht vertraute Leser findet im Anhang A die erforderlichen Grundlagen.

Definition 3.1

Bei einer **Transpositionschiffre** wird der Geheimtext durch eine Permutation der Klartextzeichen erzeugt. Die Zeichen bleiben gleich, tauschen aber ihre Plätze.

Bei einer **Substitutionschiffre** wird jedes Zeichen des Klartextes durch ein anderes ersetzt. Die Position bleibt jedoch gleich.

Eine Substitutionschiffre heißt **monoalphabetisch**, wenn jedes Klartextzeichen immer auf das gleiche Geheimtextzeichen abgebildet wird. Sie heißt **polyalphabetisch**, wenn sie nicht monoalphabetisch ist.

Über dem natürlichen Alphabet $(a, b, c, \ldots, z)$ gibt es $26! \approx 4 \cdot 10^{26}$ monoalphabetische Chiffren. Trotz dieser großen Zahl sind sie alle relativ leicht zu knacken. Beim Verschlüsseln von Binärdaten gibt es nur zwei monoalphabetische Chiffren, nämlich die Identität und die Negation aller Bits.

Eine einfache Klasse der Substitutionschiffren sind die Verschiebechiffren.

3.1 Verschiebechiffren

Julius Caesar (100 bis 44 v. Chr) hat seine geheimen Nachrichten mit der nach ihm benannten Chiffre verschlüsselt, indem er jedes Zeichen um drei Positionen im Alphabet verschob, was folgender Vorschrift entspricht:

Klartext:	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
Chiffretext:	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

Zur besseren Unterscheidung werden hier wie auch im Rest dieses Kapitels alphabetische Klartextzeichen mit Kleinbuchstaben und Geheimtextzeichen mit Großbuchstaben bezeichnet.

Definition 3.2

Bei einer Verschiebechiffre wird jedes Klartextzeichen z durch ein um k Zeichen im Alphabet verschobenes Zeichen ersetzt. Bei einem Alphabet mit n Zeichen seien die Zeichen durchnummeriert von 0 bis $n-1$. Dann gilt für eine Verschiebechiffre:

$$z \mapsto (z+k) \bmod n\,.$$

Verschiebechiffren sind sehr leicht zu knacken, indem man alle Verschiebungen systematisch durchprobiert, bis ein lesbarer Text entsteht. Ist das Alphabet zu groß zum Probieren, so bietet sich die statistische Analyse an. Die Häufigkeiten von Buchstaben in deutschen Texten variieren kaum. Dies benutzt man beim Abzählen der Häufigkeiten aller Buchstaben im Geheimtext, um diese dann mit den bekannten mittleren Häufigkeiten in deutschen Texten zu vergleichen (Tabelle 3.1).

TABELLE 3.1 Buchstabenhäufigkeiten der deutschen Sprache

Buchstabe	Häufigkeit [%]	Buchstabe	Häufigkeit [%]
a	6.51	n	9.78
b	1.89	o	2.51
c	3.06	p	0.79
d	5.08	q	0.02
e	17.40	r	7.00
f	1.66	s	7.27
g	3.01	t	6.15
h	4.76	u	4.35
i	7.55	v	0.67
j	0.27	w	1.89
k	1.21	x	0.03
l	3.44	y	0.04
m	2.53	z	1.13

Beispiel 3.1

Gegeben sei folgender Chiffretext:

GEIWEV LEX MQQIV YQ HVIM ZIVWGLSFIR

Die abgezählten Häufigkeiten aller Buchstaben im Chiffretext sind im oberen Teil von Tabelle 3.2 angegeben. *I* als häufigstes Zeichen führt zur Hypothese $e \mapsto I$. Für diese Hypothese sind im unteren Teil die zu erwartenden Klartexthäufigkeiten der deutschen Sprache angegeben. Man erkennt eine recht gute Übereinstimmung.

TABELLE 3.2 Abgezählte Buchstabenhäufigkeiten des Chiffretextes und die bei der Hypothese $e \mapsto I$ zu erwartenden Häufigkeiten im Vergleich

Geheimtext-Alph.:	E	F	G	H	I	L	M	Q	R	S	V	W	X	Y	Z
Häufigkeit:	3	1	2	1	5	2	2	3	1	1	4	2	1	1	1
Erwartungswerte falls $e \mapsto I$:	2.0	0.6	0.9	1.5	5.2	1.4	2.3	0.8	2.9	0.8	2.1	2.2	1.8	1.3	0.2
Klartext-Alphabet:	a	b	c	d	e	h	i	m	n	o	r	s	t	u	v

Außerdem erkennt man, dass die zehn häufigsten Buchstaben des Alphabets bei dieser Hypothese (Verschiebung um vier) alle im entschlüsselten Text vertreten sind. Der Klartext lautet:

caesar hat immer um drei verschoben

Man beachte, dass es sich hier um einen Ciphertext-Only-Angriff handelt, der noch dazu mit wenig Geheimtext auskommt. Die Datenkomplexität ist also minimal. Der Grund hierfür ist der sehr kleine Schlüsselraum mit einer Mächtigkeit von 26.

3.2 Multiplikative Chiffren

Definition 3.3

Bei einer multiplikativen Chiffre über dem Alphabet A wird jedes Klartextzeichen z mit einer Zahl $t \in \{0,\ldots,n\}$ multipliziert. t und $n = |A|$ müssen teilerfremd sein, d. h. es muss gelten $\mathrm{ggT}(t,n) = 1$. Die Chiffriervorschrift lautet

$$z \mapsto (z \cdot t) \bmod n\,.$$

Wegen $(z \cdot t) \bmod n = (z \bmod n) \cdot (t \bmod n) \bmod n$ müssen für t nur Werte aus $\{0,\ldots,n-1\}$ betrachtet werden. Größere Werte von t liefern keine neuen Chiffren.

Beispiel 3.2

Wir betrachten zuerst das Beispiel einer Chiffre mit nicht erlaubtem Wert von t. Für $n = 26$ und $t = 2$ ergibt sich die Abbildung wie folgt

Klartext	a	b	c	d	…	l	m	n	o	…
z	0	1	2	3	…	11	12	13	14	…
$2z \bmod 26$	0	2	4	6	…	22	24	0	2	…
Geheimtext	**A**	**C**	E	G	…	W	Y	**A**	**C**	…

Man sieht, dass diese Chiffre nicht injektiv ist. Chiffretexte können nicht mehr eindeutig decodiert werden. Setzen wir jedoch $t = 3$, so ändert sich die Tabelle zu

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
$3z \bmod 26$	**0**	3	6	9	12	15	18	21	24	**1**	4	7	10	13	16	19	22	25	**2**	5	8	11	14	17	20	23

Hier ist die eindeutige Entschlüsselung garantiert. Wir zeigen nun, dass dies kein Zufall ist, sondern für alle multiplikativen Chiffren (mit $\mathrm{ggT}(t, n) = 1$) gilt.

Satz 3.1

Zu jeder multiplikativen Chiffre E mit $\mathrm{ggT}(t,n) = 1$ gibt es eine multiplikative Dechiffrierfunktion D mit $D(E(z)) = z$ für alle $z \in A$.

Beweis

Wegen $\mathrm{ggT}(t,n) = 1$ gibt es nach Satz A.8 ein $b \in \mathbb{Z}_n$ mit $t \cdot b \equiv 1 \bmod n$. Dieses b ist also invers zu t modulo n und macht die Multiplikation mit t rückgängig. Also ist

$$D : z' \mapsto bz' \bmod n$$

die Dechiffrierfunktion zu E, denn es gilt

$$D(E(z)) = (b(zt) \bmod n) \bmod n = ((bt) \bmod n \cdot z) \bmod n = (1z) \bmod n = z\,.$$

Überlegen wir uns nun, wie viele multiplikative Chiffren es in $\mathbb{Z}_{26}$ gibt. Wir zählen einfach alle möglichen Schlüssel t mit $\mathrm{ggT}(t,26) = 1$ und $t < 26$ auf: 1, 3, 5, 7, 9, 11, 15, 17, 19, 21, 23, 25. Es gibt nur zwölf multiplikative Chiffren bei einem Alphabet mit 26 Zeichen. Der Schlüsselraum ist noch kleiner als bei den Verschiebechiffren. Als nächstes betrachten wir Chiffren, die Addition und Multiplikation kombinieren.

3.3 Tauschchiffren (Affine Chiffren)

Definition 3.4

Eine Chiffre E mit $z \mapsto (zt+k) \bmod n$ wird affine Chiffre oder Tauschchiffre genannt. Für die Invertierbarkeit wird $\text{ggT}(t,n) = 1$ vorausgesetzt.

Beispiel 3.3

Sei $t = 5$, $k = 7$, $n = 26$. Es ergibt sich folgende Wertetabelle:

z	0	1	2	3	4	5	6	7	8	9	10	11	12	…
$(5z+7) \bmod n$	7	12	17	22	1	6	11	16	21	0	5	10	15	…

Wie lautet die Dechiffrierfunktion?

Zuerst suchen wir eine multiplikative Inverse b zu 5 in $\mathbb{Z}_{26}$ durch systematisches Probieren. Für b muss gelten $5 \cdot b \equiv 1 \bmod 26$.[1] Gesucht ist also ein Vielfaches von 5, das bei Division durch 26 den Rest 1 ergibt. Wir bilden die Tabelle der Vielfachen von 26

i	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	…
$26i$	0	26	52	78	**104**	130	156	182	208	**234**	260	286	312	338	**364**	…

und suchen die Zahlen, die um 1 kleiner sind als ein Vielfaches von 5, also Zahlen, die mit den Ziffern 4 oder 9 enden. Wir finden 104, 234, 364, …, verwenden davon 104 und berechnen

$$5b = 105 \equiv 1 \bmod 26 \quad \Rightarrow \quad b = 21\,.$$

Nun können wir die Gleichung $z' = 5z + 7$ nach z auflösen

$$\begin{aligned} z' &= 5z + 7 \\ 21z' &= 21 \cdot 5 \cdot z + 21 \cdot 7 \equiv (z + 147) \bmod 26 = (z + 17) \bmod 26 \\ \Rightarrow \quad & \boxed{z \equiv (21z' - 17) \bmod 26 \equiv (21z' + 9) \bmod 26\,.} \end{aligned}$$

Damit kann die Dechiffrierfunktion angewendet werden, um aus z' wieder z zu berechnen (siehe Tabelle 3.3).

Es bleibt noch zu erwähnen, dass es $26 \cdot 12 = 312$ verschiedene Tauschchiffren auf dem Alphabet $\{a, b, c, \ldots, z\}$ gibt. Dies sind immer noch sehr wenige und daher sind Tauschchiffren auch sehr unsicher. Wir wenden uns nun wieder den monoalphabetischen Chiffren allgemein zu.

[1] Es gibt ein allgemeines Verfahren, basierend auf dem erweiterten Euklidischen Algorithmus zum Berechnen von Inversen in $\mathbb{Z}_n$. Dieses ist im Anhang A.3 zu finden.

TABELLE 3.3 Dechiffrierung der multiplikativen Chiffre. In der Tabelle wurde durch das Weglassen des mod-Operators eine etwas unexakte, aber übersichtlichere, Schreibweise benutzt. Jede Gleichung $a+b=x$ in der Tabelle steht für $(a+b) \bmod n = x \bmod n$.

z	$z' = (5z+7) \bmod 26$	$z = (21z'+9) \bmod 26$			
0	7	$147+9=$	$17+9=$	$26=$	0
1	12	$252+9=$	$18+9=$	$27=$	1
2	17	$357+9=$	$19+9=$	$28=$	2
3	22	$462+9=$	$20+9=$	$29=$	3
4	1	$21+9=$		$30=$	4
…	…	…			
12	$67=15$	$315+9=$	$3+9=$		12
…	…	…			

3.4 Kryptanalyse monoalphabetischer Chiffren

Gegeben sei folgender monoalphabetisch verschlüsselte Chiffretext:

```
WBO BUVLPH RZWB NHBOB EOHGYVTRQ UVRQY CD GUHRGBU
LBNBTBU QHBYYBU WVB MOVYVTRQBU GOXEYZLOHEQBU UVRQYT
DBMBO WBU IBOTRQKDBTTBKDULTEOZCBTT LBNDTTY
```

Das häufige Auftreten des Buchstaben „B“ legt ein Abzählen der Buchstabenhäufigkeiten nahe. Es ergibt sich:

Buchstabe	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
Häufigkeit	0	22	2	5	4	0	4	6	1	0	2	5	2	3	9	1	7	7	0	12	11	7	4	1	8	3

Zur Decodierung bieten sich unter Beachtung der mittleren Häufigkeiten deutscher Buchstaben (Tabelle 3.1) folgende Hypothesen an:

1. $\mathrm{B} \mapsto \mathrm{e}, \mathrm{T} \mapsto \mathrm{n}$
2. $\mathrm{B} \mapsto \mathrm{e}, \mathrm{U} \mapsto \mathrm{n}$.

Um zwischen diesen zu entscheiden, werden nun Bigramme, d. h. Paare von Buchstaben ausgezählt und mit den mittleren Häufigkeiten in deutschen Texten verglichen. Die häufigsten Paare sind in Tabelle 3.4 links dargestellt. Aufgrund der linken Tabelle ist **en** viel häufiger als **ne**. Daher wird Hypothese 2 (**U** $\mapsto$ **n**) weiterverfolgt. Das vierfache Vorkommen von **BO** legt **O** $\mapsto$ **r** nahe. Die auffällige Häufigkeit von **RQ** deutet auf **ch**. Nimmt man dann noch **V** und **T** als die beiden nach **e** und **n** häufigsten Buchstaben **i** und **s** dazu, so ergibt sich als eine neue Hypothese (neben einer weiteren):

$$\mathrm{B} \mapsto \mathrm{e}, \quad \mathrm{U} \mapsto \mathrm{n}, \quad \mathrm{O} \mapsto \mathrm{r}, \quad \mathrm{R} \mapsto \mathrm{c}, \quad \mathrm{Q} \mapsto \mathrm{h}, \quad \mathrm{T} \mapsto \mathrm{s}, \quad \mathrm{V} \mapsto \mathrm{i}.$$

Damit erhält man:

```
Wer eniLPH cZWe NHere ErHGYisch nichY CD GnHcGen
LeNesen hHeYYen Wie MriYischen GrXEYZLrHEhen nichYs
DeMer Wen IerschKDesseKDnLsErZCess LeNDssY
```

TABELLE 3.4 Die häufigsten Buchstabenpaare in deutschen Texten und im Beispiel

Mittlere Häufigkeit von Paaren in deutschen Texten

Paar	Häufigkeit
en	3.88 %
er	3.75 %
ch	2.75 %
te	2.26 %
de	2.00 %
nd	1.99 %
ei	1.88 %
ie	1.79 %
in	1.67 %
es	1.52 %

Paarhäufigkeiten im Beispiel

Paar	Häufigkeit
WB	3
HB	2
LB	2
TB	2
QB	2
DB	2
NB	1
YB	1
VB	1
MB	1
IB	1
CB	1
LB	1

Paar	Häufigkeit
BU	7
BO	4
BT	3
BN	2
BY	1
BK	1
BM	1
RQ	5

Für Maschinen wird nun die Arbeit schwierig, d. h. man benötigt einen wesentlich längeren Text, um genauere Häufigkeiten zu erhalten. Der Leser kann jedoch mit etwas Intuition und Probieren den verbleibenden Lückentext ergänzen:

```
_er eni___ c__e __ere _r___isch nich_ __ _n_c_en
_e_esen h_e__en _ie _ri_ischen _r_____r__hen nich_s
_e_er _en _ersch__esse__n_s_r__ess _e__ss_
```

3.5 Polyalphabetische Chiffren

Aus der Definition der monoalphabetischen Chiffre kann man ablesen, dass diese eine bijektive Abbildung von Klartext auf Chiffretext darstellt. Bei einer polyalphabetischen Chiffre trifft dies nicht mehr zu. Jedes Klartextzeichen kann je nach Kontext auf unterschiedliche Chiffretextzeichen abgebildet werden.

3.5.1 Homophone Chiffren

Homophone Chiffren sind solche, bei denen im Geheimtext jedes Zeichen (in etwa) gleich häufig vorkommt. Dadurch wird die Kryptanalyse erschwert, denn die Zeichen sind nicht mehr aufgrund ihrer Häufigkeit unterscheidbar. Beispielsweise könnte der Buchstabe **e** in 17 verschiedene Geheimtextzeichen verschlüsselt werden, wogegen **z** nur auf eines abgebildet wird. Wenn nun sichergestellt ist, dass zur Codierung von **e** jedes der 17 Zeichen etwa gleich oft verwendet wird, so wird die relative Häufigkeit von jedem dieser Zeichen

ca. 17.4 %/17 ≈ 1.02 % betragen. Das ist etwa gleich viel wie die Häufigkeit des einzigen Geheimtextzeichens von **z**. Damit lässt sich festhalten:

Homophone Chiffren verschleiern die Häufigkeiten der Klartextzeichen.

Obwohl man mit homophonen Chiffren nahezu eine Gleichverteilung der Häufigkeiten aller Geheimtextzeichen erreichen kann, sind Angriffe mit statistischen Methoden möglich. Auch hier nutzt man wieder das unterschiedlich häufige Vorkommen von Paaren aus. Zum Beispiel werden als unmittelbare Nachfolger der Geheimtextzeichen von **c** bestimmte Geheimtextzeichen besonders häufig auftreten, nämlich die zu den Klartextzeichen **h** und **k** gehörenden.

Es folgt daher, dass bei gleichem Alphabet homophone Chiffren schwerer zu knacken sind als monoalphabetische Chiffren, denn für die statistische Analyse von Paaren werden wesentlich mehr Daten benötigt.

3.6 Die Vigenère-Chiffre

Diese polyalphabetische Chiffre wurde im 16. Jahrhundert von dem Franzosen **Blaise de Vigenère** vorgeschlagen und basiert auf der Verschiebechiffre. Genau wie bei dieser wird jedes Zeichen im Alphabet verschoben, wobei der Betrag der Verschiebung von der Position des Zeichens im Text und dem Schlüsselwort abhängt.

3.6.1 Der Algorithmus

Sender und Empfänger vereinbaren ein Schlüsselwort und benutzen zum Ver- und Entschlüsseln das in Bild 3.1 dargestellte Vigenère-Quadrat. Das Verschlüsseln eines Zeichens sei zuerst an einem Beispiel erläutert:

Beispiel 3.4

Das Schlüsselwort sei `geheim`.

Schlüsselwort K:	`geheimgeheimgehei`
Klartext P:	`dieloesunglautetx`
Chiffretext C:	`JMLPWQYYUKTMAXLXF`

Um den Buchstaben `d` mit dem Schlüssel `g` zu codieren, sucht man einfach den Eintrag in Spalte d und Zeile g im Vigenère-Quadrat und findet als Chiffretextzeichen das `J`. Analog verfährt man mit den restlichen Klartextzeichen. Es wird also jedes Klartextzeichen verschoben, und zwar um den Betrag, welcher der Position des entsprechenden Schlüsselwortzeichens entspricht.