

Häufigkeitsanalyse

Knacken monoalphabetischer Verschlüsselungen

Einfache monoalphabetische Verschlüsselungsverfahren ordnen jedem Klartextzeichen genau ein Geheimtextzeichen zu. Ein Beispiel für eine solche Zuordnung haben Sie bereits kennengelernt:

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

Tabelle 1: Beispiel für eine monoalphabetische Verschlüsselung

Wenn man die Zuordnungstabelle nicht kennt, ist es auf den ersten Blick ziemlich schwierig den verschlüsselten Text zu lesen. Denn es würde viel zu lange dauern, die 403 Trilliarden möglichen Zuordnungen alle auszuprobieren.

Bei einem längeren verschlüsselten Text können wir aber ausnutzen, dass in jeder Sprache manche Buchstaben besonders häufig vorkommen. In der deutschen Sprache ist der häufigste Buchstabe das *e* und der zweithäufigste das *n*. Da das *e* bei einer monoalphabetischen Substitution wie in Tabelle 1 immer durch das gleiche Zeichen ersetzt wurde, wird dieses Geheimtextzeichen im Geheimtext ebenfalls das häufigste Zeichen sein. Die Häufigkeit der Klartextzeichen überträgt sich also auf die Geheimtextzeichen.

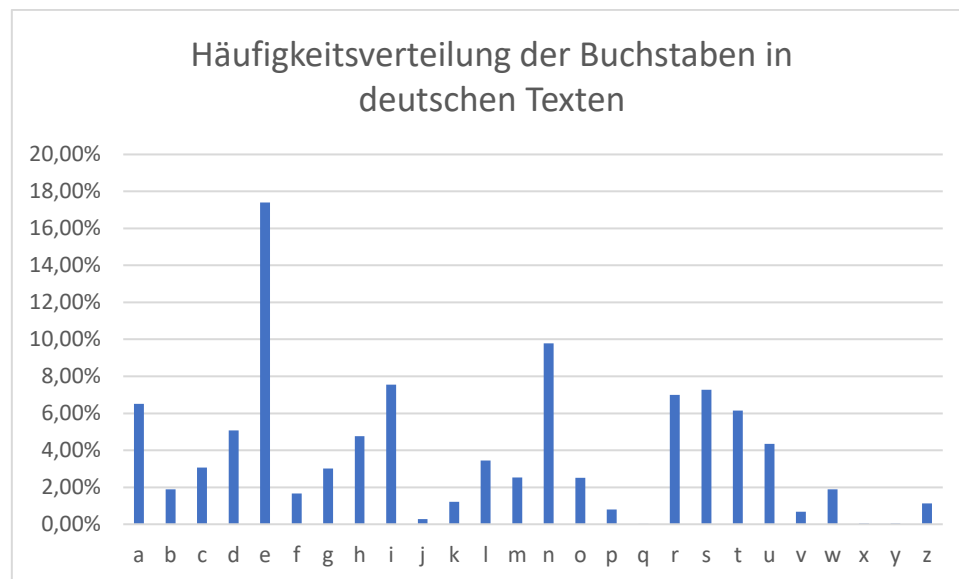


Abbildung 1: Häufigkeitsverteilung der Buchstaben in deutschen Texten¹

Aufgabe 1: Angenommen ein längerer deutscher Text wurde mit der Zuordnung in Tabelle 1 verschlüsselt. Stellen Sie eine Vermutung auf, ...

- ... welches Zeichen in dem Geheimtext am häufigsten vorkommt.
- ... welche fünf bis sechs Geheimtextzeichen sehr häufig vorkommen.
- ... welche Geheimtextzeichen kaum oder gar nicht vorkommen.

¹ Quelle: A. Beutelspacher (2009). *Kryptologie*. 9. Aufl. Wiesbaden: Vieweg+Teubner.

Indem wir das häufigste und das zweithäufigste Geheimtextzeichen finden, können wir also diese Zeichen schon mal durch das Klartextzeichen *e* bzw. das Klartextzeichen *n* ersetzen. Damit haben wir schon etwa ein Viertel des Textes entschlüsselt! Wir erhalten einen Lückentext, in dem wir die restlichen Zeichen durch intelligentes Raten zuordnen können. Dazu können wir uns zum einen die Häufigkeiten der übrigen Zeichen anschauen, um herauszufinden, ob es sich um einen eher häufigen Buchstaben wie *a*, *i*, *r*, *s* oder *t* handelt oder um einen sehr seltenen wie z. B. *q*, *x* oder *y*. Zum anderen finden wir Wörter, in denen die Lücken nur mit bestimmten Buchstaben gefüllt werden können, damit sie Sinn ergeben. In dem Wort *e?n* wird das Fragezeichen höchstwahrscheinlich für das *i* stehen. In dem Wort *?er* kann das Fragezeichen nur für die Buchstaben *d*, *h*, *p* oder *w* stehen.

Häufigkeitsanalyse mithilfe des Rechners

Aufgabe 2: Führen Sie für *geheimtext1* eine Häufigkeitsanalyse durch, indem Sie wie oben beschriebene Vorgehen, um den monoalphabetisch verschlüsselten Text zu knacken.

Hinweis zum Zählen der Zeichen: Zum Bestimmen der Häufigkeiten der einzelnen Zeichen können Sie Ihr Programm zur statistischen Auswertung von Texten verwenden.

Alternativ können Sie die Zeichen mithilfe eines Textverarbeitungsprogramms zählen, indem Sie jedes Zeichen durch sich selbst ersetzen (s. Abbildung 2).

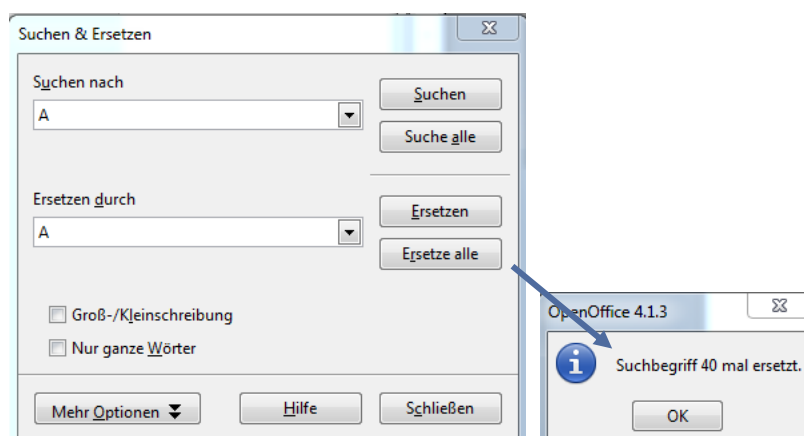


Abbildung 2: Ersetzen eines Zeichens durch sich selbst

Hinweis zum Ersetzen der Zeichen: Bei den **Geheimtextzeichen** handelt es sich um **Großbuchstaben**. Verwenden Sie für die **Klartextzeichen** **Kleinbuchstaben**, damit Sie z. B. das Geheimtextzeichen *E* nicht mit dem Klartextzeichen *e* verwechseln. Setzen Sie deshalb den Haken bei Groß-/Kleinschreibung.

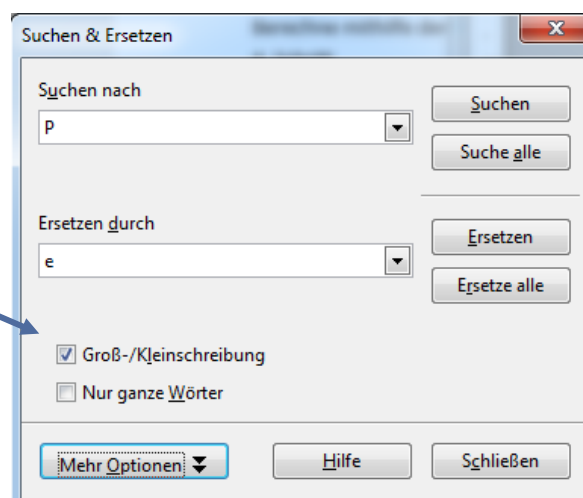


Abbildung 3: Berücksichtigung von Groß- und Kleinschreibung beim Ersetzen

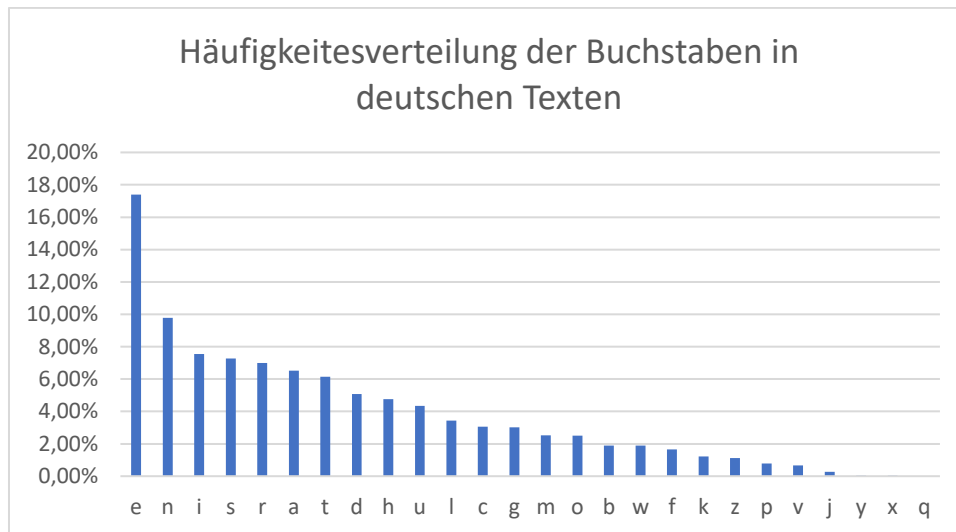


Abbildung 4: Häufigkeitesverteilung sortiert nach Häufigkeiten

Aufgabe 3 (für Schnelle):

- a) In der Datei *geheimtext2* ist ebenfalls ein monoalphabetisch verschlüsselter Text enthalten. Diesmal wurden allerdings andere Zeichen als Buchstaben für die Geheimtextzeichen verwendet. Führen Sie für diesen Text ebenfalls eine Häufigkeitsanalyse durch und versuchen Sie ihn zu knacken.

Hinweis: Die Satzzeichen Punkt, Komma, Ausrufezeichen und Fragezeichen haben ihre Bedeutung behalten und codieren keine Buchstaben.

- b) In der Datei *geheimtext3* ist ein vergleichsweise kurzer Geheimtext enthalten. Versuchen Sie auch diesen Text zu knacken.

Aufgabe 4: Erstellen Sie eine Liste von Voraussetzungen, die ein verschlüsselter Text erfüllen muss, damit er mithilfe einer Häufigkeitsanalyse geknackt werden kann.

Aufgabe 5: Diskutieren Sie, wie ein Ersetzungsverfahren sicherer gemacht werden kann, so dass ein Geheimtext nicht mehr so leicht mithilfe einer Häufigkeitsanalyse geknackt werden kann.

Lizenz

Dieses Werk ist lizenziert unter einer [Creative Commons Namensnennung - Nicht-kommerziell - Weitergabe unter gleichen Bedingungen 4.0 International Lizenz](https://creativecommons.org/licenses/by-nc-sa/4.0/). Sie erlaubt Bearbeitungen und Weiterverteilung des Werks unter Nennung meines Namens und unter gleichen Bedingungen, jedoch keinerlei kommerzielle Nutzung.

