

Die Häufigkeitsanalyse

Wie wir gesehen haben, wird in den drei Verschlüsselungsverfahren *Cäsar*, *Verschlüsselung* mit Schlüsselwort und den *Allgemeinen Substitutionschiffren* jeder Buchstabe des Klartextalphabets immer durch den gleichen Buchstaben verschlüsselt. So wird z.B. jedes e des Klartextes in genau den gleichen Buchstaben, z.B. G verschlüsselt. Solche Verschlüsselungen haben einen besonderen Namen: Man nennt sie **monoalphabetische Verschlüsselungsverfahren**.

Nun berücksichtigen wir, dass in einem deutschen Text, sofern er lang genug ist, die Buchstaben mit völlig unterschiedlichen Häufigkeiten auftreten. Das e hat z.B. – statistisch gesehen – eine mehr als 14-fach größere Häufigkeit, in einem Text aufzutreten als das k. Und die Häufigkeit der Buchstaben x und y liegt jeweils unter 0,1%. Die genaue Verteilung der Buchstaben in einem deutschen Text ist in der folgenden Tabelle angegeben.

Buchstabe	Relative Häufigkeit (in %)	Buchstabe	Relative Häufigkeit (in %)
E	17,40	M	2,53
N	9,78	O	2,51
I	7,55	B	1,89
S	7,27	W	1,89
R	7,00	F	1,66
A	6,51	K	1,21
T	6,15	Z	1,13
D	5,08	P	0,79
H	4,76	V	0,67
U	4,35	J	0,27
L	3,44	Y	0,04
C	3,06	X	0,03
G	3,01	Q	0,02

Wie können wir diese Information nun nutzen, um einen monoalphabetisch verschlüsselten Geheimtext zu knacken? Wir zählen einfach die Häufigkeiten der Buchstaben im Geheimtext und sortieren sie nach der Anzahl ihres Auftretens. Vermutlich ist dann der am häufigsten auftretende Buchstabe die Verschlüsselung des Klartextbuchstabens e. Die am zweit-, dritt- und vierthäufigsten Buchstaben des Geheimtextes stehen wahrscheinlich für einen der Buchstaben n, i, s oder r. Die restlichen Buchstaben können nun (vielleicht) aus dem Zusammenhang erschlossen werden. Hat man im Geheimtext z.B. Das Wort HWZ und vermutet man, dass Z die Verschlüsselung von e ist, so kann man versuchen, den Text HWZ durch die zu ersetzen. Natürlich muss man nun noch überprüfen, ob diese Ersetzung im restlichen Text Sinn macht oder nicht. Falls ja, kann man davon ausgehen, dass die Entschlüsselung dieser drei Buchstaben richtig war und nach weiteren solchen Anhaltspunkten im Text suchen. Vielleicht gelingt es auf diese Weise, den Geheimtext zu entschlüsseln.

Beachte: Für die Häufigkeitsanalyse ist es also nicht nötig, alle
26! = 403 291 461 126 605 635 584 000 000
Schlüssel auszuprobieren.

Aufgabe 1: Im Tausch-Verzeichnis findest du unter "knack-mich.txt" diverse Texte, die monoalphabetisch verschlüsselt worden sind. Zudem ein Tool, mit dem du die Häufigkeitsanalyse geschickt durchführen kannst. Untersuche mit Hilfe einer Häufigkeitsanalyse, welche Buchstaben des Geheimtextes besonders häufig auftreten. Versuche anschließend durch Probieren, den gesamten Geheimtext zu entschlüsseln. Versuche mindestens zwei der Texte zu entschlüsseln.

Aufgabe 2: Gegeben ist folgender Geheimtext:

FB LVA, LM LVA LBG LA LVA XPULA

- a) Führe die Häufigkeitsanalyse durch. Welche Probleme ergeben sich? Woran liegt das?
- b) Versuche trotzdem, den Geheimtext zu knacken.