Die Caesar-Verschüsselung knacken

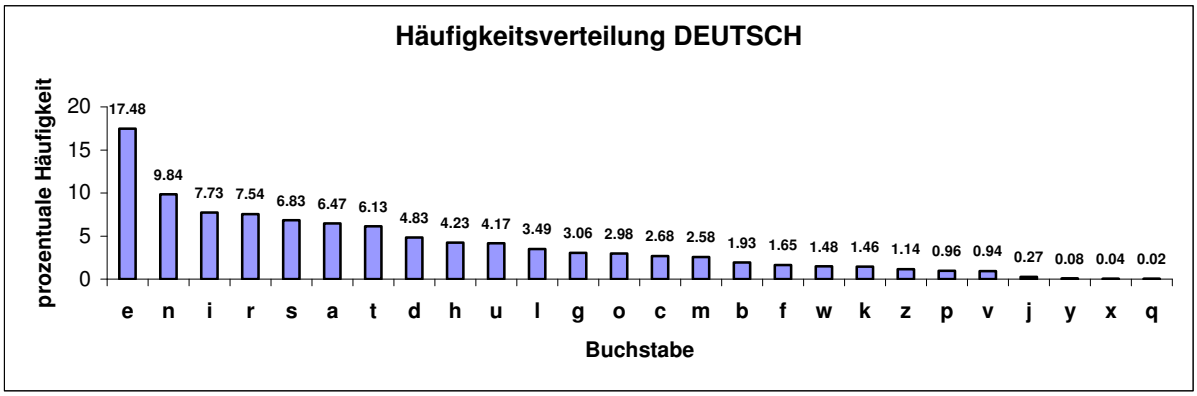
Ebenso alt wie der Wunsch den Inhalt einer Botschaft vor ungewollten Lesern zu verbergen sind die Versuche, den Inhalt von verschlüsselten Botschaften zu knacken. Zu jedem neu entwickelten Verschlüsselungsverfahren wurden sofort wieder Wege gesucht, wie mittels Kryptoanalyse die verschlüsselte Botschaft geknackt werden kann. Unter knacken verstehen wir das Herausfinden des Klartextes aus dem Geheimtext, ohne den Schlüssel zu kennen. Die Möglichkeit einen Geheimtext zu lesen, indem man sich gewaltsam oder durch List den Schlüssel verschafft, wollen wir nicht weiter untersuchen.

Das in der Folge beschriebene Verfahren eignet sich für das Knacken von Geheimtexten, die, wie bei den Caesar-Verschlüsselungen, durch monoalphabetische Verschlüsselung chiffriert wurden. Dabei besagt *monoalphabetisch*, dass bei der Chiffrierung des Geheimtextes aus dem Klartext nur ein einzelnes Geheimtextalphabet verwendet wurde.

Um eine monoalphabetisch verschlüsselte Botschaft zu entziffern müssen wir ihre Sprache kennen. Wir brauchen einen beliebigen Klartext in derselben Sprache, der lang genug ist, um ein oder zwei Blätter zu füllen. In diesem Text zählen wir, wie oft jeder Buchstabe vorkommt. Wir nennen den häufigsten Buchstaben den „ersten“, den zweithäufigsten den „zweiten“, den folgenden den „dritten“ und so weiter, bis wir alle Buchstaben im der Klartextprobe durchgezählt haben. Das Resultat dieser Analyse nennt man Häufigkeitsverteilung. Sie ist von Sprache zu Sprache verschieden.

Nun betrachten wir den Geheimtext, den wir zu entschlüsseln versuchen. Wir ordnen auch bei ihm die vorkommenden Zeichen nach deren Häufigkeit ein. Wir geben nun dem häufigsten Zeichen des Geheimtexts die Gestalt des „ersten“ Buchstabens aus unserer Klartextprobe. Dem zweithäufigsten die Gestalt des „zweiten“ usw. Dieses Vorgehen nennt man **Häufigkeitsanalyse.**

In den wenigsten Fällen wird dieses rein mechanische Vorgehen zur direkten Entschlüsselung des Geheimtextes führen, da die einzelnen Buchstaben sich in der Häufigkeit des Vorkommens oft nur wenig unterscheiden. Es gibt aber Buchstabenhäufigkeiten, die sich stark von der Häufigkeit anderer Buchstaben unterscheiden (vgl. Tabelle). So sind in der deutschen Sprache die Buchstaben E (17.48%), N (9.84%), I (7.73%), R (7.54%) und S (6.83%) die häufigsten. Es ist damit zu erwarten, dass die 5 häufigsten vorkommenden Zeichen im Geheimtext diesen Buchstaben zuzuordnen sind. Das mit Abstand am häufigsten erscheinende Zeichen wird mit grosser Wahrscheinlichkeit zum Klartextbuchstaben E gehören. Die nächsten vier lassen sich mit etwas Probieren sicher bald korrekt zuordnen. Damit kann der Geheimtext entschlüsselt werden.


Häufigkeitsverteilung DEUTSCH
prozentuale Häufigkeit
20
15
10
5
0
17.48
9.84
7.73
7.54
6.83
6.47
6.13
4.83
4.23
4.17
3.49
3.06
2.98
2.68
2.58
1.93
1.65
1.48
1.46
1.14
0.96
0.94
0.27
0.08
0.04
0.02
e n i r s a t d h u l g o c m b f w k z p v j y x q
Buchstabe