
(... die natürlich in der Praxis kombiniert werden.)

---

0. Ansatz: Exhaustion

Anwendbar, wenn die Quelle des Schlüssels (z. B. ein bestimmtes Buch) bekannt ist. Hat es die Gesamtlänge *q*, so sind bei einem Geheimtext der Länge *r* nur *q*-*r* Anfangsstellen durchzuprobieren.

---

1. Ansatz: Wahrscheinliches Wort

Im Beispiel: Das Wort `MORGEN`, über den Geheimtext geschoben und an jeder Stelle probeweise subtrahiert, ergibt ab Position 9 als Ergebnis `efreit`. Das klingt nach »`befreit`« oder »`gefreit`«. Im ersten Fall wäre der davorstehende Klartextbuchstabe `e`, im zweiten `z`. Das erste klingt plausibler und lässt schon das richtige Zitat anklingen.

Als mögliches Hilfsmittel, um das Wort »`befreit`« aufzustöbern, könnte eine vollständige Textsuche in einer Literatursammlung dienen.

---

2. Ansatz: Häufige Wortteile

Hier ist auch nützlich, dass sowohl Klartext als auch Schlüssel sinnvolle Texte ergeben sollen. Man kann also erwarten, wenn man Wörter oder Wortteile wie

```
EIN ICH NDE DIE UND DER CHE END GEN SCH UNG DAS
CHEN HEIT KEIT SICH ABER WIRD SIND ODER AUCH NACH ALSO DOCH EINS
```

über den Geheimtext schiebt, an vielen Stellen Differenzen zu finden, die plausible Trigramme, Tetragramme, ... sind und sich zu sinnvollen Texten ergänzen lassen.

Was davon zum Klartext und was zum Schlüssel gehört, muss natürlich Stück für Stück zusammengepuzzelt werden (im »Zickzack-Verfahren«).

Nützlich ist auch, sich typische Kombinationen wie

```
EIN  + EIN  = IQA
DER  + DER  = GII
HEIT + HEIT = OIQM
```

oder im Englischen

```
THE  + THE  = MOI
```

zu merken, die mit recht hoher Sicherheit nur auf diese Weise entstanden sind.

---

3. Ansatz: Häufigkeitsanalyse

Sind p_0, p_1, ... p_{n-1} die Buchstabenhäufigkeiten der (stochastischen) Sprache *M*, so haben Lauftext-Chiffrate die typischen Häufigkeiten

$$q_h = \Sigma_{i+j=h}\, p_i \cdot p_j \quad \text{für } 0 \le h \le n\text{-}1.$$

Die Häufigkeitsverteilung ist etwas abgeflacht, gibt aber doch einen Hinweis auf die Art der Verschlüsselung.

Beispiel: Deutsch (Angaben in Prozent).

```
 A   B   C   D   E   F   G   H   I   J   K   L   M
4.2 2.6 2.3 2.4 5.0 3.7 3.7 4.3 5.8 2.9 3.7 4.4 4.9

 N   O   P   Q   R   S   T   U   V   W   X   Y   Z
3.2 3.0 3.1 3.3 5.7 3.4 3.2 3.4 5.9 4.5 3.9 3.9 3.6
```

Koinzidenzindex = 0.0411.

Beispiel: Englisch (Angaben in Prozent).

```
 A   B   C   D   E   F   G   H   I   J   K   L   M
4.3 3.5 3.2 2.5 4.7 3.8 4.4 4.4 4.8 2.9 3.5 4.5 4.3

 N   O   P   Q   R   S   T   U   V   W   X   Y   Z
3.1 3.2 3.6 3.0 4.4 4.5 4.0 3.2 4.9 4.7 3.8 3.3 3.5
```

Koinzidenzindex = 0.0400.

Jeder Geheimtextbuchstabe kann auf 26 mögliche Weisen zustande gekommen sein, deren Wahrscheinlichkeit man aus den Buchstabenhäufigkeiten der Klartextsprache bestimmen kann. Auch diese Verteilung ist nicht gleichmäßig und gibt somit Hinweise auf den Klartext. Noch besser ist es, hier auch die Wahrscheinlichkeiten von Bi- und Trigrammen zu betrachten - ein Bigramm im Geheimtext kann auf $26^2 = 676$ verschiedene Weisen zustande gekommen sein, die aber völlig unterschiedliche Wahrscheinlichkeit haben, ein Trigramm sogar auf $26^3 = 17576$ Weisen, von denen sehr viele die Wahrscheinlichkeit 0 haben und daher von vornherein ausscheiden. Eine systematische Beschreibung dieses Ansatzes ist enthalten in dem Artikel

> Craig BAUER, Christian N. S. TATE: A statistical attack on the running key cipher. Cryptologia 26 (2002), 274 - 282.

---

4. Ansatz: Häufige Buchstabenkombinationen

Die Häufigkeitsanalyse (3. Ansatz) ist - zumindest bei manueller Auswertung - nicht besonders ergiebig. FRIEDMAN hat daraus aber ein systematisches Vorgehen entwickelt, das ohne bekannten Klartext auskommt und im nächsten Abschnitt behandelt wird.

Ein Programm zur vollautomatisierten Durchführung dieser Analyse (für die byteweise XOR-Verschlüsselung von ASCII-Texten) wird beschrieben in dem Artikel.

> Ed Dawson, Lauren Nielsen: Automated cryptanalysis of XOR plaintext strings. Cryptologia 20 (1996), 165 - 181.

Autor: Klaus Pommerening, 14. Januar 2000; letzte Änderung: 17. Dezember 2007.