Theoretische Sicherheit

Klaus Pommerening
Fachbereich Mathematik
der Johannes-Gutenberg-Universität
Saarstraße 21
D-55099 Mainz

Vorlesung Kryptologie
6. Februar 2000, letzte Änderung: 6. Februar 2005

Die in diesem Abschnitt vorgestellte Theorie geht auf Claude Shannon zurück [8] (mit späteren Vereinfachungen von Hellman [2]). Shannon stellte das erste allgemeine mathematische Modell der Kryptologie auf und entwickelte die Analyse von Chiffren (Kryptosystemen) mit informationstheoretischen Methoden. Die grundlegende Frage dieser Theorie ist:

> *Wieviel Information über den Klartext ist im Geheimtext enthalten?*

(ohne Rücksicht darauf, wie schwierig oder aufwendig die Gewinnung dieser Information ist). Reicht die Information nicht aus, um den Klartext zu bestimmen, wird die Chiffre als sicher betrachtet.

Shannon baute hierbei auf die Informationstheorie auf, die er zuvor [7] entwickelt hatte.

9.1 A-priori- und a-posteriori-Wahrscheinlichkeiten

Modell-Situation

Betrachtet werden:

- eine endliche Menge $M_0 \subseteq M$ von möglichen Klartexten – z. B. alle Klartexte der Länge r oder der Länge $\leq r$,
- eine endliche Menge K von Schlüsseln
- und eine Chiffre $F = (f_k)_{k \in K}$ mit $f_k \colon M \longrightarrow \Sigma^*$.

Die Beschränkung auf eine endliche Menge M_0 ermöglicht den naiven Umgang mit Wahrscheinlichkeiten und ist keine echte Einschränkung, da Klartexte der Länge $> 10^{100}$ in diesem Universum mit seinen höchstens 10^{80} Elementarteilchen extrem unwahrscheinlich sind.

Motivierendes Beispiel

Für deutsche Texte der Länge 5 kennen wir potenziell ziemlich genaue, etwa durch Auszählung gewonnene, „a-priori-Wahrscheinlichkeiten“. Ein kleiner Ausschnitt davon ist

Klartext	Wahrscheinlichkeit
hallo	$p > 0$
bauer	$q > 0$
xykph	0
...	...

Nun liege der monoalphabetisch verschlüsselte deutsche Text `XTJJA` vor. Ohne Kenntnis des Schlüssels – d. h., solange noch alle Schlüssel gleich wahrscheinlich sind – und ohne weitere Kontext-Informationen können wir den Klartexten dennoch „a-posteriori-Wahrscheinlichkeiten“ zuordnen.

Klartext	Wahrscheinlichkeit
hallo	$p_1 >> p$
bauer	0
xykph	0
...	...

Das bedeutet, dass sich alleine durch die Kenntnis des Geheimtextes (und des Verschlüsselungsverfahrens) die Information über den Klartext geändert hat.

Diese Situation wird jetzt allgemein mit einem „Bayesschen“ Ansatz modelliert.

Modell

Wahrscheinlichkeit von Klartexten. Gegeben ist eine Funktion

$$P\colon M_0 \longrightarrow [0,1] \quad \text{mit} \quad P(a) > 0 \quad \text{für alle } a \in M_0 \quad \text{und} \quad \sum_{a \in M_0} P(a) = 1.$$

(Diese soll die a-priori-Wahrscheinlichkeiten von Klartexten beschreiben.)

Wahrscheinlichkeit von Schlüsseln. Ebenso ist eine Funktion (ohne Verwechslungsgefahr gleich bezeichnet)

$$P\colon K \longrightarrow [0,1] \quad \text{mit} \quad \sum_{k \in K} P(k) = 1$$

gegeben. Hier nimmt man meist die Gleichverteilung an, d. h. $P(k) = 1/\#K$ für alle $k \in K$.

Wahrscheinlichkeit von Geheimtexten. Dadurch ist auch eine Wahrscheinlichkeit für Geheimtexte festgelegt (wobei implizit die Annahme eingeht, dass Schlüssel unabhängig von Klartexten gewählt werden):

$$P\colon \Sigma^* \longrightarrow [0,1], \quad P(c) := \sum_{a \in M_0} \sum_{k \in K_{ac}} P(a) \cdot P(k),$$

wobei $K_{ac} := \{k \in K \mid f_k(a) = c\}$ die Menge aller Schlüssel ist, die a auf c abbilden.

Bemerkungen

1. Es gibt nur endlich viele $c \in \Sigma^*$ mit $P(c) \neq 0$; diese bilden die Menge

   $$C_0 := \{c \in \Sigma^* \mid P(c) > 0\}$$

   der „möglichen Geheimtexte“.

2. Es gilt

   $$\begin{aligned} \sum_{c \in \Sigma^*} P(c) &= \sum_{c \in \Sigma^*} \sum_{a \in M_0} \sum_{k \in K_{ac}} P(a) \cdot P(k) \\ &= \sum_{a \in M_0} \sum_{k \in K} P(a) \cdot P(k) \\ &= \sum_{a \in M_0} P(a) \cdot \sum_{k \in K} P(k) \\ &= 1. \end{aligned}$$

Bedingte Wahrscheinlichkeit von Geheimtexten. Die „bedingte“ Wahrscheinlichkeit, dass ein Geheimtext aus einem bestimmten Klartext $a \in M_0$ entsteht, modelliert man durch die Funktion

$$P(\bullet|a)\colon \Sigma^* \longrightarrow [0,1], \quad P(c|a) := \sum_{k \in K_{ac}} P(k).$$

Gesprochen wird das als die „Wahrscheinlichkeit für c unter der Voraussetzung, dass a vorliegt“, oder kurz „... gegeben a“

Bemerkungen

3. $\sum_{c \in \Sigma^*} P(c|a) = \sum_{k \in K} P(k) = 1$.
4. $P(c) = \sum_{a \in M_0} P(a) \cdot P(c|a)$.

A-posteriori-Wahrscheinlichkeit von Klartexten

Der Kryptoanalytiker interessiert sich vor allem für die umgekehrte, die bedingte Wahrscheinlichkeit $P(a|c)$ für einen Klartext $a \in M_0$ bei vorliegendem Geheimtext $c \in \Sigma^*$.

Zunächst wird die Wahrscheinlichkeit für das gemeinsame Auftreten von a und c beschrieben durch

$$P\colon M_0 \times \Sigma^* \longrightarrow [0,1], \quad P(a,c) := P(a) \cdot P(c|a).$$

Bemerkungen

5. Dann ist
$$\sum_{a \in M_0} P(a,c) = \sum_{a \in M_0} P(a) \cdot P(c|a) = P(c).$$

Bedingte Wahrscheinlichkeit von Klartexten. Man definiert nun $P(\bullet|c)$ so, dass auch $P(a,c) = P(c) \cdot P(a|c)$, nämlich durch die Bayessche Formel

$$P(a|c) := \frac{P(a) \cdot P(c|a)}{P(c)}, \quad \text{falls } P(c) \neq 0,$$

und ergänzend

$$P(a|c) := 0, \quad \text{falls } P(c) = 0.$$

Bemerkungen

6. $\sum_{c \in \Sigma^*} P(c) \cdot P(a|c) = \sum_{c \in \Sigma^*} P(a) \cdot P(c|a) = P(a)$ nach Bemerkung 3.

9.2 Perfekte Sicherheit

Definition 1. Die Chiffre F heißt auf M_0 **perfekt sicher**, wenn für alle Geheimtexte $c \in \Sigma^*$ mit positiver Wahrscheinlichkeit $P(c) > 0$ gilt: $P(\bullet, c) = P$.

Deutung: Das bedeutet, dass die a-posteriori-Wahrscheinlichkeit $P(a|c)$ sich nicht von der a-priori-Wahrscheinlichkeit $P(a)$ eines Klartextes $a \in M_0$ unterscheidet. Oder, anders ausgedrückt, der Kryptoanalytiker kann aus der Kenntnis des Geheimtextes keine zusätzliche Information über den Klartext gewinnen.

Hilfssatz 1 $\#M_0 \leq \#C_0$.

Beweis. Sei $l \in K$ ein fester Schlüssel mit $P(l) > 0$. Für einen Geheimtext $c \in f_l(M_0)$, etwa $c = f_l(b)$, gilt

$$P(c) = \sum_{a \in M_0} P(a) \cdot \sum_{k \in K_{ac}} P(k) \geq P(b) \cdot P(l) > 0.$$

Also ist $c \in C_0$. Es folgt $f_l(M_0) \subseteq C_0$, und da f_l injektiv ist, $\#M_0 \leq \#C_0$. $\diamond$

Hilfssatz 2 *Ist F perfekt sicher, so $K_{ac} \neq \emptyset$ für alle $a \in M_0$ und alle $c \in C_0$.*

Beweis. Wenn $K_{ac} = \emptyset$, ist

$$P(c|a) = \sum_{k \in K_{ac}} P(k) = 0.$$

also ist $P(a|c) = 0 \neq P(a)$, Widerspruch. $\diamond$

Es kann also jeder mögliche Klartext in jeden möglichen Geheimtext verwandelt werden. Der nächste Hilfssatz sagt, dass es *sehr* viele Schlüssel geben muss.

Hilfssatz 3 *Ist F perfekt sicher, so $\#K \geq \#C_0$.*

Beweis. Da $\sum P(a) = 1$, muss $M_0 \neq \emptyset$ sein. Sei also $a \in M_0$. Wäre $\#K < \#C_0$, so gäbe es ein $c \in C_0$ mit $f_k(a) \neq c$ für alle Schlüssel $k \in K$, also $K_{ac} = \emptyset$, Widerspruch. $\diamond$

Satz 1 [Shannon] *Sei F perfekt sicher. Dann ist*

$$\#K \geq \#M_0,$$

d. h., es gibt mindestens so viele Schlüssel wie mögliche Klartexte.

Beweis. Das folgt unmittelbar aus den Hilfssätzen 1 und 3. $\diamond$

Satz 2 [SHANNON] *Sei F eine Chiffre mit*

$$P(k) = \frac{1}{\#K} \quad \textit{für alle } k \in K,$$

d. h., alle Schlüssel sind gleich wahrscheinlich, und

$$\#K_{ac} = s \quad \textit{für alle } a \in M_0 \textit{ und alle } c \in C_0.$$

mit festem $s \geq 1$. Dann ist F perfekt sicher. Ferner ist $\#K = s \cdot \#C_0$.

Beweis. Sei $c \in C_0$ ein möglicher Geheimtext. Dann gilt für einen beliebigen möglichen Klartext $a \in M_0$:

$$\begin{aligned} P(c|a) &= \sum_{k \in K_{ac}} \frac{1}{\#K} = \frac{\#K_{ac}}{\#K} = \frac{s}{\#K}, \\ P(c) &= \sum_{a \in M_0} P(a) \cdot P(c|a) = \frac{s}{\#K} \cdot \sum_{a \in M_0} P(a) = \frac{s}{\#K} = P(c|a), \\ P(a|c) &= \frac{P(c|a)}{P(c)} \cdot P(a) = P(a). \end{aligned}$$

Also ist F perfekt sicher. Der Zusatz folgt, weil

$$K = \dot{\bigcup_{c \in C_0}} K_{ac}$$

für jedes $a \in M_0$. $\diamond$

9.3 Beispiele für perfekte Sicherheit

Trivialbeispiele

Beispiel 0: $\#M_0 = 1$

Dieses Beispiel ist natürlich kryptologisch unsinnig, da der Kryptoanalytiker den einzig möglichen Klartext von vornherein kennt. Also kann er durch den Geheimtext keine zusätzliche Information über den Klartext gewinnen.

Sei $M_0 = \{a\}$. Da für alle $c \in C_0$ trivialerweise $P(a|c) = 1 = P(a)$ gilt, ist F, wie immer es auch definiert ist, perfekt sicher.

Beispiel 1: $\#M_0 = 2$

Das kleinste sinnvolle Beispiel beinhaltet zwei mögliche Klartexte. Sei (o. B. d. A.) $M_0 = \{0, 1\} = C_0 = K$. Sei f_0 die identische Abbildung auf $\{0, 1\}$ und f_1 die Vertauschung von 0 und 1. Ferner seien die beiden Schlüssel 0 und 1 gleichwahrscheinlich: $P(0) = P(1) = \frac{1}{2}$.

Dann ist $K_{00} = K_{11} = \{0\}$, $K_{01} = K_{10} = \{1\}$. Daher ist F nach Satz 2 perfekt sicher.

Die Verschiebechiffre

Hier ist $M_0 = K = C_0$ eine Gruppe und $F\colon M_0 \times K \longrightarrow C_0$ die Gruppenoperation, also $f_k(a) = a * k$. Die Mengen

$$K_{ac} = \{k \in K \mid a * k = c\} = \{a^{-1} * c\}$$

sind alle einelementig. Wird wie üblich $P(k) = \frac{1}{\#K}$ für alle Schlüssel $k \in K$ angenommen, so ist F perfekt sicher.

Die Beispiele 0 und 1 oben waren die Spezialfälle der ein- und zweielementigen Gruppe. Weitere Spezialfälle folgen als Beispiele 2 und 3.

Beispiel 2: Die CAESAR-Chiffre

Das ist die Verschiebechiffre auf der zyklischen Gruppe $\Sigma = \mathbb{Z}/n\mathbb{Z}$ der Ordnung n.

Also ist die CAESAR-Chiffre perfekt sicher, *sofern nur Nachrichten der Länge 1 verschlüsselt werden und der Schlüssel für jede Nachricht zufällig neu gewählt wird.*

Beispiel 3: Die VERNAM-Chiffre

Das ist die Vereinigung der Verschiebechiffren auf den Gruppen $\Sigma^r = M_0$ mit $\Sigma = \mathbb{Z}/n\mathbb{Z}$. Nachrichten sind also jeweils Texte der Länge r, und Schlüssel sind *zufällig gewählte* Buchstabenfolge der gleichen Länge r.

Da man insbesondere den Schlüssel für jede Nachricht neu wählen muss, heisst diese Chiffre auch **One Time Pad**. Man stellt sich einen Abreisskalender vor: Jedes Blatt enthält einen zufälligen Buchstaben und wird nach Verwendung abgerissen und vernichtet.

Die VERNAM-*Chiffre ist der Prototyp einer perfekten Chiffre.*

Im Spezialfall $\Sigma = \{0, 1\}$ erhält man die binäre VERNAM-Chiffre, die Bitstrom-Verschlüsselung mit völlig zufälliger Schlüsselbitfolge.

Gegenbeispiel: Die monoalphabetische Chiffre

Hier wird $M_0 = \Sigma^r$ gewählt und $K = \mathcal{S}(\Sigma)$. Etwa für $r = 5$ haben wir schon gesehen, dass

$$P(\texttt{bauer}|\texttt{XTJJA}) = 0 < q = P(\texttt{bauer}).$$

Die monoalphabetische Chiffre ist also (für $r \geq 2$ und $n \geq 2$) nicht perfekt. Für $r = 1$ ist sie dagegen nach Satz 2 (mit $s = (n-1)!$) perfekt.

9.4 Dichte und Redundanz einer Sprache

SHANNONs Theorie bietet nicht nur durch den Begriff der Perfektheit eine Vorstellung von einer unbrechbaren Chiffre, sondern mit der „Eindeutigkeitsdistanz“ auch ein Maß für die Nähe zur Perfektheit. Dieser Begriff greift die Erfahrung auf: Je länger ein Geheimtext ist, desto leichter ist er eindeutig zu entschlüsseln. Die Theorie wird hier nicht mathematisch exakt vorgestellt; es soll nur ein Eindruck vermittelt werden. Eine mathematisch befriedigendere Theorie der Eindeutigkeitsdistanz wird in [5] entwickelt.

Eindeutige Lösung der Verschiebechiffre

Der Geheimtext `FDHVDU` sei der Beginn einer Nachricht, die mit einer CAESAR-Chiffre erzeugt wurde. Die Methode der Exhaustion bestand darin, alle 26 möglichen Schlüssel der Reihe nach anzuwenden:

Schlüssel	Klartext	$t=1$	$t=2$	$t=3$	$t=4$	$t=5$	$t=6$
0	`fdhvdu`	+					
1	`ecguct`	+	+				
2	`dbftbs`	+					
3	`caesar`	+	+	+	+	+	+
4	`bzdrzq`	+					
5	`aycqyp`	+	?				
6	`zxbpxo`	+					
7	`ywaown`	?					
8	`xvznvm`	?					
9	`wuymul`	+	+				
10	`vtxltk`	+					
11	`uswksj`	+	+	?	?		
12	`trvjri`	+	+				
13	`squiqh`	+	?	?	?		
14	`rpthpg`	+					
15	`qosgof`	+					
16	`pnrfne`	+	+				
17	`omqemd`	+	+				
18	`nlpdlc`	+					
19	`mkockb`	+					
20	`ljnbja`	+	?				
21	`kimaiz`	+	+	+	?	?	
22	`jhlzhy`	+					
23	`igkygx`	+	+				
24	`hfjxfw`	+					
25	`geiwev`	+	+	+	?		

In der Tabelle bedeutet

- +: Bis zum t-ten Buchstaben ist der Klartext noch sinnvoll.
- ?: Bis zum t-ten Buchstaben ist der Klartext mit geringer Wahrscheinlichkeit sinnvoll.

Schon ab dem vierten Buchstaben ist mit hoher Wahrscheinlichkeit nur noch einer der getesteten Klartexte möglich. Diesen Wert 4 würde man als „Eindeutigkeitsdistanz“ der Chiffre ansehen.

Mathematisches Modell

Wir starten wieder mit unserem n-buchstabigen Alphabet Σ. Der „Informationsgehalt“ eines Buchstabens ist dann ${}^2\log n$, d. h., man braucht $\lceil {}^2\log n \rceil$ Bits, um Σ binär zu codieren.

Beispiel. Für $n = 26$ ist ${}^2\log n \approx 4.7$, man braucht 5 Bits, um alle Buchstaben zu codieren. Eine solche Codierung ist z. B. der Fernschreibercode.

Sei nun $M \subseteq \Sigma^*$ eine Sprache; $M_r = M \cap \Sigma^r$ ist dann die Menge der „sinnvollen“ Texte der Länge r, $\Sigma^r - M_r$ die Menge der „sinnlosen“ Texte. Die Anzahl der ersteren wird mit

$$t_r := \# M_r$$

bezeichnet. Dann ist ${}^2\log t_r$ der „Informationsgehalt“ eines Textes der Länge r oder die **Entropie** von M_r – so viele Bits braucht man, um die Elemente von M_r in einer binären Codierung unterscheiden zu können.

> **Anmerkung.** Die Entropie wird allgemeiner für ein Modell definiert, wo die Elemente von M_r mit Wahrscheinlichkeiten gewichtet sind. Hier wurde implizit die Gleichverteilung angenommen.

Die relative Häufigkeit sinnvoller Texte, t_r/n^r, interessiert im Moment nicht so sehr wie der **relative Informationsgehalt**,

$$\frac{{}^2\log t_r}{r \cdot {}^2\log n} :$$

Für die Codierung von Σ^r braucht man $r \cdot {}^2\log n$ Bits, für die von M_r nur ${}^2\log t_r$. Der relative Informationsgehalt gibt also den Faktor an, auf den man die Codierung von M_r im Vergleich zu Σ^r komprimieren kann; der komplementäre Anteil

$$1 - \frac{{}^2\log t_r}{r \cdot {}^2\log n}$$

ist „redundant“.

Man bezieht diese Größen üblicherweise auf ${}^2\log n$ statt auf 1 und definiert:

Definition 2. (i) Der Quotient

$$\rho_r(M) := \frac{^2\log t_r}{r}$$

heißt r**-te Dichte**, die Differenz $\delta_r(M) := {}^2\log n - \rho_r(M)$ heißt r**-te Redundanz** der Sprache M.

(ii) Existiert $\rho(M) = \lim_{r\to\infty} \rho_r(M)$, so heißt $\rho(M)$ **Dichte** von M, $\delta(M) = {}^2\log n - \rho(M)$ **Redundanz** von M.

Bemerkungen

1. Es ist $0 \le t_r \le n^r$, also $\overline{\lim}\, \rho_r(M) \le {}^2\log n$.
2. Falls $M_r \neq \emptyset$, ist $t_r \ge 1$, also $\rho_r(M) \ge 0$. Falls $M_r \neq \emptyset$ für fast alle r, ist $\underline{\lim}\, \rho_r(M) \ge 0$.
3. Existiert $\rho(M)$, so ist $t_r \approx 2^{r\rho(M)}$ für große r.

Für natürliche Sprachen ist $\rho_r(M)$ – aus empirischen Beobachtungen geschlossen – im wesentlichen monoton fallend, und somit existieren Dichte und Redundanz; ferner ist stets $t_r \ge 2^{r\rho(M)}$. Empirische Werte (mit $n = 26$) sind

M	$\rho(M) \approx$	$\delta(M) \approx$
Englisch	1.5	3.2
Deutsch	1.4	3.3

Die Redundanz der deutschen Sprache entspricht $\frac{3.3}{4.7} \approx 70\%$ [4]; man kann erwarten, dass sich deutscher Text (in den 26 Buchstaben aufgeschrieben) um diese Rate komprimieren lässt. Die entsprechende Rate für Englisch ist $\frac{3.2}{4.7} \approx 68\%$ (nach [1] jedoch 78%; siehe dazu auch [4]).

9.5 Die Eindeutigkeitsdistanz

Diese Erkenntnisse über die Redundanz werden nun auf eine vollständige Schlüsselsuche angewendet – der Aufwand wird dabei nicht berücksichtigt, nur die Durchführbarkeit. Die Herleitung folgt dem vereinfachten Ansatz von HELLMAN.

Annahmen

1. Alle sinnvollen Text der Länge r sind gleich wahrscheinlich. [Sonst werden die Formeln komplizierter; für natürliche Sprachen folgt diese Annahme für genügend große r aus den üblichen stochastischen Annahmen.]

2. Die Dichte $\rho(M)$ der Sprache M existiert. [Sonst kann man nur eine Schranke herleiten.]

3. Alle Schlüssel $k \in K$ sind gleichwahrscheinlich; es gebe $h = \#K$ Stück.

4. Alle Verschlüsselungsfunktionen f_k für $k \in K$ sind längentreu, d. h., $f(M_r) \subseteq \Sigma^r$.

Sei nun ein Geheimtext $c \in \Sigma^r$ gegeben. Dazu gibt es (im allgemeinen – falls alle Verschlüsselungsfunktionen f_k verschieden sind) h mögliche Klartexte der Länge r in Σ^r. Darunter sind längst nicht alle sinnvoll, sondern nur etwa

$$h \cdot \frac{t_r}{n^r} \approx \frac{h \cdot 2^{r\rho(M)}}{2^{r \cdot {}^2\log n}} = h \cdot 2^{-r\delta(M)}.$$

Eindeutige Dechiffrierbarkeit in M_r kann man erwarten, wenn

$$h \cdot 2^{-r\delta(M)} \leq 1, \quad {}^2\log h - r\delta(M) \leq 0, \quad r \geq \frac{{}^2\log h}{\delta(M)},$$

falls alle Verschlüsselungsfunktionen f_k verschieden sind; sonst muss man ${}^2\log h$ durch $d = d(F)$, die effektive Schlüssellänge der Chiffre F ersetzen.

Daher ist die folgende Definition motiviert:

Definition 3. Für eine Chiffre F mit effektiver Schlüssellänge $d(F)$ auf einer Sprache M mit Redundanz $\delta(M)$ heißt

$$\mathrm{ED}(F) := \frac{d(F)}{\delta(M)}$$

die **Eindeutigkeitsdistanz**.

Beispiele

Es wird stets das Alphabet $\Sigma = \{\mathtt{A}, \ldots, \mathtt{Z}\}$ mit $n = 26$ und die Sprache $M =$ „Deutsch“ angenommen.

1. Bei der Verschiebechiffre ist $d = {}^2\log 26$, ED $\approx 4.7/3.3 \approx 1.4$, nicht ungefähr 4, wie im Eingangsbeispiel vermutet. Diese Diskrepanz ist auf die vielen Ungenauigkeiten in der Herleitung zurückzuführen; insbesondere ist die Näherung $t_r \approx 2^{r\rho(M)}$ für kleine r natürlich besonders ungenau.

2. Bei der monoalphabetischen Chiffre ist $d \approx 88.4$, ED $\approx 88.4/3.3 \approx 26.8$. Dieser Wert stimmt mit empirischen Erfahrungen über die Lösbarkeit monoalphabetischer Kryptogramme recht gut überein.

3. Bei der Trithemius-Belaso-Chiffre mit Periode l ist $d \approx 4.7 \cdot l$, ED $\approx 1.4 \cdot l$.

4. Bei der Drehscheiben-Chiffre nach Porta ist $d \approx 88.4 + 4.7 \cdot l$, ED $\approx 26.8 + 1.4 \cdot l$.

5. Bei der allgemeinen polyalphabetischen Substitution der Periode l mit unabhängigen Alphabeten ist $d \approx 122 \cdot l$, ED $\approx 37 \cdot l$.

6. Bei der Vernam-Chiffre mit der Gruppe $G = \Sigma$ ist $M = K = C = \Sigma^*$, also $\#K = \infty$. Es ist aber sinnvoll, $d_r/\delta_r = r \cdot {}^2\log n/0 = \infty$ als Eindeutigkeitsdistanz anzusehen.

9.6 Kryptologische Anwendungen

Die Eindeutigkeitsdistanz ist ein sehr grobes Maß für die Qualität einer Chiffre. Sie wird in der modernen Kryptologie praktisch nicht verwendet. Unter der Annahme von bekanntem Klartext verliert sie ihre Bedeutung (außer für perfekte Chiffren, da bleibt sie ∞).

Eine große Eindeutigkeitsdistanz erreicht man durch:

- einen großen Schlüsselraum,
- Minderung der Redundanz der Sprache, z. B. durch vorherige Kompression.

Anwendung 1. Die PORTA-Drehscheiben-Chiffre ist nur unwesentlich stärker als die TRITHEMIUS-BELASO-Chiffre, da die Eindeutigkeitsdistanz nur um den konstanten Summanden 26.8 erhöht wird. Das bedeutet, dass bei langer Periode die Komplikation durch das permutierte Primäralphabet kaum zusätzliche Sicherheit bringt.

Anwendung 2. Eine weitere Anwendung der SHANNONschen Theorie betrifft die Lauftextverschlüsselung: Die Kryptoanalyse gewinnt ja aus einem Geheimtext der Länge r zwei sinnvolle Klartexte der Gesamtlänge $2r$. Damit das klappt, muss die Redundanz der Sprache mindestens 50% sein.

Stellen wir uns allgemeiner eine q-fache Lauftextverschlüsselung mit q unabhängigen Schlüsseltexten vor. Wenn die Kryptoanalyse möglich ist, ist aus einem Geheimtext der Länge r sinnvoller Text der Gesamtlänge $(q+1) \cdot r$ rekonstruierbar, die Redundanz der Sprache also $\geq \frac{q}{q+1}$. Da die Redundanz von Deutsch mit 70% kleiner als $\frac{3}{4}$ ist, kann man daraus schließen, dass eine *dreifache Lauftextverschlüsselung sicher* ist. Für Englisch mit seiner etwas geringeren Redundanz könnte sogar eine zweifache Lauftextverschlüsselung schon sicher sein.

Anwendung 3. Die Eindeutigkeitsdistanz wird auch als Anhaltspunkt dafür genommen, wieviel Geheimtext mit dem gleichen Schlüssel dem Gegner in die Hände fallen darf, ohne dass er etwas damit anfangen kann; d. h., wie häufig Schlüsselwechsel nötig sind.

Allgemein kann man die SHANNONsche Theorie zusammenfassen zu der Regel: *Eine notwendige Bedingung für die Lösung einer Chiffre ist, dass „Information im Geheimtext + Redundanz der Sprache“ ≥ „Information im Klartext + Information im Schlüssel“.*

Literatur

[1] C. A. Deavours: Unicity points in cryptanalysis. Cryptologia 1 (1977), 469–684.

[2] M. E. Hellman: An extension of the Shannon theory approach to cryptography. IEEE Trans Information Theory 23 (1977), 289–294.

[3] A. M. Jaglom, I, M, Jaglom: *Wahrscheinlichkeit und Information.* VEB Deutscher Verlag der Wissenschaften, Berlin 1967.

[4] H. Jürgensen: Language redundancy and the unicity point. Cryptologia 7 (1983), 37–48.

[5] H. Jürgensen, D. E. Matthews: Some results on the information theoretic analysis of cryptosystems. Crypto 83, 303–356.

[6] J. Reeds: Entropy calculations and particular methods of cryptanalysis. Cryptologia 1 (1977), 235–254.

[7] C. E. Shannon: A mathematical theory of communication. Bell System Technical Journal 27 (1948), 379–423, 623–656. [Online im WWW: `http://cm.bell-labs.com/cm/ms/what/shannonday/paper.html`]

[8] C. E. Shannon: Communication theory of secrecy systems. Bell System Technical Journal 28 (1949), 656–715. [Online im WWW: `http://www3.edgenet.net/dcowley/docs.html` oder: `http://www.cs.ucla.edu/~jkong/research/security/shannon.html`]

[9] C. E. Shannon: The entropy of printed english. Bell System Technical Journal 30 (1941), 50–64.