1 A-priori- und a-posteriori-Wahrscheinlichkeiten

Modell-Situation

Betrachtet werden:

- eine endliche Menge $M_0 \subseteq M$ von 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 sei eine Funktion

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

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

Wahrscheinlichkeit von Schlüsseln. Ebenso sei 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 $P(c) \neq 0$.
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.