5.1 Die Idee der linearen Kryptoanalyse

Sei

$$F\colon \mathbb{F}_2^n \times \mathbb{F}_2^l \longrightarrow \mathbb{F}_2^n$$

eine Bitblock-Chiffre. Wir stellen uns die Argumente von F als Klartexte $a \in \mathbb{F}_2^n$ und Schlüssel $k \in \mathbb{F}_2^l$, die Werte von F als Geheimtexte $c \in \mathbb{F}_2^n$ vor. Dann kann man zu zwei Linearformen

$$\alpha\colon \mathbb{F}_2^n \times \mathbb{F}_2^l \longrightarrow \mathbb{F}_2, \quad \text{und} \quad \beta\colon \mathbb{F}_2^n \longrightarrow \mathbb{F}_2$$

die Wahrscheinlichkeit der linearen Relation (α, β) beziehungsweise ihr Potenzial betrachten:

$$\begin{aligned} p_F(\alpha,\beta) &= \frac{1}{2^{n+l}} \cdot \#\{(a,k,c) \in \mathbb{F}_2^n \times \mathbb{F}_2^l \times \mathbb{F}_2^n \mid c = F(a,k), \alpha(a,k) = \beta(c)\}, \\ \lambda_F(\alpha,\beta) &= (2p_F(\alpha,\beta) - 1)^2 = \frac{1}{2^{2n+2l}} \cdot \hat{\vartheta}_F(\alpha,\beta)^2, \end{aligned}$$

wobei in der Notation nicht zwischen einer Linearform und dem zugehörigen Vektor unterschieden wird. Zerlegt man $\alpha(a,k)$ in die Summe $\alpha'(a) + \gamma(k)$ – und schreibt dann statt α' einfach in neuer Bedeutung α –, so kann man sagen, dass $p_F((\alpha,\gamma),\beta)$ die Wahrscheinlichkeit dafür angibt, dass bei bekanntem Klartext a die lineare Relation

$$\gamma(k) = \alpha(a) + \beta(c)$$

für die Schlüsselbits $k_{i_1}, \ldots, k_{i_r}$ gilt, wenn $I = (i_1, \ldots, i_r)$ die Indexmenge ist, die der Linearform γ entspricht. Dabei ist $\gamma(k) = k_{i_1} + \cdots + k_{i_r}$ ein einzelnes Bit, das die durch I definierte Summe einiger Bits des Schlüssels k darstellt. Das Potenzial $\lambda_F((\alpha,\gamma),\beta)$ misst die Abweichung der Wahrscheinlichkeit vom Wert $\frac{1}{2}$, denn eine Wahrscheinlichkeit $< \frac{1}{2}$ ist genauso gut wie eine $> \frac{1}{2}$: Sie sagt, dass die komplementäre Relation

$$\gamma(k) = \alpha(a) + \beta(c) + 1$$

überzufällig oft gilt.

Daraus leitet man folgendes Vorgehen für die Schätzung von $\gamma(k)$ ab (im Fall $p_F > \frac{1}{2}$, sonst komplementär):

1. **[Sammelphase]** Man sammelt N Klartext-Geheimtextpaare $(a_1, c_1), \ldots, (a_N, c_N)$.

2. **[Auszählung]** Man bestimmt die Anzahl

$$t := \#\{i = 1, \ldots, N \mid \alpha(a) + \beta(c) = 0\}.$$

3. **[Mehrheitsentscheidung]** aufgrund von t:

- Ist $t > \frac{N}{2}$, schätzt man $\gamma(k) = 0$.
- Ist $t = \frac{N}{2}$, „randomisiert“ man die Entscheidung, d. h., man entscheidet sich zufällig für 0 oder 1, jeweils mit Wahrscheinlichkeit $\frac{1}{2}$.
- Ist $t < \frac{N}{2}$, schätzt man $\gamma(k) = 1$.

Wenn man ein lineare Relation mit hinreichend hohem Potenzial erwischt hat, wird die Erfolgswahrscheinlichkeit dieses Verfahrens bei hinreichend großem N hinreichend gut sein.

Findet man mehrere solche lineare Relationen mit hinreichender Gewissheit, so hat man den Schlüsselraum auf einen Unter-Vektorraum eingeschränkt und kann über diesen eine Exhaustion versuchen. Das ist die Grundidee der linearen Kryptoanalyse – es gibt je nach dem konkreten Aufbau einer Chiffre verschiedene Varianten, wie in den folgenden Abschnitten deutlich wird.

Als theoretisches Ergebnis aus der Analyse einer Chiffre erhält man dadurch einen Zusammenhang zwischen der Menge von benötigtem Klartext und der Erfolgswahrscheinlichkeit oder auch der Dimension des übriggebliebenen Suchraums.

Damit das Verfahren anwendbar ist, sind folgende Fragen zu klären:

1. Wie findet man lineare Relationen von möglichst großem Potenzial?
2. Da Bitblock-Chiffren meistens aus vielen Runden zusammengestzt, sind, fragt man weiter:
   (a) Wie findet man bei einer iterierten Bitblock-Chiffre brauchbare lineare Relationen für die Rundenfunktion?
   (b) Wie setzt man diese über die Runden hinweg zu linearen Relationen für die ganze Chiffre zusammen, so dass Aussagen über Schlüsselbits resultieren?
   (c) Wie bestimmt man die Wahrscheinlichkeit einer zusammengesetzten linearen Relation für die ganze Chiffre aus der für die einzelnen Runden?
3. Wie hängt die Erfolgswahrscheinlichkeit von der Zahl N der bekannten Klartext-Blöcke ab?

Die Antwort auf die erste Frage und Teil (a) der zweiten heißt: Aus dem linearen Profil, also durch FOURIER-Analyse. Die anschließenden Teilfragen führen zur Untersuchung von „linearen Pfaden“ und „linearen Hüllen“ und der Kumulation von Wahrscheinlichkeiten.

Nun ist die FOURIER-Analyse zwar sehr effizient, wenn man den Aufwand (Zeit und Speicherplatz) als Funktion der Größe des Inputs betrachtet. Leider wächst diese Größe aber exponenziell mit der Dimension; daher wird die

FOURIER-Analyse bei aller Effizienz schon bei Dimensionen von etwas über 10 undurchführbar; die für ernsthafte Blockchiffren relevanten Dimensionen, also Blockgrößen und Schlüssellängen von 64, besser aber 128 Bits, liegen weit darüber.

Dieser Einwand gilt auch für die erste Teilfrage von Frage 2. Da die einzelnen Runden aber meist wiederum im wesentlichen aus einer parallelen Abarbeitung kleinerer Stücke, der S-Boxen, bestehen, wird man versuchen, das Problem auf die Analyse der S-Boxen zurückzuspielen, und diese ist realistisch durchführbar; selbst das AES-Verfahren verwendet nur 8-dimensionale S-Boxen.