2.3 Algebraische Kryptoanalyse

Der Angriff mit bekanntem Klartext

Sei (wie hier üblich) eine Bitblock-Chiffre durch eine Abbildung

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

beschrieben. Dann ist F ein n-Tupel $F = (F_1, \dots F_n)$ von Polynomfunktionen in $n + l$ Unbestimmten, deren sämtliche partiellen Grade ≤ 1 sind.

Ein Angriff mit bekanntem Klartext $a \in \mathbb{F}_2^n$ und Geheimtext $c \in \mathbb{F}_2^n$ ergibt ein Gleichungssystem

$$F(a, x) = c$$

von n Polynomgleichungen für den unbekannten Schlüssel $x \in \mathbb{F}_2^l$.

Solche Gleichungssysteme (über beliebigen Körpern) sind Gegenstand der Algebraischen Geometrie. Eine Faustregel besagt

Die Lösungsmenge für x ist „im allgemeinen klein“, wenn $n \geq l$.

(Andernfalls braucht man mehrere bekannte Klartextblöcke.)

Die allgemeine Theorie hierzu ist hochkompliziert, insbesondere, wenn man konkrete Lösungsverfahren haben will. Aber vielleicht hilft die Beobachtung, dass man nur partielle Grade ≤ 1 benötigt?

Beispiele

Beispiel 1, Linearität: Ist F eine *lineare* Abbildung, so ist das Gleichungssystem mit den Methoden der Linearen Algebra effizient lösbar (n lineare Gleichungen in l Unbekannten). Es reicht dazu schon, wenn F linear in x ist.

Beispiel 2: Sei $n = l = 2$, $F(T_1, T_2, X_1, X_2) = (T_1 + T_2X_1, T_2 + T_1X_2 + X_1X_2)$, $a = (0, 1)$, $c = (1, 1) \in \mathbb{F}_2^2$. Dann sieht das Gleichungssystem für den Schlüssel $(x_1, x_2) \in \mathbb{F}_2^2$ so aus:

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 + x_1 \\ 1 + 0 + x_1x_2 \end{pmatrix},$$

die Lösung ist offensichtlich $x_1 = 1$, $x_2 = 0$.

Substitution: Dass man Polynomgleichungen nicht immer auf den ersten Blick ihre Komplexität ansieht, zeigt das Beispiel (über $\mathbb{F}_2$)

$$x_1x_2x_3 + x_1x_2 + x_1x_3 + x_2x_3 + x_2 + x_3 = 0.$$

Es geht durch die Substitutionen $x_i = u_i + 1$ über in

$$u_1u_2u_3 + u_1 = 0$$

(umgekehrt sieht man das leichter) mit der Lösungsmenge

$$u_1 = 0,\ u_2, u_3 \text{ beliebig } \textit{oder } u_1 = u_2 = u_3 = 1.$$

Die vollständige Lösung der ursprünglichen Gleichung ist also

$$x_1 = 1,\ x_2, x_3 \text{ beliebig } \textit{oder } x_1 = x_2 = x_3 = 0.$$

Die Komplexität des Angriffs

Was in den Beispielen so einfach war, ist im allgemeinen zu komplex:

Satz 2 (Garey/Johnson) *Das Problem, eine gemeinsame Nullstelle eines Polynomsystems $f_1, \ldots, f_r \in \mathbb{F}_2[T_1, \ldots T_n]$ zu finden, ist* **NP***-vollständig.*

Beweis. Siehe das Buch von Garey/Johnson. ◇

Der Begriff „**NP**-vollständig“ wird später in der Vorlesung erklärt.

Deutung: Bei günstig gewählter Blockverschlüsselungsfunktion $F : \mathbb{F}_2^n \times \mathbb{F}_2^l \longrightarrow \mathbb{F}_2^n$ ist der Angriff mit bekanntem Klartext nicht effizient durchführbar.

Offene Probleme

Im Grunde besagt der Satz *gar nichts*:

1. Er bezieht sich nur auf den Fall eines Algorithmus für *beliebige* Polynome. Er macht keine Aussage für ein bestimmtes Polynomsystem.
2. Selbst wenn er das machen würde, wäre immer noch kein konkretes Beispiel eines „schwierigen“ Polynomsystems bekannt.
3. Und selbst dann würde der Satz nichts darüber sagen, ob nur einzelne, wenige Instanzen des Problems schwierig sind oder – was der Kryptologe eigentlich braucht – fast alle.

Weitere Hinweise auf die Schwierigkeit, Polynomgleichungen zu lösen, gibt der Artikel

- D. Castro, M. Giusti, J. Heintz, G. Matera, L. M. Pardo: The hardness of polynomial equation solving. Found. Comput. Math. 3 (2003), 347–420. (Für Uni-Mainz online über die EZB zugänglich.)

Interpolationsangriff

Eine Variante der algebraischen Kryptoanalyse mit bekanntem Klartext ist der Interpolationsangriff, der in

- Thomas Jakobsen, Lars R. Knudsen: The interpolation attack on block ciphers, FSE 1997,

vorgestellt wurde. Die Idee ist ganz einfach: Der Vektorraum $\mathbb{F}_2^n$ kann bei Wahl einer geeigneten Multiplikation als endlicher Körper $K = \mathbb{F}_{2^n}$ der Charakteristik 2 interpretiert werden. Bei festem Schlüssel $k \in \mathbb{F}_2^l$ ist die Bitblock-Chiffre dann einfach eine Funktion $F_k \colon K \longrightarrow K$, also ein Polynom. Ist d sein Grad, so kann es nach der Interpolationsformel aus $d+1$ bekannten Klartext-Blöcken bestimmt werden; gleiches gilt für die Umkehrfunktion. Damit kann dann ver- bzw. entschlüsselt werden, ohne dass der Schlüssel explizit bestimmt wurde.

Um diesen Angriff zu verhindern, muss man also darauf achten, dass Ver- und Entschlüsselungsfunktion auf K bei festem Schlüssel stets einen hohen Grad besitzen; das ist machbar, da ja Grade bis $2^n - 1$ möglich sind.

Allerdings funktioniert der Angriff auch bei hohem Grad, wenn das jeweilige Polynom nur wenige Koeffizienten $\neq 0$ hat, also „dünn besetzt“ ist. Daher muss auch dieses vermieden werden.

Linearisierung überbestimmter Gleichungssysteme

Gleichungssysteme höherer Ordnung kann man manchmal brechen, wenn sie so weit überbestimmt sind, dass man die Monome (oder einige davon) als eigene Unbekannte ansehen kann. Dies wird durch das folgende einfache Beispiel illustriert:

$$\begin{aligned} x^3 + xy + y^5 &= 1, \\ 2x^3 - xy &= 0, \\ xy + 3y^5 &= 3. \end{aligned}$$

Hier substituiert man alle vorkommenden Monome: $u := x^3$, $v := xy$, $w := y^5$, und erhält das lineare Gleichungssystem

$$\begin{aligned} u + v + w &= 1, \\ 2u - v &= 0, \\ v + 3w &= 3. \end{aligned}$$

aus drei Gleichungen mit drei Unbekannten. Die (in diesem Fall sehr einfach auch manuell zu erhaltende) Lösung ist $u = 0$, $v = 0$, $w = 1$; sie ist eindeutig, wenn wir einen Körper der Charakteristik $\neq 7$ annehmen. Daraus ergibt sich als vollständige Lösung des ursprünglichen Systems: $x = 0$, $y = 1$ oder irgendeine im Körper enthaltene 5. Einheitswurzel.

Dieser Angriff wurde im Jahre 2002 populär, als das Gerücht um die Welt lief, der neue AES sei für diesen Angriff anfällig. Bei genauem Hinsehen ließ sich das aber nicht bestätigen; die einzelnen Gleichungen des konstruierten (riesigen) linearen Gleichungssystems waren bei weitem nicht unabhängig.