Einführung in die Kryptographie
4. Vorlesung

18. 5. 2006

- (Pseudo)-Zufallsfolgen
- Stromchiffren

Zufallsfolgen

Theorie: Eine Folge

$$X_1, X_2, X_3, \ldots$$

von Zufallsvariablen mit Werten in einer endlichen Menge S, die

- unabhängig
- (annähernd) gleichverteilt sind.

Eigenschaft:

$$P(X_{i+1} = s \mid X_1 = s_1, \ldots, X_i = s_i) = 1/|S|.$$

Wie kann man so eine Folge in der Praxis realisieren/simulieren?

Zufallsfolgen

Wie kann man so eine Folge in der Praxis realisieren/simulieren?

- Physikalische Zufallsprozesse
- Einfluss durch Systemzustand
- Hash-Funktionen
- Pseudo-Zufallsfolge

Pseudo-Zufallsfolge

Ein *Pseudo-Zufallsgenerator* ist ein Algorithmus, der in Abhängigkeit eines Initialisierungswertes k (*seed*) eine Folge

$$s_1, s_2, s_3, \ldots$$

produziert, die 'zufällig' erscheint.

Nutzen:

- Ersatz für echten Zufallsgenerator
- Produktchiffren
- Stromchiffren
- Randomisierung von Blockchiffren

Synchrone Stromchiffren

Es gibt viele Varianten. Das Prinzip:
One-Time-Pad mit einer Pseudo-Zufallsfolge.

Sei

$$s_1, s_2, s_3, \ldots \in \{0,1\}^*$$

eine PZF mit Initialisierungswert $k \in \mathcal{K}$. Man erhält eine Verschlüsselungsfunktion

$$E_k : \{0,1\}^* \to \{0,1\}^*, \quad p_i \mapsto p_i \oplus s_i.$$

Der zugrundeliegende P.-Zufallsgenerator heisst *kryptographisch sicher* wenn das resultierende Kryptosystem sicher ist gegen Known-Plaintext-Angriffe.

Kryptographische Sicherheit

Andere Formulierung: der P.-Zufallsgenerator ist *kryptographisch sicher*, wenn ein Angreifer bei Kenntnis von

$$s_1, s_2, \ldots, s_i$$

den folgenden Wert s_{i+1} höchstens mit einer Wahrscheinlichkeit von $1/2$ vorhersagen kann.

(Man geht davon aus, dass der Angreifer den verwendeten Algorithmus kennt, aber nicht den Initialisierungswert.)

Der BBS-Generator

Sei $m = p \cdot q$ Produkt zweier (grosser) Primzahlen. Wähle einen Initialisierungswert $x_0 \in \{0, \ldots, m-1\}$, und setze

$$x_{i+1} := x_i^2 \pmod{m}$$

und

$$s_i := x_i \pmod{2}.$$

Die P.-Zufallsfolge $s_1, s_2, \ldots$ ist kryptographisch sicher, wenn

- $p, q \equiv 3 \pmod{4}, \quad x_0 \equiv x^2 \pmod{m}$,
- $\mathrm{ggT}(\varphi(p-1), \varphi(q-1))$ klein ist, und
- m nicht faktorisiert werden kann.

Kryptographische Sicherheit

Vorteil des BBS: beweisbare Sicherheit (unter der Annahme, dass Faktorisieren schwierig ist).

Nachteil des BBS: zu langsam.

Man kennt einige PZGn, die effizient *und* nach heutigem Kenntnisstand sicher gegen alle bekannten Angriffe sind, z.B.

- RC4
- ISAAC
- (gewisse) lineare Schieberegisterfolgen

Randomisieren

Für viele Anwendungen ist es nicht unbedingt nötig, dass ein PZG kryptographisch sicher ist.

Beispiel: Randomisierte affine Chiffre.

Gegeben sei eine affine Blockchiffre

$$E_k : \{0,1\}^n \to \{0,1\}^n, \quad p \mapsto A_k \cdot p + b_k.$$

Eine Folge von Klartextblöcken $p_1, p_2, \ldots$ wird verschlüsselt durch

$$c_i := E_{k_i}(p_i),$$

wobei $k_1, k_2, \ldots$ eine PZF mit Initialisierungswert k_0 ist. (Der geheime Schlüssel ist also k_0.)

Randomisierte affine Chiffre

Auch wenn man K.-Text-C.-Text-Paare

$$(p_1, c_1), \ldots, (p_m, c_m)$$

kennt, ist der in der letzte Woche diskutierte Angriff nicht erfolgreich, da in jedem Schritt ein anderer Schlüssel verwendet wird.

Hoffnung: die K.-Text-C.-Text-Paare verraten zu wenig über die Folge $k_1, k_2, \ldots, k_i$, um k_{i+1} erraten zu können.

Deshalb genügt es, wenn die Folge $k_1, k_2, \ldots$ 'pseudo-zufällig' ist, d.h. sich durch statistische Tests nicht von einer echten ZF unterscheiden lässt.

Kriterien für Pseudo-Zufälligkeit

Eine PZF sollte:

- eine grosse *Periode* haben,
- annähernd G*-zufällig* sein, und
- hohe *lineare Komplexität* haben

Periodizität

Der Algorithmus zur Erzeugung der PZF wird von einer Maschine mit *endlich vielen Zuständen* ausgeführt.

Daher kehrt die Maschine nach einer endlichen Anzahl von Schritten in einen schon früher eingenommenen Zustand zurück.

Die PZF ist also *periodisch*, d.h. $\exists\, N, m$ so, dass

$$s_{i+m} = s_i \quad \forall \;\; i \geq N.$$

Die kleinste Zahl m mit dieser Eigenschft heisst die *Periode* von $s_1, s_2, \ldots$. Sie sollte möglichst gross sein.

G-Zufälligkeit

Eine Folge $s_1, s_2, \ldots,$ mit $s_i \in \{0, 1\}$ heisst *G-zufällig*, wenn gilt:

- Ungefähr die Hälfte aller Terme s_i innerhalb eines Zyklus ist gleich 1.

- Innerhalb eines Zyklus sollte ungefähr ein Anteil von $1/n$ aller Runs die Länge n haben, für $n = 1, 2, \ldots$. (Ein *Run* der Länge n ist eine maximale Teilfolge der Form $0 \ldots 0$ oder $1 \ldots 1$.)

- Die *Autokorrelationsfunktion*

$$C(t) := \frac{A(t)}{m}, \quad t = 1, \ldots, m-1$$

  sollte annähernd konstant sein ($= 1/2$). Hierbei ist $A(t)$ die Anzahl der übereinstimmenden Glieder pro Zyklus zwischen den Folgen $s_1, s_2, \ldots$ und $s_{1+t}, s_{2+t}, \ldots$.

Schieberegisterfolgen

Eine Folge

$$s_0, s_1, s_2, \ldots \quad \in \{0,1\}^*$$

ist eine *(binäre) Schieberegisterfolge* der *Länge* n, wenn es eine Funktion

$$f : \{0,1\}^n \to \{0,1\}$$

gibt mit

$$s_{i+n} = f(s_i, \ldots, s_{i+n-1}).$$

Die Funktion f heisst die *Rückkoppelungsfunktion* der SRF.

Die SRF heisst *linear* wenn f linear ist, d.h.

$$f(s_0, \ldots, s_{n-1}) = \sum_{i=0}^{n-1} c_i s_i, \quad c_i \in \{0,1\}.$$

Die Periode einer SRF

Bemerkung 1: Eine SRF der Länge n hat eine Periode m mit

$$m \leq 2^n.$$

Bemerkung 2: Jede PZF mit Periode m kann (bis auf die Vorperiode) als lineare SRF der Länge m realisiert werden:

Setze

$$c_0 := 1, \quad c_1 = \ldots = c_{m-1} := 0.$$

Dann gilt

$$s_{i+m} = s_i.$$

Eigenschaften linearer SRF

Eigenschaft 1: Sei $s_0, s_1, \ldots$ eine lineare Schieberegisterfolge der Länge n. Kennt man $2n$ aufeinanderfolgende Glieder der Folge,

$$s_i, \ldots, s_{i+2n-1},$$

so kann man die Rückkoppelungskoeffizienten $c_0, \ldots, c_{n-1}$ effizient berechnen (*Berlekamp-Massey-Algorithmus*).

Lineare SRF sind daher *nicht* kryptographisch sicher.

Eigenschaften linearer SRF

Eigenschaft 2: Für jedes n gibt es lineare SRF der Länge n mit Periode

$$m = 2^n - 1.$$

Diese heissen *maximale* oder *m-Folgen*.

Solche m-Folgen sind annähernd G-zufällig.

Lineare Komplexität

Sei $s_0, s_1, s_2, \ldots$ eine PZF. Für $k \geq 0$ sei

$$s^{(k)} := (s_0, \ldots, s_{k-1}).$$

Die *lineare Komplexität* des Abschnittes $s^{(k)}$ ist

$$L(s^{(k)}) := \min\{\, n \mid (*)_n \,\},$$

wobei die Bedingung $(*)_n$ sagt: $\exists\, c_0, \ldots, c_{n-1}$ so, dass

$$s_{j+n} = \sum c_i s_{j+k}, \quad \forall\, j \in [0, \ldots, k-n+1].$$

Das lineare Komplexitätsprofil

Für eine PZF $s_0, s_1, \ldots$ heisst die Funktion

$$L : \mathbb{N} \to \mathbb{N}, \quad k \mapsto L(s^{(k)})$$

das *lineare Komplexitätsprofil*.

Bei einer echten Zufallsfolge ist $L(k)$ für grosse k ungefähr $k/2$.

Eine kryptographisch sichere PZF sollte daher für alle relevanten Werte von k dieselbe Eigenschaft haben.

Nichtlineare SRF

Angenommen wir haben r Pseudo-Zufallsvariablen $X_1, \ldots, X_r$. (Diese können z.B. durch lineare SRF realisiert sein.)

Sei

$$f : \{0,1\}^r \to \{0,1\}$$

ein boolsche Funktion. Dann ist

$$X := f(X_1, \ldots, X_r)$$

wieder eine PZV.

Beispiel:

$$f(X_1, X_2, X_3) := X_1X_2 + X_2X_3 + X_1X_3.$$

Nichtlineare SRF

Angenommen, die PZV X_i seien durch lineare Schieberegisterfolgen realisiert, mit grosser Periode und guter G-Zufälligkeit.

Durch geeignete Wahl der (nichtlinearen) Funktion f kann man erreichen, dass die PZV X die obigen Eigenschaften der X_i erhält, und gleichzeitig die lineare Komplexität grösser wird.

Ist $f(X_1, \ldots, X_r)$ ein Polynom mit Koeffizienten in $\{0, 1\}$, so gilt

$$L(X) \sim f(L(X_1), \ldots, L(X_r)).$$