Pseudorandom Generators

Seminar: Kryptographie

David Hobach
David.Hobach@rub.de

Ruhr-Universität Bochum

12.05.2009 / Bochum

Outline

Motivation

- viele kryptographische Algorithmen benutzen Zufallszahlen
- häufig gilt: initialisierende Zufallszahlen gefunden $\Leftrightarrow$ Krypto-Algorithmus kompromittiert
- das Generieren wirklich zufälliger Zufallszahlen ist aufwendig und benötigt Hardware (z.B. Messen kosmischer Strahlung oder Kameraflimmern)
- diese Zahlenfolgen sind nicht deterministisch $\Rightarrow$ Wiederherstellung nicht möglich

Statistische Eigenschaften

- Zufallszahlenfolgen werden in der Informatik durch statistische Tests bzgl. ihrer Güte beurteilt $\Rightarrow$ Idee: besteht eine Folge viele statistische Tests und existieren keine bekannten Angriffe, so kann sie als “sicher“ gelten

Kritik:

- recht willkürliche Definition von Sicherheit durch willkürliche Auswahl aktueller statistischer Tests
- es wird davon ausgegangen, dass ein Angreifer maximal genauso viel weiß wie der Designer des Algorithmus der Zufallszahlenfolge

$\Rightarrow$ hier wird die alternative mathematische Herangehensweise betrachtet

Idee/Ziele

- aus einer wirklich zufälligen Anfangs-Zahlenfolge (“seed“) unendlich viele pseudozufällige Zahlen erzeugen
- die Folge soll per Definition für einen effizienten Beobachter nicht unterscheidbar von einer wirklichen Zufallsfolge sein
- effiziente und deterministische Generierung der Pseudozufallsfolge

Anmerkung:
Da unsere Folge per Definition gegenüber beliebigen effizienten Algorithmen “sicher“ ist, können wir die statistischen Tests vernachlässigen. Es wird aber nirgendwo konkretisiert, was für Algorithmen der Angreifer einsetzen könnte (z.B. statistische Tests).

Probability Ensembles

Definition

Sei I eine abzählbare Index-Menge. Ein Ensemble X ist eine Folge aus Zufallsvariablen X_i, indiziert durch I, d.h. $X := \{X_i\}_{i \in I}$.

Anmerkungen:

- normalerweise: $I = \mathbb{N}$ oder $I = \{0,1\}^*$
- es gilt: $|X_n| < poly(n)$ bzw. $n < poly(|X_n|) \Rightarrow \exists$ in polynomieller Zeit arbeitende Funktionen $f_n : n \rightarrow |X_n| \quad \forall n \in \mathbb{N}$.
- $poly(n) := \mathcal{O}(n^c)$, c = konst.
- eine Zufallsvariable k wird aus einer Wahrscheinlichkeitsverteilung P gezogen $\Rightarrow k \in P = \{z_0, \ldots, z_t\}$ mit $Pr[z_i] = p_i$ und $0 \leq i \leq t$

Polynomial-Time Indistinguishability

Definition

Zwei Ensembles $X := \{X_n\}_{n\in\mathbb{N}}$ und $Y := \{Y_n\}_{n\in\mathbb{N}}$ heißen "in polynomieller Zeit nicht unterscheidbar", wenn $\forall$ probabilistischen Algorithmen D, die in polynomieller Zeit arbeiten, $\forall$ postiven Polynome p(·) und $\forall$ ausreichend großen n's gilt:
$|Pr[D(X_n) = 1] - Pr[D(Y_n) = 1]| < \frac{1}{p(n)}$

Anmerkungen:

- $D(X_n) \in \{0, 1\}$ mit $D(X_n) = 1$, wenn D X_n akzeptiert.
- $\mu : \mathbb{N} \to [0, 1]$ heißt vernachlässigbar, wenn $\forall$ positiven Polynome p und $\forall$ ausreichend großen n's gilt: $\mu(n) < \frac{1}{p(n)}$.
- $|X_n| = f_n^{-1}(n)$
- nachweisbar: $\Rightarrow$
  $|Pr[D(X_n^{(1)}, \ldots, X_n^{(poly(n))}) = 1] - Pr[D(Y_n^{(1)}, \ldots, Y_n^{(poly(n))}) = 1]| < \frac{1}{p(n)}$

Pseudorandom Ensembles

Definition

Sei $X := \{X_n\}_{n\in\mathbb{N}}$ ein Ensemble.
X ist pseudozufällig $\Leftrightarrow \exists$ Ensemble $U = \{U_{l(n)}\}_{n\in\mathbb{N}}$, so dass X und U in polynomieller Zeit nicht unterscheidbar sind.

Anmerkung:

- U: “uniformes Ensemble“/Gleichverteilung, $|U_m| := m$
- z.B. $U_2 \in \{00, 01, 10, 11\}$ mit $Pr[z_i] = \frac{1}{4} \quad \forall i$
- $l : \mathbb{N} \to \mathbb{N}$ nötig, da $|U_m| := m$, aber $|X_n| = f_n^{-1}(n)$

Pseudorandom Generators

Definition

G sei ein deterministischer, in polynomieller Zeit arbeitender Algorithmus.

1. Es existiert eine Funktion $l : \mathbb{N} \to \mathbb{N}$ mit $l(n) > n \quad \forall\, n \in \mathbb{N}$ und $|G(s)| = l(|s|) \quad \forall\, s \in \{0,1\}^*$.
2. $\{G(U_n)\}_{n\in\mathbb{N}}$ ist pseudozufällig.

Anmerkungen:

- l: Expansionsfaktor, s: Seed
- aus n Bits generiert G minimal n+1 Bits
- die Ausgabe von G muss nur für $s = U_n$ pseudozufällig sein

Beispiel: rand()

häufige Implementierung:

$$z_0 = \text{seed}$$
$$z_{i+1} = a * z_i + b \mod m$$

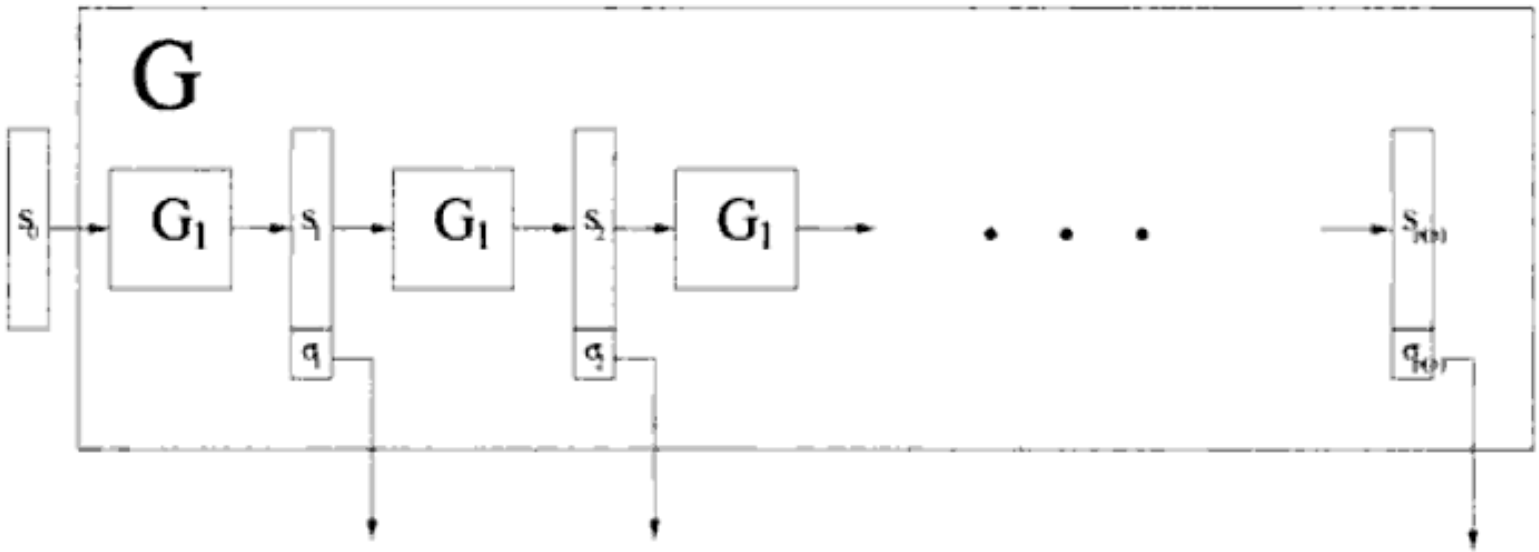
- Geg.: m; Ges.: a, b, seed (Praxis: nur seed)
- äußerst effizient
- hält bei guter Wahl der Konstanten a, b und m vielen statistischen Tests stand
- linear, d.h. bei bekannten Ausgaben leicht brechbar (LGS lösen = Algorithmus D)

$\Rightarrow \exists$ Unterscheider D $\Rightarrow$ rand() ist nicht pseudozufällig
$\Rightarrow$ rand() ist für kryptographische Zwecke ungeeignet

Anwendbarkeit allgemein

- pseudozufällige Ensembles können äquivalent zu uniformen eingesetzt werden
- der Performance-Unterschied für die empfangende Applikation ist vernachlässigbar klein, da ansonsten ein Unterscheider D existieren würde
- in polynomieller Zeit arbeitende Beobachter können die generierte Zufallsfolge nicht von einer uniformen unterscheiden
- G: n zufällige Bits $\rightarrow$ $l(n)$ zufällige Bits $\Rightarrow$ Ziel: $l(n) >>$

Increasing the Expansion Factor (1)



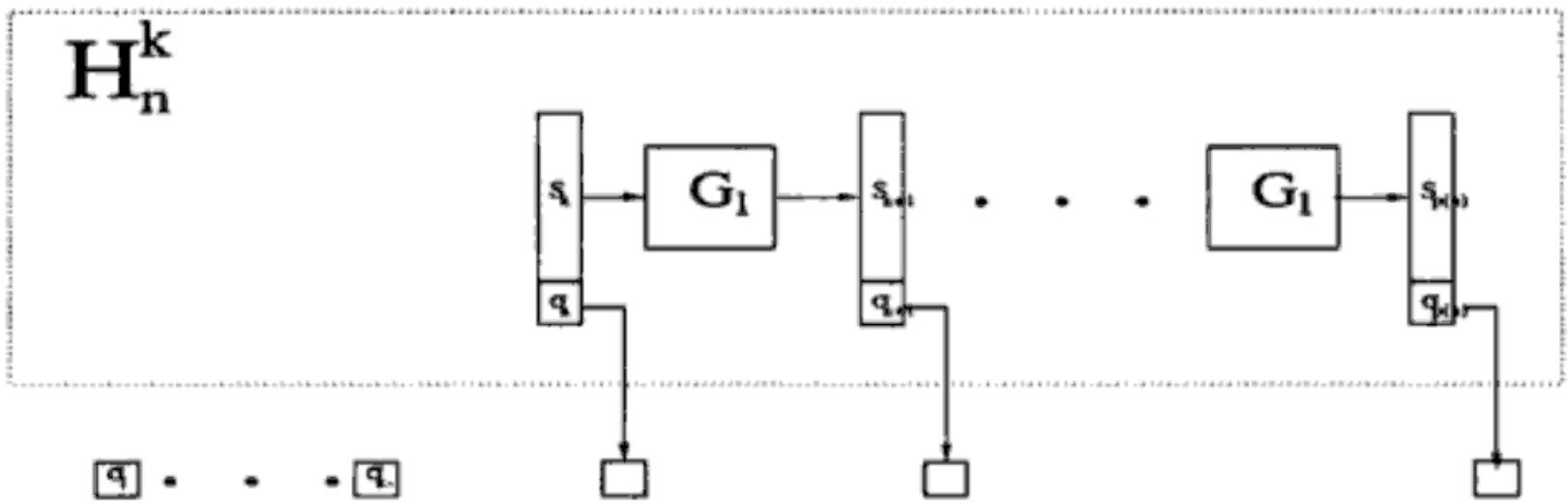
$\Rightarrow$ G: n zufällige Bits $\longrightarrow$ poly(n) zufällige Bits

Theorem

Sei G_1 ein PRG mit $l(n) = n + 1$. $\Rightarrow$ G ist ein PRG.

Increasing the Expansion Factor (2)

Beweisidee (1)



- $H_n^k := U_k^{(1)} \quad | \quad pref_{p(n)-k}(G(U_n^{(2)}))$
- $H_n^0 = G(U_n)$
- $H_n^{p(n)} = U_{p(n)}$

Increasing the Expansion Factor (3)

Beweisidee (2)

Beweis.

Ann.: G ist kein Pseudozufallszahlengenerator.
$\Rightarrow \{G(U_n)\}_{n\in\mathbb{N}}$ und $\{U_{p(n)}\}_{n\in\mathbb{N}}$ sind in polynomieller Zeit unterscheidbar
$\Rightarrow \exists$ Unterscheider D für $G(U_n) = H_n^0$ und $U_{p(n)} = H_n^{p(n)}$ ("Extrema")
$\Rightarrow \exists$ Unterscheider D' für H_n^k und $H_n^{k+1} \quad 0 \leq k \leq p(n) - 1$
$\Rightarrow \exists$ Unterscheider D'' für $\{G_1(U_n)\}_{n\in\mathbb{N}}$ und $\{U_{n+1}\}_{n\in\mathbb{N}}$ (hier nicht bew.)
$\Rightarrow$ Widerspruch zur Pseudozufälligkeit von G_1 □

Unpredictability

gewünschte Eigenschaft pseudozufälliger Ensembles:

Definition

Ein Ensemble $\{X_n\}_{n\in\mathbb{N}}$ heißt "unvorhersehbar in polynomieller Zeit", falls $\forall$ probabilistischen, in polynomieller Zeit arbeitenden Algorithmen A, $\forall$ positiven Polynome p(.) und $\forall$ ausreichend großen n's gilt:
$Pr[A(1^{|X_n|}, X_n) = next_A(X_n)] < \frac{1}{2} + \frac{1}{p(n)}$

Anmerkungen:

- A: "Vorhersehungsalgorithmus", $A(1^{|x|}, x)$ liest $i \leq |x|$ Bits von x und versucht, das nächste (i+1) zu raten
- $next_A(x)$ gibt das korrekte i+1. Bit zurück (für $i = |x|$ ein zufälliges)
- $\frac{1}{2} + \frac{1}{p(n)}$: vernachlässigbarer Vorteil über $\frac{1}{2}$
- interessant: $= \frac{1}{2}$ nicht nötig, da $\forall$ A

pseudozufällig $\Leftrightarrow$ unvorhersehbar (1)

Theorem

Pseudozufälligkeit $\Leftrightarrow$ *Unvorhersehbarkeit*

pseudozufällig $\Rightarrow$ unvorhersehbar (2)

Beweis.

"$\Rightarrow$":
das Ensemble $X := \{X_n\}_{n\in\mathbb{N}}$ sei pseudozufällig
$\Leftrightarrow$ X ist in polynomieller Zeit nicht unterscheidbar von dem uniformen Ensemble $\{U_n\}_{n\in\mathbb{N}}$
Ann.: X ist von einem Algorithmus A in polynomieller Zeit vorhersehbar
$\Leftrightarrow Pr[A(1^{|X_n|}, X_n) = next_A(X_n)] \geq \frac{1}{2} + \frac{1}{p(n)}$
$\Rightarrow \exists$ Unterscheider D, der bei korrekter Vorhersage von A 1 ausgibt, sonst 0:
$Pr[D(X_n) = 1] = Pr[A(1^{|X_n|}, X_n) = next_A(X_n)] \geq \frac{1}{2} + \frac{1}{p(n)}$
es gilt: $Pr[D(U_n) = 1] = Pr[A(1^n, U_n) = next_A(U_n)] \leq \frac{1}{2}$
$\Rightarrow Pr[D(X_n)] - Pr[D(U_n) = 1] \geq \frac{1}{p(n)}$
$\Rightarrow$ Widerspruch zur Pseudozufälligkeit von $\{X_n\}_{n\in\mathbb{N}}$

pseudozufällig $\Leftarrow$ unvorhersehbar (3)

Beweis.

"$\Leftarrow$": das Ensemble $X := \{X_n\}_{n\in\mathbb{N}}$ sei unvorhersehbar

Ann.: X ist nicht pseudozufällig

$\Leftrightarrow |Pr[D(X_n) = 1] - Pr[D(U_n) = 1]| \geq \frac{1}{p(n)}$

oBdA gelte: $Pr[D(X_n) = 1] - Pr[D(U_n) = 1] \geq \frac{1}{p(n)}$

Definiere Hybride: $H_n^i := pref_i(X_n) \quad | \quad suf_{n-i}(U_n) \quad \forall n$

(u.a. $H_n^n = X_n, H_n^0 = U_n$)

$\Rightarrow \frac{1}{n}\sum_{i=0}^{n-1}(Pr[D(H_n^{i+1}) = 1] - Pr[D(H_n^i) = 1]) \geq \frac{1}{p(n)n}$

$\Rightarrow$ Idee: $Pr[D(H_n^{i+1}) = 1] \geq Pr[D(H_n^i) = 1]$ für viele i

$\Rightarrow$ Algorithmus A zum Vorhersehen:

1. wähle i zufällig, $0 \leq i \leq n-1$
2. wähle $r_{i+1} \ldots r_n$ unabhängig und uniform in $\{0,1\}$
3. falls $D(x_1 \ldots x_i r_{i+1} \ldots r_n) = 1$, gib r_{i+1} aus, sonst $1 - r_{i+1}$

$\Rightarrow$ Widerspruch zur Unvorhersehbarkeit von X. □

Strong One-Way Functions

Erinnerung

Definition

Eine Funktion $f : \{0,1\}^* \rightarrow \{0,1\}^*$ wird als starke Einwegfunktion bezeichnet, wenn gilt:

1. einfach zu berechnen: $\exists$ ein deterministischer, in polynomieller Zeit arbeitender Algorithmus A mit $A(x) = f(x)$.
2. schwer zu invertieren: $\forall$ probabilistischen, in polynomieller Zeit arbeitenden Algorithmen A', $\forall$ positive Polynomen p(.) und $\forall$ ausreichend großen n's gilt:
   $Pr[A'(f(U_n), 1^n) \in f^{-1}(f(U_n))] < \frac{1}{p(n)}$

Anmerkung:

- $f^{-1}(f(U_n))$ kann eine Menge sein (k-1 Funktionen mit $k \geq 1$)

Bezug zu “One-Way Functions“ (1)

Theorem

$\exists$ *Pseudorandom Generators* $\Rightarrow \exists$ *One-Way-Functions*

Beh.: Sei G ein Pseudozufallszahlengenerator mit dem Expansionsfaktor $l(n) = 2n$. Dann ist die Funktion $f(x, y) := G(x)$ für jedes $|x| = |y|$ eine starke Einwegfunktion.

Bezug zu “One-Way Functions“ (2)

Beweis.

f ist wie G in polynomieller Zeit berechenbar.
Ann.: Sei A ein probabilistischer, in polynomieller Zeit arbeitender Algorithmus, der $f(U_{2n})$ mit Erfolgsws. $\geq \frac{1}{poly(n)}$ invertiert.
$\Rightarrow \exists$ Unterscheider D mit $D(\alpha) = 1$ gdw $f(A(\alpha)) = \alpha$
$\Rightarrow Pr[f(A(f(U_{2n}))) = f(U_{2n})] \geq \frac{1}{p(n)} \quad \forall$ Polynome p(.)
$\Rightarrow$ mit $\alpha = G(U_n) = f(U_{2n})$ ergibt sich:
$Pr[D(G(U_n)) = 1] = Pr[f(A(G(U_n))) = G(U_n)] \geq \frac{1}{p(n)}$
f besitzt zudem max. 2^n Urbilder, da $l(n) = 2n$, d.h. es gilt weiterhin:
$Pr[D(U_{2n}) = 1] = Pr[f(A(U_{2n})) = U_{2n}] \leq 2^{-n}$
$\Rightarrow Pr[D(G(U_n)) = 1] - Pr[D(U_{2n}) = 1] \geq \frac{1}{p(n)} - \frac{1}{2^n} \geq \frac{1}{2p(n)}$
$\Rightarrow$ Widerspruch zur Pseudozufälligkeit von G $\square$

Zusammenfassung

- PRGs können effizient Bitfolgen generieren, die ein in polynomieller Zeit arbeitender Angreifer nicht von wirklich zufälligen unterscheiden kann
- der Expansionsfaktor eines PRGs kann beliebig vergrößert werden
- pseudozufällige Ensembles sind unvorhersehbar
- $\exists$ PRGs $\Rightarrow \exists$ One-Way-Functions

Literatur

- Oded Goldreich. Foundations of Cryptography. 2001. S. 101-124
- Prof. Yehuda Lindell. Foundations of Cryptography. 2008. Lecture 4. URL: http://u.cs.biu.ac.il/~lindell/89-856/
- Prof. Christof Paar. Applied Cryptography and Data Security. 2005. S. 26 (34). URL: http://crypto.rub.de/imperia/md/content/lectures/notes.pdf
- Prof. Alexander May. Quantenalgorithmen. Vorlesung. 2008. URL: http://www.cits.rub.de/lehre/quantalws08.html