Pseudozufällige Permutationen
Konstruktion und Anwendung

Hendrik Süß

28. Januar 2003

Zusammenfassung

Es wird die Konstruktionen von Noar und Reingold für pseudozufällige Permutationen vorgestellt, die die Struktur eines Feistelnetzwerkes nutzt. Diese Konstruktion wird zu einem Schema erweitert, mit dessen Hilfe anpasste Pseudozufallspermutationen entworfen werden können. Das Schema erlaubt dann die vereinfachte Analyse der neuen Konstruktion auf Grundklage bereits bewiesener Sachverhalte.

Und schließlich wird eine Anwendung pseudozufälliger Permutationen in einem Kryptosystem vorgestellt, wobei die Pseudozufallseigenschaft der Permutation Aussagen über die Sicherheit des Kryptosystems zulässt.

Inhaltsverzeichnis

1 Einleitung **3**

2 Definitionen **3**

3 Konstruktionen **4**
3.1 Basiskonstruktion 4
3.2 Pseudozufälligkeit der Basiskonstruktionen 7
3.3 Schema zur Konstruktion Pseudozufallspermutationen 11
3.4 Einschränkung der Vorraussetzungen 12
3.5 Blockweise Anwendung 13

4 Anwendung als Kryptosysteme **13**
4.1 Definitionen 13
4.2 Analyse 15

5 Ausblick **15**

1 Einleitung

Pseudozufällige Permutationen haben gegenüber pseudozufälligen Funktionen die zusätzliche Eigenschaft injektiv zu sein. Sie können überall dort eingesetzt werden, wo auch pseudozufällige Funktionen Anwendung finden, da sie sich nicht effizient von ihnen unterscheiden lassen.

2 Definitionen

Definition 2.0.1 (Permutionsensemble) *Ein **Permutationsensemble** ist eine Folge von $P = \{P_n\}_{n\in\mathbb{N}}$ von Zufallsvariablen, so dass P_n eine Verteilung auf der Menge der Permutationen auf $\{0,1\}^n$ liefert. Das gleichverteilte Permutationsensemble wird mit $K = \{K_n\}_{n\in\mathbb{N}}$ bezeichnet*

*Ein Permutationsensemble heißt **pseudozufällig**, falls es sich nicht effizient vom gleichverteilten Ensemble unterscheiden lässt, also für jede effiziente probablistische Orakelmaschine M, jedes Polynom p(x) und genügend großes n gilt:*

$$\left|P[M^{P_n}(1^n) = 1] - P[M^{K_n}(1^n) = 1]\right| < \frac{1}{p(n)}$$

Wie wir noch sehen werden ist ein pseudozufälliges Permutationsenemble gleichzeitig ein pseudozufälliges Funktionsensemble. Die Ausgaben der Zufallsvariablen P_n lassen sich also nicht effizient als Permutationen identifizieren.

Häufig genügt die Tatsache, dass es ein eindeutig bestimmtes Urbild gibt nicht. In der Anwendung müssen die Funktionen auch effizient berechenbar sein, häufig müssen zusätzlich die eindeutig bestimmen Urbilder ebenfalls effizient bestimmbar sein.

Definition 2.0.2 *Ein Permutationsensemble, $P = \{P_n\}$ heißt **effizient berechenbar und invertierbar**, falls*

(effizient indizierbar): Es existiert ein effizienter probablistischer Algorithmus I und eine Abbildung ϕ aus $\{0,1\}^*$ in die Menge der Permutationen, so dass $\phi(I(1^n))$ und P_n die gleiche Verteilung erzeugen.

(effiziente auswertbar): Es existiert ein effizienter Algorithmus V, mit $V(i,x) = \phi(i)(x)$,

(effizient invertierbar): Es existiert ein effizienter Algorithmus N, so dass$N(i,x) = \phi(i)^{-1}(x)$

Es macht häufig Sinn zu verlangen, dass ein Ensemble auch dann nicht effizient von der Gleichverteilung unterschieden werden kann, wenn die Orakelmaschine zusätzlich Zugriff auf die Inverse der Funktion bekommt:

Definition 2.0.3 (stark pseudozufälliges Permutionsensemble) *Eine Permutationsensemble $P = \{P_n\}_{n\in\mathbb{N}}$ heißt **stark pseudozufällig**, falls für jede effiziente probablistische Orakelmaschine M, jedes Polynom p(x) und genügend großes n gilt:*

$$\left|P[M^{P_n,P_n^{-1}}(1^n) = 1] - P[M^{K_n,K_n^{-1}}(1^n) = 1]\right| < \frac{1}{p(n)}$$

Dabei verwenden wir ein leicht abgewandeltes Prinzip der Orakelmaschine. Dieses kennt nun zwei Arten von Orakelanfragen, nach dem Funktionswert $f(x)$ oder nach dem Urbild eines Wertes $f^{-1}(y)$. Wobei die Maschine keine Anfragen stellt, deren Antwort sie aus den bisherigen Anfragen bereits kennt. Angenommen M macht genau m Anfragen, dann erhalten wir zusammen mit den zugehrigen Antworten also eine Sequenz von Paaren aus einem Wert und dessen Bild unter f: $\sigma := ((x_1,y_1),\ldots,(x_m,y_m))$, wobei $y_i = f(x_i)$. Die Sequenz σ bezeichnen wir als *Transkript* der Maschine.

Im Folgenden betrachten wir bei der Analyse Orakelmaschinen, die in ihrer Laufzeit nicht polynomiell beschränkt sein müssen, wir verlangen nur, dass die Anzahl der Orakelanfragen polynomiell beschränkt ist. Dann können wir ohne Beschr'ankung der Allgemeinheit annehmen, dass die Orakelmaschine immer genau $m(n)$ Anfragen stellt, und das es sich um eine deterministische Orakelmaschine handelt (Wir fixieren das Zufallsband mit der höchsten Erfolgswahrscheinlichkeit).

Dann ist jede Orakelanfrage determinisch aus den bisherige Frage-Antwort-Paaren bestimmt, damit ist für jede Stelle des Transkripts feststellbar, ob nach $f(x_i)$ oder nach $f^{-1}(y_i)$ gefragt wurde. Wir erhalten außerdem die Ausgabe der Maschine als deterministische Funktion aus ihrem Transkript, wir bezeichne diese Funktion mit C_M. Sei f eine Permutation auf $\{0,1\}^n$, dann lässt sich ihr das Transkript zuordnen, dass M bei der Eingabe von f erzeugt wir bezeichnen dieses Transkript mit T_f, wir erhalten also: $M^{f,f^{-1}}(1^n) = C_M(T_f)$.

3 Konstruktionen

3.1 Basiskonstruktion

Ziel ist es aus pseudozufälligen Funktionen, pseudozufällige Permutationen zu erzeugen. Dazu verwendet die folgenden Konstruktionen die Struktur eines Feistelnetzwerks.

Konstruktion 3.1.1 *Sei $f_1, \ldots, f_t : \{0,1\}^n \to \{0,1\}^n$, für $(L,R) \in \{0,1\}^{2n}$ definieren wir*

$$D_{f_1}(R,L) := (R, L \oplus f_1(y))$$

$$D_{f_1,\ldots,f_t}(L,R) := D_{f_1,\ldots,f_{t-1}}(D_{f_t}(x,y))$$

Für jedes Funktionsensemble $F = \{F_n\}_{n\in N}$, und $t : N \to N$ sei $D_F^t = \{D_{F_n}^{t(n)}\}_{n\in\mathbb{N}}$ definiert durch $D_{F_n}^{t(n)} := D_{F_n^{(t(n))},\ldots,F_n^{(1)}}$

Satz 3.1.2 *Für beliebige effizient berechenbare pseudozufällige Funktionenensemble F und jede effizient berechenbare Funktion t, ist $\{D_{F_n}^{t(n)}\}_{n\in\mathbb{N}}$ ebenfalls effizient berechenbar, sowie invertierbar mit einer effizient berechenbaren Inversen.*

Beweis :
Beweis: Die effiziente Berechenbarkeit folgt sofort. Die Inverse lässt sich leicht angeben und ist offenbar ebenfalls effizient berechenbar:

$$D_{f_t}^{-1}(L,R) := (R \oplus f_1(L), L)$$

$$D_{f_1,\ldots,f_t}^{-1}(L,R) := D_{f_1}^{-1}(D_{f_2,\ldots,f_t}^{-1}(L,R))$$

□

D_F^2 ist nicht pseudozufällig, da

$$D_{f_1,f_2}(L,R)|_L \oplus D_{f_1,f_2}(L',R)|_L = L \oplus L'$$

für jede Wahl von f_1, f_2, aber

$$P_{H_n}[K_n(L,R)|_L \oplus K_n(L',R)|_L = L \oplus L'] = \frac{1}{2^n}$$

Wobei $\square|_L$, $\square|_R$, die Projektion auf die linke bzw. auf die rechte Hälfte bezeichnet, daher für $x = (L, R)$ gilt $x|_L = L$. Damit ist D_F^2 jedoch effizient von der Gleichverteilung K unterscheidbar.

Dies motiviert aber zu dem Versuch durch das Vorschalten geeigneter Permutionen den direkten Zugriff der Orakelmaschine auf die Eingaben von D_F^2 zu verhindern und die Wahrscheinlichkeit für eine solche Auswahl von Eingabe zu minimieren. Doch was heißt *geeignet* in diesem Zusammenhang?

Definition 3.1.3 *Ein Permutationsensemble $P := \{P_l\}_{n\in\mathbb{N}}$ heißt **paarweise unabhängig**, falls für $x_1 \neq x_2$ $(P_l(x_1), P_l(x_2))$ die Gleichverteilung auf $\{0,1\}^{2l}$ liefert, also $P_{P_l}[P_l(x_1) = a_1, P_l(x_2) = a_2] = \frac{1}{2^{2l}}$ gilt.*

Beispiel
Das Permutationsensemble, das durch gleichverteilte Wahl von $a, b \in \mathbb{F}_{2^n}$ und durch $f_{(a,b)} : x \mapsto a \cdot x + b$ erzeugt wird ist paarweise unabhängig.

Bemerkung 3.1.4 *Sei $\{P_{2n}\}_{n\in\mathbb{N}}$ ein paarweise unabhängiges Ensemble von Permutationen auf $\{0,1\}^{2n}$. Dann gilt für $x, y \in \{0,1\}^{2n}, x \neq y$:*

$$P[P_n(x)|_R = P_n(y)|_R] \leq 2^{-n}$$

$$P[P_n(x)|_L = P_n(y)|_L] \leq 2^{-n}$$

Beweis :
Wegen der paarweisen Unabhängigkeit gilt ist $(P_n(x), P_n(y))$ gleichverteilt in $\{0,1\}^{4n}$ damit ergibt sich

$$\begin{aligned} & P[P_n(x)|_R = P_n(y)|_R] = P[P_n(x)|_L = P_n(y)|_L] \\ = \quad & \frac{\text{günstige Fälle}}{\text{mögliche Fälle}} = \frac{2^{3n}}{2^{4n}} = 2^{-n} \end{aligned}$$

□

Konstruktion 3.1.5 (Noar-Reingold) *Es sei*

$$W(p_1, f_1, f_2) := D_{f_1 f_2} \circ p_1$$

$$S(p_1, f_1, f_2, p_2) := p_2^{-1} \circ D_{f_1, f_2} \circ p_1$$

Theorem 3.1.6 *Es sei $\{P_{2n}^1\}_{n\in\mathbb{N}}$ ein Permutationsensemble, wobei für $x \neq y$*

$$P[P_n^1(x)|_R = P_n^1(y)|_R] \leq 2^{-n},$$

$\{P_{2n}^2\}_{n\in\mathbb{N}}$ ein Permutationsensemble, wobei für $x \neq y$

$$P[P_n^2(x)|_L = P_n^2(y)|_L] \leq 2^{-n},$$

und $\{F_n\}_{n\in\mathbb{N}}$ ein pseudozufälliges Ensemble von Funktionen, dann ist
$W := \{W(P_n, F_n^{(1)}, F_n^{(2)})\}_{n\in\mathbb{N}}$ ein pseudozufälliges Permutationsensemble und
$S := \{S(P_n^{(1)}, F_n^{(1)}, F_n^{(2)}, P_n^{(1)})\}_{n\in\mathbb{N}}$ ein stark pseudozufälliges Permutationsensemble.

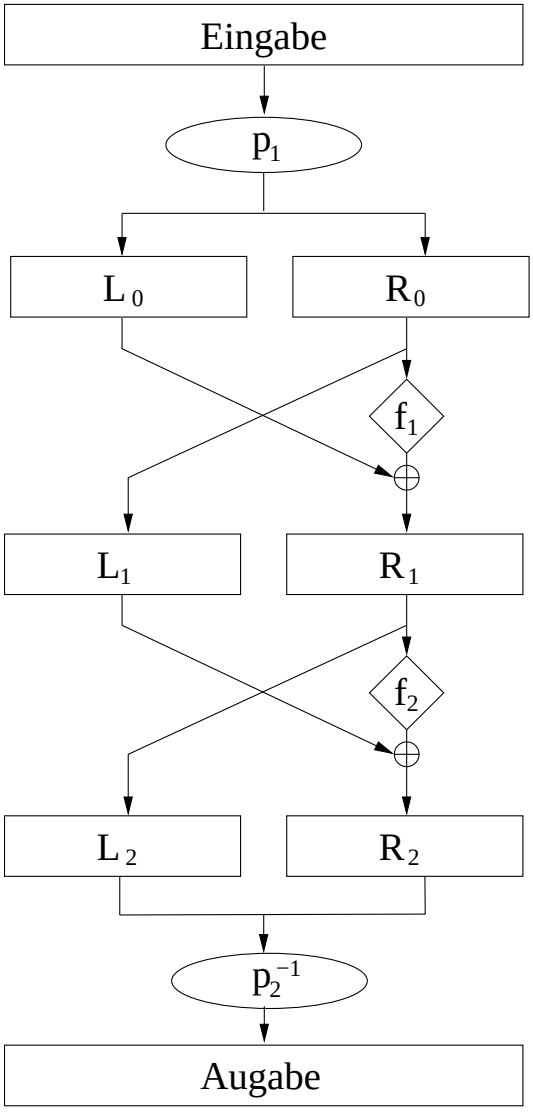

Abbildung 1: Die Konstruktion nach Noar-Reingold

3.2 Pseudozufälligkeit der Basiskonstruktionen

Wir weisen nun die Aussage über die starke Pseudozufälligkeit der Konstruktion S nach. Der Beweis für die Pseudozufälligkeit von W verläuft analog.

Theorem 3.2.1 *Es sei $\{P_n^1\}_{n\in\mathbb{N}}$ ein Permutationsensemble, wobei für $x \neq y$*

$$P[P_n^1(x)|_R = P_n^1(y)|_R] \leq 2^{-n},$$

$\{P_n^2\}_{n\in\mathbb{N}}$ ein Permutationsensemble, wobei für $x \neq y$

$$P[P_n^2(x)|_L = P_n^2(y)|_L] \leq 2^{-n},$$

und $\{H_n\}_{n\in\mathbb{N}}$ das gleichverteilte Ensemble von Funktionen, dann ist $S := \{S(P_n^{(1)}, H_n^{(1)}, H_n^{(2)}, P_n^{(1)})\}_{n\in\mathbb{N}}$ ein stark pseudozufälliges Permutationsensemble.

Theorem 3.1.6 folgt mittel eines einfachen Hybridargumentes aus Theorem 3.2.1.

Korollar 3.2.2 *D_H^4 ist stark pseudozufällig.*

Beweis :
Sei $x \neq y$, dann ist leicht zu verifizieren, dass $P[D_{H_n}(x)|_R = D_{H_n}(y)|_R] \leq 2^{-n}$ und $P[D_{H_n}^{-1}(x)|_R = D_{H_n}^{-1}(y)|_R] \leq 2^{-n}$. Demzufolge erfüllt $D_H^4 = S(D_H^{(1)}, D_H^{(2)}, D_H^{(3)}, {D^{(4)}}_H^{-1})$ die Voraussetzungen von Theorem 3.2.1.

□

Wir wollen nun zeigen, dass unser konstruiertes Ensemble S nicht effizient von der Gleichverteilung über den Permutationen unterscheidbar ist, das zeigen wir nicht direkt, sondern über den Zwischenschritt, dass beiden Ensembles nicht effizient vom gleichverteilten Funktionsensemble H unterscheidbar sind.

Definition 3.2.3 *CON sei die Menge der **konsistente Transkripte**, daher derjenigen, die von einer Permutation erzeugt werden können, daher $((x_1, y_1), \ldots, (x_m, y_m)) \in CON \Leftrightarrow \forall_{1\leq i<j\leq n} x_i \neq x_j \wedge y_i \neq y_j$*

Satz 3.2.4 *Sei, m die maximale Anzahl der Orakelanfragen von M, dann gilt:*

$$\left| P[M^{H_l,H_l}(1^l) = 1] - P[M^{K_l,K_l^{-1}}(1^l) = 1] \right| < \frac{m^2}{2^{2n+1}}$$

Beweis :
Sei σ ein konsistentes Transkript und T_{H_l}, T_{K_l} die durch die Verteilung der Funktionen erzeugte Verteuling auf den Transkripten, dann gilt offenbar $P[T_{K_n} = \sigma] = \frac{(2^l-m)!}{2^l!}$. Denn es gibt genau $(2^l - m)!$ von insgesamt $2^l!$ Permutationen auf $\{0,1\}^l$, die das Transkript σ erzeugen, da deren Werte dann auf den m Stellen von σ festgelegt sind.

Außerdem gilt $P[T_{H_n} = \sigma] = \frac{1}{2^l(2^l-1)\ldots(2^l-m+1)} = \frac{(2^l-m)!}{2^l}$. Denn bei der $(i+1)$-ten Anfrage der Orakelmaschine gibt $2^l - i$ zum bisherigen Transkript konsistente Antworten, damit gibt es insgesamt $2^l(2^l-1)\ldots(2^l-m+1)$ konsistente Transkripte, die wegen der *Unvorhersagbarkeit* von H gleichverteilt auftreten.

Damit ist die Verteilung von T_{H_l} und T_{K_l}, unter der Vorrausetzung der Konsistenz exakt die gleiche. Die Wahrscheinlichkeit des Auftretens von inkonsistenten Transkripten bei H ist jedoch gering: Ein Transkript $((x_1, y_1), \ldots, (x_m, y_m))$ ist inkonsistent, falls

$1 \leq i < j \leq n$ existieren, so dass $x_i = x_j$ oder $y_i = y_j$. Für zwei feste i, j tritt dies, höchstens mit Wahrscheinlichkeit 2^{-l} ein. Wir haben daher:

$$P[T_{H_n} \notin CON] < \left(\begin{array}{c} m \\ 2 \end{array} \right) \cdot 2^{-l} < \frac{m^2}{2^{l+1}} \tag{1}$$

Damit folgt:

$$\begin{aligned} & \left| P[M^{H_l^{(1)}, H_l^{(2)}}(1^n) - P[M^{K_l, K_l^{-1}}] \right| \\ = \; & |P[C_M(T_{H_l}) = 1] - P[C_M(T_{K_l}) = 1]| \\ \leq \; & |P[C_M(T_{H_l}) = 1 | T_{H_l} \in CON] - P[C_M(T_{K_l}) = 1]| \, P[T_{H_l} \in CON] \\ & + |P[C_M(T_{H_l}) = 1 | T_{H_l} \notin CON] - P[C_M(T_{K_l}) = 1]| \, P[T_{H_l} \notin CON] \\ \leq \; & P[T_{H_l} \notin CON] < \frac{m^2}{2^{l+1}} \end{aligned}$$

□

Korollar 3.2.5 *Das gleichverteilte Permutationsensemble K ist selbst ein pseudozufälliges Funktionsensemble.*

Beweis :
Durch Einschränkung auf Orakelmaschinen, die nur Anfragen nach $f(x)$ stellen erhält man die Behauptung.

□

Korollar 3.2.6 *Jedes pseudozufällige Permutationsensemble P ist auch ein pseudozufälliges Funktionsensemble.*

Beweis :
Sei $p(x) = \frac{q(x)}{2}$ beliebiges Polynom, dann gilt für genügend große n nach Voraussetzung und Satz 3.2.4:

$$\left| P[M^{P_l}(1^l) - P[M^{K_l}(1^l) = 1] \right| < \frac{1}{q(l)}$$

$$\left| P[M^{K_l}(1^l) = 1] - P[M^{H_l}(1^l) = 1] \right| < \frac{1}{q(l)}$$

Dann gilt weiter:

$$\begin{aligned} & \left| P[M^{P_l}(1^l) = 1] - P[M^{H_l}(1^l) = 1] \right| \\ \leq \; & \left| P[M^{P_l}(1^l) = 1] - P[M^{K_l}(1^l) = 1] \right| + \left| P[M^{K_l}(1^l) = 1] - P[M^{H_l}(1^l) = 1]] \right| \\ < \; & \frac{1}{q(l)} + \frac{1}{q(l)} = \frac{1}{p(l)} \end{aligned}$$

□

Es bleibt jetzt die Unterscheidbarkeit zwischen T_H und T_S zu untersuchen. Die Intutition dabei ist, dass für konsisente Transkripe σ, die Wahrscheinlichkeit für $T_{S_n} = \sigma$ mit der von $T_{H_n} = \sigma$, fast immer üebereinstimmt. Fast immer heißt hier bis auf eine ungünstige Wahl von von σ und einer Ausprägung von S. Die Hoffnung dabei ist, das diese ungünstige Wahl nur mit einer geringen Wahrscheinlichkeit auftritt. Nun formalisieren wir dieses Ereignis.

Definition 3.2.7 *Für p_1, p_2 Permutationen auf $\{0,1\}^{2n}$ sei $BAD(p_1, p_2)$ die Menge der Transkripte $\sigma := ((x_1, y_1), \ldots, (x_m, y_m))$, für die gilt $\exists_{1 \leq i < j \leq m} : p_1(x_i)|_R = p_1(x_j)|_R \vee p_2(y_i)|_L = p_2(y_j)|_L$.*

Lemma 3.2.8 *Sei* $\{P^1_{2n}\}_{n\in\mathbb{N}}$ *ein Permutationsensemble mit*

$$P[P^1_{2n}(x)|_R = P^1_{2n}(y)|_R] \leq 2^{-n},$$

$\{P^2_{2n}\}_{n\in\mathbb{N}}$ *ein Permutationsensemble mit*

$$P[P^2_{2n}(x)|_L = P^2_{2n}(y)|_L] \leq 2^{-n},$$

für $x \neq y$*, dann gilt für jedes M-Transkript* $\sigma = ((x_1, y_1), \ldots, (x_m, y_m))$*:*

$$P[\sigma \in BAD(P^1_n, P^2_n)] < \frac{m^2}{2^n}$$

Beweis :
Sei p_1 ein Wert von P^1 und p_2 Wert von P^2. Nach Definition ist $\sigma \in BAD(p1, p2)$, wenn $1 \leq i < j \leq m$ existieren, so dass $p1(x_j)|_R = p1(x_i)|_R$ oder $p_2(y_j)|_R = p_2(y_i)|_R$. Daraus folgt:

$$P[\sigma \in BAD(P^1_n, P^2_n)] \leq \left(\begin{array}{c} m \\ 2 \end{array} \right) \cdot 2 \cdot 2^{-n} < \frac{m^2}{2^n}$$

□

Lemma 3.2.9 *Für ein konsistentes M-Transkript* $\sigma = ((x_1, y_1), \ldots, (x_m, y_m))$ *gilt.*

$$P_S[T_{S_{2n}} = \sigma | \sigma \notin BAD(P^1_{2n}, P^2_{2n})] = P[T_{H_{2n}} = \sigma]$$

Beweis :
Eine Funktion f erzeugt das M-Transkript σ, genau dann, wenn $f(x_i) = y_i$, bzw $f(y_i) = x_i$ gilt, abhängig davon was die i-te Anfrage von M war. Jedenfalls sind durch σ m Werte von f vorgeben, damit erzeugen $(2^{2n})^{2n-m}$ Funktionen, von insgesamt $(2^{2n})^{2n}$ Funktionen das M-Transkript σ. Da H das gleichverteilt ist ergibt sich

$$P[T_{H_{2n}} = \sigma] = \frac{(2^{2n})^{2n-m}}{(2^{2n})^{2n}} = \frac{1}{2^{2nm}}$$

Sei S eine Permutation $S = S(p_1, f_1, f_2, p_2)$ und $\sigma \notin BAD(p1, p2)$. S erzeugt σ, daher $T_S = \sigma$, genau dann, wenn $\forall_{1\leq i\leq n} : y_i = S(x_i)$. Nach Definition von S gilt:

$$y_i = S(x_i) \Leftrightarrow f_1(R^2_i) = L^0_i \oplus L^2_i \wedge f_2(L^2_i) = R^0_i \oplus R^2_i$$

Für alle $1 \leq i \leq m$ gilt $R^0_i \neq R^0_j$ und $L^2_i \neq L^2_j$ (wegen $\sigma \notin BAD(p_1, p_2)$). Daher ist sowohl f_1 als auch f_2 an m Stellen durch σ festgelegt. Damit erfüllen jeweils 2^{n2^n-m} von 2^{n2^n} Funktionen die Bedingungen an f_1 bzw f_2. Damit ist wegen der Gleichverteilung der f_1, f_2 die Wahrscheinlichkeit, dass füer eine Auswahl von f_1, f_2 $T_{S(p_1,f_1,f_2,p_2)} = \sigma$ gleich $\frac{(2^n)^{2^n-m}\cdot(2^n)^{2^n-m}}{(2^n)^{2^n}\cdot(2^n)^{2^n}} = \frac{1}{2^{2nm}}$ damit gilt auch:

$$P_S[T_S = \sigma | \sigma \notin BAD(P^1_n, P^2_n)] = \frac{1}{2^{2nm}}$$

□

Beweis (Theorem 3.2.1):
Sei $\Gamma := \{\sigma | C_M(\sigma) = 1\}$ die Menge der akzeptierenden und konsistenten M-Transkripte. Dann gilt offenbar:

$$
\begin{aligned}
& \left| P[M^{S_{2n}, S_{2n}^{-1}}] - P[M^{H_{2n}^{(1)}, H_{2n}^{(2)}}(1^n)] \right| \\
= \; & |P[C_M(T_{S_{2n}} = 1] - P[C_M(T_{H_{2n}}) = 1]| \\
\leq \; & \left| \sum_{\sigma \in \Gamma} (P[T_{S_{2n}} = \sigma] - P[T_{H_{2n}} = \sigma]) \right| + P[\sigma \notin CON] && (2) \\
\leq \; & \sum_{\sigma \in \Gamma} \left| P[T_{S_{2n}} = \sigma | \sigma \notin BAD(P_{2n}^1, P_{2n}^2)] - P[T_{H_{2n}} = \sigma] \right| \cdot P[\sigma \notin BAD(P_{2n}^1, P_{2n}^2)] && (3) \\
+ \; & \left| \sum_{\sigma \in \Gamma} \left(P[T_{S_{2n}} = \sigma | \sigma \in BAD(P_{2n}^1, P_{2n}^2)] - P[T_{H_{2n}} = \sigma] \right) \cdot P[\sigma \in BAD(P_{2n}^1, P_{2n}^2)] \right| && (4) \\
+ \; & P[T_{H_{2n}} \notin CON] && (5)
\end{aligned}
$$

Aus dem Beweis von Satz 3.2.4, Ungleichung (1) wissen wir bereit, dass Ausdruck (5) kleiner als $\frac{m^2}{2^{2n+1}}$. Lemma 3.2.9 besagt, dass (3) = 0. Daher bleibt Ausdruck (4) abzuschätzen. Sei o.B.d.A.

$$
\begin{aligned}
& \sum_{\sigma \in \Gamma} P[T_{S_{2n}} = \sigma | \sigma \in BAD(P_{2n}^1, P_{2n}^2)] \cdot P[\sigma \in BAD(P_{2n}^1, P_{2n}^2)] \\
\leq \; & \sum_{\sigma \in \Gamma} P[T_{H_{2n}} = \sigma] \cdot P[\sigma \in BAD(P_{2n}^1, P_{2n}^2)]
\end{aligned}
$$

Mit Lemma 3.2.8 erhalten wir:

$$
\begin{aligned}
& + \left| \sum_{\sigma \in \Gamma} \left(P[T_{S_{2n}} = \sigma | \sigma \in BAD(P_{2n}^1, P_{2n}^2)] - P[T_{H_{2n}} = \sigma] \right) \cdot P[\sigma \in BAD(P_{2n}^1, P_{2n}^2)] \right| \\
\leq \; & \sum_{\sigma \in \Gamma} P[T_{H_{2n}} = \sigma] \cdot P[\sigma \in BAD(P_{2n}^1, P_{2n}^2)] \\
\leq \; & \max_{\sigma \in \Gamma} [\sigma \in BAD(P_{2n}^1, P_{2n}^2)] \\
< \; & \frac{m^2}{2^n}
\end{aligned}
$$

Wir erhalten also:

$$
|P[M^{S_{2n}, S_{2n}^{-1}}] - P[M^{H_{2n}^{(1)}, H_{2n}^{(2)}}(1^n)]| < \frac{m^2}{2^n} + \frac{m^2}{2^{2n+1}}
$$

Zusammen mit Satz 3.2.4 ergibt sich:

$$
\begin{aligned}
& \left| P[M^{S_{2n}, S_{2n}^{-1}} = 1] - P[M^{K_{2n}, K_{2n}^{-1}}(1^n) = 1] \right| \\
\leq \; & \left| P[M_{2n}^{S_{2n}, S^{-1}} = 1] - P[M^{H_{2n}^{(1)}, H_{2n}^{(2)}}(1^n) = 1] \right| \\
& + \left| P[M^{K_{2n}, K_{2n}^{-1}}(1^n) = 1] - P[M^{H^{(1)}, {}_{2n} H_{2n}^{(2)}}(1^n) = 1] \right| \\
< \; & \frac{m^2}{2^n} + \frac{m^2}{2^{2n}}
\end{aligned}
$$

□

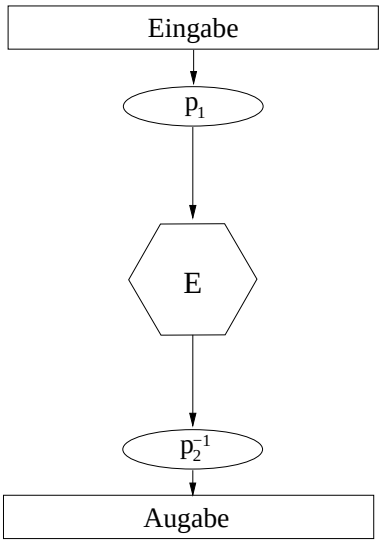

Abbildung 2: Die Struktur der allgeinen Konstruktion

3.3 Schema zur Konstruktion Pseudozufallspermutationen

Noar und Reingold [MN96] haben die bisherige Konstruktion auf mehrfache Art und Weise abgewandelt. Dabei benutzten sie eine allgemeine Struktur zum Entwurf pseudozufälliger Permuationen, die eine einfache Verallgemeinerung der bisher erbrachten Beweise ermöglicht. Die Struktur arbeitet dabei auf einer l Bits und besteht aus der drei verketteten Permutationen: $S := p_2^{-1} \circ E \circ p_1$.

E ist dabei eine Konstruktion aus pseudozu fälligen Funktionen, für die Analyse gehen wir analog dem obigen Beweisen dabei zuersteinmal von echt zufälligen Funktionen aus. E kann dabei einfach von dem gleichverteilten Ensemble unterscheidbar sein, erst durch die Verkettung mit p_2, P_2^{-1} entsteht eine pseudozufällige Permutation. Im vorherigen Abschnitt war $E = D_F^2$ gewählt, $l = 2n$ und p_1, p_2 konnten zum Beispiel aus einem paarweise unabhängigen Ensemble gewählt werden. E besitzt dabei aber die Eigenschaft, "für fast jede Sequenz" $((x_1, y_1), \ldots, (x_m, y_m))$, die Wahrscheinlichkeit, dass $\forall_{1 \leq i \leq m} : E(x_i) = y_i$ nahe der des gleichverteilten Ensembles liegt.

Definition 3.3.1 $GOOD_E := \{((x_1, y_1), \ldots, (x_m, y_m)) | P\left[\forall_{1 \leq i \leq m} : E(x_i) = y_i\right] = 2^{-lm}\}$, *sei die Menge der für E gutartigen Sequenzen.*

Wir nehmen dabei an, dass "fast alle" Sequenzen diese Bedingung erfüllen. Die Aufgabe von p_1, p_2 ist dann, dass beliebig gewählte Eingaben so transformiert werden, dass mit hoher Wahrscheinlichkeit eine für E gutartige Seqenz erzeugt wird. Betrachtet man die Beweisstruktur für Theorem 3.2.1, dann stellt man fest die Abschätzung in Satz 3.2.4 für die Distanz zwischen H und K, durch die Wahrscheinlichkeit inkonsistenter M-Transkripte weiterhin gültig bleibt. Kernaufgabe des Framworks wird es sein die "schlechten Transkripte" neu zu charakterisieren und zwar in Abhängigkeit von der Wahl der p_1, p_2 und von E.

Eine neue Menge $BAD_E(p_1, p_2)$ ersetzt dann $BAD(p_1, p_2)$. $BAD_e(p_1, p_2)$ ist eine Menge von konsistenten M-Transkripten $\sigma = ((x_1, y_1), \ldots, (x_m, y_m))$ mit der Eigenschaft $((p_1(x_1), p_2(y_1)), \ldots, (p_1(x_m), p_2(y_m)) \in GOOD_E$.

Dan beachte, dass Definition 3.2.7 einen Spezialfall der obigen Definition für $E = D_F^2$ darstellt. Genauso wie für BAD gilt nach Definition auch für BAD_E

$$P[T_{S_l} = \sigma | \sigma \notin BAD_E(p_1, p_2)] = 2^{-lm}$$

Dies impliziert, das Lemma 3.2.9 auch für BAD_E anstelle von BAD gilt:

Lemma 3.3.2 *Sei $\sigma = ((x_1, y_1), \ldots, (x_m, y_m))$ ein konsistentes M-Transkript, dann gilt:*

$$P[T_{S_l} = \sigma | \sigma \notin BAD_E(p_1, p_2)] = P[T_{H_l} = \sigma]$$

Damit wir gemäß der alten Beweisstruktur den beabsichtigten Schluss

$$\left| P[M^{S_l, S_l^{-1}} = 1] - P[M^{K_l, K_l^{-1}} = 1] \right| < \epsilon + \frac{m^2}{2^l}$$

ziehen können, benötigen wir noch

$$P[\sigma \in BAD_E(P^1, P^2)] < \epsilon \tag{6}$$

Diese Eigenschaft BAD_E muss dann im Einzelfall, äuivalent zum Lemma 3.2.8 nachgewiesen werden. Dies ist die einzige Stelle in dem Schema, an der die Definition von ϵ und von P^1, P^2 eine Rolle spielt. Im konkreten Fall gilt es hier eine zwischen der Güte der Abschätzung und der Stärke der Anforderungen an P^1, P^2 abzuwägen.

In Anwendung von (6) und Lemma 3.3.2 können wir dann analog zum Beweis von Theorem 3.2.1 schließen:

Theorem 3.3.3 *Seien P^1, P^2 Ensembles von Permutationen, $S := (P^2)^{-1} \circ E \circ P^1$ und sei K das gleichverteilte Permutationsensemble. Falls dann*

$$P[\sigma \in BAD_E(P^1, P^2)] < \epsilon,$$

so gilt:

$$\left| P[M^{S_l, S_l^{-1}} = 1] - P[M^{K_l, K_l^{-1}} = 1] \right| < \epsilon + \frac{m^2}{2^l}$$

Der entscheidende Schritt beim Beweis für eine konkrete Konstruktion ist, dass die Menge $BAD_E(p_1, p_2)$ so beschaffen ist, dass zugleich $P[\sigma \in BAD_E(P^1, P^2)] < \epsilon$ und $P[T_{S_n} = \sigma | \sigma \notin BAD_E(p_1, p_2)] = P[T_{H_n} = \sigma]$. Bei der Konstruktion gemäß dem beschriebenen Schema ist daher darauf zu achten, dass die Wahl von E eine "ausreichend" große Anzahl gutartiger Sequenzen besitzt.

3.4 Einschränkung der Vorraussetzungen

Die Tatsache, dass die Konstruktion von unabhähngig gewählten pseudozufälligen Funktionen ausgeht benötigt man eine entsprechend große Anzahl echter Zufallsbits. Es wäre daher wünschenswert eine ähnliche Konstruktion zu finden, die mit nur einer pseudozufälligen Funktion auskommt. Jedoch zeigt sich, dass beispielsweise $\{D_{F_n, F_n, F_n}\}_{n \in \mathbb{N}}$ im Gegensatz zu $D_F^3 := \{D_{F_n^{(1)}, F_n^{(2)}, F_n^{(3)}}\}_{n \in \mathbb{N}}$ nicht pseudozufällig ist. [Rue89]

Satz 3.4.1 *Sei $\{P_{2n}\}_{n \in \mathbb{N}}$ ein paarweise unabhängiges Ensemble von Permuationen und $\{F_n\}_{n \in \mathbb{N}}$ ein pseudozufälliges Ensemble von Funktionen, dann ist $S := \{S(P_{2n}^{(1)}, F_n, F_n, P_{2n}^{(1)})\}_{n \in \mathbb{N}}$ ein stark pseudozufälliges Permutationsensemble.*

Beweis (Schema):
Der Beweis richtet sich nach dem Schema aus Abschnitt 3.3, dabei ist $E_{2n} := D_{F,F}$ und

$$\begin{aligned} BAD_E(p_1, p_2) \quad := \quad & \{((x_1, y_1), \ldots, (x_m, y_m)) | \exists_{1 \le i < j \le m} : p_1(x_i)|_R = p_2(y_j)|_L\} \\ & \cup BAD(p_1, p_2) \end{aligned}$$

Es gilt dann:

$$P[T_{S_{2n}} = \sigma | \sigma \notin BAD_E(p_1, p_2)] = P[T_{H_n} = \sigma]$$

$$P[\sigma \in BAD_E(P^1, P^2)] < \frac{2m^2}{2^n}$$

damit können wir schließen:

$$\left| P[M^{S_{2n}, S_n^{-1}} = 1] - P[M^{K_{2n}, K_{2n}^{-1}} = 1] \right| < \frac{2m^2}{2^n} + \frac{m^2}{2^{2n}}$$

□

Ohne Beweis werden folgende weitere Resultate angeben:

Für alle $i, j, k \in \mathbb{N}$ ist $\{D_{F_n^i, F_n^j, F_n^k}\}_{n \in \mathbb{N}}$ nicht pseudozufällig

Für alle $i, j, k, l \in \mathbb{N}$ ist $\{D_{F_n^i, F_n^j, F_n^k, F_n^l}\}_{n \in \mathbb{N}}$ nicht starl pseudozufällig

$\{D_{F_n^2, F_n, F_n, F_n}\}_{n \in \mathbb{N}}$ ist pseudozufällig

$\{D_{F_n, I, F_n^2, F_n, I_n, F_n^2}\}_{n \in \mathbb{N}}$ ist stark pseudozufällig, wobei I_n die Identität

$\{D_{F_n \circ \xi \circ F_n, F_n, F_n}\}_{n \in \mathbb{N}}$ ist pseudozufällig, wobei ξ die bitweise Rotation

3.5 Blockweise Anwendung

kommt noch...

4 Anwendung als Kryptosysteme

Wie bereits angemerkt formalisieren Pseudozufällige Permutationen in gewisser Weise die Intuition ein Blockchiffre. Diese Festtellung soll im folgenden Ausgeführt werden. Desweiteren wird untersucht inwieweit sich die Eigenschaft des Permutationsensembles pseudozufällig beziehungsweise stark pseudozufällig zu sein auf die Sicherheit des entsprechenden Kryptosystems auswirkt.

4.1 Definitionen

Definition 4.1.1 (Symmetrische Blockchiffre) *Eine **symmetrisches Blockchiffre** ist ein Tupel (l, A, K, E, D), wobei*

$l \in \mathbb{N}$ *die Blocklänge des Systems*

A *ein beliebiges Alphabet, Wobei Klar- und Kryptotextmenge $M = A^l$*

K *eine Schlüsselmenge*

$E : K \times M \to M$ *eine Verschlüsselungsfunktion, $D : K \times M \to M$ eine Entschlüsselungsfunktion. E und D sind effizient berechenbar und es gilt $\forall_{k \in K, x \in M} : x = D(k, E(k, x))$*

Bemerkung 4.1.2 *Sei P ein effizient berechenbares und invertierbares Permutationenensemble. Sei $l \in \mathbb{N}$ beliebig, $A' := \{0, 1\}$ und $K' \subset 0, 1^p(n)$ die Menge der möglichen Werte des Algorithmus I aus 2.0.2. Sei weiter $E'(k, x) := V(k, x)$ und $D'(k, y) := N(k, y)$. Dann ist offenbar (A', l, K', E', D') eine symmetrische Blockchiffre.*

Es stellt sich nun die Frage nach der Sicherheit eines solchen Kryptosystems. Dazu muss formalisiert werden, was unter einem Angriff auf das System zu vertehen ist, um dann die Erfolgswahrscheinlichkeit eines solchen Angriffs abschätzen zu können. Die nun folgenden Definitionen sind Luby [Lub96] entlehnt.

Definition 4.1.3 (Adaptiver Choosen-Plaintext-Angriff) *Unter einem allgemeinen* ***Adaptiv-Choosen-Plaintext-Angriff*** *versteht man drei probablistische Algorithmen:*

$$M : \mathbb{N} \times M^* \times M^* \to M^*$$
$$P : M^* \times M^* \to M^2$$
$$A : M^* \times M^* \times M^2 \times M \to \{0,1\}$$

Der Angriff läuft folgendermaßen ab:

(Schlüsselwahl) Es sei ein $k \in K$ *gleichverteilt gewählt*

(Plaintext-Angriff) Definiere e_j *und* m_j *induktiv:* $m_0 = e_0 = \varepsilon$, $m_{j+1} := (m_j, M(j, m_j, e_j))$, $e_{j+1} = (e_j, E(k, m_{j+1}))$ *und* $m := m_{p(n)}$ *und* $e := e_{p(n)}$.

(Wähle geheime Nachricht) Sei $(m^0, m^1) = P(m, e)$ *ein Paar von Nachrichten, das von* P *erzeugt wurde und das nicht in* m *enthalten ist.* $b \in \{0,1\}$ *sei gleichverteilt gewählt und* $e^b := E(k, m^b)$.

(Vorhersage) Die Erfolgswahrscheinlichkeit des Angriffs ist dann

$$P[A(m, e, m^0, m^1, e^b) = b]$$

und es sei

$$\delta(n) := P[A(m, e, m^0, m^1, e^b) = b] - \frac{1}{2}$$

Die Intuition hinter diesem Angriff ist, dass der Angreifer in der Lage ist ohne Kenntnis des Schüssels Nachrichten zu verschlüsseln. Abhängig von den bisherigen Ergebnissen kann er neue Klartexte wählen. Nach einer Anzahl von $p(n)$ Schritten. Darf sich der Angreifer zwei neue Klartexte vorgeben. Und erhält einen Kryptotext. Nun muß er versuchen diesen Kryptotext einem der beiden Klartexte zuzuordnen.

Definition 4.1.4 (Simultaner Plaintext/Cryptotext-Angriff) *Unter einem allgemeinen* ***Simultaner Plaintext/Cryptotext-Angriff*** *versteht man drei probablistische Algorithmen:*

$$M : \mathbb{N} \times M^* \times M^* \to \{+,-\} \times M^*$$
$$P : M^* \times M^* \to M^2$$
$$A : M^* \times M^* \times M^2 \times M \to \{0,1\}$$

Der Angriff läuft folgendermaßen ab:

(Schlüsselwahl) Es sei ein $k \in K$ *gleichverteilt gewählt*

(Plaintext-Angriff) Definiere e_j *und* m_j *induktiv:* $m_0 = e_0 = \varepsilon$, *wenn dann* $M(j, m_j, e_j) = (+, x)$, *so ist* $m_{j+1} = (m_j, x), e_{j+1} := (e_j, E(k, x))$, *wenn* $M(j, m_j, e_j) = (-, y)$ *ist* $e_{j+1} := (e_j, y), m_{j+1} := (m_j, D(k, y))$ *Es sei dann weiter* $m := m_{p(n)}$ *und* $e := e_{p(n)}$

(Wähle geheime Nachricht) Sei $(m^0, m^1) = P(m, e)$ *ein Paar von Nachrichten, das von* P *erzeugt wurde und das nicht in* m *enthalten ist.* $b \in \{0,1\}$ *sei gleichverteilt gewählt und* $e^b := E(k, m^b)$.

(Vorhersage) Die Erfolgswahrscheinlichkeit des Angriffs ist dann

$$P[A(m, e, m^0, m^1, e^b) = b]$$

und es sei

$$\delta(n) := P[A(m, e, m^0, m^1, e^b) = b] - \frac{1}{2}$$

Dieser Angriff entspricht weitgehend dem Adaptiv-Choosen-Plaintext-Angriff aus Deinition 4.1.3 mit dem wesentlichen Unterschied, dass der Angreifer nun sowohl entschlüssen als auch verschlüssen kann.

4.2 Analyse

Wir wollen die Sicherheit von pseudozufälligen Permutationen gegen die oben definierten Angriffstypen analysieren. Eine sinnvolle Analyse ergibt sich nur dann, wenn man die Erfolgswahrscheinlichkeit in Abhängigkeit von der verwendeten Blocklänge untersucht. Wir erhoffen uns dann, dass für genügend große $n \in \mathbb{N}$ und effiziente (daher polynomielle) Angriffe die Erfolgswahrscheinlichkeit sich nicht mehr wesentlich vom vom zufälligen Raten, daher $\frac{1}{2}$ unterscheidet. Versuchen wir dies zu formalisieren.

Definition 4.2.1 *Ein Kryptosystem heißt* ***S(n)-sicher*** *gegen einen der oben definierten Angriffe, wenn für genügend große* $n\mathbb{N}$ $\frac{Time(n)}{|\delta(n)|} > S(n)$ *wobei* $Time(n)$ *die Gesamtlaufzeit des Angriffs, daher* $Time(n) = p(n) \cdot Time_M(n) + Time_P(n) + Time_A(n)$.
Ein Kryptosystem heißt ***sicher****, wenn es S(n)-sicher für jedes Polynom S ist.*

Satz 4.2.2 *Sei P ein (stark) pseudozufälliges Permutationsensemble, dann ist das Kryptosystem der durch P induzierten Blockchiffren sicher gegen adaptive Choosen-Plaintext (simultane Plaintext/Cryptotext)-Angriffe.*

Beweis (Skizze):
Sei das Kryptosystem nicht sicher, dann gibt es ein Polynom q, mit $q(n) \geq \frac{Time(n)}{|\delta(n)|}$ für unendlich viele n, daher ist $|\delta(n)| \geq \frac{Time(n)}{q(n)}$. Sei nun M eine probablistische Orakelmmaschine, die folgendermaßen arbeitet:

Wähle $b \in \{0,1\}$ gleichverteilt.

Falls $b = A(m, e, m^0, m^1, e^b)$ gib 1 aus, sonst 0.

dann ergibt sich wegen

$$P[M^{K_n}_{M,P,A}(1^n) = 1] = \frac{1}{2}$$

$$\left| P[M^{P_n}_{M,P,A}(1^n) = 1] - P[M^{K_n}_{M,P,A}(1^n) = 1] \right| = \delta(n) \geq \frac{Time(n)}{q(n)}$$

Dies steht jedoch im Widerspruch zu der Annahme, dass P pseudozufällig.

□

5 Ausblick

kommt noch...

Literatur

[Gol95] O. Goldreich. *Foundations of Cryptography(fragments of a book).* Online-Publikation: http://www.ecccuni-trier.de/eccc/info/ECCC-Books/eccc-books.html, 1995.

[Lub96] M. Luby. *Pseudorandomness and Applications.* Princton University Press, 1996.

[MN96] O. Reingold M. Noar. On the construction of pseudo-random permutations: Luby-rackoff revisited. Technical report, Weizmann Institute of Science, 1996.

[Rue89] R. A. Rueppel. On security of schnorr's pseudo random generator. Lecture Notes in Computer Science. Springer-Verlag, 1989. 432 ff.