Blockchiffren

Def: Eine Blockchiffre mit Blocklänge n ist ein symmetrisches Verschlüsselungssystem mit $M = C = A^n$ und A endlich.

Prop: Die Algorithmen $\mathcal{E}$ und $\mathcal{D}$ einer Blockchiffre definieren bijektive Funktionen.

Bew: Wegen $\mathcal{D}(k, \mathcal{E}(k, m)) = m$ wird keinen zwei Nachrichten der gleiche Chiffretext zugeordnet. Wegen $C = M$ und A endlich wird jeder Nachricht daher genau ein Chiffretext zugeordnet. □

Bezeichnet $S(A^n)$ die Menge aller bijektiven Abbildungen (Permutationen) von A^n in sich, so gilt also $\mathcal{E}(k, \cdot), \mathcal{D}(k, \cdot) \in S(A^n)$.

Weiter erhalten wir Abbildungen $K \to S(A^n)$, $k \mapsto \mathcal{E}(k, \cdot)$ bzw. $k \mapsto \mathcal{D}(k, \cdot)$.

Blockchiffren

Beispiele:

- Die affin-linearen Chiffren (historisch).
- Substitutionchiffren (definiert durch koordinatenweise Anwendung eines Elements aus $S(A)$ auf A^n, historisch).
- DES (Data Encryption Standard, 1977, veraltet).
- AES (Advanced Encryption Standard, 2001, aktuell).

Ideale Blockchiffren

Designziel / Konstruktion eines idealen Blockchiffre:

- Für jedes $k \in K$ wähle $f_k \in S(A^n)$ gleichverteilt zufällig und definiere $\mathcal{E}(k, \cdot) := f_k$.
- $\mathcal{E}$ hat dann keine besondere (mathematische) Struktur, die durch einen Angreifer ausgenutzt werden kann.

Problem: Algorithmisch nicht praktikabel / machbar.

- Eine Beschreibung beliebiger Elemente aus $S(A^n)$ benötigt mindestens $\log_2(\#S(A^n))$ Bits für jedes Element. Es gilt $\log_2(\#S(A^n)) \geq \#A^n(\log_2(\#A^n) - 1/\log(2))$. Zu groß!
- Beliebige Elemente aus $S(A^n)$ können daher im Prinzip nur durch die Angabe aller Urbild-Bild Paare beschrieben werden.
- Ebenfalls zu viele k!

Ideale Blockchiffren

Blockchiffren in der Praxis:

- Man beschränkt sich auf eine relativ gesehen kleine Teilmenge von $S(A^n)$, welche als Ver- und Entschlüsselungsfunktionen vorkommen ($\#A^n$ viele).
- Diese Teilmenge muß durch den Schlüssel k algorithmisch (effizient) definierbar sein, d.h. aus k muß sich effizient die Funktion $\mathcal{E}(k, \cdot)$ bestimmen lassen, und die Berechnung von Bildwerten unter $\mathcal{E}(k, \cdot)$ muß ebenfalls effizient möglich sein.
- Teilmenge erfordert also eine ganz spezielle Beschreibung, soll aber trotzdem möglichst „zufällig aus $S(A^n)$ gewählt aussehen“.

Ideale Blockchiffren

Idealisierte Emulation einer zufällig gewählten Funktion aus $f \in S(A^n)$ durch ein Orakel:

- Algorithmen befragen das Orakel nach Funktionswerten $f(m)$.
- Nach Eingabe von m wählt das Orakel einen zufälligen Wert c, speichert $f(m) = c$ in einer Liste und liefert c als $f(m)$ zurück.
- Falls $f(m)$ ein zweites Mal erfragt wird, verwendet das Orakel den Wert aus der Liste.

Genauere Erläuterung von „zufällig aus $S(A^n)$ gewählt aussehen“:

Ein Angreifer A kommuniziert mit Orakel O im folgenden Spiel.

- O wählt ein zufälliges Bit b und ein zufälliges k.
- Danach wählt A eine Reihe beliebiger m_i und erfragt die Funktionswerte $\mathcal{E}(k, m_i)$ und $f(m_i)$ von O.

Ideale Blockchiffren

- O liefert die die Antworten $(\mathcal{E}(k, m_i), f(m_i))$ falls $b = 0$ und $(f(m_i), \mathcal{E}(k, m_i))$ falls $b = 1$.
- A gewinnt, wenn A mit Wahrscheinlichkeit $\neq 1/2$ den Wert von b herausfinden kann.

Wenn es keinen erfolgreichen Angreifer gibt, dann können die Funktionen $\mathcal{E}(k, \cdot)$ für zufällig gewähltes k nicht von echt zufällig gewählten Funktionen unterschieden werden.

Entsprechend können auch keine speziellen Eigenschaften in der Definition von $\mathcal{E}$ durch Angreifer ausgenutzt werden.

A soll hierbei natürlich wieder polynomielle Laufzeit in der Schlüssellänge von k haben.

One-Time Pad

Entspricht der einmaligen Anwendung einer allgemeinen, zufällig gewählten Permutation aus $S(\{0,1\}^n)$. Liefert daher ein „Fragment“ eines idealen Blockchiffre.

Vernam’s One-Time Pad (1917) ist wie folgt definiert:

- $M = C = K = \{0,1\}^n$.
- $c = \mathcal{E}(k, m) := m \oplus k$ (bitweises XOR = bitweise Addition in $\mathbb{Z}/2\mathbb{Z}$).
- $m = \mathcal{D}(k, c) := c \oplus k$.
- k wird zufällig gewählt und nur einmal (!) verwendet.

Known-Plaintext Angriff liefert $k = m \oplus c$. Unsicher unter mehrfacher Verwendung von k.

Aber: Unter einmaliger Anwendung vollkommen sicher!

One-Time Pad

Sicherheit durch Shannon bewiesen (1949).

Problem:

- Hoher Schlüsselverbrauch, daher ineffizient.
- Wie Schlüssel erzeugen?

Wurde angeblich für die Verbindung zwischen Washington und Moskau verwendet (und andere militärische/diplomatische Anwendungen).

Das One-time pad ist ein Spezialfall der Chiffre von Vignére, wenn nur einmal verwendet.

Perfekte Sicherheit

Theorie durch Shannon (1949). Sicherheit in Anwesenheit von Angreifern mit unbeschränkter Rechenleistung.

Modell:

1. Betrachten Klartexte m als Zufallsvariablen mit Werten in M.
   („HEUTE“ kommt häufiger vor als „XZYQR“.)
2. Betrachten Schlüssel k als Zufallsvariablen mit Werten in K.
   (Die Schlüssel werden irgendwie zufällig gewählt.)
3. Annahme: m und k sind unabhängig.
4. Definieren Chiffretexte als Zufallsvariable $c = \mathcal{E}(k, m)$.

Als Wahrscheinlichkeitsraum können wir $M \times K$ nehmen.
Können $\Pr(c = c_0) > 0$ für alle $c_0 \in C$ annehmen (sonst C verkleinern), ebenso $\Pr(m = m_0) > 0$ für alle $m_0 \in M$.

Perfekte Sicherheit

Def: Ein Verschlüsselungssystem heißt perfekt sicher, wenn $\Pr(m = m_0 \mid c = c_0) = \Pr(m = m_0)$ für alle $m_0 \in M$, $c_0 \in C$ gilt.

Kenntnis von $c = c_0$ ergibt also keine weiteren Hinweise über den Wert von m.

Prop: Für ein perfekt sicheres Verschlüsselungssystem gilt $\#K \geq \#C$.
Genauer gibt es für jedes $m_0 \in M$, $c_0 \in C$ ein $k_0 \in K$ mit $c_0 = \mathcal{E}(k_0, m_0)$.

Thm (Shannon): Ein Verschlüsselungssystem mit $\#K = \#C = \#M$ ist genau dann perfekt sicher, wenn $\Pr(k = k_0) = 1/\#K$ für alle $k_0 \in K$ und wenn es für jedes $m_0 \in M$, $c_0 \in C$ genau ein $k_0 \in K$ gibt mit $c_0 = \mathcal{E}(k_0, m_0)$.

Folg: Das One-Time Pad ist bei gv. Schlüsselwahl perfekt sicher.

Perfekte Sicherheit

Bew Prop: Nach dem Satz von Bayes gilt

$$\Pr(m = m_0 \mid c = c_0) \Pr(c = c_0) = \Pr(c = c_0 \mid m = m_0) \Pr(m = m_0).$$

Wegen $\Pr(m = m_0 \mid c = c_0) = \Pr(m = m_0) > 0$ nach Voraussetzung folgt durch Kürzen $\Pr(c = c_0 \mid m = m_0) = \Pr(c = c_0)$ für jedes $m_0 \in M$ und $c_0 \in C$.
Wegen $\Pr(c = c_0) > 0$ ist c_0 Chiffretext für jedes m_0. Daher gibt es ein $k_0 \in K$ mit $c_0 = \mathcal{E}(k_0, m_0)$. □

Bew Thm:

$\Rightarrow$. Die Existenz von k_0 und die Gleichung $C = \{\mathcal{E}(k_0, m_0) \mid k_0 \in K\}$ folgen aus der Prop. Gibt es m_0 und $k_1 \neq k_2$ mit $\mathcal{E}(k_1, m_0) = \mathcal{E}(k_2, m_0)$, so folgt damit $\#C < \#K$ im Widerspruch zur Annahme. Also muß $k_1 = k_2$ gelten, und die Eindeutigkeit von k_0 ergibt sich.
Für $m_0 \in M$ und $c_0 \in C$ sei $k(m_0, c_0) \in K$ mit $\mathcal{E}(k(m_0, c_0), m_0) = c_0$.

Perfekte Sicherheit

Dann gilt

$$\begin{aligned}\Pr(m = m_0 \mid c = c_0) \Pr(c = c_0) &= \Pr(c = c_0 \mid m = m_0) \Pr(m = m_0) \\ &= \Pr(c = c_0 \text{ und } m = m_0) \\ &= \Pr(k = k(m_0, c_0) \text{ und } m = m_0) \\ &= \Pr(k = k(m_0, c_0)) \Pr(m = m_0)\end{aligned}$$

unter Verwendung der Definitionen und der Annahme, daß m und k unabhängig sind. Aus $\Pr(m = m_0 \mid c = c_0) = \Pr(m = m_0) > 0$ ergibt sich daraus durch Kürzen

$$\Pr(k = k(m_0, c_0)) = \Pr(c = c_0).$$

Die rechte Seite ist unabhängig von m_0. Auf der linken Seite ist jedes Element von K von der Form $k(m_1, c_0)$ für ein $m_1 \in M$, denn $m \mapsto k(m, c_0)$ ist injektiv und wegen $\#M = \#K$ auch surjektiv. Folglich sind alle Wahrscheinlichkeiten gleich und betragen $1/\#K = 1/\#C$.

Perfekte Sicherheit

$\Leftarrow$. Sei wieder $k(m_0, c_0) \in K$ mit $\mathcal{E}(k(m_0, c_0), m_0) = c_0$. Dann gilt

$$\begin{aligned}\Pr(c = c_0) &= \sum_{m_1 \in M} \Pr(m = m_1 \text{ und } k = k(m_1, c_0)) \\ &= \sum_{m_1 \in M} \Pr(m = m_1) \Pr(k = k(m_1, c_0)) \\ &= \sum_{m_1 \in M} \Pr(m = m_1)(1/\#K) \\ &= 1/\#K\end{aligned}$$

unter Verwendung der Definitionen und der Annahme. Ähnlich wie in $\Rightarrow$ ergibt sich damit

$$\begin{aligned}\Pr(m = m_0 \,|\, c = c_0) &= \Pr(k = k(m_0, c_0)) \Pr(m = m_0) / \Pr(c = c_0) \\ &= (1/\#K) \Pr(m = m_0) / (1/\#K) \\ &= \Pr(m = m_0). \quad \square\end{aligned}$$

Perfekte Sicherheit

Bew Folg: Nach Definition des One-Time Pads gilt $\#M = \#K = \#C$ und $\Pr(k = k_0) = 1/\#K$. Für jedes m und jedes c ist $k := c \oplus m$ ein Schlüssel mit $c = k \oplus m = \mathcal{E}(k, m)$. Gilt $c = \mathcal{E}(k_1, m) = \mathcal{E}(k_2, m)$, so folgt $k_1 \oplus m = k_2 \oplus m$ und daraus durch Addition von m auch $k_1 = k_2$. Daher gibt es zu jedem m und jedem c genau einen Schlüssel k mit $c = \mathcal{E}(k, m)$. Die Voraussetzungen des Thm (Shannon) sind also erfüllt. $\square$.

Bemerkungen

In einem perfekt sicheren Verschlüsselungssystem erhält ein Angreifer aus dem Chiffretext keine Information über den Klartext oder den Schlüssel.

Für nicht perfekt sichere Systeme hat Shannon die folgenden Fragen untersucht (mehrfache Verschlüsselung mit dem gleichen Schlüssel):

1. Sei $c = \mathcal{E}(k, m)$. Ist $\mathcal{D}(k_1, c)$ ein „sinnvoller“ Klartext (z.B. deutsch) und $k_1 \neq k$, so heißt k_1 Nebenschlüssel. Wieviel Nebenschlüssel gibt es durchschnittlich, basierend auf der Redundanz der Klartexte?

Gibt es Nebenschlüssel, so kann ein Angreifer den Klartext nicht eindeutig reproduzieren.

Bemerkungen

2. Wieviel Chiffretext ist durchschnittlich erforderlich, damit es keine Nebenschlüssel mehr geben kann?

Dann kann ein Angreifer den Schlüssel bestimmen.

Für Details siehe Buch von D. Stinson.

Konsequenzen

Für ein perfekt sicheres Verschlüsselungsystem gilt notwendigerweise $\#K \geq \#C$, so daß für $K = \{0,1\}^r$ und $C = \{0,1\}^s$ die Schlüssellänge r mindestens gleich der Chiffretextlänge s sein muß.

Man kann also keine perfekte (informationstheoretische) Sicherheit erreichen, wenn man viel Klartext mit einem kleinen Schlüssel verschlüsseln will bzw. wenn man ein praktikables Verschlüsselungssystem haben will.

Man kann aber versuchen, komplexitätstheoretische Sicherheit zu erreichen.

Die Philosophie hierbei ist, daß $\mathcal{E}(k,\cdot)$ „sich von einer zufälligen Funktion“ nicht in vertretbarer Laufzeit unterscheiden lassen sollte.

Entropie

Schlüsselkonzept zu Shannon's Untersuchung ist die Entropie einer Zufallsvariablen.

Sind die x_i für $1 \leq i \leq n$ die Werte der ZV X, so wird definiert $H(X) = -\sum_{\Pr(X=x_i)>0} \Pr(X = x_i) \log_2(\Pr(X = x_i))$.

$-\log_2(\Pr(X = x_i))$ ist Länge in Bits der Information, daß x_i eintritt.
Also $H(X)$ erwartete (durchschnittliche) Information oder Unsicherheit.
$0 \leq H(X) \leq \log_2(n)$,
$H(X) = 0$ für $n = 1$, $H(X) = \log_2(n)$ für $\Pr(X = x_i) = 1/n$.
...
Beispiel: Würfeln mit fairem und „frisiertem“ Würfel.
Exhaustive Search Aufwand $\approx 2^{H(k)}$ statt $\#K$.