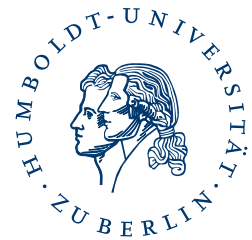
Einführung in die Kryptologie

Johannes Köbler



Institut für Informatik
Humboldt-Universität zu Berlin

SS 2022

Sicherheit von Kryptosystemen

Informationstheoretische Sicherheit

- Claude E. Shannon untersuchte die Sicherheit kryptografischer Systeme auf informationstheoretischer Basis
- Seinen Untersuchungen liegt das Modell einer Nachrichtenquelle $\mathcal{X}$ zugrunde, die einzelne Klartextnachrichten x unter einer bestimmten Wahrscheinlichkeitsverteilung $p(x) = \Pr[\mathcal{X} = x]$, $x \in M$, generiert
- Zudem nehmen wir an, dass der Schlüssel $k \in K$ von einem Schlüsselgenerator $\mathcal{S}$ unter einer Verteilung $p(k) = \Pr[\mathcal{S} = k]$ erzeugt wird
- Da der Schlüssel unabhängig vom Klartext gewählt wird, ist $p(k, x) = p(k)p(x)$ die Wahrscheinlichkeit, dass $\mathcal{X}$ den Klartext x generiert und dieser mit dem Schlüssel k verschlüsselt wird
- Dabei gehen wir davon aus, dass für jede Nachricht $x \in M$ ein neuer Schlüssel generiert wird
- Dies bedeutet, dass wir beispielsweise bei der additiven Chiffre den Klartextraum auf $M = A^n$ vergrößern müssen, falls der Schlüssel nach n Zeichen gewechselt wird

Sicherheit von Kryptosystemen

Informationstheoretische Sicherheit

- Die Zufallsvariablen $\mathcal{X}$ und $\mathcal{S}$ induzieren eine Verteilung auf dem Kryptotextraum C, die wir durch die Zufallsvariable $\mathcal{Y}$ beschreiben
- Die Wahrscheinlichkeit eines Kryptotextes y berechnet sich zu

$$p(y) = \Pr[\mathcal{Y} = y] = \sum_{k,x:E(k,x)=y} p(k,x)$$

und für einen beobachteten Kryptotext y (mit $p(y) > 0$) ist

$$p(x|y) = \frac{p(x,y)}{p(y)} = \sum_{k:E(k,x)=y} \frac{p(k,x)}{p(y)}$$

die (bedingte) Wahrscheinlichkeit, dass sich hinter dem Kryptotext y der Klartext x verbirgt

- Da der Schlüsselgenerator für die Sicherheit eine wichtige Rolle spielt, nehmen wir bei Sicherheitsbetrachtungen die Schlüsselverteilung $\mathcal{S}$ als sechste Komponente eines Kryptosystems hinzu

Informationstheoretische Sicherheit

Definition

Ein Kryptosystem $KS = (M, C, E, D, K, \mathcal{S})$ mit Schlüsselverteilung $\mathcal{S}$ heißt informationstheoretisch (oder absolut) sicher, falls jede Klartextverteilung $\mathcal{X}$ auf M unabhängig von der zugehörigen Kryptotextverteilung $\mathcal{Y}$ auf C ist

- Bei einem absolut sicheren Kryptosystem ist also die A-posteriori-Wahrscheinlichkeit $p(x|y)$ von x gleich der A-priori-Wahrscheinlichkeit $p(x)$
- Die Wahrscheinlichkeit von x bleibt somit gleich, ob nun der Kryptotext y bekannt ist oder nicht, d.h. die Kenntnis von y erlaubt keinerlei Rückschlüsse auf die gesendete Nachricht
- Dies bedeutet, dass es dem Gegner nicht möglich ist, das System zu brechen; auch nicht mit unbegrenzten Rechenressourcen
- Wie wir sehen werden, lässt sich dieses Maß an Sicherheit nur mit einem sehr hohen Aufwand erreichen

- Sind $p(x), p(y) > 0$, so gilt wegen

$$p(x|y)p(y) = p(x, y) = p(y|x)p(x)$$

die Gleichheit

$$p(x|y) = \frac{p(y|x)p(x)}{p(y)} \quad \text{(Satz von Bayes)}$$

- Daher ist die Bedingung $p(x) = p(x|y)$ äquivalent zu $p(y) = p(y|x)$, was wiederum mit der Unabhängigkeit der Ereignisse $\mathcal{X} = x$ und $\mathcal{Y} = y$ gleichbedeutend ist

Beispiel

- Sei $KS = (M, C, E, D, K)$ ein Kryptosystem mit $M = \{x_1, \ldots, x_4\}$, $K = \{k_1, \ldots, k_4\}$, $C = \{y_1, \ldots, y_4\}$ und der Verschlüsselungsfunktion

E	x_1	x_2	x_3	x_4
k_1	y_1	y_4	y_3	y_2
k_2	y_2	y_1	y_4	y_3
k_3	y_3	y_2	y_1	y_4
k_4	y_4	y_3	y_2	y_1

- Zudem betrachten wir die folgenden Schlüssel- und Klartextverteilungen

k_i	k_1	k_2	k_3	k_4
$p(k_i)$	1/2	1/4	1/8	1/8

bzw.

x_i	x_1	x_2	x_3	x_4
$p(x_i)$	1/2	1/6	1/6	1/6

Beispiel (Fortsetzung)

		$1/2$	$1/6$	$1/6$	$1/6$
	E	x_1	x_2	x_3	x_4
$1/2$	k_1	y_1	y_4	y_3	y_2
$1/4$	k_2	y_2	y_1	y_4	y_3
$1/8$	k_3	y_3	y_2	y_1	y_4
$1/8$	k_4	y_4	y_3	y_2	y_1

- Wegen $p(y) = \sum_{k,x:E(k,x)=y} p(k,x)$ ergibt sich die Kryptotextverteilung

$$p(y_1) = 1/2 \cdot 1/2 + (1/4 + 1/8 + 1/8) \cdot 1/6 = 1/3$$
$$p(y_2) = 1/4 \cdot 1/2 + (1/8 + 1/8 + 1/2) \cdot 1/6 = 1/4$$
$$p(y_3) = 1/8 \cdot 1/2 + (1/8 + 1/2 + 1/4) \cdot 1/6 = 5/24$$
$$p(y_4) = 1/8 \cdot 1/2 + (1/2 + 1/4 + 1/8) \cdot 1/6 = 5/24$$

i	1	2	3	4
$p(y_i)$	$1/3$	$1/4$	$5/24$	$5/24$

Beispiel (Schluss)

			$1/2$	$1/6$	$1/6$	$1/6$
	E		x_1	x_2	x_3	x_4
$1/2$	k_1		y_1	y_4	y_3	y_2
$1/4$	k_2		y_2	y_1	y_4	y_3
$1/8$	k_3		y_3	y_2	y_1	y_4
$1/8$	k_4		y_4	y_3	y_2	y_1

- Die bedingten Wahrscheinlichkeiten $p(x|y_1)$ berechnen sich wie folgt:

$$
\begin{aligned}
p(x_1|y_1) &= p(k_1, x_1)/p(y_1) = p(k_1)p(x_1)/p(y_1) = 1/2 \cdot 1/2 \cdot 3 = 3/4 \\
p(x_2|y_1) &= p(k_2, x_2)/p(y_1) = p(k_2)p(x_2)/p(y_1) = 1/4 \cdot 1/6 \cdot 3 = 1/8 \\
p(x_3|y_1) &= p(k_3, x_3)/p(y_1) = p(k_3)p(x_3)/p(y_1) = 1/8 \cdot 1/6 \cdot 3 = 1/16 \\
p(x_4|y_1) &= p(k_4, x_4)/p(y_1) = p(k_4)p(x_4)/p(y_1) = 1/8 \cdot 1/6 \cdot 3 = 1/16
\end{aligned}
$$

- Wegen $p(x_1) = 1/2 \neq 3/4 = p(x_1|y_1)$ ist das Kryptosystem nicht absolut sicher

◁

Informationstheoretische Sicherheit

Frage

Lässt sich das Kryptosystem KS aus obigem Beispiel durch Verwendung eines anderen Schlüsselgenerators absolut sicher machen?

- KS ist genau dann absolut sicher, wenn $p(y_j) = p(y_j|x_i)$ für alle $(x_i, y_j) \in M \times C$ gilt
- Da es in KS für jedes Paar (x_i, y_j) genau einen Schlüssel $k = k_{i,j} \in K$ mit $E(k, x_i) = y_j$ gibt, gilt $p(y_j|x_i) = p(k_{i,j})$
- Somit gilt die Äquivalenz $p(y_j) = p(y_j|x_i) \Leftrightarrow p(y_j) = p(k_{i,j})$
- Für $j = 1$ und $i = 1, \ldots, 4$ implizieren die Gleichungen $p(y_1) = p(k_{i,1})$, dass alle vier Schlüssel $k_{i,1}$ $(= k_i)$ gleichwahrscheinlich sein müssen
- Wegen

$$p(y_j) = \sum_{i=1}^{4} p(x_i) \underbrace{p(y_j|x_i)}_{p(k_{i,j})=1/4} = 1/4 \sum_{i=1}^{4} p(x_i) = 1/4 = p(k_{i,j}) = p(y_j|x_i)$$

  ist das System in diesem Fall tatsächlich absolut sicher

In Verallgemeinerung dieses Beispiels lässt sich für eine wichtige Klasse von Kryptosystemen die absolute Sicherheit wie folgt charakterisieren

Satz

- Sei $KS = (M, C, E, D, K, \mathcal{S})$ ein Kryptosystem mit $|M| = |C| = |K|$, dessen Schlüsselraum K für jedes Paar $(x, y) \in M \times C$ genau einen Schlüssel k mit $E(k, x) = y$ enthält
- Dann ist KS genau dann absolut sicher, wenn $\mathcal{S}$ auf K gleichverteilt ist

Beweis.

- Bezeichne $k_{x,y}$ den eindeutigen Schlüssel, der den Klartext x auf den Kryptotext y abbildet
- Falls $\mathcal{S}$ auf K gleichverteilt ist, folgt wegen $p(k_{x,y}) = |K|^{-1}$ für alle x, y mit $p(x) > 0$ zunächst

$$p(y|x) = \sum_{k:E(k,x)=y} p(k) = p(k_{x,y}) = |K|^{-1}$$

und

$$p(y) = \sum_{x:p(x)>0} p(x) \underbrace{p(y|x)}_{|K|^{-1}} = |K|^{-1} \sum_{x} p(x) = |K|^{-1}$$

- Somit folgt $p(x,y) = p(x)\underbrace{p(y|x)}_{p(y)} = p(x)p(y)$, d.h. KS ist absolut sicher
- Die Umkehrung wird in den Übungen gezeigt □

Ist der One-Time-Pad absolut sicher?

- Verwendet man beim One-Time-Pad nur Klartexte einer festen Länge n, d.h. $M = C = K = A^n$, so ist dieser nach obigem Satz genau dann absolut sicher, wenn die Schlüssel gleichwahrscheinlich sind
- Variiert die Klartextlänge, so kann ein Gegner aus y nur die Länge des zugehörigen Klartextes x ableiten
- Wird jedoch dieselbe zufällige Schlüsselsequenz k zweimal verwendet, so kann aus den Kryptotexten die Differenz der zugehörigen Klartexte ermittelt werden:

$$\left.\begin{array}{rcl} y_1 & = & E(x_1, k) = x_1 + k \\ y_2 & = & E(x_2, k) = x_2 + k \end{array}\right\} \rightsquigarrow y_1 - y_2 = x_1 - x_2$$

- Sind die Klartexte natürlichsprachig, so können aus $y_1 - y_2$ die beiden Nachrichten x_1 und x_2 ähnlich wie bei der Analyse einer Lauftextverschlüsselung rekonstruiert werden

Verwendung von Pseudo-Zufallsgeneratoren

- In einem absolut sicheren Kryptosystem muss der Schlüsselraum mindestens so groß wie der Klartextraum sein (siehe Übungen)
- Daher erfordert die absolute Sicherheit einen extrem hohen Aufwand
- Vor der Kommunikation muss ein Schlüssel, der mindestens so lang wie der zu übertragende Klartext ist, zufällig generiert und zwischen den Partnern auf einem sicheren Kanal ausgetauscht werden
- Für die meisten Anwendungen ist jedoch keine absolute Sicherheit erforderlich
- Wie wir bei der Betrachtung von Stromsystemen gesehen haben, kann der Schlüsselstrom auch von einem Pseudo-Zufallsgenerator erzeugt werden
- Dieser erhält als Eingabe eine Zufallszahl s_0 (den sogenannten Keim) und erzeugt daraus eine lange Folge $v_0\, v_1 \ldots$ von Pseudo-Zufallszahlen
- Als Schlüssel muss jetzt nur noch der Keim ausgetauscht werden

Der Entropiebegriff

Ein Maß für den Informationsgehalt

- In der Informationstheorie wird die Unsicherheit über eine Nachrichtenquelle $\mathcal{X}$ nach ihrer Entropie bemessen
- Dabei entspricht die Unsicherheit über $\mathcal{X}$ genau dem Informationsgewinn, der sich aus der Beobachtung der Quelle $\mathcal{X}$ ergibt
- Intuitiv ist die in einer einzelnen Nachricht x steckende Information umso größer, desto unwahrscheinlicher sie ist
- Tritt eine Nachricht x mit der Wahrscheinlichkeit $p(x) = \Pr[\mathcal{X} = x] > 0$ auf, dann ist ihr Informationsgehalt definiert als

$$Inf_{\mathcal{X}}(x) = \log_2(1/p(x)) = -\log_2 p(x)$$

- Im Fall $p(x) = 0$ sei $Inf_{\mathcal{X}}(x) = 0$

Der Entropiebegriff

Ein Maß für den Informationsgehalt

- Diese Definition des Informationsgehalts ergibt sich zwangsläufig aus den beiden folgenden Axiomen:
  - Der (gemeinsame) Informationsgehalt $Inf_{\mathcal{X},\mathcal{Y}}(x, y)$ von zwei Nachrichten x und y, die aus unabhängigen Quellen $\mathcal{X}$ und $\mathcal{Y}$ stammen, ist $Inf_{\mathcal{X}}(x) + Inf_{\mathcal{Y}}(y)$
  - Eine Nachricht x, die mit Wahrscheinlichkeit $\Pr[\mathcal{X} = x] = 1/2$ auftritt, hat den Informationsgehalt $Inf_{\mathcal{X}}(x) = 1$
- Die Einheit des Informationsgehalts ist bit (basic indissoluble information unit)
- Die Entropie von $\mathcal{X}$ ist nun der erwartete Informationsgehalt einer von $\mathcal{X}$ generierten Nachricht

Der Entropiebegriff

Definition

- Sei $\mathcal{X}$ eine Zufallsvariable mit Wertebereich $W(\mathcal{X}) = \{x_1, \ldots, x_n\}$ und sei $p_i = \Pr[\mathcal{X} = x_i]$
- Dann ist die Entropie von $\mathcal{X}$ definiert als

$$\mathcal{H}(\mathcal{X}) = \sum_{i=1}^{n} p_i \, Inf_{\mathcal{X}}(x_i) = \sum_{i=1}^{n} p_i \log_2(1/p_i) = -\sum_{i=1}^{n} p_i \log_2(p_i)$$

Beispiel

- Sei $\mathcal{X}$ eine Zufallsvariable mit der Verteilung

x_i	sonnig	leicht bewölkt	bewölkt	stark bewölkt	Regen	Schnee	Nebel
p_i	$1/4$	$1/4$	$1/8$	$1/8$	$1/8$	$1/16$	$1/16$

- Dann ergibt sich die Entropie von $\mathcal{X}$ zu

$$\mathcal{H}(\mathcal{X}) = 1/4 \cdot (2+2) + 1/8 \cdot (3+3+3) + 1/16 \cdot (4+4) = 2{,}625$$

◁

Der Entropiebegriff

- Die Entropie nimmt im Fall der Gleichverteilung $p_1 = \cdots = p_n = 1/n$ den Wert $\log_2(n)$ an, während sie für jede andere Verteilung auf einer n-elementigen Menge einen Wert $\mathcal{H}(\mathcal{X}) < \log_2(n)$ hat (siehe unten)
- Die Unsicherheit über eine Zufallsvariable $\mathcal{X}$ ist um so größer, je größer der Wertebereich und je gleichmäßiger die Verteilung von $\mathcal{X}$ ist
- Bringt $\mathcal{X}$ zum Beispiel nur einen einzigen Wert mit positiver Wahrscheinlichkeit hervor, dann (und nur dann) hat $\mathcal{H}(\mathcal{X})$ den Wert 0
- Für den Nachweis von oberen Schranken für die Entropie benutzen wir folgende Hilfsmittel aus der Analysis

Definition Sei $I \subseteq \mathbb{R}$ ein Intervall.

- Eine Funktion $f : I \to \mathbb{R}$ heißt konkav auf I, falls für alle $x \neq y \in I$ und $0 \leq t \leq 1$ gilt:

$$f(tx + (1-t)y) \quad \geq \quad tf(x) + (1-t)f(y)$$

- Gilt sogar „$>$“ anstelle von „$\geq$“, so heißt f streng konkav auf I

Beispiel

Die Funktion $f(x) = \log_2(x)$ ist streng konkav auf $(0, \infty)$ ◁

Für den Beweis des nächsten Satzes benötigen wir die Jensensche Ungleichung, die wir ohne Beweis angeben

Jensensche Ungleichung

- Sei f eine streng konkave Funktion auf I und seien $0 < a_1, \ldots, a_n < 1$ reelle Zahlen mit $\sum_{i=1}^{n} a_i = 1$
- Dann gilt für alle $x_1, \ldots, x_n \in I$,

$$f\left(\sum_{i=1}^{n} a_i x_i\right) \geq \sum_{i=1}^{n} a_i f(x_i)$$

- Hierbei tritt Gleichheit genau dann ein, wenn alle x_i den gleichen Wert haben

Der Entropiebegriff

Satz

- Sei $\mathcal{X}$ eine Zufallsvariable mit Wertebereich $W(\mathcal{X}) = \{x_1, \ldots, x_n\}$ und Verteilung $p_i = \Pr[\mathcal{X} = x_i]$ für $i = 1, \ldots, n$
- Dann gilt $\mathcal{H}(\mathcal{X}) \leq \log_2(n)$, wobei Gleichheit genau im Fall $p_i = 1/n$ für $i = 1, \ldots, n$ eintritt

Beweis.

- Aufgrund der Jensenschen Ungleichung gilt

$$\mathcal{H}(\mathcal{X}) = \sum_{i=1}^{n} p_i \log_2(1/p_i) \leq \log_2 \sum_{i=1}^{n} (p_i/p_i) = \log_2 n,$$

  wobei Gleichheit genau im Fall $1/p_1 = \cdots = 1/p_n$ eintritt
- Letzteres ist mit der Bedingung $p_i = 1/n$ für $i = 1, \ldots, n$ gleichbedeutend

□

Redundanz von Codes

- Die Entropie liefert eine sehr gute untere Schranke für die mittlere Codewortlänge von Binärcodes
- Ein Binärcode für eine Zufallsvariable $\mathcal{X}$ mit $W(\mathcal{X}) = \{x_1, \ldots, x_n\}$ ist eine (geordnete) Menge $C = \{y_1, \ldots, y_n\}$ von binären Codewörtern $y_i \in \{0,1\}^*$ mit der Eigenschaft, dass die Abbildung $c : \{1, \ldots, n\}^* \to \{0,1\}^*$ mit $c(i_1 \cdots i_k) = y_{i_1} \cdots y_{i_k}$ injektiv ist
- Die Injektivität von c stellt sicher, dass jede Folge $y_{i_1} \cdots y_{i_k}$ von Codewörtern eindeutig decodierbar ist
- Die mittlere Codewortlänge von $C = \{y_1, \ldots, y_n\}$ unter $\mathcal{X}$ ist

$$L(C) = \sum_{i=1}^{n} p_i \cdot |y_i|$$

- Ein Binärcode C heißt optimal, wenn kein anderer Binärcode für $\mathcal{X}$ eine kürzere mittlere Codewortlänge besitzt
- Für einen optimalen Binärcode C für $\mathcal{X}$ gilt

$$\mathcal{H}(\mathcal{X}) \leq L(C) < \mathcal{H}(\mathcal{X}) + 1 \qquad \text{(ohne Beweis)}$$

Redundanz von Codes

Beispiel

- Sei $\mathcal{X}$ die Zufallsvariable aus dem letzten Beispiel mit der Verteilung

i	1	2	3	4	5	6	7
p_i	$1/4$	$1/4$	$1/8$	$1/8$	$1/8$	$1/16$	$1/16$

- Betrachte die Codes $C = \{y_1, \ldots, y_7\}$ und $C' = \{y'_1, \ldots, y'_7\}$ mit

i	1	2	3	4	5	6	7
y_i	001	010	011	100	101	110	111
y'_i	00	01	100	101	110	1110	1111

- Dann hat C die mittlere Codewortlänge $L(C) = 3$
- Dies ist nicht optimal, während der Code C' wegen

  $$|y'_i| = \log_2(1/p_i) \text{ für } i = 1, \ldots, 7$$

  den minimalen Wert $L(C') = \mathcal{H}(\mathcal{X}) = 2{,}625$ erreicht ◁

Redundanz von Codes

- Die Redundanz eines Codes für $\mathcal{X}$ ist um so höher, je größer seine mittlere Codewortlänge im Vergleich zur Entropie von $\mathcal{X}$ ist
- Um auch Codes über unterschiedlichen Alphabeten vergleichen zu können, sollte die Codewortlänge das Alphabet berücksichtigen
- Hierzu definiert man die Bitlänge eines Wortes x über einem Alphabet A mit $m > 2$ Zeichen zu $|x|_2 = |x| \log_2(m)$
- Beispielsweise das Wort `GOLD` über dem lateinischen Alphabet die Bitlänge $|\texttt{GOLD}|_2 = 4 \log_2(26) = 18,8$
- Entsprechend ist die mittlere Codewortlänge (in bit) eines Codes $C = \{y_1, \ldots, y_n\}$ für $\mathcal{X}$ über einem beliebigen Alphabet

$$L_2(C) = \sum_{i=1}^{n} p_i \cdot |y_i|_2$$

- Mit dem Huffman-Verfahren lässt sich ein optimaler Code über einem beliebigem Alphabet mit $m \geq 2$ Zeichen effizient berechnen

Redundanz von Codes

Nun können wir die Redundanz eines Codes als den mittleren Anteil der Codewortbuchstaben definieren, die keine Information tragen

Definition

Die (relative) Redundanz eines Codes C für $\mathcal{X}$ ist definiert als

$$\mathcal{R}(C) = \frac{L_2(C) - \mathcal{H}(\mathcal{X})}{L_2(C)}$$

Beispiel (Fortsetzung)

- Während eine von $\mathcal{X}$ generierte Nachricht im Durchschnitt $\mathcal{H}(\mathcal{X}) = 2{,}625$ bit an Information enthält, haben alle Codewörter von C die Bitlänge $L_2(C) = 3$
- Der Anteil an „überflüssigen“ Zeichen pro Codewort beträgt also

$$\mathcal{R}(C) = \frac{3 - 2{,}625}{3} = 12{,}5\%,$$

wogegen C' wegen $L_2(C') = \mathcal{H}(\mathcal{X})$ keine Redundanz besitzt ◁

Redundanz von Sprachen

- Auch Schriftsprachen wie Deutsch oder Englisch und Programmiersprachen wie C++ können als eine Art Code aufgefasst werden
- Eine solche Sprache enthält umso mehr Redundanz, desto restriktiver die Gesetze sind, unter denen in ihr Worte und Sätze gebildet werden
- Um die statistischen Eigenschaften einer Sprache L zu erforschen, erweist es sich als zweckmäßig, die Textstücke der Länge n (n-Gramme) von L für $n \geq 1$ getrennt voneinander zu betrachten
- Sei also $\mathcal{L}_n$ die Zufallsvariable, die die Verteilung aller n-Gramme in L beschreibt
- Betrachten wir die Menge A^n aller n-Gramme als einen Code für $\mathcal{L}_n$, so hat dieser die Redundanz

$$\mathcal{R}(\mathcal{L}_n) = \frac{n \log_2 m - \mathcal{H}(\mathcal{L}_n)}{n \log_2 m}$$

Redundanz von Sprachen

Definition

- Für eine Sprache L über einem Alphabet A mit $|A| = m$ ist $\mathcal{H}(\mathcal{L}_n)/n$ die n-Gramm-Entropie von L (pro Buchstabe)
- Falls dieser Wert für n gegen ∞ gegen einen Grenzwert

  $$\mathcal{H}(L) = \lim_{n\to\infty} \mathcal{H}(\mathcal{L}_n)/n$$

  konvergiert, so wird dieser Grenzwert als die Entropie von L bezeichnet
- In diesem Fall konvergiert $\mathcal{R}(\mathcal{L}_n)$ gegen den Grenzwert

  $$\mathcal{R}(L) = \lim_{n\to\infty} \mathcal{R}(\mathcal{L}_n) = \frac{\log_2 m - \mathcal{H}(L)}{\log_2 m},$$

  der als die (relative) Redundanz von L bezeichnet wird
- Der Zähler $\mathcal{R}_{abs}(L) = \log_2 m - \mathcal{H}(L) = \mathcal{R}(L)\log_2 m$ wird auch als absolute Redundanz von L bezeichnet (gemessen in bit/Zeichen)

Redundanz von Sprachen

Beispiel

- Die Redundanz von (natürlichen) Sprachen lässt sich näherungsweise bestimmen, indem man die Entropien $\mathcal{H}(\mathcal{L}_n)$ ihrer n-Gramme empirisch ermittelt
- In der deutschen Sprache D hat die Verteilung der Einzelzeichen in A_{lat} eine Entropie von $\mathcal{H}(\mathcal{D}_1) = 4{,}1$ bit
- Im Vergleich hierzu hat eine auf A_{lat} gleichverteilte Zufallsvariable $\mathcal{U}$ den maximalen Entropiewert von $\mathcal{H}(\mathcal{U}) = \log(26) = 4{,}7$ bit
- Für die Bigramme ergibt sich ein Entropiewert von $\mathcal{H}(\mathcal{D}_2)/2 = 3{,}5$ bit pro Buchstabe

Redundanz von Sprachen

Beispiel (Fortsetzung)

- Mit wachsender Länge sinkt die Entropie weiter ab und strebt gegen einen Grenzwert $\mathcal{H}(D)$ von 1,5 bit pro Buchstabe

n	$\mathcal{H}(\mathcal{D}_n)$	$\mathcal{H}(\mathcal{D}_n)/n$	$\mathcal{R}_{abs}(\mathcal{D}_n)/n$	$\mathcal{R}(\mathcal{D}_n)$
1	$4,1$	$4,1$	$0,6$	13%
2	$7,0$	$3,5$	$1,2$	26%
3	$9,6$	$3,2$	$1,5$	32%
6	$12,2$	$2,0$	$2,7$	57%
15	$27,6$	$1,8$	$2,9$	62%
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
∞	∞	$\mathcal{H}(D) = 1,5$	$\mathcal{R}_{abs}(D) = 3,2$	$\mathcal{R}(D) = 67\%$

- Deutsche Texte hinreichender Länge besitzen also eine durchschnittliche Redundanz von ca. 67%, so dass ihre Länge bei optimaler Kodierung auf ca. $1/3$ komprimierbar ist

Die Eindeutigkeitsdistanz eines Kryptosystems

- Wir betrachten nun den Fall, dass Klartexte einer variablen Länge n verschlüsselt werden, ohne den Schlüssel zu wechseln
- Die Chiffrierfunktion hat also die Form

$$E_n : K \times A^n \to C_n$$

- Dabei nehmen wir an, dass die Menge C_n der zugehörigen Kryptotexte die gleiche Kardinalität $|C_n| = |A^n| = m^n$ wie der Klartextraum hat
- Ist y ein abgefangener Kryptotext, so ist

$$K(y) = \{k \in K \mid \exists x \in A^n : E_n(k, x) = y \wedge p(x) > 0\}$$

die Menge aller infrage kommenden Schlüssel

- Die Menge $K(y)$ besteht aus einem „echten“ Schlüssel (mit dem y erzeugt wurde) und $|K(y)| - 1$ so genannten „unechten“ Schlüsseln

Die Eindeutigkeitsdistanz eines Kryptosystems

- Aus informationstheoretischer Sicht ist das Kryptosystem unter einer Klartextverteilung $\mathcal{X}$ umso sicherer, desto größer die erwartete Anzahl

$$\bar{s}_n = \sum_{y \in C_n} p(y) \cdot (|K(y)| - 1) = \sum_{y \in C_n} p(y) \cdot |K(y)| - 1$$

  der unechten Schlüssel ist
- Im besten Fall kommen für jeden Kryptotext alle Schlüssel infrage, d.h. $\bar{s}_n = |K| - 1$
- Ist dagegen $\bar{s}_n$ gleich 0, so liefert der abgefangene Kryptotext dem Gegner genügend Information, um den benutzten Schlüssel und somit den zugehörigen Klartext eindeutig bestimmen zu können (sofern er über genügend Ressourcen verfügt)

Die Eindeutigkeitsdistanz

Definition

Die Eindeutigkeitsdistanz n_0 eines Kryptosystems unter einer Klartextverteilung $\mathcal{X}$ ist der kleinste Wert von n, für den $\bar{s}_n = 0$ wird (falls $\bar{s}_n > 0$ für alle n gilt, sei $n_0 = \infty$)

- Als nächstes wollen wir eine untere Schranke für $\bar{s}_n$ (und damit für n_0) herleiten
- Hierzu benötigen wir den Begriff der bedingten Entropie $\mathcal{H}(\mathcal{X}|\mathcal{Y})$ einer Zufallsvariablen $\mathcal{X}$, wenn der Wert von $\mathcal{Y}$ bereits bekannt ist

Definition. Seien $\mathcal{X}, \mathcal{Y}$ Zufallsvariablen

Die bedingte Entropie von $\mathcal{X}$ unter $\mathcal{Y}$ ist definiert als

$$\mathcal{H}(\mathcal{X}|\mathcal{Y}) = \sum_{y \in W(\mathcal{Y})} p(y)\mathcal{H}(\mathcal{X}|y) = \sum_y p(y) \sum_x p(x|y) \log_2(1/p(x|y)),$$

wobei $\mathcal{X}|y$ die Zufallsvariable mit der Verteilung $p_y(x) = p(x|y)$ ist

Die Eindeutigkeitsdistanz

Satz

1. $\mathcal{H}(\mathcal{X},\mathcal{Y}) = \mathcal{H}(\mathcal{Y}) + \mathcal{H}(\mathcal{X}|\mathcal{Y})$
2. $\mathcal{H}(\mathcal{X},\mathcal{Y}) \leq \mathcal{H}(\mathcal{X}) + \mathcal{H}(\mathcal{Y})$, wobei Gleichheit genau dann eintritt, wenn $\mathcal{X}$ und $\mathcal{Y}$ unabhängig sind

Beweis. Siehe Übungen.

Korollar

Es gilt $\mathcal{H}(\mathcal{X}|\mathcal{Y}) \leq \mathcal{H}(\mathcal{X})$, wobei Gleichheit genau dann eintritt, wenn $\mathcal{X}$ und $\mathcal{Y}$ unabhängig sind

Satz

In jedem Kryptosystem gilt für die Klartextentropie $\mathcal{H}(\mathcal{X})$, die Schlüsselentropie $\mathcal{H}(\mathcal{S})$ und die Kryptotextentropie $\mathcal{H}(\mathcal{Y})$ die Gleichung

$$\mathcal{H}(\mathcal{S}|\mathcal{Y}) = \mathcal{H}(\mathcal{S}) + \mathcal{H}(\mathcal{X}) - \mathcal{H}(\mathcal{Y})$$

Die Eindeutigkeitsdistanz

Satz

In jedem Kryptosystem gilt die Gleichung

$$\mathcal{H}(\mathcal{S}|\mathcal{Y}) = \mathcal{H}(\mathcal{S}) + \mathcal{H}(\mathcal{X}) - \mathcal{H}(\mathcal{Y})$$

Beweis.

- Zunächst ist $\mathcal{H}(\mathcal{S}|\mathcal{Y}) = \mathcal{H}(\mathcal{S},\mathcal{Y}) - \mathcal{H}(\mathcal{Y})$
- Es reicht also zu zeigen, dass $\mathcal{H}(\mathcal{S},\mathcal{Y}) = \mathcal{H}(\mathcal{S}) + \mathcal{H}(\mathcal{X})$ ist
- Bei Kenntnis des Schlüssels ist der Wert von $\mathcal{X}$ eindeutig durch $\mathcal{Y}$ und der Wert von $\mathcal{Y}$ eindeutig durch $\mathcal{X}$ festgelegt
- Da $\mathcal{X}$ und $\mathcal{S}$ unabhängig sind, folgt daher

$$\begin{aligned}\mathcal{H}(\mathcal{S},\mathcal{Y}) &= \mathcal{H}(\mathcal{S},\mathcal{X},\mathcal{Y}) - \underbrace{\mathcal{H}(\mathcal{X}|\mathcal{S},\mathcal{Y})}_{=0} = \mathcal{H}(\mathcal{S},\mathcal{X}) + \underbrace{\mathcal{H}(\mathcal{Y}|\mathcal{S},\mathcal{X})}_{=0} \\ &= \mathcal{H}(\mathcal{S}) + \mathcal{H}(\mathcal{X})\end{aligned}$$

◁

Die Eindeutigkeitsdistanz

Jetzt verfügen wir über alle Hilfsmittel, um eine untere Schranke für die erwartete Anzahl

$$\bar{s}_n = \sum_{y \in C_n} p(y) \cdot |K(y)| - 1$$

der unechten Schlüssel herleiten zu können

Lemma

- Seien $\mathcal{X}_n$ und $\mathcal{Y}_n$ die Zufallsvariablen, die die Verteilungen der n-Gramme der Klartextsprache und der zugehörigen Kryptotexte beschreiben
- Dann gilt
  - $\mathcal{H}(\mathcal{S}|\mathcal{Y}_n) \leq \log_2(\bar{s}_n + 1)$
  - $\mathcal{H}(\mathcal{S}|\mathcal{Y}_n) \geq \mathcal{H}(\mathcal{S}) - n\mathcal{R}(\mathcal{L}_n)\log_2 m$

Die Eindeutigkeitsdistanz

Beweis von $\mathcal{H}(\mathcal{S}|\mathcal{Y}_n) \leq \log_2(\bar{s}_n + 1)$

Unter Verwendung der Jensenschen Ungleichung folgt

$$\begin{aligned}\mathcal{H}(\mathcal{S}|\mathcal{Y}_n) &= \textstyle\sum_{y \in C_n} p(y) \cdot \mathcal{H}(\mathcal{S}|y) \\ &\leq \textstyle\sum_{y \in C_n} p(y) \cdot \log_2 |K(y)| \\ &\leq \log_2 \textstyle\sum_{y \in C_n} p(y) \cdot |K(y)| \\ &= \log_2(\bar{s}_n + 1)\end{aligned}$$

□

Beweis von $\mathcal{H}(\mathcal{S}|\mathcal{Y}_n) \geq \mathcal{H}(\mathcal{S}) - n\mathcal{R}(\mathcal{L}_n)\log_2 m$

- Wegen $\mathcal{X}_n = \mathcal{L}_n$ folgt

  $$\mathcal{H}(\mathcal{X}_n) = (1 - \mathcal{R}(\mathcal{L}_n))n \log_2 m, \text{ wobei } m = |A| \text{ ist}$$

- Wegen $W(\mathcal{Y}_n) = C_n$ und $|C_n| = m^n$ gilt für die Kryptotextentropie

  $$\mathcal{H}(\mathcal{Y}_n) \leq n \log_2 m \text{ und somit } \mathcal{H}(\mathcal{X}_n) - \mathcal{H}(\mathcal{Y}_n) \geq -n\mathcal{R}(\mathcal{L}_n)\log_2 m$$

- Damit folgt

  $$\mathcal{H}(\mathcal{S}|\mathcal{Y}_n) = \mathcal{H}(\mathcal{S}) + \mathcal{H}(\mathcal{X}_n) - \mathcal{H}(\mathcal{Y}_n) \geq \mathcal{H}(\mathcal{S}) - n\mathcal{R}(\mathcal{L}_n)\log_2 m$$

□

Die Eindeutigkeitsdistanz

- Zusammen ergibt sich

$$\log_2(\bar{s}_n + 1) \geq \mathcal{H}(\mathcal{S}|\mathcal{Y}_n) \geq \mathcal{H}(\mathcal{S}) - n\mathcal{R}(\mathcal{L}_n)\log_2 m$$

- Bei gleichverteiltem Schlüssel erreicht $\mathcal{H}(\mathcal{S})$ das Maximum $\log_2 |K|$ und wir erhalten in diesem Fall die Abschätzung

$$\log_2(\bar{s}_n + 1) \geq \log_2 |K| - n\mathcal{R}(\mathcal{L}_n)\log_2 m$$

bzw.

$$\bar{s}_n + 1 \geq |K| m^{-n\mathcal{R}(\mathcal{L}_n)}$$

Satz. Sei $(M, C, E, D, K, \mathcal{S})$ ein Kryptosystem mit $M = A^n$ und $|C| = m^n$.

Falls $\mathcal{S}$ auf K gleichverteilt ist, gilt

$$\bar{s}_n \geq |K| m^{-n\mathcal{R}(\mathcal{L}_n)} - 1 \geq |K| m^{-n\mathcal{R}(\mathcal{L})} - 1$$

für die erwartete Anzahl $\bar{s}_n$ der unechten Schlüssel

Die Eindeutigkeitsdistanz

Setzen wir in obiger Abschätzung $\bar{s}_n = 0$, so erhalten wir folgende untere Schranke für die Eindeutigkeitsdistanz n_0 eines Kryptosystems

Korollar

Unter den Bedingungen des obigen Satzes gilt

$$n_0 \geq \frac{\log_2 |K|}{\mathcal{R}(L) \log_2 m} = \frac{\log_2 |K|}{\mathcal{R}_{abs}(L)}$$

- Man beachte, dass dies nur untere Schranken für die Menge an Kryptotext sind, die zur eindeutigen Bestimmung des Schlüssels benötigt wird, und die tatsächlich benötigte Menge deutlich größer sein kann
- Natürlich erlaubt die eindeutige Bestimmung des Schlüssels auch die eindeutige Bestimmung des Klartexts
- Unter Umständen kann aber der Klartext auch schon aus einer wesentlich geringeren Menge an Kryptotext rekonstruierbar sein

Die Eindeutigkeitsdistanz

Beispiel (Untere Schranken für n_0 bei klassischen Chiffren)

Für Substitutionen bei deutschsprachigem Klartext ergeben sich folgende unteren Schranken für die Eindeutigkeitsdistanz n_0:

Kryptosystem	Schlüsselanzahl $\|K\|$	$\log_2 \|K\|$	$\log_2 \|K\|/\mathcal{R}_{abs}(D)$
additive Chiffre	26	$4,7$	$4,7/3,2 \approx 1,5$
affine Chiffre	$12 \cdot 26 = 312$	$8,3$	$2,6$
einfache Substitution	$26!$	$88,4$	$27,6$
Vigenère-Chiffre	26^d	$4,7 \cdot d$	$1,5 \cdot d$

Für Blocktranspositionen erhalten wir die folgenden unteren Schranken für die Menge an Kryptotext, die zur eindeutigen Bestimmung des Schlüssels benötigt wird

Die Eindeutigkeitsdistanz

- Wird der Kryptotext einer Blocktransposition nur auf der Basis von n-Grammen analysiert, so erhalten wir folgende untere Schranke für n_0

$$n_0 \geq \log_2 |K| / \mathcal{R}(\mathcal{L}_n) \log_2 m = \log_2 \ell! / \mathcal{R}(\mathcal{L}_n) \log_2 m$$

- Dies entspricht der Situation, dass die Wahrscheinlichkeiten der Zeichen im Klartext höchstens von den $n-1$ vorausgehenden Zeichen abhängen

Beispiel (Untere Schranken für n_0 bei Blocktranspositionen)

Untere Schranken für n_0 bei einer Häufigkeitsanalyse auf der Basis von		Blocklänge ℓ 10	20	50	100	1 000
Einzelzeichen	$\mathcal{R}(\mathcal{D}_1) = 13\%$	36	102	357	875	14 216
Bigrammen	$\mathcal{R}(\mathcal{D}_2) = 26\%$	18	51	179	437	7 108
Trigrammen	$\mathcal{R}(\mathcal{D}_3) = 32\%$	15	41	143	350	5 686
n-Grammen, $n \to \infty$,	$\mathcal{R}(D) = 67\%$	7	19	67	164	2 665

◁

Weitere Sicherheitsbegriffe

- Da die Benutzung eines informationstheoretisch sicheren Kryptosystems einen immensen Aufwand erfordert, begnügt man sich in der Praxis meist mit schwächeren Sicherheitsanforderungen
- Ein Kryptosystem gilt als komplexitätstheoretisch sicher oder als berechnungssicher (computationally secure), falls es höchstens mit einem unverhältnismäßig hohen Aufwand zu brechen ist
- Die Kosten für einen erfolgreichen Angriff (sofern er überhaupt möglich ist) übersteigen also den potentiellen Nutzen bei weitem
- Ein Kryptosystem gilt als nachweisbar sicher, wenn seine Sicherheit auf der Basis von bekannten komplexitätstheoretischen Annahmen bewiesen werden kann, die gemeinhin als zutreffend gelten
- Als praktisch sicher wird ein Kryptosystem eingestuft, wenn es über mehrere Jahre allen Versuchen einer erfolgreichen Kryptoanalyse widerstehen konnte, obwohl es weit verbreitet ist und allein schon deshalb ein attraktives Ziel für einen Angriff darstellt

Komplexitätstheoretische Sicherheit

- Die komplexitätstheoretische Analyse eines Kryptosystems ist äußerst schwierig, da der Aufwand für einen erfolgreichen Angriff unabhängig von der dabei benutzten Technik abgeschätzt werden muss
- Es reicht also nicht, alle bekannten kryptoanalytischen Ansätze in Betracht zu ziehen, sondern alle möglichen
- Dabei kommt als Angriffsziel nicht nur die vollständige Rekonstruktion des Klartextes in Betracht, da bereits kleine Unterschiede zwischen A-posteriori- und A-priori-Wahrscheinlichkeit einen lohnenswerten Vorteil für den Angreifer bedeuten könnten
- Aus diesen Gründen ist noch für kein praktikables Kryptosystem der Nachweis gelungen, dass es komplexitätstheoretisch sicher ist
- Damit ist auch nicht zu rechnen, solange fundamentale komplexitätstheoretische Fragen wie etwa das $\mathrm{P} \stackrel{?}{=} \mathrm{NP}$-Problem nicht gelöst sind
- Dagegen gibt es eine ganze Reihe praktikabler Kryptosysteme, die als nachweisbar sicher oder praktisch sicher gelten

Sicherheit gegen einen IND-CPA Angriff

- Eine mögliche Präzisierung der komplexitätstheoretischen Sicherheit ist die Ununterscheidbarkeit unter einem gewählten Klartextangriff (kurz IND-CPA; indistinguishability under a chosen plaintext attack)
- Konkret läuft ein IND-CPA-Angriff wie folgt ab
  1. Zuerst wählt der Gegner zwei Klartexte $x_0 \neq x_1 \in M$
  2. Dann wird x_0 oder x_1 zufällig ausgewählt und der zugehörige Kryptotext y gebildet
  3. Dem Gegner wird der Kryptotext y vorgelegt und er muss raten, welcher der beiden Klartexte sich hinter y verbirgt
  4. Der Angriff ist erfolgreich, falls der Gegner richtig rät
- Um die Erfolgsaussichten des Gegners zu formalisieren, stellen wir das gewünschte Maß an Sicherheit über einen Parameter $s \in \mathbb{N}$ ein
- Typischerweise wird für s die Schlüssellänge $s = |k|$ gewählt
- Aus Praktikabilitätsgründen sollten legale Operationen (wie die Chiffrierung oder die Schlüsselgenerierung) effizient (in Zeit $s^{O(1)}$) ablaufen
- Natürlich darf dann auch der Aufwand des Gegners in Abhängigkeit von s steigen, weshalb er zusätzlich den Parameterwert s erhält

Ein Maß für den Erfolg eines IND-CPA Angriffs

Definition

- Sei $(M, C, E, D, K, \mathcal{S})$ ein Kryptosystem mit Sicherheitsparameter s
- Ein IND-CPA-Gegner (oder kurz Gegner) ist ein Paar $G = (\mathcal{X}, \mathcal{V})$ von probabilistischen Algorithmen, wobei
  - $\mathcal{X}$ bei Eingabe $s \in \mathbb{N}$ ein Klartextpaar $\mathcal{X}(s) = (\mathcal{X}_0(s), \mathcal{X}_1(s)) \in M^2$ ausgibt und
  - $\mathcal{V}$ bei Eingabe $s \in \mathbb{N}$, $x_0, x_1 \in M$ und $y \in C$ ein Bit $\mathcal{V}(s, x_0, x_1, y)$ ausgibt
- Der Vorteil von G bei Parameterwert $s \in \mathbb{N}$ ist

$$\alpha_G(s) = 2\,(\Pr[\mathcal{V}(s, \mathcal{X}(s), E(\mathcal{S}, \mathcal{X}_{\mathcal{B}}(s))) = \mathcal{B}] - 1/2),$$

wobei $\mathcal{B}$ auf $\{0, 1\}$ gleichverteilt und von $\mathcal{S}, \mathcal{X}, \mathcal{V}$ unabhängig ist

Ist der Wert des Sicherheitsparameters s irrelevant, fest vorgegeben oder aus dem Kontext ersichtlich, so verzichten wir meist auf seine explizite Angabe

IND-CPA Angriff auf eine Blockchiffre im ECB-Modus

Beispiel

- Im ECB-Modus (electronic code book mode) werden Klartextblöcke $a_1, a_2, \ldots$ unabhängig voneinander zu einer Folge $b_1, b_2, \ldots$ von Kryptotextblöcken $b_i = E(k, a_i)$ verschlüsselt
- In diesem Fall kann ein Gegner wie folgt einen Vorteil von 1 erzielen (d.h. mit Wahrscheinlichkeit 1 den richtigen Klartext raten)
- Er wählt (deterministisch) zwei beliebige Klartexte $x_0 = a_1 a_2 \ldots$ und $x_1 = a'_1 a'_2 \ldots$ mit der Eigenschaft $a_1 = a_2$ und $a'_1 \neq a'_2$
- Dann kann er bei Vorlage eines Kryptotextes $y = b_1 b_2 \ldots$ leicht erkennen, aus welchem Klartext y generiert wurde:

$$\mathcal{V}(x_0, x_1, y) = \begin{cases} 0, & b_1 = b_2 \\ 1, & \text{sonst} \end{cases}$$

Vergleich mit absoluter Sicherheit

Erwartungsgemäß sind absolut sichere Kryptosysteme gegen IND-CPA-Angriffe resistent

Satz

Der Vorteil jedes beliebigen Gegners G gegenüber einem absolut sicheren Kryptosystem ist $\alpha_G(s) = 0$

- Ein Gegner kann also bei einem absolut sicheren Kryptosystem höchstens mit Wahrscheinlichkeit $1/2$ den richtigen Klartext raten, auch wenn er über unbeschränkte Rechenressourcen verfügt
- In den Übungen wird auch die umgekehrte Implikation bewiesen
- Ein Kryptosystem ist somit genau dann absolut sicher, wenn kein Gegner G einen Vorteil $\alpha_G(s) > 0$ erzielen kann

Vergleich mit absoluter Sicherheit

Satz

Der Vorteil jedes beliebigen Gegners G gegenüber einem absolut sicheren Kryptosystem ist $\alpha_G(s) = 0$

Beweis

- Sei $(\mathcal{X}, \mathcal{V})$ ein beliebiger Gegner und sei $\mathcal{Y}_b = E(\mathcal{S}, \mathcal{X}_b)$, $b \in \{0, 1\}$
- Da KS absolut sicher ist, ist der Kryptotext $\mathcal{Y}_\mathcal{B} = E(\mathcal{S}, \mathcal{X}_\mathcal{B})$ vom Klartext $\mathcal{X}_\mathcal{B}$ (und somit auch von $\mathcal{B}$) stochastisch unabhängig
- Daher folgt

$$
\begin{aligned}
&\Pr[\mathcal{V}(\mathcal{X}_0, \mathcal{X}_1, \mathcal{Y}_\mathcal{B}) = \mathcal{B}] \\
&\quad = \Pr[\mathcal{V}(\mathcal{X}_0, \mathcal{X}_1, \mathcal{Y}_\mathcal{B}) = \mathcal{B} = 0] + \Pr[\mathcal{V}(\mathcal{X}_0, \mathcal{X}_1, \mathcal{Y}_\mathcal{B}) = \mathcal{B} = 1] \\
&\quad = \Pr[\mathcal{V}(\mathcal{X}_0, \mathcal{X}_1, \mathcal{Y}_\mathcal{B}) = 0] \cdot \underbrace{\Pr[\mathcal{B} = 0 \mid \mathcal{V}(\mathcal{X}_0, \mathcal{X}_1, \mathcal{Y}_\mathcal{B}) = 0]}_{= \Pr[\mathcal{B}=0] = 1/2} \\
&\qquad + \Pr[\mathcal{V}(\mathcal{X}_0, \mathcal{X}_1, \mathcal{Y}_\mathcal{B}) = 1] \cdot \underbrace{\Pr[\mathcal{B} = 1 \mid \mathcal{V}(\mathcal{X}_0, \mathcal{X}_1, \mathcal{Y}_\mathcal{B}) = 1]}_{= \Pr[\mathcal{B}=1] = 1/2} \\
&\quad = 1/2
\end{aligned}
$$

□

Präzisierung der komplexitätstheoretischen Sicherheit

- Um zu einer formalen Definition zu kommen, müssen wir nun folgende Fragen beantworten:
  1. Über welche Rechenressourcen verfügt ein Gegner realistischerweise?
  2. Wie groß darf der vom Gegner erzielbare Vorteil höchstens sein, ohne die Vertraulichkeit der verschlüsselten Nachricht zu verletzen?
- Bezüglich Frage 1 geht man typischerweise davon aus, dass der Gegner über probabilistische Schaltkreise polynomieller Größe verfügt (siehe nächste Folie)
- Bezüglich Frage 2 verlangt man, dass der Gegner für jedes Polynom p höchstens für endlich viele Parameterwerte s einen Vorteil größer gleich $1/p(s)$ erzielen darf
- Andernfalls wäre die Sicherheit gefährdet, da er für jedes solche s nach polynomiell vielen Wiederholungen der probabilistischen Berechnung von $\mathcal{V}(s, x_0, x_1, y)$ fast sicher den richtigen Klartext ausfindig machen könnte, indem er das mehrheitlich berechnete Bit ausgibt

Definition

- Ein boolescher Schaltkreis der Größe m mit n Eingängen $x_1, \ldots, x_n$ und l Ausgängen $i_1, \ldots, i_l \in [m]$ ist eine Folge $c = (g_1, \ldots, g_m)$ von Gattern

$$g_l \in \{0, 1, x_1, \ldots, x_n, (\neg, j), (\wedge, j, k), (\vee, j, k)\} \text{ mit } 1 \leq j, k < l$$

- Der von c bei Eingabe $a \in \{0,1\}^n$ am Gatter g_l berechnete Wert $g_l(a) \in \{0,1\}$ ist induktiv wie folgt definiert:

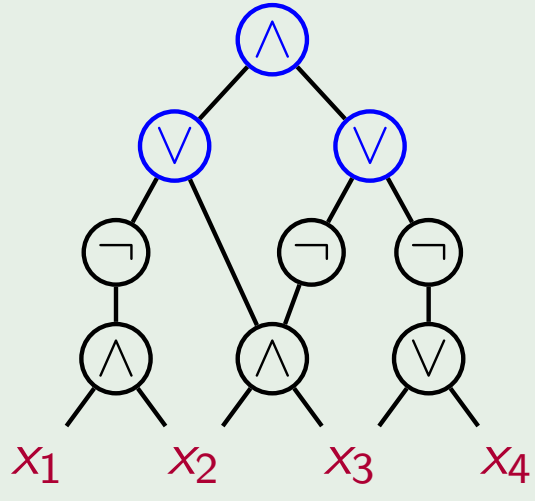
g_l	0	1	x_i	$(\neg, j)$	$(\wedge, j, k)$	$(\vee, j, k)$
$g_l(a)$	0	1	a_i	$1 - g_j(a)$	$g_j(a)g_k(a)$	$g_j(a) + g_k(a) - g_j(a)g_k(a)$

- Die Ausgabe von c bei Eingabe $a \in \{0,1\}^n$ ist die Bitfolge $c(a) = g_{i_1}(a) \ldots g_{i_l}(a)$

Beispiel

Der Schaltkreis

$$c = (x_1, x_2, x_3, x_4, (\wedge, 1, 2), (\wedge, 2, 3), (\vee, 3, 4), (\neg, 5), (\neg, 6), (\neg, 7), (\vee, 6, 8), (\vee, 9, 10), (\wedge, 11, 12))$$



mit den Eingängen x_1, x_2, x_3, x_4 und Ausgabegattern g_{11}, g_{12}, g_{13} gibt bei Eingabe $a = 0110$ die Bitfolge $c(0110) = 100$ aus

◁

Präzisierung der komplexitätstheoretischen Sicherheit

- Ein probabilistischer Schaltkreis c hat neben den regulären Eingabegattern $x_1, \ldots, x_n$ noch eine Reihe von Zufallsgattern $z_1, \ldots, z_m$
- Hierbei werden die Eingabegatter x_i wie bisher mit den Bits a_i eines Eingabevektors $a = a_1 \ldots a_n \in \{0,1\}^n$ belegt, während die m Zufallsgatter unabhängig gleichverteilte Bits $\mathcal{Z}_1, \ldots, \mathcal{Z}_m$ erzeugen (d.h. es gilt $\Pr[\mathcal{Z}_1 \ldots \mathcal{Z}_m = b] = 2^{-m}$ für alle $b \in \{0,1\}^m$)
- Dadurch wird die Ausgabe $c(a, \mathcal{Z}_1, \ldots, \mathcal{Z}_m)$ zu einer Zufallsvariablen, die wir auch kurz mit $\mathcal{C}(a)$ bezeichnen

Präzisierung der komplexitätstheoretischen Sicherheit

Definition

Sei KS ein Kryptosystem mit variablem Sicherheitsparameter $s \in \mathbb{N}$

- Eine Funktion $\varepsilon : \mathbb{N} \to \mathbb{R}$ heißt vernachlässigbar, wenn für jedes Polynom p eine Zahl $n_0 \in \mathbb{N}$ existiert, so dass $\varepsilon(n) < 1/p(n)$ für alle $n \geq n_0$ gilt
- Ein Gegner $G = (\mathcal{X}, \mathcal{V})$ heißt effizient, wenn für jedes s probabilistische Schaltkreise c_s und c'_s der Größe $s^{O(1)}$ mit $\mathcal{C}_s = (\mathcal{X}_0(s), \mathcal{X}_1(s))$ und $\mathcal{C}'_s(x_0, x_1, y) = \mathcal{V}(s, x_0, x_1, y)$ existieren, wobei die Ein- und Ausgaben von c_s und c'_s binär kodiert sind
- KS heißt komplexitätstheoretisch sicher, wenn jeder effiziente Gegner G nur einen vernachlässigbaren Vorteil erzielen kann (d.h. die Funktion $\alpha_G(s)$ ist vernachlässigbar)