Kryptographie und Komplexität

Einheit 2.2

Blockbasierte Kryptosysteme

1. Einfache Permutationschiffre
2. Affin-Lineare Chiffren
3. Methoden zur Steigerung der Sicherheit

Verringere Anfälligkeit gegen Häufigkeitsanalysen

- **Bisherige Kryptosysteme ersetzen Symbole**
  - Substitution von Buchstaben durch andere Symbole
  - Position von ersetzten Klartextsymbolen bleibt im Prinzip erhalten
  - Originalsymbole sind rekonstruierbar durch Häufigkeitsanalysen
- **Ersetzung von Textblöcken erhöht Sicherheit**
  - Blöcke von Klartextsymbolen werden in Schlüsseltextblöcke umgewandelt
  - Statistische Verteilung im Schlüsseltext sagt wenig aus (**Konfusion**)
  - Ein Klartextsymbol beeinflußt viele Schlüsseltextsymbole (**Diffusion**)
  - Zusammenhang zwischen Klar- und Schlüsseltexten wird komplexer
- **Unterschiedlich aufwendige Varianten**
  - Permutationschiffre: Vertauschung der Reihenfolge im Block
  - Affin-Lineare Chiffren: Anwendung von Matrizenmultiplikation
  - Substitutions-Permutations-Netzwerke: Kombination von Techniken
  - Zahlentheoretische Funktionen: Komplexe Abbildungen auf Σ^m

Blockchiffren und Permutationen

- **Blockchiffren sind Permutationen auf Σ^m**
  - Grund: Verschlüsselungsfunktionen sind invertierbar, also injektiv
    Klar- und Schlüsseltextraum sind üblicherweise identisch
  - Prinzipiell gibt es $(|\Sigma|^m)!$ mögliche Schlüssel (Permutationen)
    Anzahl wächst schon bei kleinen Blockgröße ins Unermessliche
- **Allgemeinstes Verfahren wäre zu komplex**
  - Fülle Text auf, so daß Textlänge Vielfaches der Blocklänge m wird
  - Ersetze Block $x_1..x_m$ durch $\pi(x_1..x_m)$, wobei π Permutation auf Σ^m
  - Tabelle zur Darstellung von π hat $|\Sigma|^m$ viele Einträge
  - Aufwand für Ersetzung eines Textblocks zu hoch
- **Realistische Verfahren müssen einfacher sein**
  - Schnell auszuführende Rechenvorschrift ersetzt Tabelle
  - Wichtiges Ziel ist hohe Konfusion und Diffusion

Permutationschiffre

- **Vertausche Elemente eines Buchstabenblocks**
  – Verwende Permutation π der Zahlen $1..m$ (Liste $[\pi(1),...,\pi(m)]$)
  · $e_K(x_1,..,x_m) = (x_{\pi(1)},...,x_{\pi(m)})$, ($\pi$ gibt an, wo die Elemente herkommen!)
  · $d_K(y_1,..,y_m) = (y_{\pi^{-1}(1)},...,y_{\pi^{-1}(m)})$
  Aus `ENDE UM ELF` wird mit π = [2 4 3 1] `NEDEU M L FE`
  (Letzter Viererblock durch Leerzeichen vervollständigt)

- **Geringer Aufwand für Ver- und Entschlüsselung**
  – Plazierung eines Symbols im Block (Tabellennachschlag) $\mathcal{O}(m)$
  bei geschickter Programmierung auch Aufwand $\mathcal{O}(1)$ möglich
  – Gesamtaufwand für Verschlüsselung eines Klartextwortes w $\mathcal{O}(|w|)$

- **Sicher gegen Brute-Force Attacken**
  – Es gibt $m!$ verschiedene Permutationen – Blocklänge 20 reicht
  – Gute Konfusion: Häufigkeitsverteilung entspricht der im Klartext
  – Geringe Diffusion: Ein Klartextsymbol beeinflußt ein Schlüsseltextsymbol

HILL CHIFFRE

Aufwendigere Codierung von Buchstabenblöcken

- **Linearkombinationen der Blockelemente**
  - Für $(y_1, .., y_m) = e_K(x_1, .., x_m)$ gilt $y_j = \sum_{i=1}^{m} x_i \cdot_n k_{i,j}$
  - Schlüsselelemente $k_{i,j}$ bilden eine $m{\times}m$ Matrix K
  - `ENDE UM ELF` $\hat{=}$ [4;13;3;4;26;20;12;26;4;11;5;26] wird mit $K = \begin{pmatrix} 11 & 8 \\ 3 & 7 \end{pmatrix}$ zu [2;22;18;25;22;24;21;8;23;1;25;6] $\hat{=}$ `CWSZWYUIXBZG`

    (Letzter Zweierblock durch Leerzeichen vervollständigt)
  - Diffusion: Jedes Klartextsymbol beeinflußt jedes Schlüsseltextsymbol

- **Entschlüsselung benötigt inverse Matrix K^{-1}**
  - Für $(x_1, .., x_m) = d_K(y_1, .., y_m)$ gilt $x_i = \sum_{j=1}^{m} k_{i,j}^{-1} \cdot_n y_j$
  - Für $K = \begin{pmatrix} 11 & 8 \\ 3 & 7 \end{pmatrix}$ ist $K^{-1} = \begin{pmatrix} 20 & 8 \\ 3 & 16 \end{pmatrix}$

- **Permutationschiffre ist Spezialfall der Hill-Chiffre**
  - Permutationsmatrix zu π ist $k_{i,j} = \begin{cases} 1 & \text{falls } i{=}\pi(j) \\ 0 & \text{sonst} \end{cases}$

- **$k\times m$ Matrix über einem Ring R**
  - Schema $A = (a_{i,j}) := \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,m} \\ \vdots & \vdots & \vdots & \vdots \\ a_{k,1} & a_{k,2} & \dots & a_{k,m} \end{pmatrix}$ mit Elementen aus R
  - Definition gilt für beliebige Ringe wie $\mathbb{R}$, $\mathbb{Q}$, oder $\mathbb{Z}/n\mathbb{Z}$ bzw. $\mathbb{Z}_n$
  - $\boldsymbol{R^{(k,m)}}$ ist die Menge aller $k\times m$ Matrizen über R
  - **Vektoren** sind Elemente von $R^{(1,m)}$ oder $R^{(k,1)}$
  - Computerdarstellung: Arrays oder Listen von Listen über R

- **Produkt von Matrizen und Vektoren**
  - Für $A = (a_{i,j}) \in R^{(k,m)}$ und $x = (x_j) \in R^{(1,k)}$
    ist $\boldsymbol{x \star A} = (y_1, .., y_m) \in R^{(1,m)}$ mit $y_j = \sum_{i=1}^{k} x_i \cdot a_{i,j}$
  - Komplexität der Berechnung über $\mathbb{Z}_n$ ist $\mathcal{O}(k{\cdot}m{\cdot}\|n\|^2)$
    - · Für jedes y_j sind k Multiplikationen und Additionen erforderlich
    - · Zugriff auf die $a_{i,j}$ ist jeweils in konstanter Zeit möglich

Mathematik: Operationen auf Matrizen

- **Addition von Matrizen**
  - Für $A, B \in R^{(k,m)}$ ist $\boldsymbol{A+B} = (c_{i,j}) \in R^{(k,m)}$ mit $c_{i,j} = a_{i,j} + b_{i,j}$
  - Komplexität der Berechnung über $\mathbb{Z}_n$ ist $\mathcal{O}(k{\cdot}m{\cdot}\|n\|)$
    - Für jedes $c_{i,j}$ ist jeweils eine Addition erforderlich
- **Produkt von Matrizen**
  - Für $A \in R^{(k,m)}$ und $B \in R^{(m,q)}$ ist $\boldsymbol{A \star B} = (c_{i,j}) \in R^{(k,q)}$ mit $c_{i,j} = \sum_{l=1}^{m} a_{i,l} \cdot b_{l,j}$
  - Komplexität der Berechnung über $\mathbb{Z}_n$ ist $\mathcal{O}(k{\cdot}m{\cdot}q{\cdot}\|n\|^2)$
    - Für jedes $c_{i,j}$ sind m Multiplikationen und Additionen erforderlich
    - Zugriff auf die $a_{i,l}$ und $b_{l,j}$ in jeweils konstanter Zeit möglich
- **Null- und Einheitsmatrix in $\boldsymbol{R^{(m,m)}}$**
  - $m{\times}m$ Nullmatrix **(0)**: Matrix, deren sämtliche Elemente $0 \in R$ sind
  - $m{\times}m$ Einheitsmatrix $\boldsymbol{E_m}$: Matrix $(e_{i,j})$ mit $e_{i,j} = \begin{cases} 1 \in R & \text{falls } j{=}i \\ 0 \in R & \text{sonst} \end{cases}$

$(\boldsymbol{R^{(m,m)}}, +, \star)$ ist ein (nichtkommutativer) Ring mit Einselement $\boldsymbol{E_m}$

Mathematik: Operationen auf Matrizen (II)

- **Determinante einer Matrix in $R^{(m,m)}$**
  – **det A** = $\begin{cases} a_{1,1} & \text{falls } m{=}1 \\ \sum_{j=1}^{m}(-1)^{i+j}a_{i,j}\det A_{i,j} & \text{sonst} \end{cases}$
  – Dabei ist i beliebig und $\boldsymbol{A_{i,j}}$ die Matrix A ohne Zeile i und Spalte j
  – Wichtig für Analysen und Konstruktion inverser Matrizen
  – Komplexität der Berechnung über $\mathbb{Z}_n$ ist $\mathcal{O}(m^3{\cdot}\|n\|^2)$
    · Explizite Definition benötigt $m!$ Additionen und Multiplikationen
    · Schnellere Algorithmen verwenden Gauß-Elimination und benötigen m^3 Additionen, Multiplikationen, Subtraktionen und Invertierungen

- **Inverse einer Matrix in $R^{(m,m)}$**
  – Für $A \in R^{(m,m)}$ ist $\boldsymbol{A^{-1}} = (\det A)^{-1} * A^*$
  – Dabei ist $\boldsymbol{A^*} = (\,(-1)^{i+j}\det A_{j,i}\,)$ die Adjunkte von A
  – **In $\mathbb{Z}_n$ ist A genau dann invertierbar wenn $gcd(\det A, n){=}1$**
  – Komplexität der Berechnung über $\mathbb{Z}_n$ ist $\mathcal{O}(m^3{\cdot}\|n\|^2)$

Mathematik: Affin lineare Abbildungen

Verbinde Lineare Algebra mit Modulararithmetik

- **Affin lineare Funktion über einem Ring R**
  - Funktion $f{:}R^m \to R^k$ mit $f(x) = x{\star}A + b$ für ein $A \in R^{(m,k)}$, $b \in R^{(k,1)}$
  - f heißt linear, wenn $b \equiv 0$
  - f ist genau dann bijektiv, wenn $m = k$ und $\det A$ invertierbar in R
  - Voraussetzung für Verwendung als Verschlüsselungsfunktion
- **Affin lineare Blockchiffre**
  - Blockchiffre über $\mathbb{Z}_n$, deren Verschlüsselungsfunktion affin linear ist
  - Vigenere Chiffre ist eine einfache affin lineare Blockchiffre
    $e_K(x_1, .., x_m) = (x_1, .., x_m) \star_n E_m + k$ mit Schlüssel $k \in R^{(m,1)}$
- **Hill Chiffren sind lineare Blockchiffren**
  - $\boldsymbol{e_K(x_1, .., x_m)} = (x_1, .., x_m) \star_n K$,
    $\boldsymbol{d_K(y_1, .., y_m)} = (y_1, .., y_m) \star_n K^{-1}$
  - Schlüssel ist $m{\times}m$ Matrix K mit $gcd(\det K, n){=}1$
  - Permutationschiffre ist sehr einfache lineare Blockchiffre

Hill Chiffre am Beispiel

- **Auswahl eines invertierbaren Schlüssels**
  - $K = \begin{pmatrix} 12 & 23 & 4 \\ 5 & 11 & 25 \\ 9 & 17 & 3 \end{pmatrix}$ liefert $K^{-1} = \begin{pmatrix} 16 & 5 & 18 \\ 3 & 0 & 23 \\ 16 & 12 & 23 \end{pmatrix}$
- **Verschlüsselung eines Klartextes**

  `FEST GEMAUERT IN DER ERDEN STEHT DIE FORM AUS LEHM GEBRANNT`
  - Umwandlung in Liste von Zahlen, aufgeteilt in Dreierblöcke

    [[5;4;18]; [19;26;6]; [4;12;0]; [20;4;17]; [19;26;8]; [13;26;3]; [4;17;26]; [4;17;3]; [4;13;26]; [18;19;4]; [7;19;26]; [3;8;4]; [26;5;14]; [17;12;26]; [0;20;18]; [26;11;4]; [7;12;26]; [6;4;1]; [17;0;13]; [13;19;26]]
  - Multiplikation jedes Blocks mit dem Schlüssel

    [[26;6;12]; [7;15;15]; [0;8;19]; [8;10;15]; [25;22;21]; [16;15;9]; [16;19;6]; [25;6;18]; [23;2;14]; [23;16;19]; [8;2;14]; [4;9;8]; [4;0;1]; [12;20;14]; [19;13;14]; [25;4;13]; [0;6;1]; [20;10;19]; [24;18;26]; [26;5;11]]
  - Rückumwandlung in Text

    `GMHPPAITIKPZWVQPJQTGZGSXCOXQTICOEJIEABMUOTNOZENAGBUKTYS  FL`
- **Entschlüsselung analog**
  - Multiplikation der entstehenden Blöcke mit K^{-1} liefert Originaltext

Berechnungsaufwand der Hill Chiffre

- **Aufwand für Auswahl des Schlüssels (einmalig)**
  - Alice wählt $K \in \mathbb{Z}_n^{(m,m)}$ mit $gcd(\det K, n){=}1$ pro test $\mathcal{O}(m^3{\cdot}\|n\|^2)$
    Bob bestimmt die inverse Matrix $K^{-1} \in \mathbb{Z}_n^{(m,m)}$ $\mathcal{O}(m^3{\cdot}\|n\|^2)$
- **Aufwand für Ver- und Entschlüsselung**
  - Umwandlung zwischen Buchstaben und Zahlen $\mathcal{O}(|w|)$
  - Multiplikation von $|w|/m$ Blöcken $\mathcal{O}(|w|{\cdot}m^2{\cdot}\|n\|^2)$
  - Gesamtaufwand bei festem Alphabet/kleinen Blöcken $\mathcal{O}(|w|)$
- **Relativ sicher gegen Ciphertext only Attacken**
  - Pro Versuch Aufwand $\mathcal{O}(|w|)$ für die Entschlüsselung
  - Anzahl der möglichen Schlüssel ist in $\mathcal{O}(n^{m^2})$
  - Brute-Force Attacke schon für $m{\geq}3$ undurchführbar
  - Häufigkeitsanalysen nur für maximale Blocklänge 3 möglich
    Es gibt keine statistischen Erhebungen für längere Buchstabenblöcke

Kryptoanalyse der Hill Chiffre

- **Einfache known plaintext Attacke**
  - Angreifer benötigt m Klar-/Schlüsseltextpaare (x_j, y_j) der Länge m
    · Es reicht eine Nachricht der Länge m^2 und ihre Verschlüsselung
  - Bilde zwei Matrizen $X := (x_{i,j})$ und $Y := (y_{i,j})$
  - Wegen $Y = X \star_n K$ ist $K = X^{-1} \star_n Y$ in $\mathcal{O}(m^3 \cdot \|n\|^2)$ zu berechnen und kann anhand weiterer Paare überprüft werden
  - Für allgemeine affin lineare Chiffre benötigt man $m+1$ Paare und bildet $X := (x_{i,j} - x_{m+1,j})$ und $Y := (y_{i,j} - y_{m+1,j})$
- **Attacke kann iterativ vorgehen**
  ... solange genügend Klar-/Schlüsseltextpaare gebildet werden können
  - Ist $gcd(\det X, n) \neq 1$ so wähle andere Kombination der Paare
  - Bei unbekannter Schlüsselgröße erhöhe m schrittweise
- **Ciphertext only Attacke für $m \leq 3$**
  - Analysiere häufigste Vorkommen von Bi-/Trigrammen im Schlüsseltext
  - Ordne diese den häufigsten 10 deutschen Bi-/Trigrammen zu
  - Maximal 90 bzw. 720 known plaintext Attacken durchzuführen

Kryptoanalyse der Hill Chiffre am Beispiel

- **Identifiziere Klar-/Schlüsseltextpaar**
  - Schlüsseltext `CIPHERTXT` $\hat{=}$ [2;8;15;7;4;17;19;23;19] wurde als Klartext `PLAINTEXT` $\hat{=}$ [15;11;0;8;13;19;4;23;19] identifiziert
  - Schlüssellänge 3 ist bekannt
- **Kombiniere erste drei Dreierblöcke**
  - Liefert $X := \begin{pmatrix} 15 & 11 & 0 \\ 8 & 13 & 19 \\ 4 & 23 & 19 \end{pmatrix}$, $\det X = 13$, und $Y := \begin{pmatrix} 2 & 8 & 15 \\ 7 & 4 & 17 \\ 19 & 23 & 19 \end{pmatrix}$
  - Berechne $X^{-1} = \begin{pmatrix} 26 & 7 & 20 \\ 5 & 15 & 12 \\ 24 & 23 & 26 \end{pmatrix}$ und $K = X^{-1} \star_n Y = \begin{pmatrix} 10 & 12 & 4 \\ 16 & 4 & 18 \\ 25 & 18 & 21 \end{pmatrix}$
- **Anwendung auf andere Schlüsseltexte**
  - Entschlüssele `PFYFZUSXFIHFPIII EAEPXWONQU QC BUEUKHBN` mit K^{-1} zu `DIE HILL CHIFFRE IST LEICHT ZU BRECHEN`

Erhöhung der Diffusion von Blockchiffren

- **Mehrfachverschlüsselung**
  - Vielfache Ver- und Entschlüsselung mit verschiedenen Schlüsseln
    - $\boldsymbol{e_K(x)} = e_{K_1}(d_{K_2}(e_{K_3}(x)))$
  - Vergrößert Schlüsselraum und Diffusion, ohne große Erhöhung des Berechnungsaufwands (z.B. in Subtitutions-Permutations Netzwerken)
  - Sinnlos bei linearen Chiffren, da $e_{K_1}(d_{K_2}(e_{K_3}(x))) = e_{K_1 \star K_2^{-1} \star K_3}(x)$
- **Komplexere Verschlüsselung langer Texte**
  - Standard ECB-Mode wandelt gleiche Klartextblöcke in dieselben Schlüsseltextblöcke um (ermöglicht Häufigkeitsanalysen / Fälschungen)
  - CBC-Mode macht Verschlüsselung auch von Vorgängerblöcken abhängig
  - CFB- und OFB-Modi erlauben mit der Entschlüsselung zu beginnen, bevor der gesamte Block empfangen wurde
    (effizienter bei sehr großen Blöcken, nur für symmetrische Verfahren geeignet)
- **Neue mathematische Theorien**
  - Zahlentheoretische Funktionen mit großer Diffusion (z.B. ECC)