Seminar Kryptographie und Datensicherheit

Einfache Kryptosysteme und ihre Analyse

Christoph Kreitz

1. Grundlagen von Kryptosystemen
2. Buchstabenorientierte Systeme
3. Blockbasierte Verschlüsselung
4. Strombasierte Codierungen

Wozu Kryptosysteme?

Sichere Übertragung geheimer Botschaften

- **Übertragungskanäle sind oft unsicher**
  - Nachricht könnte evtl. abgehört werden
  - Originaltext zu übertragen ist unsicher
- **Verschlüsselung macht Botschaft unlesbar**
  - Unbefugte sollen abgehörte Nachricht nicht decodieren können
  - Zieladressat muß Originaltext leicht wiederherstellen können
- **Kryptographie kann mehr als nur das**
  - Vertraulichkeit: Nachricht kann von Dritten nicht gelesen werden
  - Integrität: Fälschung/Manipulation der Nachricht ist nicht möglich
  - Authentizität: Nachweis, daß Nachricht vom angegebenen Sender stammt
  - Verbindlichkeit: Sender kann Urheberschaft nicht nachträglich leugnen

Grundbestandteile von Kryptosystemen

- **Klartexte**
  - Endliche Menge $\mathcal{P}$ ("plaintexts") von Wörtern eines Alphabets
  - Nachrichten, die vom Sender zum Empfänger gehen sollen
- **Chiffretexte (Schlüsseltexte)**
  - Endliche Menge $\mathcal{C}$ ("ciphertexts") von Wörtern eines Alphabets
  - Texte, die tatsächlich übermittelt werden
- **Schlüssel**
  - Endliche Menge $\mathcal{K}$ ("keyspace") möglicher Schlüssel
  - Daten, die essentiell für Codierung und Decodierung sind
- **Ver- und Entschlüsselungsregeln**
  - Funktionen $e_K : \mathcal{P} \rightarrow \mathcal{C}$ und $d_K : \mathcal{C} \rightarrow \mathcal{P}$ für jeden Schlüssel $K \in \mathcal{K}$ mit der Eigenschaft $d_K(e_K(x)) = x$ für jeden Klartext $x \in \mathcal{P}$

Einfache Kryptosysteme

Verschlüsselungen, die 'von Hand' durchführbar sind

- **Verschiebungschiffre** **(Shift Cipher)**
  – Zyklische Verschiebung der Buchstaben im Alphabet
- **Affin-Lineare Chiffre** **(Affine Cipher)**
  – Sprunghafte Verschiebung durch Verwendung affiner Funktionen
- **Substitutionschiffre** **(Substitution Cipher)**
  – Ersetzung von Buchstaben durch Permutation des Alphabets
- **Vigenere Chiffre** **(Vigenere Cipher)**
  – Verschiebung von Buchstabengruppen mit Schlüsselwort
- **Hill Chiffre** **(Hill Cipher)**
  – Codierung eines Buchstabenblocks durch Linearkombinationen
- **Permutationschiffre** **(Permutation Cipher)**
  – Permutiere Elemente von Buchstabengruppen innerhalb des Textes
- **Strom Chiffren** **(Stream Ciphers)**
  – Verschiebungschiffren mit ständig wechselnden Schlüsseln

Verschiebungschiffre

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z				
X	Y	Z		A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z

- **Buchstaben werden um festen Betrag verschoben**
  - Chiffretext entsteht durch Vorwärtsverschieben

    Aus `ENDE UM ELF` wird `AJ AWQIWAHB`
  - Originaltext durch einfaches Rückwärtsschieben ermittelbar
- **Programmierbar mit Modulararithmetik**
  - Bei n Buchstaben erhält jeder Buchstabe eine Zahl zwischen 0 und $n-1$
  - Schlüssel K ist Zahl zwischen 0 und $n-1$
  - Ver- und Entschlüsselung wird Addition/Subtraktion modulo n

    $e_K(x) = x + K \bmod n, \quad d_K(y) = y - K \bmod n$
  - Im Beispiel: $n = 27$, $K = 23$

Verschiebungschiffren sind leicht zu knacken

- **Brute-force Attacke: Ausprobieren aller Schlüssel**
  - Einfache, aber gefährliche Attacke auf Verschlüsselungssysteme
  - Computer ermöglichen Überprüfung von Milliarden von Schlüsseln
  - Sicherheit nur bei extrem großer Anzahl möglicher Schlüssel ($\geq$128bit)
- **Verschiebungschiffre hat maximal n Schlüssel**
  - Austesten aller n Entschlüsselungen erzeugt Texte
  - Der Originaltext ist einer der erzeugten Texte
  - Nur wenige erzeugte Texte sind sprachlich akzeptabel (Wörterbuch)
  - Akzeptable Texte können "von Hand" untersucht werden
  - Aus `AJ AWQIWAHB` wird schrittweise `ENDE UM ELF` ($K = 23$)

Kryptoanalyse

Systematische Techniken zur Codeentschlüsselung

- **Annahme: Verschlüsselungsverfahren ist bekannt**
  - Ansonsten überprüft Angreifer schrittweise alle gängigen Verfahren
- **Angreifer versucht Schlüssel aufzufdecken**
  - Das ist mehr als Decodierung einer Nachricht
  - Im Erfolgsfall sind alle zukünftigen Nachrichten ungeschützt
- **Viele Angriffsszenarien**
  - Brute Force: Austesten aller Möglichkeiten
  - Ciphertext only: Entschlüsselung durch Statistische Analyse
  - Known plaintext: Paare von Klar-/Schlüsseltexten sind bekannt
    Bei vielen Paaren wird verwendeter Schlüssel ableitbar
  - Chosen Plaintext: Verwendung des Verschlüsselungsverfahrens
    Liefert Paare von Klar-/Schlüsseltexten und ggf Schlüssel
  - Chosen Ciphertext: Verwendung des Entschlüsselungsverfahrens

Affin-Lineare Chiffre

Erweiterung der Verschiebung auf affine Funktionen

- **Modulare Addition, Multiplikation und Inverse**
  - Schlüsselpaar (a, b) liefert $e(x) = ax + b \bmod n$,
    und $d(y) = a^{-1}(y - b) \bmod n$
  - Verschiebungschiffre ist Spezialfall $a = 1$
- **Eindeutig genau dann, wenn a und n teilerfremd**
  - Es gibt $\phi(n) * n$ mögliche Schlüssel (Brute Force Attacke möglich)
    $\phi(n)$ = Anzahl der zu n teilerfremden Zahlen in $\mathbb{Z}_n$
  - Satz: ist $p_1^{e_1}..p_k^{e_k}$ die Primfaktorzerlegung von n
    so ist $\phi(n) = (p_1^{e_1} - p_1^{e_1-1}) * .. * (p_k^{e_k} - p_k^{e_k-1})$
  - Für Primzahlen p ist $\phi(p) = p-1$
- **Beispiel $n = 27$** ($\phi(27) = 18$, also 486 mögliche Schlüssel)
  - $K = (8, 3)$ liefert $e_K(x) = 8x + 3 \bmod 27$, $d_K(y) = 17(y-3) \bmod 27$
  - Aus `ENDE UM ELF` $\hat{=}$ (4 13 3 4 26 20 12 26 4 11 5)
    wird (8 26 0 8 22 1 18 22 8 10 16) $\hat{=}$ `I AIWBSWIKQ`

Substitutionschiffre

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	
X	N	Y	A	H	P	O	G	Z	Q		W	B	T	S	F	L	R	C	V	M	U	E	K	J	D	I

- **Buchstaben werden permutiert**
  - Chiffretext entsteht durch entsprechende Ersetzung der Buchstaben

    Aus `ENDE UM ELF` wird `HTAHIMBIHWP`
  - Originaltext durch inverse Permutation ermittelbar
- **Etwas sicherer als die Verschiebungchiffre**
  - Es gibt $n!$ verschiedene Permutationen
  - Brute-Force Attacke bei mehr als 26 Buchstaben nicht möglich
- **Anfällig für statistische Analysen**
  - Bei langen Texten erlaubt Buchstabenhäufigkeit Rückschlüsse
  - Deutsch: E$\widehat{=}$14.7%, N$\widehat{=}$8.8%, R$\widehat{=}$6.8%, ... Q$\widehat{=}$0.014%, X$\widehat{=}$0.013%
  - Im Beispiel ist die Zuordnung $H \equiv E$ sofort erkennbar

Vigenere Chiffre

Einfaches ‘polyalphabetisches’ Kryptosystem

- **Verschiebe m Buchstaben asynchron**
  - Chiffretext entsteht durch Vorwärtsverschieben von Buchstabenblöcken
  - Schlüssel verschiebt jedes Element des Blocks unterschiedlich
  - `ENDE UM ELF` wird mit Schlüssel `KEY` (10 4 24) zu `PRAPDRWDBVJ`
  - Originaltext durch Rückwärtsschieben der Blöcke ermittelbar
- **Programmierung ähnlich zur Verschiebungschiffre**
  - Bei n Buchstaben erhält jeder Buchstabe eine Zahl aus $\mathbb{Z}_n$
  - Schlüssel K ist Zahlentupel aus $(\mathbb{Z}_n)^m$
  - Simultane Ver-/Entschlüsselung durch Addition/Subtraktion modulo n
    - $e_K(x_1, .., x_m) = (x_1 + k_1, ..., x_m + k_m) \bmod n$,
    - $d_K(y_1, .., y_m) = (y_1 - k_1, ..., y_m - k_m) \bmod n$
- **Brute Force Attacken schwer**
  - Es gibt $62^8 = 2.2 * 10^{14}$ achtbuchstabige alphanumerische Schlüssel
  - Zu einfache Schlüssel ermöglichen Wörterbuchattacke (200.000 Tests)

Kryptoanalyse des Vigenere Chiffre

Aufwendige Wahrscheinlichkeitsanalysen erforderlich

- **Kasaki Test**
  - Identische Schlüsseltextsegmente haben gleichen Klartext, wenn Abstand Vielfaches der Blocklänge ist
  - Suche identische Segmente und bestimme mögliche Abstände (Mindestgröße der Segmente sollte 3 sein)
  - Liefert Informationen über mögliche Blocklängen
- **Koinzidenzanalyse**
  - Unterschiedliche Wahrscheinlichkeiten einzelner Buchstaben im Klartext liefern Wahrscheinlichkeiten für verwendete Teilschlüssel
  - Analysen ohne Computerunterstützung kaum durchführbar

Hill Chiffre

Aufwendige Codierung von Buchstabenblöcken

- **Verschlüsselung durch Matrixmultiplikation**
  - i-tes Element des Schlüsseltextblocks ist Linearkombination der x_j
  - Schlüssel muß invertierbare $m \times m$ Matrix K sein
    - $e_K(x_1, .., x_m) = (x_1, .., x_m) \star K \bmod n$,
    - $d_K(y_1, .., y_m) = (y_1, .., y_m) \star K^{-1} \bmod n$
- **Verbinde Lineare Algebra mit Modulararithmetik**
  - Spezialgesetze erleichtern Invertierung von Matrizen modulo n
  - K^{-1} ist $(\det K)^{-1} * K^* \bmod n$ (inverse Determinante * adjungierte Matrix)

    z.B. $K = \begin{pmatrix} 11 & 8 \\ 3 & 7 \end{pmatrix}$ liefert $K^{-1} = \begin{pmatrix} 20 & 8 \\ 3 & 16 \end{pmatrix}$
  - `ENDE UM ELF` $\hat{=}$ (4 13 3 4 26 20 12 26 4 11 5 26) wird mit Schlüssel K

    zu (2 22 18 25 22 24 21 8 23 1 25 6) $\hat{=}$ `CWSZWYUIXBZG`

    (Letzter Zweierblock wurde durch Leerzeichen vervollständigt)

Relativ sicher gegen “Ciphertext only” Attacken

- **Known Plaintext Attacke relativ einfach**
  - Angreifer hat mindestens m Klar-/Schlüsseltextpaare (x_j, y_j) erhalten
  - Bilde zwei Matrizen $X := (x_{i,j})_{i,j=1}^m$ und $Y := (y_{i,j})_{i,j=1}^m$
    (Länge der Texte ist mindestens Blocklänge m)
  - Wegen $Y = X \star K \bmod n$ ist $K = X^{-1} \star Y \bmod n$
- **Attacke kann iterativ vorgehen**
  - Bei unbekannter Schlüssellänge erhöhe m schrittweise solange genügend Paare vorhanden
  - Lange Klar-/Schlüsseltextpaare liefern viele Blöcke der Länge m
  - Ist X nicht invertierbar, wähle andere Klar-/Schlüsseltextpaare

Permutationschiffre

Codierung ohne Verschiebung im Alphabet

- **Vertausche Elemente eines Buchstabenblocks**
  - Verwende Permutation π der Zahlen $1..m$
    - $e_K(x_1, .., x_m) = (x_{\pi(1)}, ..., x_{\pi(m)})$,
    - $d_K(y_1, .., y_m) = (x_{\pi^{-1}(1)}, ..., x_{\pi^{-1}(m)})$

    Aus `ENDE UM ELF` wird mit $\pi = (2\ 4\ 3\ 1)$ `NEDEU M L FE`

    (Letzter Viererblock wurde durch Leerzeichen vervollständigt)
  - Ver- und Entschlüsselung sehr leicht durchzuführen
- **Als Spezialfall des Hill-Chiffre darstellbar**
  - Beschreibe Vertauschung von Blockelementen in Permutationsmatrix
- **Sicher nur bei sehr großen Blöcken**
  - Es gibt $m!$ verschiedene Permutationen
  - Blocklänge unter 15 erlaubt Brute-Force Attacke mit PC's

Codierung mit “unendlicher” Blocklänge

- **Generiere Strom von Schlüsseln $k_1k_2k_3$....**
  - Codiere Klartext $x_1x_2x_3$... als $e_{k_1}(x_1)e_{k_2}(x_2)e_{k_3}(x_3)$...
  - Schlüsselstrom muß systematisch erzeugbar sein
  - Empfänger muß Schlüsselstrom für Decodierung erzeugen können
- **Synchrone Erzeugung des Schlüsselstroms**
  - Schlüsselstrom wird ausschließlich aus einem Basisschlüssel K erzeugt

    z.B. Fibonaccizahlen modulo n: 1 2 3 5 8 13 21 7 1 8 9 17 26 16 15 ...
  - Blockchiffren entsprechen periodischen Schlüsselströmen
- **Asynchrone Erzeugung des Schlüsselstroms**
  - Klartext wird in Schlüsselerzeugung mit einbezogen

    z.B. letzer Klartextbuchstabe wird Schlüssel für nächsten Buchstaben

Einfacher Asynchroner Strom Chiffre

- **Vigenere Chiffre mit "Klartext als Schlüssel"**
  - Wähle $k_1 = K$ (der geheime Schlüssel) und als $k_{i+1} = x_i$

    $e_K(x_i) = x_i + k_i \bmod n, \quad d_K(y_i) = y_i - k_i \bmod n$
  - `ENDE UM ELF` $\hat{=}$ (4 13 3 4 26 20 12 26 4 11 5) wird mit $K = 3$

    zu (7 17 16 7 3 19 5 11 3 15 16) $\hat{=}$ `HRQHDTFLDPQ`
- **Relativ sicher gegenüber statistischen Analysen**
  - Regelmäßigkeit des Alphabets wird aufgehoben
  - Brute-Force Attacke nur durch lange Startschlüssel vermeidbar

Einfache Kryptosysteme im Rückblick

- **Buchstabenorientierte Systeme**
  – Verschiebung und Substitution im Alphabet
  – Anfällig für Brute-Force oder statistische Attacken
- **Blockbasierte Verschlüsselung**
  – Vigenere & Hill-Chiffren, Permutationen
  – Zu brechen durch Spezialverfahren mit genügend Information
- **Strombasierte Verschlüsselung**
  – Macht statistische Analysen schwer
  – Synchrone und asynchrone Varianten
  – Zu brechen, da Schlüsselstrom generierbar sein muß

Keine Sicherheit im Computerzeitalter