

Security Lab – Entropie und Block Cipher Modes

VMware

Dieses Lab können Sie mit dem **Ubuntu-Image** durchführen. In der Aufgabenstellung wird angenommen, dass Sie mit diesem Image arbeiten.

Alternativ können Sie das Lab auch auf Ihrem eigenen Laptop bearbeiten. Dazu benötigen Sie Java und eine Entwicklungsumgebung. Weiter unten ist erklärt, wie Sie das bereitgestellte Projekt in *NetBeans* oder *Eclipse* importieren können. Ebenfalls benötigen Sie ein Programm zum Entpacken von ZIP Files.

1 Einleitung

In diesem Praktikum studieren Sie zudem einen zentralen Unterschied der beiden Betriebsmodi Cipher Block Chaining (CBC) und Electronic Code Book (ECB) von blockbasierten Verschlüsselungsverfahren (Block Cipher). Diese Modi werden von den meisten symmetrischen Verschlüsselungsalgorithmen angeboten.

Hinweis: Ist in diesem Praktikum eine Zahl ohne Kommastellen angegeben, handelt es sich um den exakten Wert. Beispiel: 0, 4, 17/2. Ist eine Zahl mit Kommastellen angegeben, ist der korrekte Wert auf die angegebene Anzahl Kommastellen gerundet. Beispiel: korrekter Wert 1.9994872 wird zu 2.00.

2 Grundlagen für dieses Lab

- Laden Sie *Seclab_BlockCipherModes.zip* von OLAT herunter. Die Datei enthält das Softwareprojekt, mit dem Sie in diesem Praktikum arbeiten werden, sowie einige weitere benötigte Dateien.
- Verschieben Sie die Datei an einen geeigneten Ort (auf dem Ubuntu-Image z.B. in ein Verzeichnis *securitylabs* in */home/user*) und entzippen Sie sie. Dies erzeugt ein Verzeichnis *Seclab_BlockCipherModes*.
- Um das Projekt in *NetBeans* (ist auf dem Ubuntu-Image installiert) zu importieren und zu bauen, gehen Sie wie folgt vor:
 - Wählen Sie *File* → *New Project...* → *Java with Ant* → *Java Free-Form Project*, dann *Next*.
 - *Name and Location*: Wählen Sie bei *Location* das Unterverzeichnis *code* im oben beim Entzippen erzeugten Verzeichnis *Seclab_BlockCipherModes* (d.h. *Seclab_BlockCipherModes/code*) aus. Die anderen Felder füllen sich automatisch. Klicken Sie *Next*.
 - *Build and Run Actions*: Klicken Sie *Next*.
 - *Source Package Folders*: Wählen Sie bei *Source Package Folders* mit dem Button *Add Folder...* das Verzeichnis *src* (im Verzeichnis *code*) aus. Wählen Sie bei *Source Level* zudem die neueste JDK Version. Klicken Sie *Next*.
 - *Java Sources Classpath*: Klicken Sie *Add JAR/Folder...* und wählen Sie *lib/jmathplot.jar* (im Verzeichnis *code*) aus. Klicken Sie *Next*.
 - *Project Output*: Klicken Sie *Add JAR/Folder...* und wählen Sie das Verzeichnis *jar* (im Verzeichnis *code*) aus. Klicken Sie *Finish*
 - Um das Projekt zu bauen: Rechts-Klick auf *build.xml* → *Run Target* → *JarFiles*. Funktioniert alles, befinden sich vier jar-Files in *Seclab_BlockCipherModes/code/jar*.
- Um das Projekt in *Eclipse* zu importieren und zu bauen, gehen Sie wie folgt vor:
 - Wählen Sie *File* → *Import...* → *General* → *Existing Projects into Workspace*, dann *Next*.

- Wählen Sie bei *Select root directory* das Unterverzeichnis *code* im oben beim Entzippen erzeugten Verzeichnis *SecLab_BlockCipherModes* (d.h. *SecLab_BlockCipherModes/code*) aus. Klicken Sie *Finish*.
- Um das Projekt zu bauen: Rechts-Klick auf *build.xml* → *Run as* → *Ant Build*. Funktioniert alles, befinden sich vier jar-Files in *SecLab_BlockCipherModes/code/jar*.
- Starten Sie die erzeugten jar-Files jeweils via Kommandozeile (Windows) resp. einem Terminal (Linux/macOS) wie folgt:
 - `java -jar jarfile <arguments>`

3 Entropie und Work Factor

Der Begriff *Entropie* stammt eigentlich aus der Thermodynamik, wird aber in der Kryptographie mit einer bestimmten Bedeutung verwendet, die eng mit dem Begriff des *work factor* verknüpft ist. Formal sei $X = \{x_1, \dots, x_n\}$ eine endliche Menge möglicher Ergebnisse eines Experiments. Das können beispielsweise die sechs Seiten eines Würfels sein, die nach einem Wurf oben liegen können; die zwei Seiten einer Münze; die verwendeten Schlüssel eines Kryptosystems oder die verschiedenen Blöcke eines Ciphertexts. Laut random oracle model sollten diese Blöcke ja zufällig gewählt sein. Es sei nun weiter für $1 \leq i \leq n$ mit p_i die Wahrscheinlichkeit bezeichnet, dass x_i als Ergebnis des Experiments auftaucht. Für einen fairen Würfels ist beispielsweise $X = \{1, 2, 3, 4, 5, 6\}$ und es ist wegen der Fairness $p_i = 1/6$. Ist der Würfel nicht fair, ist zwar X unverändert, aber die p_i sind dann nicht mehr alle gleich. Unter der *Entropie von X* versteht man nun den Ausdruck

$$H(x) = - \sum_{i=1}^n p_i \log p_i .$$

Die Entropie wird in bit angegeben. Manchmal (so z.B. in Abschnitt 4) wird Entropie *pro bit* angegeben. In diesem Fall ist die Entropie ohne Einheit, bzw. hat die Einheit bit/bit.

Im Folgenden betrachten wir eine nicht näher ausgeführte Spielzeug-Verschlüsselung mit 4 bit Schlüssellänge. Sie machen nun mit zwei verschiedenen Systemen G (für *good*) und B (für *bad*) Experimente und stellen fest, dass die verschiedenen Schlüssel bei den beiden Systemen mit verschiedenen Wahrscheinlichkeiten $p_{i,G}$ und $p_{i,B}$ ausgewählt werden:

Schlüssel	$p_{i,G}$	$p_{i,B}$	Schlüssel	$p_{i,G}$	$p_{i,B}$
0000	1/16	$1/2^1$	1000	1/16	$1/2^9$
0001	1/16	$1/2^2$	1001	1/16	$1/2^{10}$
0010	1/16	$1/2^3$	1010	1/16	$1/2^{11}$
0011	1/16	$1/2^4$	1011	1/16	$1/2^{12}$
0100	1/16	$1/2^5$	1100	1/16	$1/2^{13}$
0101	1/16	$1/2^6$	1101	1/16	$1/2^{14}$
0110	1/16	$1/2^7$	1110	1/16	$1/2^{15}$
0111	1/16	$1/2^8$	1111	1/16	$1/2^{15}$

Wie bei jeder Wahrscheinlichkeitsverteilung sollte auch hier die Summe der Wahrscheinlichkeiten Eins ergeben. Verifizieren Sie das für G und B.

16 * 1/16 = 1
1/2 + 1/4 + 1/8 + ... + 1/32768 = 1

Versetzen Sie sich jetzt in die Lage eines Angreifers auf System G. Sie wollen das System knacken, müssen dazu aber Schlüssel einen nach dem anderen ausprobieren. Begründen Sie, warum die Reihenfolge, in der Sie die Schlüssel ausprobieren, keinen Einfluss auf die zu erwartende Anzahl der Proben hat, die Sie brauchen, bis Sie den richtigen Schlüssel herausgefunden haben.

Da jeder Schlüssel die identische Wahrscheinlichkeit hat. So ist es gleich wahrscheinlich dass Schlüssel "0000" gewählt wurde, wie dass Schlüssel "1111" gewählt wurde.

Berechnen Sie nun den work factor von G (korrekte Antwort: 17/2). Berechnen Sie diesen work factor auch in bit (korrekte Antwort: 3.09).

work factor = average number of keys to try -> Wenn alle keys gleichwahrscheinlich sind, ist der work factor in etwa 0.5*key space size
Anzahl keys = 16
16 / 2 = 8
8 + 0.5 = 8.5 = 17/2
 $\log_2(8.5) = 3.087$

Berechnen Sie nun die Entropie von G (korrekte Antwort: 4)

16 * 1/16 * $\log_2(1/(1/16)) = 4$

$$H = \sum_{i=1}^n p(i) \log_2 \left(\frac{1}{p(i)} \right)$$

Die Schlüssel von G haben also etwa 3.0 bit work factor und 4 bit Entropie.

Versetzen Sie sich jetzt in die Lage eines Angreifers auf System B. Sie wollen wieder das System knacken, müssen dazu aber wieder Schlüssel einen nach dem anderen ausprobieren. Begründen Sie, warum die Strategie guten Erfolg verspricht, Schlüssel in absteigender Reihenfolge ihrer Wahrscheinlichkeit auszuprobieren.

Da der Schlüssel "0000" die grösste Wahrscheinlichkeit hat -> $p(0000) = 1/2$, wenn man absteigend durchgeht, überprüft man immer den Schlüssel (von denen noch nicht getestet wurden) mit der höchsten Wahrscheinlichkeit

Berechnen Sie nun den work factor von B (korrekte Antwort: 2.00). Berechnen Sie diesen work factor auch in bit (korrekte Antwort: 1.00).

$I * (1/2^A)$
 $1 * (1/2^1) + 2 * (1/2^2) + \dots + 15 * (1/2^{15}) + 16 * (1/2^{16}) = 1.9999 = 2.00$
In Bit: $\log_2(2) = 1$

Berechnen Sie nun die Entropie von G (korrekte Antwort: 2.00)

$0.5*\log(1/0.5) + 0.25*\log(1/0.25) + 0.125*\log(1/0.125) + \dots + (0.00003051*\log(1/0.00003051) = 2$

$$H = \sum_{i=1}^n p(i) \log_2 \left(\frac{1}{p(i)} \right)$$

Das System B hat also nur noch etwa 1.0 bit work factor und 2.0 bit Entropie.

Wir haben also gesehen, dass je nach Verteilung der Schlüssel und daher natürlich auch je nach Wissenstand des Angreifers ein System einen deutlich niedrigeren work factor (und auch Entropie) haben kann, als es nach der Schlüssellänge allein eigentlich zu erwarten war.

Stellen Sie nun aufgrund dieser (zugegeben etwas dünnen) Datenlage eine Vermutung auf, wie sich Entropie und work factor in etwa zueinander verhalten könnten. Eine Begründung ist nicht nötig, aber Ihre Vermutung muss zu den beobachteten Fakten passen.

Grundsätzlich kann man sagen, dass der Work-Factor ca. 0.5 der Entropy entspricht. Dies stimmt jedoch nicht in jedem Fall, sondern ist abhängig von der Strategie des Angreifers.

Stellen Sie nun ebenfalls eine Vermutung auf, bei welcher Verteilung der Schlüssel der work factor (und damit auch die Entropie) am grössten sind. Eine Begründung ist nicht nötig, aber Ihre Vermutung muss zu den beobachteten Fakten passen.

Wenn die Wahrscheinlichkeiten pro Schlüssel gleich gross sind, so kommen wir auf eine grosse Entropy, wie auch einen grossen Work-factor. Dies hat den einfachen Grund, dass die Entropy aussagt, wie viel Informationen (oder Überraschungswert) enthalten sind. Wenn jeder Schlüssel die selbe Wahrscheinlichkeit hat, so ist der Informationsgehalt pro Schlüssel maximiert, da man am wenigsten aussagen kann, welcher Schlüssel vorkommen wird.

Bei einer grossen Entropie, kann der Angreifer schlechter vorhersagen welches Bit gesetzt ist, somit benötigt er einen grossen Aufwand (work factor) für die Entschlüsselung.

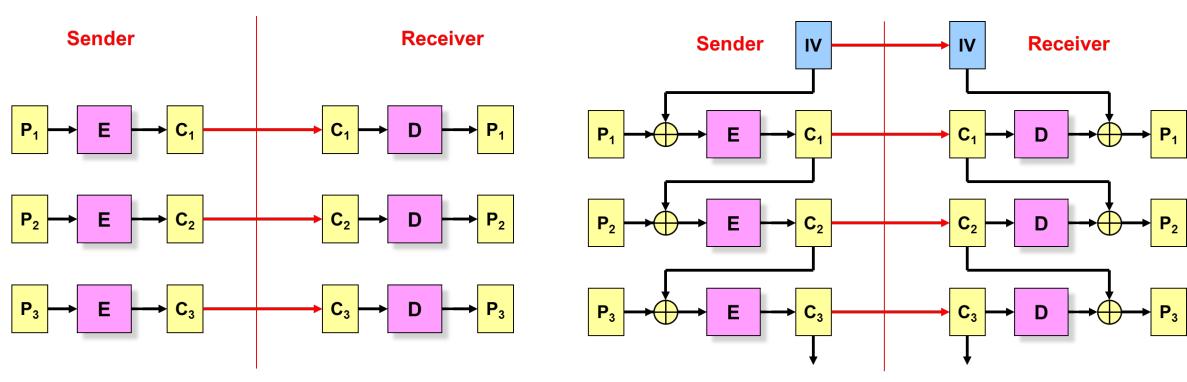
Im Idealfall verwandelt eine Verschlüsselung einen Plaintext in einen Ciphertext, der von rein zufälligem Text nicht zu unterscheiden ist. In diesem Fall sei $X = \{0,1\}$ die Menge der im Text auftretenden bits. Was gilt für die Wahrscheinlichkeiten p_0 und p_1 , mit denen ein Null- bzw Eins-bit auftritt? Welche Entropie hat also ein bit des Klartexts in diesem Fall? Das Ergebnis wird in Abschnitt 4 noch gebraucht. (Korrekte Ergebnisse: 1 bit.)

Die Wahrscheinlichkeit beträgt 0.5

$2 * 0.5 * \log_2(1/0.5)$

4 Block Cipher Modes

In der Vorlesung haben wir als eine informationstheoretisch sichere Verschlüsselung das One-Time-Pad kennengelernt. Das ist aber reichlich impraktikabel, da die Schlüssellänge so lang wie der Plaintext sein muss und zudem nur für eine Nachricht verwendbar ist. Deshalb werden typischerweise blockbasierte Verschlüsselungsverfahren (*block ciphers*) eingesetzt. Eine Block Cipher ist ein deterministisches Verschlüsselungsverfahren, bei dem ein Plaintext fester Länge auf einen Ciphertext fester Länge abgebildet wird. Die genaue Transformation wird dabei durch einen Schlüssel bestimmt und lässt sich durch das *random oracle model* beschreiben, bei dem durch den Schlüssel eine zufällige Permutation der Eingabewerte auf die Ausgabewerte bestimmt wird. Im Gegensatz zu einer *Stream Cipher* kann ein Block Cipher nur ganze Blöcke verschlüsseln. Zur Verschlüsselung grösserer Datenmengen wird ein Betriebsmodus verwendet, der festlegt, wie die Block Cipher wiederholt anzuwenden ist. Zwei dieser Modi sind Electronic Code Book (ECB) und Cipher Block Chaining (CBC). Die nachfolgenden Grafiken zeigen die grundlegenden Funktionsweisen:



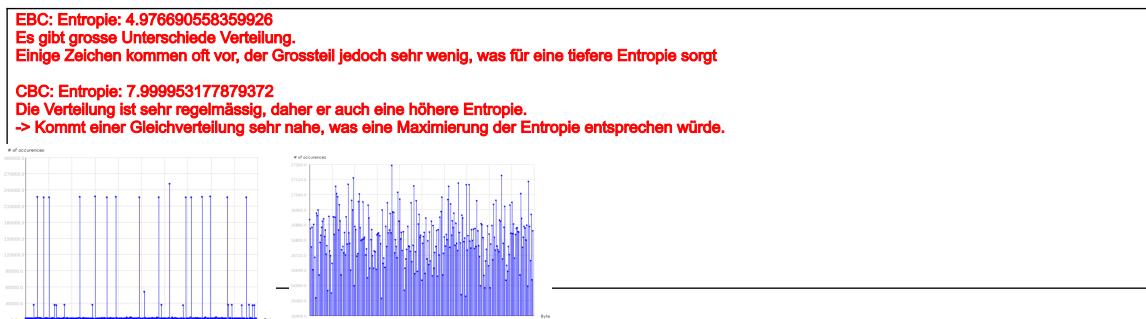
Electronic Code Book Mode

Cipher Block Chaining Mode

Sicherheitsexperten raten typischerweise dringend von der Verwendung von Block Ciphers im ECB Mode ab. Einer der Hauptgründe für diesen Rat sollte Ihnen nach der Durchführung des folgenden Experiments klar werden. **Hinweis:** Je nach Plaintext und je nach Angreifer kann auch die Verwendung von CBC problematisch sein. Näheres entnehmen Sie bitte der Vorlesung.

Verwenden Sie die ausführbare jar-Datei *AesTool.jar* (diese finden Sie im Verzeichnis *SecLab_BlockCipherModes*, das Sie zu Beginn des Praktikums beim Entpacken des ZIP Files erzeugt haben) um die Bilddatei *image.bmp* (diese finden Sie im Unterverzeichnis *files*) einmal im ECB Mode und einmal im CBC Mode zu verschlüsseln. Die verschlüsselte Datei enthält dabei den Namen *image.bmp.enc* und wird im Unterverzeichnis *files* abgelegt. Benennen Sie diese Datei nach dem Verschlüsseln jeweils um in *image.bmp.ecb.enc* bzw. *image.bmp.cbc.enc*. Wenn Sie das Tool ohne Argumente aufrufen erhalten Sie Informationen, wie es zu verwenden ist.

Analysieren Sie nach der Verschlüsselung die Originaldatei und die beiden verschlüsselten Varianten zuerst mit *HistogramApp*. Vergleichen Sie die Histogramme und Entropien. Welche Beobachtungen geben Ihnen zu denken? (Hinweis: Die Entropien sind *pro bit* angegeben; vergleiche dazu den Text am Anfang und die letzte Aufgabe von Abschnitt 3 in diesem Lab.)



Als nächstes modifizieren Sie die beiden verschlüsselten Dateien so, dass diese von einem Programm, das BMP Dateien anzeigen kann, angezeigt werden können. Dazu überschreiben Sie den nun verschlüsselten BMP Header wieder mit dem originalen BMP Header. Für die Beispieldatei müssen Sie hierzu die ersten 54 Bytes der verschlüsselten Dateien mit den ersten 54 Bytes der Originaldatei ersetzen. Die Datei *image.header.bmp* (ebenfalls im Unterverzeichnis *files*) enthält genau diese 54 Bytes. Das Ersetzen können Sie je nach Plattform auf folgende Weisen erledigen (hier für den ECB Mode gezeigt, CBC funktioniert analog):

- Linux (und Mac):

- Mit Hilfe eines Hex Editors für Linux (resp. Mac), z.B. *bless*.
- Empfohlen: Mit Bordmitteln mittels dem Tool *dd*:

```
cp image.bmp.ecb.enc image.bmp.ecb.enc.bmp
```

```
dd conv=notrunc if=image.header.bmp of=image.bmp.ecb.enc.bmp
```

- Windows:

- Mit Hilfe eines Hex Editors. Z.B. mittels der freien Version des Hex Editors Neo:
<http://www.hhdsoftware.com/free-hex-editor>
- Mit Bordmitteln. Dann allerdings „unsauber“. Anstatt die ersten 54 Bytes der verschlüsselten Datei durch den unverschlüsselten Header zu ersetzen, hängen Sie den unverschlüsselten Header vorne an die verschlüsselte Datei an. Dies können Sie mit folgenden Aufrufen tun:

```
copy /B image.header.bmp + image.bmp.ecb.enc  
image.bmp.ecb.enc.bmp
```

Bemerkung: Dies führt zu einer etwas „verzerrten“ Darstellung, da der verschlüsselte

Header nun auch als Bildinformation interpretiert wird. Den gewünschten Effekt werden Sie aber dennoch gut erkennen können.

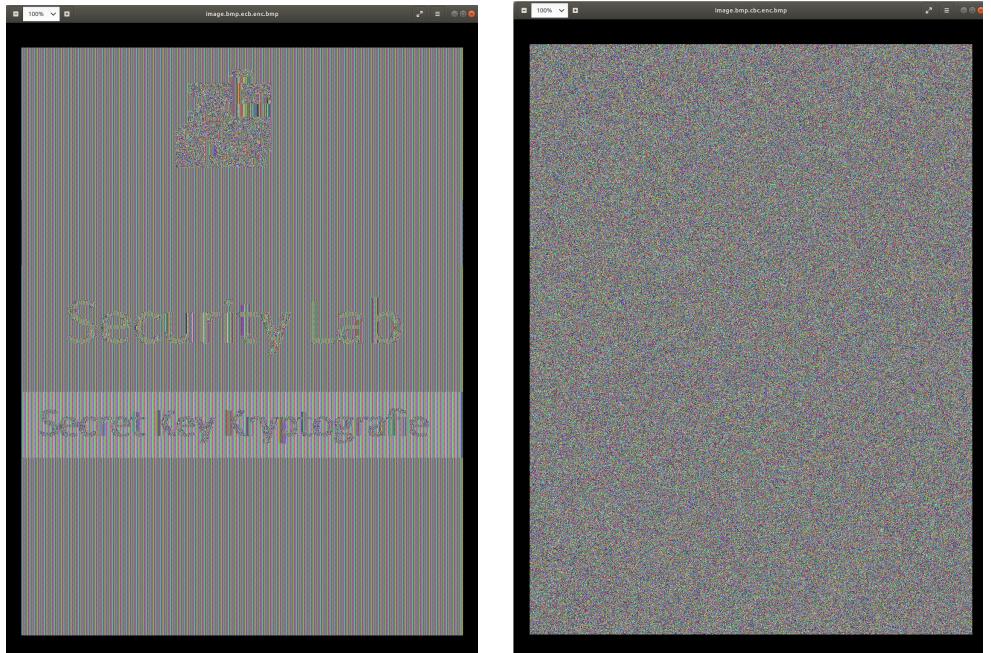
Betrachten Sie anschliessend die beiden so erzeugten Bitmapdateien. Was beobachten Sie? Haben Sie eine Erklärung für die Ursache Ihrer Beobachtung?

EBC:

Der Inhalt ist abgesehen vom ZHAW Logo sehr gut ersichtlich.
Es kann sein, dass das ZHAW-Logo mit einem bereits komprimierten Verfahren (bspw. JPG) abgespeichert wurde. Aus diesem Grund sind nicht mehr alle Informationen der Bild-Datei vorhanden -> Es kann nicht richtig angezeigt werden.

CBC:

CBC ist mittels XOR-Verknüpfung des Vorgängers abgehandelt worden. Aus diesem Grund ist der Inhalt nicht mehr erkennlich.



Praktikumspunkte

In diesem Praktikum können Sie **2Praktikumspunkte** erreichen:

- Zwei Punkte erhalten Sie, wenn Sie dem Betreuer Ihre Antworten auf die Fragen in der Praktikumsanleitung zeigen und diese Antworten mehrheitlich korrekt sind. Ebenfalls müssen Sie allfällige Kontrollfragen des Betreuers richtig beantworten. Zudem müssen Sie die beiden ECB- und CBC-verschlüsselten Bilder aus der letzten Aufgabe zeigen.