Authentische und zuverlässige Mobilkommunikation für sicherheitsrelevante Anwendungen Teil I: Sicherheitsanforderungen und grundlegende Verfahren

Authentic and Reliable Mobile Communication for Safety-Related Applications Part I: Security Requirements and Basic Algorithms

Von Bernd Friedrichs

Übersicht:

Als Beispiel einer sicherheitsrelevanten Anwendung wird eine auf das GSM-Mobilfunksystem gestützte Kommunikation zur Steuerung und Kontrolle von Zügen betrachtet. Dabei ist die Datenübertragung sowohl stochastischen Fehlern wie intelligenten kryptographischen Angriffen ausgesetzt. Zur Verhinderung bzw. zur Erkennung derartiger Verfälschungen sind Verfahren sowohl aus Kryptographie wie aus Kanalcodierung erforderlich, wobei der Message Authentication Code (MAC) ein wesentlicher Baustein ist. Schwerpunkt von Teil I ist die informationstheoretische Analyse des MAC, insbesondere seine Interpretation als Fehlererkennungscode und die Dimensionierung unter den Randbedingungen Sicherheit und Zuverlässigkeit. In Teil II [1] wird das Gesamtsystem zur sicherheitsrelevanten Kommunikation in Form einer geschichteten Architektur sowie die Einbettung in das GSM-System dargestellt.

Abstract:

A communication system for train control based on the GSM mobile radio standard is introduced as an example of a safety-related application. However, data transmission is exposed to stochastic errors as well as intelligent cryptographical attacks. In order to prevent or detect such falsifications, methods of cryptography and error control coding are required, the message authentication code (MAC) beeing the most important component. Part I deals with the analysis of the MAC by means of information theory. Topics are its interpretation as an error detection code and its design and dimensioning in accordance with the constraints of security and reliability. Part II [1] presents the general concept for the safety-related application in the form of a layered architecture and the adaptation to the GSM system.

Für die Dokumentation:

Kryptographie / Authentizität / Message Authentication Code / Codierungstheorie / Zugbeeinflussung / Mobilfunk

1. Einführung: Zugbeeinflussung als sicherheitsrelevante Mobilkommunikation

Als Beispiel einer sicherheitsrelevanten Kommunikation wird ein Mobilfunk-gestütztes System zur Steuerung und Kontrolle von Zügen betrachtet. Die sogenannte Funk-Zugbeeinflussung (FZB) basiert, wie in **Bild 1** dargestellt, auf einer bidirektionalen Kommunikation zwischen dem Fahrzeugrechner in der Lok und einer Streckenzentrale.

Die Streckenzentrale ist Teil des sogenannten Computer Integrated Railroading (CIR)-Netzes, wobei hier lediglich von Interesse ist, daß die dort zusammengefaßten Anwendungen und Dienste wie CIROD (Computer Integrated Railroading Operation and Dispatching), RZÜ (Rechnerunterstützte Zugüberwachung) und ESTW (Elektronisches Stellwerk) eine offene Netzarchitektur mit vielfältigen Zugängen erfordern. Netztechnisch ist die Streckenzentrale eine ISDN-Nebenstellenanlage (ISDN-PABX), die über das ISDN-Bahn-Netz an das Mobilfunknetz angeschlossen ist, das die bekannten Festnetz-Komponenten MSC (Mobile Switching Center), BCE (Base Station Subsystem Central Equipment) und BTS (Base Transceiver Station) enthält. Von geringerer Bedeutung hier sind die weiteren Komponenten wie TCE (Transcoding Equipment), VLR (Visitor Location Register), AC (Authentication Center) und EIR (Equipment Identity Register). Daneben gibt es noch Übergänge zu anderen Netzen wie ISDN, PSTN (Public Switched Telephone Network) und PDN (Public Data Network). Mit GSM-U_m wird die Luftschnittstelle zwischen der Basisstation einer Funkzelle und der Mobilstation (MS) im Zug bezeichnet.

Zwischen Zentrale und Zug werden Datenblöcke ausgetauscht, die hier als *Telegramme* bezeichnet werden. Diese enthalten wichtige Informationen zur Steuerung und Kontrolle, die Vorrang vor den ortsfesten Signalen an der Strecke haben. Diese Kommunikation muß also höchsten Anforderungen an Zuverlässigkeit, Sicherheit und Schutz vor Manipulationen genügen.

Damit werden spezielle Verfahren aus den Bereichen Kanalcodierung und Kryptographie erforderlich, denn ohne entsprechende Maßnahmen ist die Übertragung nur sehr unzureichend geschützt. Erstens ist eine Mobilfunkübertragung prinzipiell relativ unzuverlässig sowohl aufgrund der zeitvarianten Kanaleigenschaften wie der Verfügbarkeit von Funk- und Vermittlungsebene. Zweitens ist die FZB-Kommunikation in besonderer Weise durch Manipulationen gefährdet, da die drei beteiligten Netze offen sind mit vielfältigen und nicht kontrollierbaren Zugängen, so daß Manipulationen von jedem Zugangspunkt aus möglich sind. Insbesondere die Mobilfunkübertragung ist relativ einfach von jedem Punkt der Ausleuchtung aus manipulierbar.

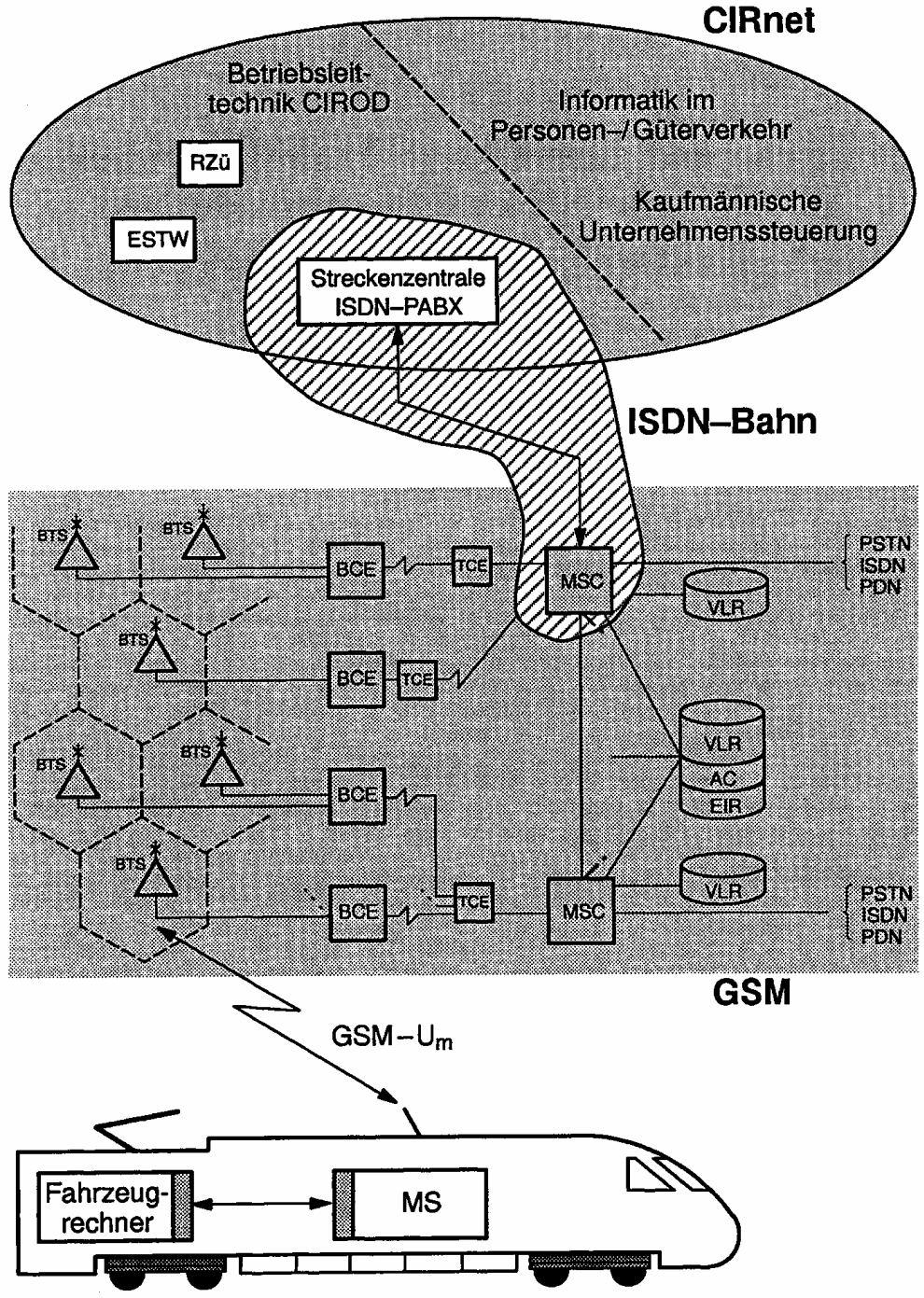

Bild 1: Funk-Zugbeeinflussung (FZB)

GSM als Teil des Übertragungssystems beeinflußt natürlich die Übertragungseigenschaften, aber die in GSM standardmäßig vorgesehenen Sicherheitsfunktionen sind ohne Einfluß auf die FZB, denn einerseits sind diese Funktionen unzureichend, weil sie sich nur auf die Funkebene und nicht auf das Festnetz erstrecken und anderseits sind für die FZB teilweise andere Sicherheitsfeatures erforderlich, wie noch gezeigt wird.

Die FZB-Kommunikation ist Teil des deutschen Forschungsprojektes DIBMOF (Dienste integrierender Bahnmobilfunk), in dem die Integration verschiedener bisher getrennt betriebener Mobilkommunikationsdienste angestrebt wird. Für die Bahn ergibt sich dann ein Mobilfunknetz „D3" entlang ihrer Linien mit ähnlicher Technik wie in den beiden öffentlichen Netzen D1 und D2. Eine ausführlichere Darstellung von DIBMOF und des bisherigen Systems zur Zugbeeinflussung findet sich in [2].

2. Anforderungen der sicherheitsrelevanten Kommunikation

Neben den prinzipiellen Vorteilen integrierter Konzepte erfordert die Einbettung der FZB-Kommunikation in offene Netze aber besondere Überlegungen und Vorkehrungen, die sich an den folgenden potientiellen Bedrohungen zu orientieren haben:

- Unbefugter Informationsgewinn (Verlust der Vertraulichkeit):
  - *Abhören* von übertragenen Daten und der Vergabe von kryptographischen Schlüsseln.
  - *Maskerade*, d.h. Vorspiegeln einer falschen Identität, um bestimmte Privilegien und damit verbundene Informationen zu erhalten.
- Unbefugte Modifikation von Informationen (Verlust der Integrität):

- Unbeabsichtigte Modifikationen durch
  - *stochastische Übertragungsfehler,* die primär durch die Funkübertrtagung verursacht werden;
  - *determinierte bzw. systematische Fehler,* z.B. durch falsche Synchronisation oder Softwarefehler.
- Mutwillige Modifikationen („intelligente Fehler, kryptographische Angriffe") durch
  - *Replay-Attack,* d.h. Wiedereinspielen von abgehörten und aufgezeichneten Telegrammen;
  - *Verzögerung, Vertauschung, Unterdrückung* von Telegrammen (Verlust der Zeit- bzw. Sequenzrichtigkeit);
  - *Einspielen selbst generierter Telegramme.*

- Unbefugte Beeinträchtigung der Funktionalität (Verlust der Verfügbarkeit bzw. Zuverlässigkeit):
  - *Blockade* sowohl durch unbefugte Modifikationen wie durch Überlastung (Funk, Vermittlung), unbefugte Nutzung der Ressourcen, Sabotage.
  - *Unzureichende Qualität des Übertragungskanals.*
  - *Fehler durch das Bedienungspersonal,* z.B. bei falscher Eingabe von Adressen.

Mit informationstechnischen Verfahren wie Codierung und Verschlüsselung allein ist es natürlich unmöglich, die vorangehend dargestellten Bedrohungen alle prinzipiell zu verhindern. Ziel ist aber nicht die Verhinderung, sondern die sofortige Erkennung der dargestellten Bedrohungen mit extrem hoher Sicherheit. In einem solchen Fall kann dann beispielsweise eine Notbremsung des Zuges ausgelöst werden, so daß nur die Verfügbarkeit des Eisenbahnbetriebes reduziert wird – aber die Sicherheit bleibt dabei voll erhalten.

Die Erkennung bzw. Abwehr mutwilliger Modifikationen erfordert den Einsatz kryptographischer Verfahren. Wer die entsprechenden Schlüssel nicht kennt, soll nicht in der Lage sein, unentdeckbare Manipulationen vornehmen zu können. Mit dem Einsatz kryptographischer Verfahren erfolgt eine logische Trennung der Welt in berechtigte und unberechtigte Benutzer des Systems. Wesentlich für die FZB sind folgende *Sicherheitsfunktionen,* die in **Bild 2** mit den in GSM auf der Luftschnittstelle verfügbaren Sicherheitsfunktionen verglichen werden:

Zunächst ist der Unterschied in den Kommunikationsbeziehungen hervorzuheben, die bei der FZB zwischen dem mobilen Teilnehmer Fahrzeug und dem stationären Teilnehmer Zentrale definiert sind, so daß die kryptographischen Verfahren Ende-zu-Ende wirksam sind und nicht nur auf der Luftschnittstelle zwischen Teilnehmer und Netz wie bei GSM.

	FZB	GSM
Kommunikationsbeziehung	**Tln(F) – Tln(Z)**	**Tln – Netz**
Vertraulichkeit der Nutzdaten	**optional**	**ja**
Integrität / Authentizität der Nutzdaten	**ja**	**schwach**
Identifikation und Authentifikation	**Tln(F) ↔ Tln(Z)**	**Tln → Netz**
Vertraulichkeit der Identifikation	**nein**	**ja**

Bild 2: Erforderliche Sicherheitsfunktionen pro Kommunikationsbeziehung

Eine direkte Verbindung zwischen zwei mobilen Teilnehmern ist für die FZB nicht erforderlich. Auch die Vertraulichkeit der Nutzdaten ist für die FZB nicht unbedingt notwendig, wird aber dennoch optional vorgesehen.

Ganz wesentlich für die FZB sind aber die Integrität (Unverfälschtheit) und die Authentizität (Echtheit bzgl. des Absenders) der Nutzdaten, was bei GSM nur vergleichsweise schwach gewährleistet ist. Identifikation und Authentifikation (Beweis der behaupteten Identität) müssen bei der FZB gegenseitig erfolgen. Bei GSM muß sich das Netz nicht authentifizieren, aber dafür wird auch auf die Vertraulichkeit der Identifikation Wert gelegt, so daß Bewegungs- oder Verkehrsanalysen für Abhörer unmöglich sein sollen. Für die FZB ist das dagegen belanglos.

Generell erfolgt für die FZB wie für GSM eine Beschränkung auf symmetrische Secret-Key-Systeme mit einem geheimen Schlüssel pro Kommunikationsbeziehung. Von wesentlicher Bedeutung ist dabei das sichere Management der kryptographischen Schlüssel, das in Teil II dargestellt wird.

3. Chiffrierung zur Vertraulichkeit

Vertraulichkeit („Wer darf eine Nachricht empfangen") und Authentizität („Wer darf eine Nachricht senden") sind zwei voneinander unabhängige Attribute eines Kryptosystems, die auch mit voneinander unabhängigen Verfahren erreicht werden sollten. Die gleichzeitige Erfüllung beider Attribute mit dem gleichen Verfahren kann zu gefährlichen Sicherheitseinbußen führen, wie noch gezeigt wird. Grundsätzlich ist die Autentizität ein wesentlich subtileres Thema als die Vertraulichkeit [7].

Vertraulichkeit wird in diesem und Authentizität wird im nächsten Abschnitt behandelt.

3.1 ECB-Modus (Electronic Codebook Mode)

Das Grundprinzip der symmetrischen Blockverschlüsselung (Block Cipher) zur Vertraulichkeit zeigt **Bild 3.** Der Klartext X wird im Encrypter E_Z umgesetzt in den Chiffretext

$$Y = E_Z(X). \tag{1}$$

Empfangsseitig wird im Decrypter D_Z daraus wieder der Klartext X zurückgewonnen. X und Y werden als binäre Blöcke gleicher Länge $L_x = L_y$ angenommen, so daß durch die Verschlüsselung keine Redundanz hinzuge-

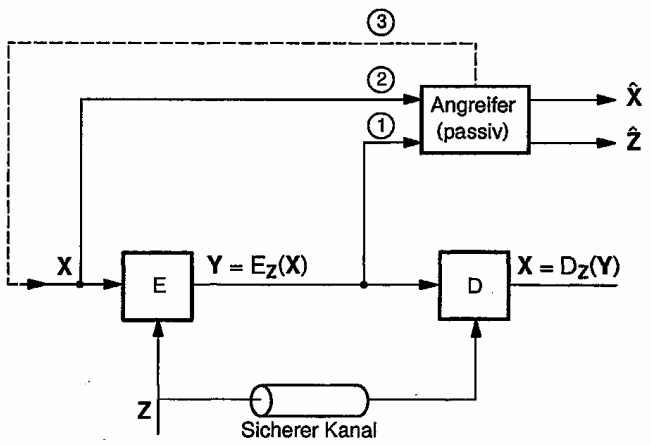

Bild 3: Informationsfluß in einem Chiffriersystem zur Geheimhaltung

fügt wird. Die Bezeichnung ECB ist leicht verständlich, da Encrypter und Decrypter praktisch Codebuch-Tabellen mit Klartext- und Chiffretext-Blöcken realisieren. E_Z und D_Z sind vom Schlüssel Z abhängige deterministische, bijektive und zueinander inverse Funktionen.

Dabei ist der Schlüssel Z ein binärer Block der Länge L_z, der sowohl bei Sender wie Empfänger vorliegen muß und dem potentiellen Abhörer der Übertragung nicht bekannt sein darf. Somit muß die Schlüsselvereinbarung über einen sicheren Kanal erfolgen.

Nach dem *Kerkhoff'schen Prinzip* [5, 9] sollte die Sicherheit eines Chiffriersystems auch dann gewährleistet sein, wenn der Angreifer sämtliche Details des Ver- und Entschlüsselungsprozesses kennt, also die formalen Zuordnungen

$$Z \mapsto E_Z \qquad Z \mapsto D_Z. \tag{2}$$

Das Geheimnis steckt nicht in diesen Algorithmen, sondern allein in den Schlüsseln als Parameter. Es ist mathematisch nicht beweisbar, daß durch das Verheimlichen von Algorithmen die Sicherheit erhöht werden kann.

Das Ziel des Angreifers ist die Ermittlung des Schlüssels bzw. des Klartextes. Die kryptographischen Angriffe können wie folgt klassifiziert werden (die Nummern beziehen sich auf Bild 3):

1. *Ciphertext-Only Attack* (Angriff nur mit Chiffretext): Hier kennt der Angreifer nur den Chiffretext und keinen Klartext.

   In den allermeisten Fällen wird damit der Angreifer gefährlich unterschätzt. Für die Erleichterung eines Angriffs ist es nämlich schon ausreichend, wenn der Klartext einige Redundanz enthält, so daß vernünftiger Klartext von unvernünftigem Klartext leicht unterschieden werden kann. Insbesondere bei der Anwendung FZB ist dies der Fall, da der Klartext immer wieder gleiche (oder fast gleiche) Befehle und immer wieder gleiche (nicht geheime) Adressen enthält.

2. *Known-Plaintext Attack* (Angriff mit bekanntem Klartext): Hier kennt der Angreifer einige Paare Klartext–Chiffretext.

   Diese für den Angreifer einfachere Situation wird bei der Analyse moderner Chiffriersysteme üblicherweise immer unterstellt.

3. *Chosen-Plaintext Attack* (Angriff mit frei wählbarem Klartext): Hier kann der Angreifer Chiffretext zu selbstgewähltem Klartext erzeugen.

   Natürlich hängt es vom gesamten Kommunikationssystem ab, ob diese konservative Annahme realistisch ist. Ein Angreifer hat es hier offensichtlich am einfachsten, da er die Reaktion des Chiffriersystems auf selbstgewählte Testsequenzen beobachten kann.

3.2 Theoretische und praktische Vertraulichkeit

Für die informationstheoretische Beschreibung mit Entropien werden für X, Y und Z jeweils Wahrscheinlichkeitsverteilungen unterstellt. Ferner können X und Z als statistisch unabhängig vorausgesetzt werden. Der Schlüssel wird zufällig mit gleichmäßiger Verteilung gewählt, so daß

$$H(Z) = L_z \tag{3}$$

für die Schlüsselentropie gilt. Zu unterscheiden ist, ob für jede neue Verschlüsselung ein neuer Schlüssel gewählt wird oder ob mehrere Klartexte mit gleichem Schlüssel chiffriert werden.

Nachfolgende werden drei Fälle betrachtet:

1) *Theoretisch perfekte Vertraulichkeit* ist nach Shannon definiert durch [6]

$$H(X \mid Y) = H(X), \tag{4}$$

d.h. der Chiffretext vermittelt keine Information (bzw. vermindert nicht die Unsicherheit) über den Klartext. Das ist äquivalent damit, daß X und Y statistisch unabhängig sind. Nach der sogenannten *pessimistischen Ungleichung* von Shannon muß

$$H(X) \leq H(Z) \tag{5}$$

für perfekte Vertraulichkeit gelten. Für den allgemeinen Fall erfordert das die sichere Verteilung von Schlüsseln, die mindestens so lang wie der Klartext sind (Vernam Cipher, One-Time Pad), d.h. $L_z \geq L_x$ und jeder neue Klartext erfordert einen neuen Schlüssel. Für die Anwendung FZB ist das natürlich unmöglich und somit kann theoretisch perfekte Vertraulichkeit nicht erreicht werden.

2) Zur Situation beim Known-Plaintext Attack: Der Angreifer kennt zumindest ein Paar X, Y und ist an der Bestimmung von Z interessiert, weil mit diesem Z noch weitere Klartexte verschlüsselt werden. Für die Entropie des Schlüssels bei einem bekannten Paar Klartext–Chiffretext gilt näherungsweise

$$H(Z \mid X, Y) = \max\{L_z - L_y, 0\}. \tag{6}$$

Wenn also (wie normalerweise) der Klartext bzw. Chiffretext länger als der Schlüssel ist, so ist Z schon durch ein einziges Paar (X, Y) eindeutig bestimmt. Auch bei $L_z > L_y$ bedarf es nur mehrerer verschiedener (X, Y)-Paare, um den Schlüssel eindeutig zu bestimmen. Wenn nun weitere Klartexte mit gleichem Z verschlüsselt werden, kann der Angreifer erfolgreich den Chiffretext dechiffrieren.

Da Gleichung (6) auch im Zusammenhang mit Authentizität wichtig ist, erfolgt hier noch eine ausführliche Begründung. Algebraisch kann die Beziehung $Y = E_Z(X)$ interpretiert werden als System von L_y Gleichungen mit L_z Unbekannten. Bei $L_z \leq L_y$ existiert höchstens eine Lösung, bei $L_z \geq L_y$ ist der Lösungsraum $(L_z - L_y)$-dimensional.

Gleichung (6) kann auch stochastisch begründet werden. Bei $L_z \ll L_y$ ist in $E_Z(X)$ bei festem X die volle Information über Z schon enthalten; $H(Z \mid X, E_Z(X)) = 0$. Bei $L_z \gg L_y$ vermindert sich die Entropie von Z bei L_y Nebenbedingungen der Form $Y = E_Z(X)$ um eben L_y; $H(Z \mid X, E_Z(X)) = L_z - L_y$. Die Kombination ergibt (6).

Trivialerweise gilt $H(E_Z(X) \mid X) \leq H(Z \mid X) = H(Z)$ und $H(E_Z(X) \mid X) \leq H(E_Z(X)) \leq L_y$. Bei $L_z \ll L_y$ ist wie oben in $E_Z(X)$ bei festem X die maximale Information über Z enthalten; $H(E_Z(X) \mid X) = H(Z) = L_z$. Bei $L_z \gg L_y$ ist $E_Z(X)$ bei festem X gleichmäßig verteilt mit der maximalen Entropie eines Binärwortes der Länge L_y; $H(E_Z(X) \mid X) = L_y$. Die Kombination ergibt nun näherungsweise

$$H(Y \mid X) = \min\{L_z, L_y\}. \tag{7}$$

Auch für den Spezialfall $L_x = L_y = L_z$ mit $Y = X + Z$ gelten (6) und (7) exakt. Obwohl (6) und (7) unabhängig voneinander hergeleitet wurden, sind beide Aussagen den-

noch gleichwertig. Dazu werden die folgenden für bedingte Entropien allgemein gültigen Beziehungen betrachtet:

$$H(Z, Y \mid X) = \underbrace{H(Z \mid X)}_{=H(Z)} + \underbrace{H(Y \mid Z, X)}_{=0 \text{(Enc.)}}$$

$$H(Z, Y \mid X) = H(Y \mid X) + H(Z \mid Y, X).$$

Somit folgt

$$H(Z \mid X, Y) = H(Z) - H(Y \mid X). \tag{8}$$

Anhand von Gleichung (8) folgt jetzt (6) aus (7) bzw. (7) aus (6).

3) Zur Situation beim Ciphertext-Only Attack: Für bedingte Entropien gilt allgemein

$$\underbrace{H(Z, X \mid Y)}_{\geq H(X \mid Y)} = H(Z \mid Y) + \underbrace{H(X \mid Z, Y)}_{=0 \text{(Dec.)}}.$$

Somit ergibt sich das bekannte Ergebnis [9], daß bei bekanntem Chiffretext die Unsicherheit über den Schlüssel generell größer als über den Klartext ist,

$$H(Z \mid Y) \geq H(X \mid Y), \tag{9}$$

d.h. die *Schlüsseläquivokation* übersteigt die *Klartextäquivokation*. Mit ähnlichen Argumenten wie zuvor gilt

$$H(X, Y \mid Z) = H(Y \mid Z) + \underbrace{H(X \mid Y, Z)}_{=0} \tag{10}$$

und

$$H(X, Y \mid Z) = H(X \mid Z) + \underbrace{H(Y \mid X, Z)}_{=0}. \tag{11}$$

Zusammen mit der Unabhängigkeit von X und Z folgt

$$\begin{aligned} H(X) + H(Z) &= H(Z) + H(X \mid Z) \\ &= H(Z) + H(Y \mid Z) \\ &= H(Y, Z) \\ &= \underbrace{H(Y)}_{\leq L_y = L_x} + H(Z \mid Y). \end{aligned} \tag{12}$$

Mit der Redundanz eines Klartext-Blocks

$$R(X) = L_x - H(X) \tag{13}$$

ergibt sich eine untere Grenze für die Schlüsseläquivokation [9] (die obere Grenze ist trivial)

$$H(Z) \geq H(Z \mid Y) \geq H(Z) - R(X). \tag{14}$$

Je kleiner $R(X)$ wird, desto größer ist zwangsläufig die Schlüsseläquivokation. Im redundanzfreien Fall mit $H(X) = L_x$ gilt $H(Z \mid Y) = H(Z)$ und somit vermittelt der Chiffretext keine Information über den Schlüssel. Für die Anwendung FZB ist jedoch der gegenteilige Fall $R(X) = L_x$ bzw. $H(X) = 0$ realistischer. In diesem Fall gilt

$$\underbrace{H(Y) + H(X \mid Y)}_{=0} = H(X, Y) = \underbrace{H(X)}_{=0} + H(Y \mid X), \tag{15}$$

und somit folgt aus (12) und (7)

$$\begin{aligned} H(Z \mid Y) &= \underbrace{H(X)}_{=0} + \underbrace{H(Z)}_{=L_z} - \underbrace{H(Y)}_{=\min\{L_z, L_y\}} \\ &= \max\{L_z - L_y, 0\}. \end{aligned} \tag{16}$$

Für den Normalfall $L_z \leq L_y$ und $H(X) = 0$ gilt also nicht nur $L_z \geq H(Z \mid Y) \geq 0$ nach (12), sondern sogar $H(Z \mid Y) = 0$.

Fazit zu 1) bis 3): Theoretisch perfekte Vertraulichkeit oder maximale Schlüsseläquivokation ist nicht erreichbar. Bestenfalls ist durch einen Chiffretext der Schlüssel nicht schon theoretisch eindeutig bestimmt. Diese Überlegungen setzen allerdings einen Angreifer mit unbeschränkter Rechenkapazität voraus. Unter realistischen Gesichtspunkten ist jedoch die sogenannte *praktische Sicherheit* ausreichend, bei der die Berechnung des Schlüssels in vernünftiger Zeit aus Aufwandsgründen unmöglich sein muß. In diesem Fall kann ein theoretisch absolut unsicheres Verfahren dennoch praktisch sicher sein.

Konkreter bedeutet praktische Sicherheit, daß der Zusammenhang zwischen Klartext, Chiffretext und Schlüssel so „kompliziert" ist, daß es zur Berechnung des Schlüssels keine kryptoanalytische Methode geben soll, die schneller ist als die *Probiermethode* (Brute Force Attack, Exhaustive Cryptanalysis). Dabei werden auf den Klartext alle möglichen Schlüssel angewendet, bis sich der bekannte Chiffretext ergibt. Um diesen Angriff abzuwehren, sollte die Anzahl der Schlüssel also möglichst groß sein, so daß die Probiermethode auch mit den schnellsten Rechnern noch sehr lange Rechenzeit benötigt.

Die praktische Sicherheit bzw. Vertraulichkeit ist dann gewährleistet, wenn die beiden folgenden von Shannon formulierten Entwurfskriterien erfüllt sind [5]:

Diffusions-Prinzip: Die Änderung eines Bits in der Nachricht bzw. die Änderung eines Bits im Schlüssel sollten zu einer Änderung von rund 50% der Chiffretext-Bits führen (Avalanche-Effekt),

Konfusions-Prinzip: Die Chiffretext-Statistiken sollten in sehr komplizierter Weise von den Klartext-Statistiken abhängig sein.

3.3 Data Encryption Standard (DES)

Eine weitverbreitete Methode zur Blockchiffrierung ist der Data Encryption Standard (DES), der bereits 1977 genormt wurde und trotz einiger Kritik noch heute von überragender Bedeutung ist. Das Format des DES mit $L_x = L_y = 64$ und $L_z = 56$ zeigt **Bild 4.**

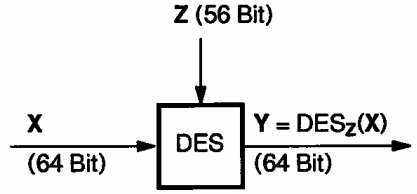

Bild 4: Data Encryption Standard (DES)

Zwischen den 2^{64} verschiedenen Klartext-Blöcken und den 2^{64} Chiffretext-Blöcken gibt es mit $(2^{64})! \approx 10^{3{,}5 \cdot 10^{20}}$ astronomisch viele verschiedene bijektive Abbildungen. Mit dem 56 Bit langen Schlüssel kann jedoch mit $2^{56} \approx 10^{17}$ nur ein winziger Teil davon als Chiffrierabbildungen aktiviert werden. Der DES ist nun so konzipiert, daß bei Kenntnis des Schlüssels Encrypter und Decrypter relativ leicht berechnet werden können (mit beinah identischen Algorithmen) und daß die Shannon'schen Entwurfskriterien nahezu ideal erfüllt sind, wie ausführliche Untersuchungen in der Literatur zeigen [10, 12]. Allerdings weist der DES dennoch einige algebraische Eigen-

schaften auf, die die Kyptoanalyse leider etwas vereinfachen, und es gibt auch einige bekannte schwache Shlüssel, die nicht verwendet werden sollten [2].

Eine moderne Altenative zum DES-Standard ist der 1991 vorgestellte IDEA (International Data Encryption Algorithm) mit $L_x = L_y = 64$ und $L_z = 128$, für den das Diffusions-Prinzip ideal erfüllt ist [8].

3.4 CBC-Modus (Cipher Block Chaining Mode)

Mit dem ECB-Modus aus Bild 3 sind zwei erhebliche Nachteile verbunden:

- Keine Fehlertransparenz. Wenn durch einen Übertragungsfehler im Chiffretext nur ein einziges Bit verfälscht wird, so wird mit dem Entschlüsseln der gesamte Klartext-Block verfälscht, so daß im Mittel die Hälfte aller Bits falsch sein wird (Diffusions-Prinzip).
- Nur lokale Vertraulichkeit. Wenn immer wieder der gleiche Klartext zu übertragen wäre, so entstehen auch immer wieder gleiche Chiffretexte. Ein Angreifer kann daraus eventuell erhebliche Schlüsse ziehen.
- Das Vertauschen von Chiffretexten oder der Replay-Attack werden im Empfänger nicht bemerkt, sofern die Klartext-Blöcke mit keiner eindeutigen Kennzeichnung (wie beispielsweise einem Zeitstempel) versehen sind.

Beim Cipher Block Chaining Mode (CBC) werden die Chiffretexte miteinander verkettet, so daß ein Chiffretext-Block von allen vorangehenden Klartext-Blöcken beeinflußt wird. Das Prinzip ist in **Bild 5** dargestellt. Im Sender werden die Chiffretext-Blöcke einen Blocktakt verzögert zurückgeführt und zu den Klartext-Blöcken addiert (modulo 2):

$$Y_i = E_Z(X_i + Y_{i-1}). \tag{17}$$

Die Chiffrierung eines Klartext-Blocks ist von der Stelle dieses Blocks in der gesamten Nachricht abhängig. Gleiche Klartext-Blöcke X_i haben nicht gleiche Chiffretext-Blöcke Y_i zur Folge, da der aktuelle Chiffretext von der gesamten Klartext-Vergangenheit abhängt.

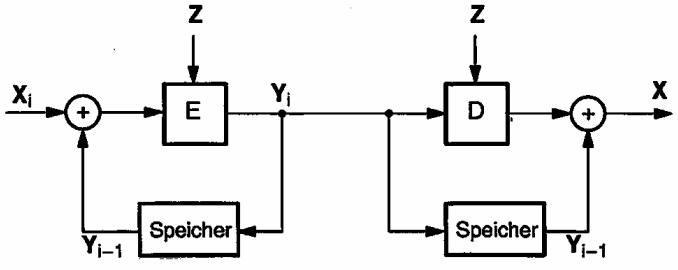

Bild 5: Cipher Block Chaining Mode (CBC)

Der Rückkopplung im Sender entspricht eine Vorwärtskopplung im Empfänger,

$$D_Z(Y_i) + Y_{i-1} = X_i. \tag{18}$$

Ein Übertragungsfehler im CBC-Modus erscheint zunächst als harmlos, da damit keine katastrophalen Auswirkungen verbunden sind, sondern nur zwei Klartext-Blöcke zerstört werden. Ein Fehler in Y_i bewirkt aufgrund des Diffusions-Prinzips eine Fehlerrate von 50% in X_i sowie einen weiteren Fehler in X_{i+1}. Bei der Möglichkeit aktiver Angriffe (siehe Abschnitt 4) darf X_{i+1} natürlich nicht akzeptiert werden, da ein Angreifer durch Modifikation von Y_i eine gezielte Manipulation von X_{i+1} vornehmen kann. Somit muß erstens durch zusätzliche Redundanz im Klartext gesichert sein, daß Klartext-Blöcke mit 50% Fehlerrate sicher als falsch erkannt werden, und zweitens darf der dann folgende Block nicht akzeptiert werden. Folglich können kleine Fehlerraten auf dem Kanal die effektive Übertragungsrate drastisch reduzieren.

4. Chiffrierung zur Authentizität

4.1 Grundprinzip

Das in Bild 3 dargestellte Prinzip der Verschlüsselung zur Vertraulichkeit sicher nur gegen das Abhören (passiver Angriff). Wenn ein Angreifer beispielsweise den Übertragungsweg auftrennen und selber Nachrichten einspielen kann (aktiver Angriff), so liegt die in **Bild 6** dargestellte Situation vor (zur Unterscheidung zwischen Authentizität und Vertraulichkeit wird hier $\tilde{E}$, $\tilde{D}$ statt E, D geschrieben).

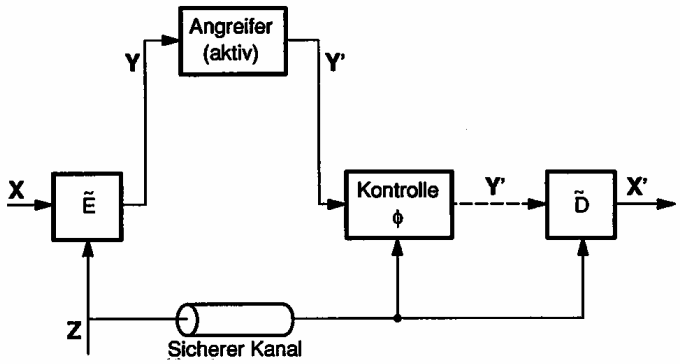

Bild 6: Informationsfluß in einem Chiffriersystem zur Integrität und Authentizität

Der Klartext X wird im Encrypter $\tilde{E}_Z$ umgesetzt in den Chiffretext $Y = \tilde{E}_Z(X)$ wie in Gleichung (1). Empfangsseitig wird Y' zunächst mit der Kontrollfunktion ϕ auf Authentizität kontrolliert. Bei positivem Ergebnis, d.h. $\phi(Y', Z) = 1$, wird im Decrypter $\tilde{D}_Z$ wieder der Klartext zurückgewonnen. Bei negativem Ergebnis, d.h. $\phi(Y', Z) = 0$, wird Y' nicht akzeptiert und verworfen.

X und Y sind binäre Blöcke der Längen L_x und L_y, wobei jetzt aber $L_y \geq L_x$ gilt, so daß durch $\tilde{E}_Z$ Redundanz hinzugefügt wird. Wie in Abschnitt 3 hat der Schlüssel Z die Länge L_z und ist über einen sicheren Kanal zu vereinbaren.

Der passive Angreifer hat nur Zugriff auf Y und versucht den Schlüssel bzw. den Klartext zu ermitteln. Der aktive Angreifer kann dagegen den Übertragungsweg auftrennen und selber Nachrichten einspielen. Mit Kenntnis des Schlüssels kann der Angreifer dem Empfänger jeden beliebigen Klartext unterschieben, ohne daß der Empfänger davon etwas merkt. Ohne Kenntnis des Schlüssels sind folgende Szenarien zu unterscheiden, deren Relevanz stark von der Betriebsart (zyklisch oder ereignisgesteuert) abhängt:

Beim *Impersonation Attack* kennt der Angreifer Y nicht oder Y existiert nicht:

$$P_I = P \text{ (Empfänger akzeptiert } Y'). \tag{19}$$

Beim *Substitution Attack* kennt der Angreifer Y:

$$P_S = P \text{ (Empfänger akzeptiert } Y' \text{ mit } Y \neq Y'). \tag{20}$$

Beim *Deception Attack* kann der Angreifer die Strategie frei wählen:

$$P_D = \max\{P_I, P_S\}. \tag{21}$$

Der Replay-Attack kann nur abgesichert werden, wenn zwischenzeitlich der Schlüssel gewechselt wurde oder wenn der Klartext eine eindeutige Redundanz wie beispielsweise eine Uhrzeit enthält.

4.2 Message Authentication Code (MAC)

Mit der in **Bild 7** dargestellten Methode des Message Authentication Code (MAC) [4, 5, 9] wird ausschließlich die Integrität und Authentizität gesichert – die Nachricht selbst wird offen übertragen:

$$Y = \tilde{E}_Z(X) = (X, A_Z(X)) = (X, MAC). \tag{22}$$

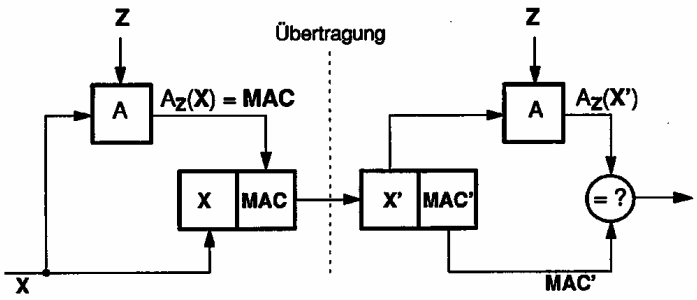

Bild 7: Message Authentication (MAC)

Es gilt also $L_y = L_x + L_{mac}$, wobei L_{mac} die Länge von MAC bezeichnet. Die MAC-Bildung basiert auf dem CBC-Modus eines Blockverschlüsslers, wobei der Data Encryption Standard (DES) verwendet wird. Bei einer aus drei Klartext-Blöcken bestehenden Nachricht $X = (X_1, X_2, X_3)$ gilt beispielsweise

$$\begin{aligned} MAC &= A_Z(X) \\ &= E_Z(X_3 + E_Z(X_2 + E_Z(X_1))) \\ &= DES_Z(X_3 + DES_Z(X_2 + DES_Z(X_1))). \end{aligned} \tag{23}$$

Die Initialisierung im CBC-Modus wird dabei mit Null angenommen, $Y_0 = \mathbf{0}$. Bei $L_x > L_{mac}$ haben gleiche Nachrichten zwangsläufig den gleichen MAC – das ist aber unproblematisch, da es nicht das Ziel ist, X aus MAC zu rekonstruieren.

Bei der Übertragung wird $Y = (X, MAC)$ verfälscht zu $Y' = (X', MAC')$. Im Empfänger ist ebenfalls der Authentikator realisiert, um die Kontrollfunktion

$$\phi(Y', Z) = \begin{cases} 1 & A_Z(X') = MAC' \\ 0 & \text{sonst} \end{cases} \tag{24}$$

zu berechnen. Nur bei $\phi(Y', Z) = 1$ (d.h. der MAC der empfangenen Nachricht entspricht dem empfangenen MAC) akzeptiert der Empfänger die Nachricht X. Anderenfalls wird die Nachricht als ungültig eingestuft und nicht akzeptiert. Da eine Korrektur der Nachricht unmöglich ist, muß somit eine Wiederholung der Nachricht angefordert werden.

Das MAC-Verfahren erfordert ein gegenseitiges Vertrauen der beiden Kommunikationspartner. Wer einen MAC überprüfen kann, kann natürlich auch einen MAC fälschen. Deshalb ist auch hier die Qualität des Verfahrens nur so gut wie die Qualität der Schlüsselverteilung.

Die Sicherheit des MAC-Verfahrens wird nicht dadurch gemindert, daß die Nachricht im Klartext übertragen wird, da dies nur dem Known-Plaintext Attack entspricht, der ja grundsätzlich ungefährlich sein soll. Ganz im Gegenteil vermindert sich die Sicherheit, wenn der MAC mit einer CBC-Verschlüsselung kombiniert wird. Es sind dann nämlich Modifikationen der übertragenen Daten möglich, die bei der MAC-Prüfung nicht erkannt werden. Als Beispiel wird dazu die Nachricht $X = (X_1, X_2)$ betrachtet, die als $Y = (Y_1, Y_2)$ mit

$$\begin{aligned} Y_1 &= E_Z(X_1 + Y_0) \\ Y_2 &= E_Z(X_2 + Y_1) = MAC = A_Z(X_1, X_2) \end{aligned}$$

übertragen wird (dabei kann die Initialisierung Y_0 beliebig festgelegt werden). Nach einer möglichen Manipulation wird $Y' = (Y_1', Y_2') = (Y_2, Y_1)$ empfangen. Der Empfänger dechiffriert

$$\begin{aligned} X_1' &= D_Z(Y_1') + Y_0 = D_Z(Y_2) + Y_0 \\ &= X_2 + Y_1 + Y_0 \neq X_1 \end{aligned}$$

und

$$\begin{aligned} X_2' &= D_Z(Y_2') + Y_1' = D_Z(Y_1) + Y_2 \\ &= X_1 + Y_0 + Y_2 \neq X_2. \end{aligned}$$

Also ist der Chiffretext zu einem falschen Klartext dechiffriert worden. Die MAC-Bildung im Empfänger ergibt

$$\begin{aligned} A_Z(X') &= A_Z(X_1', X_2') \\ &= E_Z(X_2' + E_Z(X_1' + Y_0)) \\ &= E_Z(X_1 + Y_0 + Y_2 + \underbrace{E_Z(X_2 + Y_1)}_{= Y_2}) \\ &= E_Z(X_1 + Y_0) = Y_1 = Y_2'. \end{aligned}$$

Die MAC-Prüfung führt also zu einer positiven Authentifizierung, d.h. die Manipulation wird nicht erkannt. Durch Verschlüsselung zur Vertraulichkeit wird also hier die Integrität aufgehoben.

Fazit: Aus Sicherheitsgründen sollte sich der MAC direkt auf die übertragenen Daten beziehen. Wenn zusätzlich Vertraulichkeit gefordert wird, muß der Klartext erst verschlüsselt werden mit einem von der MAC-Bildung getrennten Verfahren und anschließend wird der MAC von den verschlüsselten Daten berechnet.

Die Shannon'schen Entwurfsprinzipien übertragen sich von der ECB-Blockchiffrierung auf das MAC-CBC-Verfahren:

Diffusions-Prinzip: Die Änderung eines Bits in der Nachricht bzw. die Änderung eines Bits im Schlüssel sollten zu einer Änderung von rund 50% der MAC-Bits führen,

Konfusions-Prinzip: Die MAC-Statistiken sollten in sehr komplizierter Weise von den Klartext-Statistiken abhängig sein.

4.3 Theoretische und praktische Authentizität

Zunächst wird die Transinformation zwischen Chiffretext und Schlüssel, also die Differenz zwischen Schlüsselentropie und Schlüsseläquivokation, in verschiedenen Darstellungen betrachtet, wobei nur (28) spezifisch für Authentizität bzw. das MAC-Verfahren ist:

$$I(Y;Z) = H(Z) - H(Z \mid Y) \quad (25)$$
$$= H(Y) - H(Y \mid Z) \quad (26)$$
$$= H(Y) - H(X) \quad (27)$$
$$= H(A_Z(X) \mid X). \quad (28)$$

(25 und (26) entsprechen direkt der üblichen Definition der Transinformation. Aus (10) folgt

$$H(Y \mid Z) = H(X, Y \mid Z) = \underbrace{H(X \mid Z)}_{=H(X)} + \underbrace{H(Y \mid X, Z)}_{=0}$$

die eigentlich selbstverständliche Beziehung $H(Y \mid Z) = H(X)$ und damit (27). (28) folgt schließlich aus

$$H(Y) = H(X, A_Z(X)) = H(X) + H(A_Z(X) \mid X).$$

Grundlage für die Theorie der Authentizität ist die Simmons-Bound für den Impersonation Attack [9]

$$P_I \geq 2^{-I(Y;Z)}. \quad (29)$$

Perfekte Authentizität ist durch $P_D = 2^{-I(Y;Z)}$ definiert, auch wenn dabei $P_D = 1$ gelten sollte. Offensichtlich ist $P_I = 0$ oder sogar $P_D = 0$ nicht erreichbar, d. h. es gibt keine direkte Entsprechung zur perfekten Vertraulichkeit. Je mehr Information der Chiffretext über den Schlüssel vermittelt, desto einfacher ist zwar gemäß Abschnitt 3.2 der Schlüssel zu ermitteln, aber desto schwieriger die Situation für den aktiven Angreifer. Überraschend ist das nicht, denn auf den statistischen Bindungen zwischen Chiffretext und Schlüssel basiert die Authentizitäts-Kontrolle im Empfänger.

Die Bedingungen für Gleichheit in (29) sind für das MAC-Verfahren näherungsweise erfüllt, wie am Ende dieses Abschnitts noch gezeigt wird. In diesem Fall besteht die beste Wahl des Angreifers beim Impersonation Attack in einer zufälligen Wahl von Y' mit gleichmäßiger Verteilung.

Der aktive Angriff beim MAC-Verfahren entspricht direkt dem Known-Plaintext Attack bei der Chiffrierung zur Vertraulichkeit. Die Eigenschaften von X spielen keine Rolle. Nach (6) und (7) gilt mit $L_y = \text{Länge}(E_Z(X))$ näherungsweise

$$H(Z \mid X, E_Z(X)) = \max\{L_z - L_y, 0\}$$
$$H(E_Z(X) \mid X) = \min\{L_z, L_y\}.$$

Wenn nun $E_Z(X)$ durch $A_Z(X)$ der Länge L_{mac} ersetzt wird, so gilt unter Beachtung von (22) in direkter Analogie näherungsweise

$$H(Z \mid Y) = \max\{L_z - L_{mac}, 0\} \quad (30)$$
$$I(Y;Z) = H(A_Z(X) \mid X) = \min\{L_z, L_{mac}\}. \quad (31)$$

Auch für den (unsinnigen) Spezialfall $A_Z(X) = X + Z$ mit $L_x = L_z = L_{mac}$ gelten (30) und (31) exakt. Anhand von (26) und (28) folgt (31) aus (30) bzw. (30) aus (31).

Prinzipiell sollte ein gutes MAC-Verfahren folgende Bedingungen erfüllen:

- Aus (X, ***MAC***) darf der Schlüssel Z nicht berechenbar sein. Theoretisch erfordert das $L_z \gg L_{mac}$ gemäß (30).
- Ohne Kenntnis des Schlüssels Z soll es unmöglich sein, Paare (X, ***MAC***) zu berechnen, d. h. zu X darf ***MAC*** und zu ***MAC*** darf X nicht berechenbar sein. Theoretisch erfordert das sowohl großes L_z wie großes L_{mac} gemäß (31).

Der Schlüssel sollte also möglichst lang sein, damit nicht durch zufällige Wahl von Z ein zu X passendes $A_Z(X)$ erraten wird. Ferner sollte ***MAC*** möglichst lang sein, um das direkte Erraten von ***MAC*** zu vorgegebenem X zu verhindern. Das gleiche gilt, wenn X zu ***MAC*** erraten wird. Die Länge von X spielt dabei keine Rolle, das gilt herunter bis zu $L_x = 1$.

Die Situation beim Substitution Attack wird jetzt genauer betrachtet. Bei $L_z \leq L_{mac}$ ist durch Y der Schlüssel eindeutig bestimmt und somit gilt $P_S = 1$. Bei $L_z \geq L_{mac}$ gilt $H(Z \mid Y) = L_z - L_{mac}$ nach (30) und somit passen zu Y näherungsweise $2^{L_z - L_{mac}}$ Schlüssel. Von diesen wird ein Schlüssel Z' zufällig gewählt, d. h. $L'_z := H(Z') = L_z - L_{mac}$, und mit Z' wird ein Y' mit $Y' \neq Y$ berechnet. Dann gilt

$$P_S = P(Y' \text{ gültig}) = 2^{-I(Y';Z')}.$$

Nach (31) gilt

$$I(Y';Z') = \min\{L'_z, L_{mac}\} = \min\{L_z - L_{mac}, L_{mac}\},$$

und somit folgt insgesamt

$$P_S \geq \begin{cases} 1 & \text{für } L_z \leq L_{mac} \\ 2^{-(L_z - L_{mac})} & \text{für } L_{mac} \leq L_z \leq 2L_{mac} \\ 2^{-L_{mac}} & \text{für } 2L_{mac} \leq L_z \end{cases} \quad (32)$$

$$\geq 2^{-L_z/2} = \frac{1}{\sqrt{2^{L_z}}}. \quad (33)$$

Die letzte Abschätzung ist unmittelbar klar und wohlbekannt [5, 11].

Beispiel: Für die Kombination $L_x = 3 \cdot 64$, $L_{mac} = 64$, $L_z = 56$ gilt $P_I = 2^{-56}$, wobei der Angreifer den Schlüssel erraten muß (bei zufälliger Wahl von ***MAC*** ist der Angreifer nur mit Wahrscheinlichkeit 2^{-64} erfolgreich). Wegen $H(Z \mid Y) = 0$ ist durch ein Paar (X, ***MAC***) der Schlüssel eindeutig bestimmt, und der Angreifer kann in jedem Fall ein (X', ***MAC'***) nach Wahl erzeugen, d. h. $P_S = 1$.

Für kleines P_S muß $1 \ll L_{mac} \ll L_z$ gemäß (32) gelten. Zur Verlängerung des Schlüssels auf DES-Basis gibt es verschiedene Methoden wie Triple-DES [6] oder bidirektionaler MAC [5], die $L_{mac} = 64$ mit $L_z = 112$ kombinieren. In diesem Fall gilt $P_I = 2^{-64}$, $P_S = 2^{-48}$. Für den IDEA gilt $P_I = P_S = 2^{-64}$.

Allerdings kann auch ein längerer Schlüssel aus der Beobachtung mehrerer Paare Klartext-MAC doch wieder bestimmt werden, beispielsweise ist beim DES

$$0 = H(Z \mid Y_1, Y_2) < H(Z \mid Y_1) = 112 - 64 = 48 \quad (34)$$

möglich, aber nur, wenn sich X_1, X_2 wesentlich unterscheiden, was bei der Anwendung FZB i. a. nicht der Fall sein wird. In jedem Fall wird jedoch der erforderliche Rechenaufwand für den Angreifer enorm gesteigert.

Fazit: Wie bei der Chiffrierung zur Vertraulichkeit ist theoretische Sicherheit bzw. perfekte Authentizität kaum erreichbar, sofern ein Angreifer mit unbeschränkter Rechenkapazität vorausgesetzt wird. Allerdings zeigt das obige Beispiel, daß ein theoretisch absolut unsicheres Verfahren praktisch dennoch ziemlich sicher ist. Praktische Authentizität ist entsprechend zur praktischen Vertraulichkeit zu verstehen und ist dann gewährleistet, wenn

die Shannon'schen Entwurfskriterien erfüllt sind, was gemäß Abschnitt 4.2 vorausgesetzt werden kann.

Für Gleichheit in (29) müssen folgende Bedingungen erfüllt sein [5, 9]:

B1) Für alle Y mit $P(Y)>0$ soll $P(Y \text{ gültig})$ den gleichen Wert haben.

B2) Es soll $P(Y \mid Z)=c \cdot \phi(Y, Z)$ mit einer von Z unabhängigen Konstanten c gelten, d.h. wenn Y gültig bei Z ist, dann soll $P(Y \mid Z)$ für alle Z den gleichen Wert haben.

Es sei $Y=(X, MAC)$: Für $\phi(Y, Z)=0$ bzw. äquivalent $A_Z(X) \neq MAC$ gilt $P(Y \mid Z)=P(X, MAC \mid Z)=0$ und für $\phi(Y, Z)=1$ gilt $P(Y \mid Z)=P(X \mid Z)=P(X)$. Insgesamt gilt also

$$P(Y \mid Z)=P(X) \cdot \phi(Y, Z), \tag{35}$$

und damit ist B2) nachgewiesen. Weiter gilt

$$\begin{aligned} P(Y \text{ gültig}) &= \sum_Z P(Z) \cdot P(Y \text{ gültig} \mid Z) \\ &= \sum_Z P(Z) \cdot \phi(Y, Z), \end{aligned} \tag{36}$$

sowie mit (35)

$$\begin{aligned} P(Y) &= \sum_Z P(Z) \cdot P(Y \mid Z) \\ &= P(X) \cdot \sum_Z P(Z) \cdot \phi(Y, Z) \\ &= P(X) \cdot P(Y \text{ gültig}). \end{aligned} \tag{37}$$

$\sum_Z \phi(Y, Z)$ entspricht der Anzahl der Z mit der Eigenschaft $A_Z(X)=MAC$. Es sei $\mathfrak{Y}$ die Menge aller möglichen Chiffretexte:

$$\begin{aligned} \mathfrak{Y} &= \{Y \mid P(Y)>0\} \\ &= \left\{ Y \,\middle|\, \sum_Z \phi(Y, Z) \geq 1 \right\}. \end{aligned} \tag{38}$$

Klar ist

$$|\mathfrak{Y}| \leq \min\{2^{L_x} \cdot 2^{L_z}, 2^{L_y}\} = 2^{L_x + \min\{L_z, L_{mac}\}}.$$

Für $Y \notin \mathfrak{Y}$ gilt $P(Y)=0$ und $P(Y \text{ gültig})=0$. Dieser Fall ist normalerweise nur bei $L_z < L_{mac}$ möglich. Wenn es dagegen mindestens ein Z gibt, so beträgt die Anzahl nach (30) näherungsweise $2^{\max\{L_z - L_{mac}, 0\}}$, und nach (36) gilt dann

$$P(Y \text{ gültig}) = \begin{Bmatrix} 2^{-\min\{L_z, L_{mac}\}} = P_I & \text{für } Y \in \mathfrak{Y} \\ 0 & \text{für } Y \notin \mathfrak{Y} \end{Bmatrix}. \tag{39}$$

Damit ist auch B1) nachgewiesen.

Bei gleichmäßig verteiltem X mit $P(X)=2^{-L_x}$ ist aufgrund des Diffusions-Prinzips auch MAC bzw. Y gleichmäßig verteilt, und aus (37) folgt

$$P(Y) = \begin{Bmatrix} 2^{-\min\{L_x + L_z, L_y\}} = \dfrac{1}{|\mathfrak{Y}|} & \text{für } Y \in \mathfrak{Y} \\ 0 & \text{für } Y \notin \mathfrak{Y} \end{Bmatrix}. \tag{40}$$

4.4 Codierungstheoretische Interpretation des MAC-Verfahrens

Theorem 1: Das MAC-Verfahren entspricht einem binären, systematischen, nichtlinearen $(n, k)=(L_x+L_{mac}, L_x)$-Fehlererkennungscode (EDC, Error Detection Code), der mit dem Schlüssel Z parametrisiert ist. L_x und L_{mac} entsprechen der Anzahl der Infobits bzw. Prüfbits, und L_y entspricht der Blocklänge. Die Codemenge ist gegeben durch

$$\Gamma_Z = \{(X, MAC) \mid A_Z(X)=MAC\} \tag{41}$$

und ein Encoder durch

$$X \mapsto (X, A_Z(X)). \tag{42}$$

Die Nichtlinearität folgt aus der DES-Nichtlinearität bzw. aus dem Konfusions-Prinzip. Bei einem linearen Code könnte der Zusammenhang zwischen X und MAC über lineare Gleichungssysteme sofort berechnet werden, und damit wäre das MAC-Verfahren kryptographisch wertlos.

Die Erkennung von Verfälschungen ist natürlich unabhängig davon, ob das Verfahren als MAC oder als EDC interpretiert wird. Für

$$(X', MAC') = (X, MAC) + (e_1, e_2) \tag{43}$$

mit einem entsprechenden partionierten Fehlermuster $e=(e_1, e_2)$ der Länge $n=L_y=L_x+L_{mac}$ gilt

$$(X', MAC') \in \Gamma_Z \Leftrightarrow \underbrace{A_Z(X+e_1)}_{A_Z(X')} = \underbrace{A_Z(X)+e_2}_{MAC'}. \tag{44}$$

Die Erkennung eines Fehlermusters ist also nicht nur vom Fehlermuster selbst abhängig, sondern auch vom gesendeten Infowort X. Nur wenn A_Z linear wäre, würde sich (44) auf $A_Z(e_1)=e_2$ reduzieren.

Codierungstheorie und Kryptographie wachsen hier zusammen [3], und somit stellen sich für das MAC-Verfahren die gleichen Fragen wie für einen Code zur Fehlererkennung bzw. Fehlerkorrektur (ECC, Error Control Code): Wie gut eignet sich das MAC-Verfahren zur Erkennung stochastischer Fehler? Die Güte eines Codes wird wesentlich durch die Minimaldistanz

$$d_{\min} = \min\{d_H(Y_1, Y_2) \mid Y_1, Y_2 \in \Gamma_Z, Y_1 \neq Y_2\} \tag{45}$$

(d_H = Hammingabstand) geprägt, die natürlich von Z abhängt. Jedes Fehlermuster $e \neq 0$ mit Hamminggewicht $w_H(e) < d_{\min}$ wird prinzipiell erkannt, auch bei nichtlinearen Codes. Eine vollständige Erfassung der Güte eines Codes erfolgt mit der Gewichtsverteilung A_d (= Anzahl der Codewörter vom Hamminggewicht d, $d=0, \ldots, n$).

Eine direkte Berechnung von $d_{\min}$ oder A_d ist zwar aufgrund der komplexen Nichtlinearitäten unmöglich, aber dennoch ist eine stochastische Interpretation auf der Grundlage des Diffusions-Prinzips möglich. Voraussetzung dafür sind aber gleiche apriori-Wahrscheinlichkeiten für die Informationswörter, d.h. eine gleichmäßige Verteilung von X.

Theorem 2: Der Message Authentication Code Γ_Z kann als *Random Code* interpretiert werden, bei dem die 2^{L_x} Codewörter zufällig und unabhängig gewählt werden und 0 und 1 mit jeweils 50% Wahrscheinlich-

keit auftreten. Im Mittel ergibt sich eine binomiale Gewichtsverteilung mit

$$A_d = 2^{-L_{mac}} \binom{L_y}{d} = 2^{-(n-k)} \binom{n}{d}. \tag{46}$$

Dabei wird ein zufällig gewählter Schlüssel unterstellt, wobei aber die Verteilung von $\boldsymbol{Z}$ sowie L_z unwesentlich sind.

Begründung: Ein Random Code hat im Mittel die angegebene Gewichtsverteilung, da das Gewicht jedes einzelnen Codewortes binomialverteilt ist. Ferner ist $\sum_d A_d = 2^k$ die Anzahl der Codewörter.

Bei gleichmäßig verteiltem $\boldsymbol{X}$ ist das Hamminggewicht $w_H(\boldsymbol{X})$ binomialverteilt. Bei $L_x \gg L_{mac}$ ist nach dem Diffusions-Prinzip auch $\boldsymbol{MAC}$ und als Folge davon auch $\boldsymbol{Y}$ gleichmäßig verteilt, so daß $w_H(\boldsymbol{Y})$ binomialverteilt ist. Bei $L_x \leq L_{mac}$ werden allerdings von den $2^{L_{mac}}$ möglichen $\boldsymbol{MAC}$-Werten nur noch 2^{L_x} Werte angenommen, aber auch hier ist $w_H(\boldsymbol{MAC})$ bzw. $w_H(\boldsymbol{Y})$ binomialverteilt.

In der Gewichtsverteilung sind also MAC und Random Code identisch, und somit kann der MAC als Random Code interpretiert werden. Dabei ist es unwesentlich, daß der MAC ein systematischer Code ist oder daß beim Random Code zwei Codewörter identisch sein können, was beim MAC ausgeschlossen ist.

Beispiel: Für die Kombination $L_x = 3 \cdot 64 = 192$, $L_{mac} = 64$, $L_y = 256$ ergibt sich (bei $A_{d_{\min}} \approx 1$) eine Minimaldistanz $d_{\min} \approx 11$. Nach der Hamming-Schranke muß $d_{\min} \leq 23$ gelten, und nach der Gilbert-Varshamov-Schranke existieren Codes mit $d_{\min} \geq 13$. Schließlich existiert auch ein (255, 191)-BCH-Code mit $d_{\min} = 17$ [13]. Bekanntlich liegt die Mehrzahl der Random Codes in der Nähe der GV-Schranke, so daß damit die Minimaldistanz von $\Gamma_{\boldsymbol{Z}}$ einigermaßen genau bekannt ist. Allerdings könnte es dennoch besonders ungünstige Schlüssel mit wesentlich kleinerem $d_{\min}$ geben. Es ist dann eine Bewertungsfrage, ob der „schlechteste" Schlüssel oder der Mittelwert über alle Schlüssel maßgebend sein soll.

5. Sicherheit und Zuverlässigkeit

Wie bereits in Abschnitt 2 ausgeführt wurde, ist Sicherheit über die sofortige Erkennung von Verfälschungen (beabsichtigte Manipulationen, unbeabsichtigte Modifikationen, insbesondere stochastische Übertragungsfehler) definiert. Hohe Zuverlässigkeit liegt dann vor, wenn derartige Störungen möglichst selten auftreten. Mit dem überlagernden Fehlermuster gemäß (43) ergeben sich folgende formale Definitionen in bezug auf ein einzelnes Telegramm:

Zuverlässigkeit bedeutet, daß die Wahrscheinlichkeit P_{ee} (ee = error event) eines Fehlermusters möglichst klein ist:

$$\begin{aligned} P_{ee} &= P \text{ (Fehlermuster tritt auf)} \\ &= P(\boldsymbol{e} \neq \boldsymbol{0}). \end{aligned} \tag{47}$$

Sicherheit bedeutet, daß die Wahrscheinlichkeit P_{ue} (ue = undetected error) eines unerkannten Fehlermusters extrem klein ist:

$$\begin{aligned} P_{ue} &= P \text{ (Fehlermuster wird nicht erkannt)} \\ &= P(\boldsymbol{e} \notin \varepsilon(\boldsymbol{X}, \boldsymbol{Z}) \wedge \boldsymbol{e} \neq \boldsymbol{0}). \end{aligned} \tag{48}$$

Dabei bezeichnet $\varepsilon(\boldsymbol{X}, \boldsymbol{Z})$ die Menge der erkennbaren Fehlermuster $\boldsymbol{e} = (e_1, e_2)$, wenn $\boldsymbol{Y} = (\boldsymbol{X}, A_{\boldsymbol{Z}}(\boldsymbol{X}))$ gesendet wurde:

$$\begin{aligned} \varepsilon(\boldsymbol{X}, \boldsymbol{Z}) &= \{\boldsymbol{e} \mid \boldsymbol{Y} + \boldsymbol{e} \notin \Gamma_{\boldsymbol{Z}}\} \\ &= \{\boldsymbol{e} \mid A_{\boldsymbol{Z}}(\boldsymbol{X} + \boldsymbol{e}_1) \neq A_{\boldsymbol{Z}}(\boldsymbol{X}) + \boldsymbol{e}_2\}. \end{aligned} \tag{49}$$

Unabhängig von $\boldsymbol{X}$ und $\boldsymbol{Z}$ sind allerdings stets $|\varepsilon(\boldsymbol{X}, \boldsymbol{Z})| = 2^n - |\Gamma_{\boldsymbol{Z}}| = 2^n - 2^k$ Fehlermuster erkennbar.

Der schlechteste Fall ist $P_{ee} = 1$, d.h. jedes Empfangswort ist mit einem Fehlermuster überlagert (extrem unzuverlässiger Kanal). Auch in diesem Fall darf P_{ue} den vorgegebenen Wert nicht überschreiten.

Durch entsprechend viele Prüfstellen kann die Sicherheit auch bei schlechten Kanälen beliebig weit gesteigert werden. Gleichzeitig wird es aber immer wahrscheinlicher, daß ein sehr langer Codeblock mit vielen Prüfstellen auch tatsächlich Fehler aufweisen wird, d.h. die Zuverlässigkeit sinkt. Formal gilt bei fester Infowortlänge $k = L_x$ und einer Prüfwortlänge $n - k = L_{mac} \to \infty$

$$P_{ue} \to 0 \quad \text{und} \quad P_{ee} \to 1. \tag{50}$$

Sicherheit und Zuverlässigkeit sind also gegenläufige Qualitätskriterien einer Übertragung. Je höher man die Sicherheitsanforderungen ansetzt, desto größer wird ein anderes wichtiges praktisches Problem, nämlich, daß wichtige Steuerungsbefehle eventuell nicht mehr zeitgerecht übertragen werden können. Eine gemeinsame Verbesserung von Sicherheit und Zuverlässigkeit ist nur dadurch möglich, daß sowohl Fehler erkannt wie auch Fehler korrigiert werden.

Ein mögliches und einfaches Verfahren dazu wäre die sogenannte Typ-I-Hybrid-ARQ [14]. Bei einem Code mit der Minimaldistanz $d_{\min}$ sind gleichzeitig t Fehler erkennbar und e Fehler korrigierbar ($0 \leq e \leq t$), wenn

$$e + t + 1 \leq d_{\min} \tag{51}$$

gilt.

Extremfälle sind $t + 1 \leq d_{\min}$ (nur Fehlererkennung, also hohe Sicherheit und geringe Zuverlässigkeit) sowie $2e + 1 \leq d_{\min}$ (nur Fehlerkorrektur, also hohe Zuverlässigkeit und geringe Sicherheit). Je mehr Fehler korrigiert werden (Zuverlässigkeit steigt), desto weniger Fehler sind erkennbar (Sicherheit sinkt). In Teil II wird allerdings dargestellt, daß eine direkte Kombination von Fehlererkennung und Fehlerkorrektur in einem Verfahren ungünstig ist.

Die Wahrscheinlichkeit P_{ue} hängt natürlich vom stochastischen Kanalmodell ab. Da im Gegensatz zu den „gewöhnlichen" Übertragungssystemen dieses Modell bei der Anwendung FZB aufgrund der vielfältigen potentiellen Bedrohungen nicht a priori klar ist, ist P_{ue} robust zu dimensionieren unter Berücksichtigung verschiedener Modelle für die stochastische Verteilung der Fehlermuster. Dazu werden nun drei Fälle betrachtet.

1) Alle Fehlermuster treten mit gleicher Wahrscheinlichkeit auf. Dann gilt

$$\begin{aligned} P_{ue} &= P(\boldsymbol{e} \text{ nicht erkennbar} \mid \boldsymbol{e} \neq \boldsymbol{0}) \cdot P(\boldsymbol{e} \neq \boldsymbol{0}) \\ &= \frac{\text{Anzahl unerkennbare Fehlermuster}}{\text{Anzahl Fehlermuster}} \cdot P_{ee} \\ &= \frac{2^n - |\varepsilon(\boldsymbol{X}, \boldsymbol{Z})| - 1}{2^n - 1} \cdot P_{ee} = \frac{2^k - 1}{2^n - 1} \cdot P_{ee} \\ &< 2^{-(n-k)} \cdot P_{ee}. \end{aligned} \tag{52}$$

Die Eigenschaften des Codes spielen hierbei keine Rolle; als Prüfstellen könnten anstelle ***MAC*** auch L_{mac} Nullen verwendet werden.

Dieser Fall ist übrigens nicht so unrealistisch, wie er zunächst erscheinen mag. Wenn auf dem Übertragungsweg eine Blockverschlüsselung zur Vertraulichkeit enthalten ist, die wiederum dem Diffusions-Prinzip genügt, dann erzeugen beliebig verteilte Fehler zwischen Verschlüssler und Entschlüssler hinter dem Entschlüssler tatsächlich gleichmäßig verteilte Fehlermuster.

2) Fehlermuster niedrigen Gewichts treten häufiger auf als Fehlermuster hohen Gewichts. Der einfachste Fall ist der binäre symmetrische Kanal (BSC), bei dem die einzelnen Fehler statistisch unabhängig voneinander mit Wahrscheinlichkeit p auftreten,

$$P_{ee} = 1-(1-p)^n. \tag{53}$$

Da generell alle Fehlermuster vom Gewicht $w_H(e) < d_{min}$ erkannt werden. gilt die Abschätzung

$$P_{ue} \le P(w_H(e) \ge d_{min}) = 1 - \sum_{d=0}^{d_{min}-1} \binom{n}{d} p^d (1-p)^{n-d}. \tag{54}$$

Für $p = 0{,}5$ folgt hieraus nur das schwache Resultat $P_{ee} \approx 1$, $P_{ue} \approx 1$. Für einen linearen Code gilt mit der Gewichtsverteilung die Gleichheit [14]

$$P_{ue} = \sum_{d=1}^{n} A_d p^d (1-p)^{n-d}. \tag{55}$$

Aus (55) folgt unabhängig von p, aber mit der Näherung (46) für die Gewichtsverteilung

$$P_{ue} = 2^{-(n-k)} \cdot \sum_{d=1}^{n} \binom{n}{d} p^d (1-p)^{n-d} \le 2^{-(n-k)}. \tag{56}$$

Aus (55) folgt für $p = 0{,}5$, aber unabhängig von den Codeeigenschaften

$$P_{ue} = 2^{-n} \cdot \sum_{d=1}^{n} A_d \le 2^{-(n-k)}. \tag{57}$$

Dies stimmt natürlich mit (52) überein, da der BSC für $p = 0{,}5$ genau Fall 1 entspricht. Bekanntlich gibt es aber besonders ungünstige Codes, bei denen P_{ue} nicht für $p = 0{,}5$, sondern für $p < 0{,}5$ maximal wird [15].

3) Fehlermuster treten in Form von Bündelfehlern auf, d.h. die einzelnen Fehler sind nicht statistisch unabhängig. Die Formeln (52) und (57) gelten unabhängig von den Codeeigenschaften, also auch wenn die $n-k = L_{mac}$ Prüfstellen alle konstant Null sind. In diesem Fall gilt bei Bündelfehlern jedoch das katastrophale Resultat

$$P_{ue} \approx \frac{k}{n} \cdot P_{ee}, \tag{58}$$

denn jeder Bündelfehler außerhalb der Prüfstellen wird nicht erkannt. Für einen vernünftigen Code ist jedoch auch hier garantiert, daß alle Bündelfehler bis zum Gewicht $d_{min} - 1$ erkannt werden.

Fazit zu 3): Die Erkennung stochastischer Fehler erfordert bei robuster Dimensionierung einen Code mit möglichst großer Minimaldistanz. Da d_{min} beim MAC-Verfahren im Mittel über die Schlüssel schlechter ist als bei den besten linearen Codes, ist die Fehlererkennungsfähigkeit des MAC geringer als bei guten Fehlererkennungscodes. Das widerspricht nicht der „philosophischen" Überlegung, daß die Erkennung stochastischer Fehler kein schwierigeres Problem als die Erkennung intelligenter kryptographischer Angriffe sein kann. Ein Widerspruch läge nur dann vor, wenn es eine Verteilung der stochastischen Fehlermuster derart gäbe, daß $P_{ue} > P_I$ ist; denn dann könnte der Angreifer seine Strategie beim Impersonation Attack dadurch verbessern, daß er Y' entsprechend dieser Verteilung wählt. Jedoch gilt die Simmons-Bound für die optimale Strategie beim Impersonation Attack, so daß $P_{ue} \le P_I$ gelten muß.

Abschließend erfolgt nun ein expliziter Vergleich von P_I und P_{ue}. Gleichwahrscheinliche stochastische Fehlermuster e korrespondieren mit der gleichmäßigen Verteilung von Y' beim Impersonation Attack. Nach (29), (31) und (52) gilt für $e \neq 0$

$$P_I = 2^{-\min\{L_z, L_{mac}\}} \ge 2^{-L_{mac}} \ge P_{ue}. \tag{59}$$

Interpretation: Für unerkennbare Fehlermuster e gilt $A_Z(X+e_1) = A_Z(X) + e_2$. Mit $e' = e + (X, A_Z(X))$ gilt dann $A_Z(e_1') = e_2'$, d.h. $Y' = e'$ bewirkt einen erfolgreichen Impersonation Attack. Bei $L_z \ge L_{mac}$ haben e und e' die gleiche Auftrittswahrscheinlichkeit, und somit muß $P_I = P_{ue}$ gelten. Bei $L_z < L_{mac}$ weiß der Angreifer natürlich, daß zu X von den $2^{L_{mac}}$ möglichen ***MAC***-Werten nur 2^{L_z} tatsächlich auftreten können und trifft seine Wahl entsprechend mit einer Erfolgswahrscheinlichkeit von $P_I = 2^{-L_z}$. Bei den gleichmäßig verteilten Fehlermustern paßt ***MAC*** zu X natürlich nur mit Wahrscheinlichkeit $2^{-L_{mac}}$, und somit gilt hier also $P_I > P_{ue}$.

Für das Beispiel $L_x = 3 \cdot 64$, $L_{mac} = 64$ gilt somit

$$P_I = 2^{-56} > 2^{-64} = P_{ue} \quad \text{für } L_z = 56,$$
$$P_I = 2^{-64} = P_{ue} \quad \text{für } L_z = 112.$$

In Teil II [1] erfolgt die Integration der Verfahren zur sicherheitsrelevanten Kommunikation in Form einer geschichteten Architektur. Schwerpunkte sind eine geeignete Schlüsselhierarchie, Protokolle zur Sicherheit des gesamten Kommunikationsprozesses und die Einbettung in das GSM-Mobilfunksystem.

(Fortsetzung folgt im nächsten Heft)

Literatur:

[1] Friedrichs, B.: Authentische und zuverlässige Mobilkommunikation für sicherheitsrelevante Anwendungen. Teil II: Systemarchitektur und Einbettung in GSM.
[2] Friedrichs, B.: Sicherungsverfahren für die Datenübertragung über „Bedrohte Kanäle" bei sicherheitsrelevanten Diensten. ANT Nachrichtentechnische Berichte 10 (1993) S. 47—63.
[3] Friedrichs, B.: Verfahren zur Kanalcodierung und Authentizität für sicherheitsrelevante Mobilkommunikation. 8. Aachener Kolloquium Signaltheorie (1994).
[4] Davies, D. W.; Price, W. L.: Security for Computer Networks. New York: John Wiley 2. Ed. 1989.
[5] Simmons, G. J. (Ed.): Contemporary Cryptology. New York: IEEE Press 1992.
[6] Fumy, W.; Rieß, H. P.: Kryptographie. München: R. Oldenbourg Verlag 1988.
[7] Massey, J. L.: Cryptography – Fundamentals and Applications. Engelberg (Schweiz): Advanced Technology Seminars 1992.
[8] Lai, X.: On the Design and Security of Block Ciphers. ETH Series in Information Processing Vol. 1. Konstanz: Hartung-Gorre 1992.
[9] Johanesson, R.: Informationstheorie. Lund: Addison-Wesley „Studentlitteratur" 1992.
[10] Heider, F.-P.; Kraus, D.; Welschenbach, M.: Mathematische Methoden der Kryptoanalyse. Braunschweig: Vieweg 1985.
[11] Beutelspacher, A.; Rosenbaum, U.: Projektive Geometrie. Braunschweig: Vieweg 1992.
[12] Kalinski, B. S.; Rivest, R. L.; Sherman, A. T.: Is the Data Encryption Standard a Group? Jornal of Cryptology (1988) 1, S. 3—36.
[13] Blahut, R. E.: Theory and Practice of Error Control Codes. Reading: Addison-Wesley 1983.
[14] Lin, S.; Costello, D. J.: Error Control Coding. Englewood Cliffs: Prentice-Hall 1983.
[15] Michelson, A. M.; Levesque, A. H.: Error-Control Techniques for Digital Communication. New York: John Wiley 1985.

Dr.-Ing. B. Friedrichs, ANT Nachrichtentechnik GmbH, Gerberstraße 33, D-71522 Backnang.

(Eingegangen am 11. 3. 1994)