Naturwissenschaft

Christian Schiestl

Pseudozufallszahlen in der Kryptographie

Diplomarbeit

Diplom.de

Bibliografische Information der Deutschen Nationalbibliothek:

Bibliografische Information der Deutschen Nationalbibliothek: Die Deutsche Bibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über http://dnb.d-nb.de/ abrufbar.

ISBN: 9783832441494

Christian Schiestl

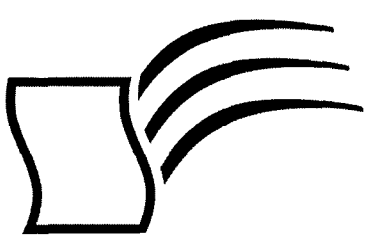
Pseudozufallszahlen in der Kryptographie

Diplom.de

Christian Schiestl

Pseudozufallszahlen in der Kryptographie

Diplomarbeit
an der Universität Klagenfurt, 7
Fachbereich Systemsicherheit
Mai 1999 Abgabe



Diplom.de

Diplomica GmbH
Hermannstal 119k
22119 Hamburg

Fon: 040 / 655 99 20
Fax: 040 / 655 99 222

agentur@diplom.de
www.diplom.de

ID 4149

ID 4149
Schiestl, Christian: Pseudozufallszahlen in der Kryptographie / Christian Schiestl - Hamburg: Diplomica GmbH, 2001
Zugl.: Klagenfurt, Universität, Diplom, 1999

Diplomica GmbH
http://www.diplom.de, Hamburg 2001
Printed in Germany

Inhaltsverzeichnis

1 Einleitung

The Generation of Random Numbers
is too important to be left to chance.
— Robert R. Coveyou

Schon seit Jahrzehnten stellt die auf den ersten Blick trivial erscheinende Aufgabe mittels Computer Zufallszahlen zu erzeugen sowohl MathematikerInnen als auch InformatikerInnen vor große Probleme. Obwohl Taschenrechner, Betriebssysteme und Programmiersprachen über entsprechende Zufallsfunktionen verfügen, konnte diese komplexe Problemstellung noch nicht zufriedenstellend gelöst werden und stellt somit auch heute noch einen aktuellen Forschungsgegenstand dar.

Die Aufgabenstellung, mittels Computer Zufallszahlen zu erzeugen, scheint für viele auf den ersten Blick einfach lösbar zu sein, was regelrecht zu einer Flut an Verfahren für diese Problemstellung führt. Ständig schlagen WissenschaftlerInnen aus den verschiedensten Bereichen (wie Mathematik, Informatik, Physik, usw.) Methoden zur Erzeugung von Zufallszahlen vor. Wie zahlreiche mißlungene Lösungsansätze jedoch zeigen, handelt es sich bei der Generierung von Zufallszahlen auf keinen Fall um eine triviale Aufgabe. Ein Computer ist und bleibt eine deterministische Maschine, dazu geschaffen, auf eine konkrete Eingabe eine definierte Ausgabe zu liefern. Wie kann man nun einer deterministischen Maschine – wie sie der Computer nun mal ist – beibringen, Zufall zu generieren?

Viele Bereiche der Informatik sind in zunehmenden Ausmaß auf Zufallszahlen angewiesen. Man denke nur an Monte-Carlo-Simulationen, Optimierungen mittels genetischer Algorithmen oder aber an Computerspiele, die ohne „intelligente" Monster wohl nur halb so interessant wären. Durch die zunehmende weltweite Vernetzung von Rechnern haben Sicherheitsaspekte in den letzten Jahren an Bedeutung gewonnen. Schutzmechanismen gegen unbefugten Zugriff auf vertrauliche Daten sowie zur Authentifizierung und Identifikation von Kommunikationspartnern spielen eine immer größer werdende Rolle. Kryptographische Verfahren wie symmetrische Verschlüsselungs-, Public-Key- und Signaturverfahren bieten Möglichkeiten, diese Sicherheitsrisiken zu verringern.

Gerade kryptographische Basismechanismen kommen heutzutage kaum noch ohne Zufallszahlen aus. Beinahe jedes Kryptosystem benötigt irgendwann geheime, nicht vorhersagbare Zufallszahlen bzw. Zufallsfolgen. Ohne Zufallsgeneratoren gäbe es keine Kryptographie! Man denke nur an folgende, exemplarische Einsatzgebiete (siehe [Cro94], [Men97], [Sch97b], [Sch96], [Sta98]):

- Schlüsselerzeugung
  Symmetrische und asymmetrische Kryptosysteme benötigen für die sichere Datenverschlüsselung zufällige Schlüssel.

- Parametererzeugung
  Ein weiteres wichtiges Einsatzgebiet für Zufallszahlen ist die Erzeugung von Parametern für asymmetrische Verschlüsselungsverfahren (z.B. die Generierung großer Primzahlen im RSA-Verfahren). Symmetrische Blockchiffren im CBC-Mode erfordern in Form von Initialisierungsvektoren ebenfalls zufällige Parameter.

- Identifikationsprotokolle
  Bei Challenge-Response-Verfahren wird auf einer Seite eine zufällige Challenge erzeugt, die die Gegenseite mit ihrem geheimen Schlüssel in signierter Form retourniert. Im Zuge einseitiger Challenge-Response-Verfahren empfängt und verarbeitet beispielsweise jedes GSM-Handy bei jeder Netzanmeldung eine Zufallszahl.

- Signatur-Verfahren
  Bestimmte digitale Signatur-Verfahren (wie z. B. DSA, ElGamal) benötigen bei jedem Signiervorgang einen neuen Zufallswert. Bei Zero-Knowledge-Signatur-Verfahren setzt der Signierende eine Zufallszahl ein, um sein Geheimnis zu verbergen.

- Protokolle zur Schlüsselverteilung
  Zu Beginn einer Sitzung müssen zufällige Sitzungsschlüssel erzeugt und verteilt werden. Bei Diffie-Hellman ähnlichen Protokollen benötigen dazu beide Parteien jeweils einen zufälligen Startwert.

- Verschlüsselung
  Zufallsfolgen können unmittelbar zur Verschlüsselung eingesetzt werden. Man denke dabei an das informationstheoretisch sichere Verschlüsselungsverfahren „One-Time-Pad" oder an das Vernam-Verfahren. Das Übermittlungsproblem der langen Zufallsfolgen auf einem sicheren Kanal führte zur Entwicklung von Stromchiffren. Dabei werden ausgehend von einem kurzen zufälligen Startwert mittels Schlüsselstromgenerator auf deterministische Art und Weise reproduzierbare Zufallswerte erzeugt und diese zur Nachrichten-Verschlüsselung verwenden. Der Vorteil von Stromchiffren gegenüber einem One-Time-Pad liegt darin, daß nur noch der kurze, zufällige Initialisierungswert dem Kommunikationspartner auf geheimen Wege mitgeteilt werden muß. Der Einsatz von deterministischen Schlüsselstromgeneratoren hat aber den Verlust der informationstheoretisch beweisbaren Sicherheit zur Folge.

Die Güte der in Kryptosystemen verwendeten Zufallszahlen wirkt sich unmittelbar auf deren Sicherheit aus. Wie Beispiele aus der jüngsten Vergangenheit zeigen (z. B. die aufsehenerregenden Angriffe zweier Berkeley-Studenten auf den Netscape-Browser im Jahre 1995), ermöglicht der Einsatz naiver Zufallszahlengeneratoren in Kryptosystemen diese mit relativ geringem Aufwand zu brechen. Schwache Zufallsgeneratoren stellen somit eine Gefährdung der Sicherheit von Kryptosystemen dar (siehe [Gol96a], [Wie96]). Selbst Jeff Schiller, Co-Autor des 1994 veröffentlichten Internet RFC mit dem Titel „*Randomness Recommendations for Security*" [Cro94] und Mitentwickler von Kerberos V4, hat mit dem Zufall seine liebe Not, konnten doch 1996 Cyberpunks aufgrund einer Schwäche im Kerberos-Zufallszahlengenerator das weltweit eingesetzte System erfolgreich attackieren und die erzeugten Sitzungsschlüssel kompromittieren (siehe [Gut98], [Wie96]).

Viele der in der Literatur beschriebenen Verfahren zur Erzeugung von Zufallszahlen taugen zwar für unterschiedlichste Anwendungen (z.B. für die Monte-Carlo-Simulation), erweisen sich jedoch für kryptographische Zwecke oft als ungeeignet. Eine zufällige Folge, die zwar verschiedenste statistische Tests besteht, gilt noch lange nicht als kryptographisch sicher, sondern ist vielmehr nur dann für die Anwendung in der Kryptographie geeignet, wenn sie unter keinen Umständen vorhersagbar ist. Es stellt sich also die Frage: Welche Generatoren können in der Kryptographie verwendet werden und von welchen Generatoren sollte man lieber die Finger lassen?

Jerry Dwyer und K. B. Williams in [Dwy96b]:

> *„It is fair to say that much of the literature on random number generators is heavy stuff. It's heavy in two ways. First, it's heavy in the sense that many of the articles require uncommon mathematical skills and can take hours to read and understand. Second, it's heavy in the sense that even a fraction of the articles on the subject would make a big pile."*

Diese vorliegende Literaturstudie versucht, die umfassende Literatur auch für mathematisch interessierte „Laien" auf verständliche Art und Weise aufzuarbeiten und einen Überblick über Verfahren zur Erzeugung von Pseudozufallszahlen für die Kryptographie zu geben.

- In Kapitel 2 werden zuerst allgemein die Problematik rund um Zufall im Computer, Definitionen von Zufall sowie die Unterscheidung von echtem Zufall und Pseudozufall herausgearbeitet. Es werden allgemeine Begriffe und Definitionen, die für das generelle Verständnis der Arbeit notwendig sind, eingeführt. Eine Diskussion allgemeiner Forderungen an Pseudozufallsgeneratoren bildet den Abschluß des Kapitels.

- Von Pseudozufallsfolgen wird gefordert, daß sie über dieselben statistischen Eigenschaften wie echte Zufallsfolgen verfügen. Kapitel 3 gibt einen Überblick über empirisch-statistische Testverfahren zur Überprüfung der Gleichverteilung bzw. der Unabhängigkeit von Pseudozufallsfolgen. Nach einem in die statistische Testtheorie einführenden Abschnitt, der den Leser kurz mit einigen Grundlagen vertraut machen soll, werden die historisch interessanten Golomb'schen Zufallskriterien präsentiert. Anschließend wird eine Beschreibung ausgewählter Tests der klassischen Standard-Tests von Donald Knuth sowie der von George Marsaglia entwickelten Stringent- und Monkey-Tests gegeben. Abschließend wird ein universeller Test von Ueli Maurer vorgestellt.

- Zufall kann nicht einfach durch zufälliges Aneinanderreihen komplexer Funktionen erzeugt werden. Ein historischer Rückblick auf „naive" Zufallsgeneratoren in Kapitel 4 soll dem Leser die Notwendigkeit vor Augen führen, daß nur auf solider Mathematik basierende Pseudozufallsgeneratoren zur Erzeugung von kryptographischen Zufallswerten geeignet sind. Beginnend mit Glücksspielen, Glücksrädern und Tabellen werden ältere Generatoren wie die britische Lotto-Maschine ERNIE oder der EUMEL-Generator vorgestellt. Auch die als gescheitert zu betrachtenden Ansätze, aus transzendenten und irrationalen Zahlen Zufallswerte zu extrahieren, werden kurz beleuchtet. Den Abschluß des historischen Rückblicks bilden die ersten arithmetischen Verfahren zur Erzeugung

von Pseudozufallszahlen wie die beiden Zufallszahlengeneratoren von Apple und Commodore, die Middle-Square-Methode von Neumann sowie der komplexe Super-Random-Number-Generator von Knuth.

- Kapitel 5 beschreibt häufig verwendete klassische Pseudozufallsgeneratoren. Zum einen handelt es sich dabei um Generatoren, die rekursiv nach der linearen Kongruenzen-Methode arbeiten. Zum anderen wird auf linear bzw. nicht-linear rückgekoppelte Schieberegister näher eingegangen. Sowohl für Kongruenzgeneratoren wie auch für rückgekoppelte Schieberegister werden kryptoanalytische Angriffe präsentiert und bekannte Schwachpunkte dieser klassischen Generatoren aufgezeigt. Auf diese Art und Weise soll vor dem bedenkenlosen Einsatz dieser für die Kryptographie nur beschränkt geeigneten Generatoren gewarnt werden.

- Starke Pseudozufallsgeneratoren werden in Kapitel 6 diskutiert. Diese in diversen Standards (ANSI, FIPS-PUB, ISO) publizierten und für die Kryptographie konzipierten Generatoren verwenden zur Erzeugung von Pseudozufallszahlen Einwegfunktionen, um damit einfach vorherzusagende Operationen zu verschleiern. Diese einfachen Operationen reichen vom Hochzählen eines Zählers bis hin zu klassischen Pseudozufallsgeneratoren. Als Einwegfunktionen werden kryptographische Primitive wie kryptographische Hashfunktionen oder Blockchiffren eingesetzt.

- Kapitel 7 ist kryptographisch sicheren Pseudozufallsgeneratoren (z. B. Blum-Blum-Shub-, Blum-Micali-, RSA-Generator, etc.) gewidmet. Diese Pseudozufallsgeneratoren wurden eigens für die Kryptographie entwickelt und ermöglichen die Konstruktion von Kryptosystemen mit höherer Sicherheit. Die Klasse der kryptographisch sicheren Pseudozufallsgeneratoren beruht auf Annahmen der Komplexitätstheorie, d. h. auf der Nicht-Existenz effizienter Algorithmen für wohlbekannte zahlentheoretische Probleme (wie z. B. der Faktorisierung großer Zahlen). Unter diesen Annahmen sind Angreifer nachweislich nicht in der Lage, aus der Kenntnis einiger Zufallswerte zukünftige Zufallszahlen vorherzusagen und auf diese Art und Weise das System anzugreifen.

- Auf die Erzeugung echter Zufallszahlen wird in Kapitel 8 eingegangen. Hochwertige, auf speziellen physikalischen Apparate und Standard-Hardware basierende Zufallsquellen liefern Bits, die von Angreifern nur schwer zu erraten sind. Neben diesen hochwertigen Zufallsquellen werden auch einige minderwertige Zufallsquellen, die Systempuffer, Netzwerkstatistiken oder den Menschen als Zufallsquelle nutzen, diskutiert. Angreifer können unter Umständen bei diesen Verfahren mit geringem Aufwand die gewonnenen Zufallsbits vorhersagen oder sogar aktiv den Generierungsprozeß beeinflussen und manipulieren. Minderwertige Zufallsquellen dürfen daher nicht als alleinige Zufallsquelle, sondern nur in Kombination mit weiteren Zufallsquellen zur Erzeugung von Zufallsbits eingesetzt werden. Da die gelieferten Zufallsbits nur in den seltensten Fällen gleichverteilt und unabhängig sind, bedarf es einer entsprechenden Nachbearbeitung. Dazu werden sogenannte triviale bzw. starke Mischfunktionen eingesetzt. Mit einer Beschreibung der bekanntesten Mischfunktionen endet das Kapitel 8.

- Kapitel 9 bildet mit einigen Schlußbemerkungen den Abschluß der Diplomarbeit.

2 Was ist Zufall?

Zufall ist vielleicht das Pseudonym Gottes,
wenn er nicht unterschreiben will!
— Anatole France

Zufall begegnet uns im Alltag immer und überall. Sei es in Form von Lottoziehungen, Auslosung der Gegner im Sport, in einarmigen Banditen im Spiel-Casino oder aber in CD-Playern beim Shuffle-Mode. Jeder Mensch hat intuitiv ein Gefühl dafür, was mit Zufall gemeint ist, und verwendet den Begriff „Zufall" in der Alltagssprache ständig. Mark Kac [Kac83, S. 405]:

> *„I am convinced that the vast majority of my readers, and in fact the vast majority of scientists and even nonscientists, are convinced that they know what ‚random' is. A toss of a coin is random; (...), and so is the emission of an alpha particle. (...) Simple, isn't it?"*

Der Versuch einer exakten Definition von Zufall erweist sich allerdings als äußerst schwieriges Unternehmen (siehe [Cha66], [Kac83], [Knu97], [ML66], [May97], [Pop71]). Donald Knuth [Knu97, S. 2]:

> *„People who think about this topic almost invariably get into philosophical discussions about what the word ‚random' means. In a sense, there is no such thing as a random number; for example, is 2 a random number?"*

Die „Zufälligkeit" ist eine Eigenschaft eines sehr abstrakten Modells. Ob es sich bei diesem abstrakten Modell auch um eine tatsächliche, exakte Beschreibung der Realität handelt, ist eine philosophische, viel diskutierte Frage. Sind die Gesetze, denen das Universum gehorcht, deterministisch oder nicht?

Schon seit Jahrhunderten diskutieren PhilosophInnen und NaturwissenschaftlerInnen also darüber, ob es Zufall überhaupt gibt. Albert Einstein vertrat die Ansicht, daß Gott nicht mit dem Universum würfelt, und für den Philosophen Hegel galt der Zufall als unwesentlich und unwirklich. Der klassische Determinismus ging noch Anfang dieses Jahrhunderts davon aus, daß prinzipiell alles berechenbar sei. Vertreter dieser Richtung nahmen an, daß alle Ereignisse durch ihre Ursachen exakt bestimmt werden können. In der klassischen Physik kann alles vorberechnet werden, sofern brauchbare Gesetze vorhanden sind und die Rahmenbedinungen genau genug bekannt sind. Zufall ist nur in Ermangelung geeigneter Gesetze und im Zuge unzureichender Information möglich. Damit is Zufall der Ausdruck unseres Unwissens über die Ursachen (siehe [VR98, S. 149ff]). Nehmen wir als Beispiel das klassische Zufallsexperiment: den Münzwurf. Hierbei fehlen offenbar die Randbedingungen. Wenn alle physikalischen

Informationen wie Abwurfgeschwindigkeit, Gewicht und Form der Münze hinreichend genau gemessen werden können, könnte der Ausgang exakt prognostiziert werden. Lediglich das Schütteln der Würfel in einem Becher sowie die komplexen Wurfbewegungen stellen den Physiker vor Probleme. Da die Münze nie auf die gleiche Art und Weise von einem Menschen geschüttelt und geworfen wird, tritt somit ein Informationsdefizit auf, das die Berechnung des exakten Ausgangs des Münzwurfes verhindert (siehe [Pop71, S. 158f]). Dieser Standpunkt wird als „Laplace'scher Dämon" bezeichnet, da nach Laplace jedes Wesen, das über vollständige Information verfügt, jedes zukünftige Ereignis mit Gewißheit vorhersagen könnte [VR98, S. 150]. Karl Popper [Pop71, S. 159]:

> *„Wir sprechen von Zufall, wenn wir nach dem Stand unserer Kenntnisse mit Prognosen nicht zurechtkommen. (...) Darüber, ob es Zufallsfolgen gibt, deren Glieder in keiner Weise prognostiziert werden können, stellen wir keine Behauptungen auf. Wir dürfen ja aus dem zufallsartigen Charakter einer Folge nicht einmal darauf schließen, daß [ihre Glieder nicht prognostizierbar sind oder daß] ein ‚Zufall' im subjektiven Sinn mangelnder Kenntnisse vorliegt, geschweige denn auf das objektive Fehlen von Gesetzen (im metaphysischen Sinn)."*

Das Gegenstück zum Determinismus bildet der Indeterminismus. Der erste Philosoph des 19. Jahrhunderts, der sich gründlich mit dem Zufall auseinandersetzte, war C. S. Peirce. Für die Vertreter des Indeterminismus regiert der Zufall die gesamte Welt oder zumindest einige Bereiche der Realität. Wie der Physiker J. C. Maxwell treffend feststellte, müssen viele Gesetze der Physik als Wahrscheinlichkeitsgesetze gedacht werden. Das Verhalten der winzigen Einzelteile mag ja deterministisch sein, doch da die Fülle der Teilchenbewegungen nicht exakt verfolgt werden kann, können für das Gesamtverhalten lediglich statistische Gesetze gebildet werden (siehe [VR98, S. 150ff]). Damit nahm z.B. mit der Quantenmechanik echter Zufall Einzug in die Physik (siehe [Pop71, 167ff], [Tar98, S. 149ff]). Die Beschreibung der Quantenrealitäten mittels Schrödinger-Gleichung ist zwar vollkommen deterministisch, Unschärfen treten aber dann auf, wenn der Aufenthaltsort und der Impuls kleiner Teilchen bestimmt werden soll. Jedem Ereignis wird dabei eine bestimmte Wahrscheinlichkeit zugeordnet, und ein atomares Teilchen befindet sich nur mehr mit einer gewissen Wahrscheinlichkeit an einem bestimmten Ort im Raum (siehe [VR98, S. 165f]). Es ist prinzipiell nicht mehr möglich vorherzusagen, zu welchem genauen Zeitpunkt ein Atomkern zerfallen wird – nur die durchschnittliche Halbwertszeit ist bekannt. Man spricht in diesem Zusammenhang von einem zufälligen Prozeß, den sich z. B. echte Zufallsgeneratoren zu nutze machen.

2.1 Eigenschaften von Zufallsfolgen

Wie entscheiden wir im Alltag, ob ein Ereignis zufällig ist? Wie sehen die Kriterien für eine derartige Entscheidung aus? Können wir beobachten, wie ein zufälliges Ereignis zustande kommt, so überzeugt uns dies meist davon, daß wirklich Zufall im Spiel ist. Kaum jemand wird die Zufälligkeit einer Kopf-Zahl-Folge anzweifeln, die vor seinen Augen mit einer aus seiner Brieftasche stammenden Münze erzeugte wurde.

Wie verhalten wir uns jedoch, wenn wir den Erzeugungsprozeß der Sequenz nicht verfolgen können und lediglich mit der produzierten Folge konfrontiert werden? Betrachten wir den

wiederholten Münzwurf mit den möglichen Ergebnissen „Kopf“ oder „Zahl“. Jede mögliche durch Münzwurf erzeugte Folge einer gegebenen Länge k kann mit der Wahrscheinlichkeit 2^{-k} als Ergebnis von k Münzwürfen auftreten. Trotzdem hält der Mensch für gewöhnlich die Sequenz 0110110101001 für „zufälliger“ als die Folge 00000000000, obwohl zu erwarten ist, daß beide Folgen in Münzexperimenten gleich häufig auftreten (siehe [Fum94], [L'E98a]). Warum erachten wir die zweite Folge als spezieller?

Der Mensch entwickelte – um Überleben zu können – im Laufe der Evolution die bemerkenswerte Fähigkeit, in einer Folge von Ereignissen Regelmäßigkeiten zu entdecken. Hierzu ein kleines Beispiel. Ein Steinzeitmensch sammelte rote Beeren, verspeiste diese und bekam Magenschmerzen. Mehrere Tage wiederholte sich dieses Szenario. Nach einigen Tagen konnte der Steinzeitmensch die Regelmäßigkeit von „Sammeln roter Beeren – Essen – Magenschmerzen“ erkennen, und er verzichtete in der Folge auf das Verzerren roter Beeren. Diese über Jahrtausende angeeignete Fähigkeit des Menschen zur Mustererkennung nutzt der Mensch auch heute zur Beurteilung der Zufälligkeit von Zahlenfolgen. Findet er beim Betrachten einer Sequenz kein auffälliges Muster, so wird er die Folge als zufällig bezeichnen. Umgekehrt werden wir mißtrauisch, wenn wir Eigenschaften in der Sequenz erkennen, die wir nicht in einer Zufallsfolge erwarten würden. Das bedeutet aber, daß wir bei der Diskussion über die Zufälligkeit einer gegebenen Folge stets von deren Zufälligkeit ausgehen und lediglich mittels Mustererkennung Gründe suchen, die Sequenz als nicht zufällig qualifizieren zu können. Es ist offensichtlich, daß die Beurteilung längerer Folgen leichter fällt als bei kurzen Sequenzen. Eine Zufallsfolge kann sich uns auf vielen Arten und Weisen verdächtig machen. Selbst wenn wir feststellen, daß eine Folge alle in vorher untersuchten und abgelehnten Sequenzen entdeckten Muster nicht aufweist, ist das keine Garantie dafür, daß nicht eine neue Art von Regelmäßigkeit in der Folge steckt (siehe [Len96, S. 17ff]). Karl Popper [Pop71, S. 308]:

> *„Wir können die Zufälligkeit als ‚Nichtvorhandensein von Regelmäßigkeit‘ erklären, was uns aber (...) nicht weiterhilft. Denn es gibt keine Möglichkeit, das Vorhandensein oder Nichtvorhandensein von Regelmäßigkeiten im allgemeinen nachzuprüfen, man kann nur das Vorhandensein oder Nichtvorhandensein von gegebenen oder behaupteten spezifischen Regelmäßigkeiten nachprüfen.“*

Wie wir bereits gesehen haben ist eine wesentliche Eigenschaft von Zufallsfolgen, daß sie nach keinem erkennbaren Muster aufgebaut sind und über keinerlei Regelmäßigkeit verfügen. Ist die Zufallsfolge jedoch genügend lange, so wird jedes beliebige Muster in der Sequenz auftreten und unter Umständen auch für eine Zufallsfolge äußerst untypische Teilfolgen enthalten. Geben wir beispielsweise einem unsterblichen Schimpansen eine Schreibmaschine, und lassen wir ihn blind wild drauf lostippen, so wird der Affe irgendwann todsicher als Teilfolge diese Diplomarbeit geschrieben haben (siehe [Krä96, S. 49], [VR98, S. 139f]). Knuth [Knu97, S. 167] meint zu diesem interessanten Verhalten von Zufallsfolgen folgendes:

> *„Any given finite sequence is as likely as any other. Still, nearly everyone would agree that the sequence* 011101001 *is ‚more random‘ than* 101010101*, and even the latter sequence is ‚more random‘ than* 000000000*. Although it is true that truly random sequences will exhibit locally nonrandom behavior, we would expect such behavior only in a long finite sequence, not in a short one.“*

Wie schon Knuth in [Knu97] feststellt, macht es nur Sinn, über die Zufälligkeit von Zufallsfolgen – nicht jedoch über die Zufälligkeit einzelner Zahlen (wie z. B. 17) – zu „philosophieren". Knuth spricht von einer Zufallsfolge, wenn die Folge unabhängiger Zufallszahlen eine bestimmte Verteilung aufweist und wenn jede Zahl „zufällig" und „unabhängig" von allen anderen Zahlen der Folge gefunden wird. Jeder Zahl können durch die vorgegebene Verteilung Wahrscheinlichkeiten, daß die Zahl in ein bestimmtes Intervall von Werten fällt, zugeordnet werden. Als sinnvolle Verteilung im Zusammenhang mit Zufallszahlen nimmt man die Gleichverteilung an. Zufällig und unabhängig ist nach Knuth so zu verstehen, daß die zur Auswahl konkreter Zahlenwerte führenden Ereignisse in ihrem Wesen und Zusammenspiel zu komplex sind, um ihren Einfluß mittels statistischer Tests nachweisen zu können. Vage formuliert bedeutet Zufall also, daß jede Zahl einer Folge unter denselben Bedingungen erhalten wird und daß die einzelnen Zahlen nichts mit den übrigen Zahlen in der Sequenz zu tun haben.

Zufallsfolgen müssen folgenden Kriterien genügen:

- *Gleichverteilung*
  Die erzeugten Zufallszahlen sollen im Intervall $[0, 1)$ gleichverteilt sein. Die Häufigkeit des Auftretens der einzelnen Zahlen in der Folge soll in etwa gleich hoch sein. Es existieren zahlreiche, wohldefinierte Testverfahren, die die Gleichverteilung einer Zufallsfolge nachweisen können. Zum Testen der empirischen Verteilung wird das Intervall $[0, 1)$ in Abschnitte aufgeteilt, und danach wird die Anzahl der Elemente pro Abschnitt mit der Größe dieses Intervalls in Bezug gesetzt. Schlußendlich wird noch ein Kriterium zur Interpretation der Differenz zwischen gezählten und erwarteten Elementen je Abschnitt benötigt. Dazu werden χ^2-Tests und Kolmogorov-Smirnov-Tests verwendet.

- *Unabhängigkeit*
  Kein Wert in der Folge darf aus anderen Werten gefolgert werden können, d. h. es darf kein Zusammenhang zwischen den Realisierungen der einzelnen Zufallszahlen existieren. Leider gibt es keinen Test, der die Unabhängigkeit beweisen kann. Tests sind nur in der Lage, Abhängigkeiten aufzuzeigen. Besteht eine Folge eine Anzahl von Test erfolgreich, so hilft dies, unser Vertrauen in die Unabhängigkeit der Folge zu erhöhen.

Zahlenfolgen, die verschiedenste statistische Tests auf Gleichverteilung und Unabhängigkeit bestehen, gelten noch nicht unbedingt als kryptographisch sicher. Die Kryptographie stellt noch eine zusätzliche, für die Sicherheit von Kryptosystemen entscheidende Anforderung an Zufallsfolgen. Carl Ellision [Ell95]:

> *„There are several definitions of randomness used by cryptographers, but in general there is only one criterion for a random source – that any adversary with full knowledge of your software and hardware, the money to build a matching computer and run tests with it, the ability to plant bugs in your site, etc., must not know anything about the bits you are to use next even if he knows all the bits you have used so far."*

- *Unvorhersagbarkeit*:
  Bestimmte kryptographische Anwendungen fordern nicht so sehr die statistische Zufälligkeit von Zufallsfolgen, sondern daß aus Kenntnis von aufeinanderfolgenden Realisierun-

gen von Zufallswerten unter keinen Umständen zukünftige oder vergangene Zufallszahlen vorhergesagt werden können. Auch bei Kenntnis der zuvor oder danach generierten Zahlen darf es nicht möglich sein, auf das fehlende Element schließen zu können.

2.2 Turing-Kolmogorov-Chaitin Komplexität

Der russische Mathematiker Kolmogorov nähert sich dem Begriff „Zufall" über die Komplexität und den Informationsgehalt der zu untersuchenden Sequenz. Ist die Entropie einer Folge hoch, so akzeptiert Kolmogorov diese als Zufallsfolge. Kann die Zufallsfolge durch eine kürzere Sequenz beschrieben werden, so lehnt er die Folge ab. Sequenzen werden also als zufällig bezeichnet, wenn sie nur als vollständiger Ablauf ohne Verkürzung mitgeteilt werden können, also nicht komprimierbar sind.

Diese Idee greift U. Maurer mit seinem Universal-Test für Zufallsfolgen auf (vgl. Abschnitt 3.6). Kann eine Datei mit Zufallszahlen noch gut komprimiert werden, so weisen die Zahlen höchstwahrscheinlich eine hohe Korrelation auf. Umgekehrt, ist eine Datei mit einem Komprimierverfahren nicht weiter komprimierbar, heißt das noch lange nicht, daß der Datei-Inhalt als zufällig zu betrachten ist. Extrem untypische Zufallsfolgen können bereits mit Standardkomprimierverfahren wie `gzip` und `compress` aufgedeckt werden.

Für einen Test auf Zufälligkeit einer Folge ist jedoch die Entwicklung von für diese Aufgabe spezialisierten Algorithmen nötig. Grundsätzlich ist jede mit einem Algorithmus erzeugte Zufallssequenz komprimierbar und somit nach Kolmogorov nicht zufällig, sofern eine genügend lange Folge verfügbar ist und das Komprimierverfahren den Produktionsalgorithmus mitberücksichtigt.

Ende der 60er Jahre versuchten Per Martin-Löf [ML66] und G. Chaitin [Cha66] unabhängig von einander den Begriff „Zufall" unter einem anderen Gesichtspunkt zu definieren (siehe [Knu97, S. 169-170]). Die Komplexität oder die Unordnung einer Folge wird dabei an der Länge des kürzesten Programms, das eine gegebene Folge reproduzieren kann, gemessen. Ein Programm (sprich eine Turing Maschine), das eine unregelmäßige Zahlenfolge erzeugt, muß mindestens so lange wie die generierte Folge selbst sein (siehe [May97]).

Diese Forderung kann von Computerprogrammen als Zufallsgeneratoren nicht erfüllt werden, da diese im allgemeinen deutlich kürzer als die damit erzeugten Zufallssequenzen sind. Besteht also beispielsweise eine Folge von Zahlen aus k Bits, so kann diese durch Eingabe von k oder auch weniger Bits in einem Computer erhalten werden. Je größer die Ordnung der Zahlen, umso kleiner ist das dafür benötigte Programm. So erfordert die Sequenz 12121212121212121212 höchst wahrscheinlich ein viel kürzeres Programm (z. B. $(12)^{10}$) als eine Folge von 20 mittels Münzwurf ermittelter Zufallswerte. Folgen, die keinerlei Muster enthalten, erzwingen ein Programm von maximaler Länge. Sind kürzere Programme möglich (d.h. existieren Abhängigkeiten in der Folge), so gibt dies Auskunft über den Mangel an Zufälligkeit in der Sequenz (siehe [Sip83], [Wil83]).

Leider sind die Ansätze von Kolmogorov, Martin-Löf und Chaitin nicht sonderlich konstruktiv (siehe [Gol86]). Sie helfen uns nicht, geeignete Zufallsgeneratoren zu finden und somit zufällige Folgen zu erzeugen. Aus der Theorie der Berechenbarkeit ist bekannt, daß das Erstellen eines Programms minimaler Länge zur Berechnung einer beliebigen Funktion algorithmisch