RSA & Co. in der Schule

Moderne Kryptologie, alte Mathematik, raffinierte Protokolle

Teil 1: Sprache und Statistik

von Helmut Witten, Irmgard Letzner und Ralph-Hardo Schulz

Elektronische Post (E-Mail) und das World Wide Web (WWW) werden immer stärker genutzt, das Internet wird zunehmend auch geschäftlich verwendet (E-Commerce), Banken wären ohne den elektronischen Zahlungsverkehr nicht mehr handlungsfähig. Damit wird die Sicherung von Computernetzen gegen unbefugte Zugriffe ein immer dringenderes Problem. „Hacker" sind im Begriff, die modernen Nachfolger von Räubern und Spionen zu werden. Wenn sie dabei teilweise das positiven Image eines Robin Hood genießen, liegt das wohl in der Tatsache begründet, daß sie „anonymen" Mächten wie Großbanken, Telekom oder Microsoft das Fürchten lehren. Aber auch Orwells Big Brother erscheint erschreckend aktuell, wenn man liest, in welchem Umfang Geheimdienste den Austausch von Informationen über das Internet abhören (siehe Ruhmann/Schulzki-Haddouti, 1998).

So werden Fragen des Datenschutzes bei elektronischer Post und Probleme der Sicherheit bei Home-Banking, bei EC-Kassen mit Standleitung zur Bank und bei „Plastikgeld" immer wichtiger für die Lebenswelt unserer Schülerinnen und Schüler. Nimmt man den Anspruch ernst, die jetzige und künftige Lebenswirklichkeit für die Lernenden „durchschaubar, verstehbar, und den sich in ihr entwickelnden Menschen in ihr urteilsfähig, kritikfähig, handlungsfähig werden zu lassen" (Klafki, ^{2}1991, S. 166), muß sich die Schule auch diesen Inhalten zuwenden.

Kryptologie – eine „Schlüssel"-Technik!

Ein wesentlicher Teil der möglichen Sicherungsmaßnahmen in Datennetzen wird mit modernen kryptographischen Verfahren und Protokollen realisiert. *Kryptographie* bedeutet wörtlich übersetzt *Geheimschrift*. Dabei wird eine Nachricht mit Hilfe eines *Schlüssels* chiffriert. Die verschlüsselte Nachricht, die auch als *Chiffrat* bezeichnet wird, kann im Idealfall nur von dem berechtigten Empfänger dechiffriert werden.

Um die Sicherheit kryptographischer Verfahren beurteilen zu können, muß man untersuchen, mit welchem Aufwand sich die chiffrierte Nachricht auch ohne Kenntnis des Schlüssels dechiffrieren läßt. Untersuchungen dieser Art werden als *Kryptoanalyse* (manchmal auch *Kryptanalyse*) bezeichnet, zusammen mit der Kryptographie bildet sie die *Kryptologie*. Dieses Teilgebiet der *Codierungstheorie* (vgl. Schulz, 1991; Witten, 1994) kann man daher als Wissenschaft von der Verschlüsselung und (evtl. auch unbefugten) Entschlüsselung bezeichnen.

Wenn man verbergen will, daß überhaupt eine Nachricht übermittelt wird, spricht man von *Steganographie*. Eine klassische steganographische Methode ist die Verwendung von Geheimtinte (z. B. Zitronensaft oder Urin), die erst nach einer speziellen Behandlung (z. B. Erwärmung mit einem Bügeleisen) wieder sichtbar wird. Im Computerzeitalter verwendet man digitalisierte Bilder oder Töne, in deren Grauschleier bzw. Rauschen die eigentliche Nachricht verborgen wird. Die Steganographie wird nicht zur Kryptologie gerechnet, bei letzterer sind die Chiffrate auch als solche zu erkennen. Natürlich können diese beiden Methoden auch kombiniert werden.

In der *klassischen* Kryptographie ist der Schüssel zum Dechiffrieren leicht aus dem Schlüssel zum Chiffrieren zu bestimmen. Aus diesem Grund spricht man auch von *symmetrischen* Kryptosystemen. Will man die Kommunikation mit solchen Verfahren schützen, stehen heute sehr leistungsfähige Verfahren zur Verfügung. Allerdings ist die Übermittlung des Schlüssels häufig eine Schwachstelle. Außerdem braucht man für n Teilnehmer $n(n-1)/2$ Schlüssel; die Zahl der benötigten Schlüssel wächst also quadratisch, wenn für jeden möglichen Kommunikationskanal ein eigener Schlüssel verwendet werden soll.

Kennzeichnend für die *moderne* Kryptographie sind die erst in den 70er Jahren entwickelten *asymmetrischen* oder *öffentlichen* Chiffriersysteme (***Public Key Cryptography (PKC)***). Solche Verschlüsselungsverfah-

ren wurden 1976 von Diffie und Hellmann vorgeschlagen und zuerst 1978 von ***R**ivest*, ***S**hamir* und ***A**dleman* realisiert (daher *RSA*-Algorithmus). Inzwischen hat sich herausgestellt, daß Ralph Merkle schon 1974 ein entsprechendes Verfahren publizieren wollte. Weil die Redaktion der wissenschaftlichen Zeitschrift die Bedeutung nicht erkannte, blieb die Arbeit bis 1978 liegen (Weis, 1998). Damit nicht genug: Mitarbeiter der britischen Regierungsbehörde ***C**ommunications-**E**lectronics **S**ecurity **G**roup (CESG)* hatten diese Möglichkeiten schon 1970-1973 entwickelt, durften sie aber nicht veröffentlichen. Bislang unbestätigten Gerüchten zufolge soll das amerikanische Pendant der CESG, die ***N**ational **S**ecurity **A**gency (NSA)* bereits Mitte der 60er Jahre PKC entdeckt haben. „Und natürlich fragt man sich unwillkürlich: Wenn sie das schon vorher wußten, ohne es zu verraten – was wissen sie dann jetzt?" (Gröndahl, 1998).

Wie funktionieren nun diese berühmten Systeme mit den vielen „Vätern"? Jeder Teilnehmer an der Kommunikation erhält ein eigenes Schlüsselpaar: einen *öffentlichen* Schlüssel, der z. B. in einer Art Telefonbuch verzeichnet ist, und einen *privaten* Schlüssel, den nur der Teilnehmer selbst kennt und für andere unzugänglich aufbewahrt. Damit braucht man nur noch so viele Schlüsselpaare wie Teilnehmer.

Will man eine Nachricht übermitteln, die vor unbefugtem Lesen geschützt sein soll, so verschlüsselt man sie mit dem öffentlichen Schlüssel des Empfängers; nur dieser kann sie dann mit seinem privaten Schlüssel entschlüsseln. Umgekehrt kann man mit den asymmetrischen Systemen auch das Problem der elektronischen Unterschrift lösen: Der Teilnehmer verschlüsselt einen Teil seiner Nachricht mit seinem geheimen Schlüssel und jeder kann mit dem öffentlichen Schlüssel nachprüfen, ob diese wirklich von dem Betreffenden stammt. Kombiniert man beide Methoden, erhält man eine vertrauliche und authentische Kommunikation, die vor Abhören und Verfälschen (relativ) sicher ist.

Solche Anwendungen kryptographischer Verfahren nach einem vorgegebenen Schema nennt man *Protokoll*. Es gibt aber noch weitaus raffiniertere Protokolle: Eine Person A will sich z. B. gegenüber einer Person Z dadurch identifizieren, indem sie zeigt, daß sie einen bestimmten geheimen Zahlencode kennt. Idealerweise sollte das so erfolgen, daß A ihr Geheimnis zur Identifikation nicht preisgeben muß, damit Z während des Prozesses nichts darüber erfährt und so keine Chance erhält, sich Dritten gegenüber als A auszugeben. Dies leistet der berühmte „Null-Ahnung-Beweis" (zero knowledge protocol), der 1985 entdeckt wurde und asymmetrische Kryptographie nutzt. Hierbei müssen die Teilnehmer nicht nur etwas berechnen, sondern sie müssen sich gemäß genau festgelegter, ausgeklügelter Regeln unterhalten, d. h. das entsprechende Protokoll beachten (vgl. z. B. Beutelspacher u. a., 1995, S. 37 ff.).

Wie kann man nun asymmetrische Kryptosysteme realisieren? Da die Dechiffrierfunktion die Inverse der Chiffrierfunktion ist, bedeutet dies, daß die Chiffrierung mit Funktionen erfolgen muß, deren Inverse nur mit einem inakzeptabel hohen Aufwand oder zusätzlichen Kenntnissen bestimmt werden kann. Solche Funktionen werden auch als *Einweg-Funktionen mit Hintertür* bezeichnet (engl. *trap-door one-way functions*, wörtlich übersetzt also Einweg-Funktionen mit Falltür; „Hintertür" scheint uns aber treffender zu sein). Der Informationsgehalt ist für den unberechtigten Lauscher nach der Verschlüsselung verschwunden und kann nur mit einer zusätzlichen Information wieder ans Tageslicht befördert werden – man muß eben die Hintertür kennen!

27 60	–**Erich** Cauer-10	3 41 24 01	Meie
35 79	–**Erich** Gontermann-26	7 85 81 28	(Sbg
62 47	–**Erich** Kol.Britzer Wiesen Kirschenweg 3	6 01 24 50	Meie
89 16	–**Erich** Rüdersdorfer -14	2 91 90 30	–Gür
67 44	–**Erich** Schneller-70a	6 71 73 27	–Gür
99 81	–**Erika** KGA Familiengärten 7	4 00 20 74	–Gür
15 16	–**Erika** Kornblumenring 99	6 64 36 86	–Gür
38 55	–**Erika** Prignitz-140	5 41 95 00	–Gür
33 83	–**Erika** Schwatlo-6	7 12 71 52	–Gür
80 48	–**Erika** Stolberg-9	7 51 72 82	(Lir)
52 10	–**Erika** Wisbyer -45A	4 71 67 63	–Gür
58 40	–**Erna** Am Falkenberg 96	6 72 89 88	–Gür
79 87	–**Erna** Danziger -109	4 23 31 37	–Gür
74 92	–**Erna** Sundgauer -61	8 11 62 59	–Gür
95 54	–**Ernst** Belfaster -21	4 52 38 84	–Gür
01 46	–**Erwin** Mutziger -1	9 27 36 47	Auto
46 52	–**Erwin** (Buk) Rudower -111	6 02 45 30	–Gür
33 09	–**Erwin** Simmel-27	4 92 25 82	–Gür
80 10	–**Esther** Bärenstein-31	5 42 33 38	–Gür
72 32	–**Eva** Gleditsch-68	2 15 42 46	–Gür
35 19	–**Eva** u. Jens-A. Gontermann-9	7 86 67 22	–Gür
97 02	–**Eva** Marienfelder Chaussee 46	7 43 11 16	–Gür
27 93	D1Fu 0 17 13 30 26 08		–Gür
13 96	–**Eva** (Pre) Metzer -3	4 41 82 27	–Gür
14 50	–**Eva-Maria** Kurfürsten-106	2 11 65 75	–Gür
84 98	–**Eveline** Schneller-63	6 71 85 87	–Gür
98 84	–**Evi** Bochumer -9	3 91 24 02	–Gür
42 31	–**Ewald** Stuttgarter Platz 13	3 23 26 98	–Gur
22 36	–F. u. M.	3 01 65 95	–H. D
11 98	–F. Paster-Behrens--	6 01 64 44	
	–F. Wiclef-	3 95 12 01	–H. D
15 13	**Meier-Fleschner Christina**	7 44 04 75	–H. P
82 14	Marienfelder -85		–H. R
	Meier Flora Ahrenshooper -35	9 29 56 36	–H.-I
08 00	–**Frank** Ahrenshooper -55	9 20 04 89	Gem
55 89	–**Frank** (Tem) Blumenthal-10	7 52 92 22	Joac
	D2-Nr 0 17 23 94 01 51		–Hal
23 41	–**Frank** Essener -20	3 91 82 73	–Har
44 25	–**Frank** Grottkauer -24	5 62 69 72	–Har
08 40	–**Frank** u. Antje (Adl) Handjery-28B	6 71 55 32	–Har

Bild 1: Das Telefonbuch – ein alltägliches Beispiel für eine Einwegfunktion.

Ein alltägliches Beispiel für eine Einwegfunktion bietet jedes Telefonbuch (Bild 1). Während es einfach ist, zu einem bekannten Namen und Adresse die Telefonnummer zu finden, ist es umgekehrt (fast) unmöglich, zu einer Telefonnummer den passenden Eintrag zu finden – es sei denn, man kennt den Namen oder man benutzt eine der datenschutzrechtlich problematischen CDs, die eine entsprechende Suchfunktion besitzen.

Will man computertaugliche Einwegfunktionen finden, kommt die im Untertitel erwähnte *alte Mathematik* ins Spiel: Die Erfinder des RSA-Verfahrens haben ein Stück klassischer elementarer Zahlentheorie verwendet, auf die wir in einer späteren Folge dieser Artikelserie noch genauer eingehen werden. Um die Einwegfunktion zu konstruieren, werden bei RSA Produkte sehr großer Primzahlen verwendet, die zwar schnell zu berechnen sind, sich bei einer entsprechenden Länge aber nur mit ungeheuer großem Rechenaufwand wieder in ihre Faktoren zerlegen lassen (vgl. Schulz, 1996). Weitere Einwegfunktionen erhält man durch das Quadrieren bzw. Potenzieren in endlichen Körpern: Während Potenzen leicht zu berechnen sind, gibt es (soweit bekannt) keine vergleichbar effizienten Verfahren zur Bestimmung der Wurzeln bzw. Logarithmen.

In den letzten 25 Jahren hat sich die moderne (zivile) Kryptologie stark entwickelt. In diesem Grenzgebiet zwischen Mathematik und Informatik wird intensiv geforscht und manchmal auch viel Geld verdient. Wir wollen in dieser Artikelserie darlegen, welche Inhalte der Kryptologie für die Schule geeignet sind. Dabei können wir auf Artikel aus fachdidaktischen Zeitschriften zurückgreifen, auf die wir an den entsprechenden

Stellen hinweisen. Dennoch ist die Erfahrungsbasis noch ziemlich schmal, und viele didaktische Fragen sind nach wie vor offen. Wir glauben, daß wir mit unseren Erfahrungen aus dem Unterricht, aus Schülerseminaren an der Universität, aus der Lehrplanarbeit und aus gemeinsam durchgeführten Lehrerfortbildungen einige Akzente neu setzen können.

Kryptologie in der Schule?

Wenn man sich näher mit der Frage beschäftigt, wie man Kryptologie im Unterricht behandeln kann, stellt man fest, daß sie ganz unterschiedliche Aspekte miteinander verbindet: mathematische (z. B. Statistik, elementare Zahlentheorie, Geschichte der Mathematik), speziell algorithmische (Langzahlarithmetik, schnelles Potenzieren, erweiterter Euklidischer Algorithmus zur Bestimmung der modularen Inversen, Primzahltests, ...), raffinierte kryptographische Protokolle („Cybercash" (vgl. Baumann, 1997), „Null-Ahnung-Beweise", ...) und nicht zuletzt die gesellschaftlichen Auswirkungen (z. B. Kryptopolitik: Datenschutz versus Kontrollmöglichkeiten des Staates – vgl. Heibey/Pfitzmann/Sandl, 1996; Kippenhahn, 1997, S. 262 ff.; Schneier, 1996, S. 677 ff.; Wobst, 1997, S. 307 ff.).

Insofern ergeben sich bei der Kryptologie Bezüge zur Mathematik, zur Informatik bzw. ITG und zur politischen Bildung. Während beim Mathematikunterricht die inhaltliche Abgrenzung traditionell eher eng verstanden wird, ist die Situation im Informatik-/ITG-Unterricht in dieser Hinsicht günstiger: So sehen z. B. die Berliner Rahmenpläne vor, ca. 25 % der Unterrichtszeit auf den Bereich „Anwendungen und Auswirkungen der Informationstechnik" zu verwenden; hierbei müssen u. a. Datenschutz und Datensicherheit behandelt werden.

Aus dem unterschiedlichen Fachverständnis ergibt sich ein interessantes Spannungsfeld. In den neueren Mathematikbüchern zur elementaren Zahlentheorie wird die moderne Kryptologie als Anwendung aufgeführt, wobei als historische Pointe angemerkt werden kann, daß für viele Mathematiker gerade die Zahlentheorie als Musterbeispiel der „Reinen" Mathematik galt, die nicht durch „Anwendungen beschmutzt" war.

Der Informatik- und ITG-Unterricht ist umgekehrt an der Anwendung der Kryptologie im Bereich der Sicherheit in Netzen interessiert und kann daher die benötigte Mathematik im Prinzip als „black box" übernehmen, vielleicht noch mit Beispielen anschaulich machen. Natürlich ist ein solches Vorgehen unbefriedigend, weil die Frage, warum das funktioniert, weitgehend unbeantwortet bleibt.

Wir wollen uns nicht mit der Diskussion aufhalten, welches Unterrichtsfach als „Heimatdisziplin" für die Kryptologie „zuständig" ist (vgl. Baumann, [2]1996, S. 375). Es wird mit Sicherheit kein eigenes Fach „Kryptologie" geben, daher werden diese Inhalte nur in unterschiedlichen Fächern behandelt werden können – am besten natürlich fachübergreifend. Wir folgen in dieser Artikelserie einem pragmatischen Ansatz und geben Anregungen, wie Fragen der Computersicherheit und der Kryptologie im Mathematikunterricht (der Sekundarstufen I und II, mit und ohne Computer- bzw. Taschenrechner-Einsatz), im Informatik- bzw. ITG-Unterricht und nicht zuletzt im fächerübergreifenden Projektunterricht behandelt werden können. Wir freuen uns bei der knappen Unterrichtszeit über jede Stunde, die diesen wichtigen Inhalten gewidmet werden kann.

Das Entziffern von einfachen Geheimschriften ist für Schülerinnen und Schüler eine motivierende Aufgabe. Hierbei wird man – zumindest am Anfang – sog. *monoalphabetische* Chiffrierungen betrachten. Zwischen den Buchstaben aus dem Klartext und den Zeichen der Geheimschrift besteht hier eine umkehrbar eindeutige Zuordnung. Obwohl es beliebig viele solcher Geheimalphabete gibt (vgl. z. B. Gardner, 1981, S. 30 ff.; Kippenhahn, 1997, S. 13 ff.; Beutelspacher, 1997, S. 15 ff.; Sgarro/Würmli, 1991, S. 16 ff.; Niederdrenk-Felgner, 1988, S. 95 ff.), ist die Kryptoanalyse dieser Geheimschriften elementar möglich, wenn der Klartext in natürlicher Sprache formuliert ist. Damit erhält man einen sinnvollen und schülergerechten Einstieg zu einer intensiveren Beschäftigung mit der Kryptologie. Wir wollen zur Kryptoanalyse monoalphabetischer Verschlüsselungen einem berühmten Beispiel aus der Literatur folgen.

Das Gold des Gehenkten

Edgar Allan Poe ist nicht nur ein bedeutender und nach wie vor populärer Schriftsteller des 19. Jahrhunderts, sondern war auch ein passionierter Hobby-Kryptologe. Als Mitarbeiter von „Graham's Magazine" hatte er seine Leser aufgefordert, ihm monoalphabetische Kryptogramme zur Entschlüsselung zu schicken. Bei hundert Einsendungen gelang ihm das auch in 97 Fällen, in zwei Fällen konnte er nachweisen, daß es sich nur um eine sinnlose Zeichenfolge handelte (Kippenhahn, 1997, S. 103 ff.).

Später hat er diese Fähigkeit in der Erzählung „Der Goldkäfer" seinem *alter ego* Legrand verliehen. Um den Schatz des 150 Jahre zuvor gehenkten Seeräubers Kidd zu finden, mußte der scharfsinnige Held der Erzählung das folgende Kryptogramm entziffern, nachdem er die mit unsichtbarer Tinte geschriebenen Zeichen sichtbar gemacht hatte:

53##+305))6*;4826)4#.)4#);806*;48+8|60))85;;]8*;:
#*8+83(88)5*+;46(;88*96*?;8)*#(;485);5*+2:*#(;4
956*2(5*-4)8|8*;4069285);)6+8)4##;1(#9;48081;8:8
#1;48+85;4)485+528806*81(#9;48;(88;4(#?34;48)4#
;161;:188;#?;

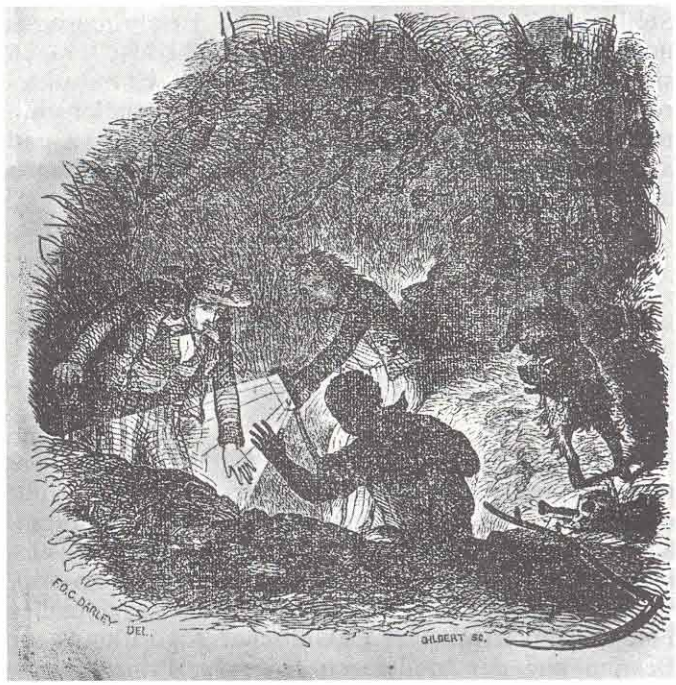
Legrand überreicht dem Ich-Erzähler der Geschichte das Pergament mit diesen Geheimzeichen, der steht aber immer noch im Dunkeln: „Und warteten meiner bei Lösung dieses Rätsels auch alle Juwelen von Golkonda, ich wär's, des' bin ich mir sicher, doch nicht imstande, sie mir zu verdienen."

„Und dennoch", sagte Legrand, „ist die Lösung keineswegs so schwierig, wie es Ihnen beim ersten eiligen Überblicken der Zeichen wohl erscheinen mag. [...] Nach allem, was von Kidd bekannt ist, konnte ich auch wieder nicht annehmen, daß er sich auf besonders raffinierte Chiffrierkünste verstanden habe [...]."

„Und Sie fanden wirklich die Lösung?"

„Mit Leichtigkeit; ich habe schon andere Chiffreschriften gelöst, die zehntausendmal komplizierter waren" (Poe, 1966, S. 901 ff.).

Warum konnte Legrand diese Aufgabe so leicht lösen? Zunächst einmal konnte er annehmen, daß es sich um einen englischsprachigen Text handeln müsse, weil Kidd mit einem gezeichneten Zicklein (engl.: *kid*) unterzeichnet hatte und dieses Wortspiel nur im Englischen funktioniert. Wir werden weiter unten sehen, daß man im Computerzeitalter nicht mehr auf solche Hinweise angewiesen ist. Die zweite Annahme von Legrand war, daß es sich um eine monoalphabetische Chiffrierung handeln müsse, weil eine andere Chiffre für Kidd zu kompliziert gewesen wäre.



Quelle: Lenning, 1959, S. 117

Bild 2: Der Schatz von Käpt'n Kidd ist gefunden! Originalillustration zu E. A. Poes Erzählung „Der Goldkäfer".

Wie man eine monoalphabetische Chiffrierung knackt

Die Kryptoanalyse von Legrand/Poe ist auch nach mehr als 150 Jahren so interessant, daß ihr im Kryptologie-Standardwerk von F. L. Bauer ein eigener Abschnitt gewidmet wird („Freistil-Methoden" in Bauer, 1995, S. 244 f.).

Legrand stellt zunächst eine Häufigkeitstabelle auf und stellt fest, daß das Zeichen „8" am häufigsten vorkommt, nämlich 33mal. Das nächsthäufigste ist das Semikolon (26mal), andere Zeichen wie der Punkt und der Gedankenstrich kommen nur einmal vor. Er nimmt daher an, daß es sich bei „8" um „e" handeln muß, weil dieser Buchstabe im Englischen am häufigsten vorkommt. Diese These wird noch dadurch gestärkt, daß dieses Zeichen fünfmal doppelt auftritt – und tatsächlich sind im Englischen Worte mit Doppel-„e" im Gegensatz zum Deutschen sehr häufig (z. B. *meet, fleet, speed, seen, been, agree* usw.).

Als nächstes macht sich Legrand auf die Suche nach dem häufigsten Wort der englischen Sprache: *the*. Er findet siebenmal das Muster *;48* und hat damit die ersten drei Buchstaben gefunden – „damit ist ein großer Schritt getan" (Poe, 1966, S. 905), denn man kann damit auch Wortgrenzen finden. Durch Einsetzen der gefundenen Buchstaben werden anhand wahrscheinlicher Wörter weitere Zeichen entziffert: *the t(ee* liefert z. B. „r". So findet er nach und nach die Tabelle:

5	+	8	3	4	6	*	#	(	;
a	d	e	g	h	i	n	o	r	t

Mit diesen 10 Buchstaben ist die restliche Entschlüsselung nicht mehr schwer. Die Lösung des Kryptogramms ist im übrigen immer noch ziemlich kryptisch, aber Legrands Scharfsinn soll sich ja noch auf anderen Gebieten entfalten:

> „A good glass in the bishop's hostel in the devil's seat – twenty-one degrees and thirteen minutes – northeast and by north – main branch seventh limb east side – shoot from the left eye of the death's head – a bee line from the tree through the shot fifty feet out".
>
> „Ein gut Glas im Bishop's Hotel auf dem Teufelssitz – einundzwanzig Grad und dreizehn Minuten – Nordnordost – Hauptast siebter Zweig Ostseite – schieße vom linken Auge des Totenkopfes – eine gerade Linie vom Baum durch den Schuß fünfzig Fuß fort" (Poe, 1966, S. 908).

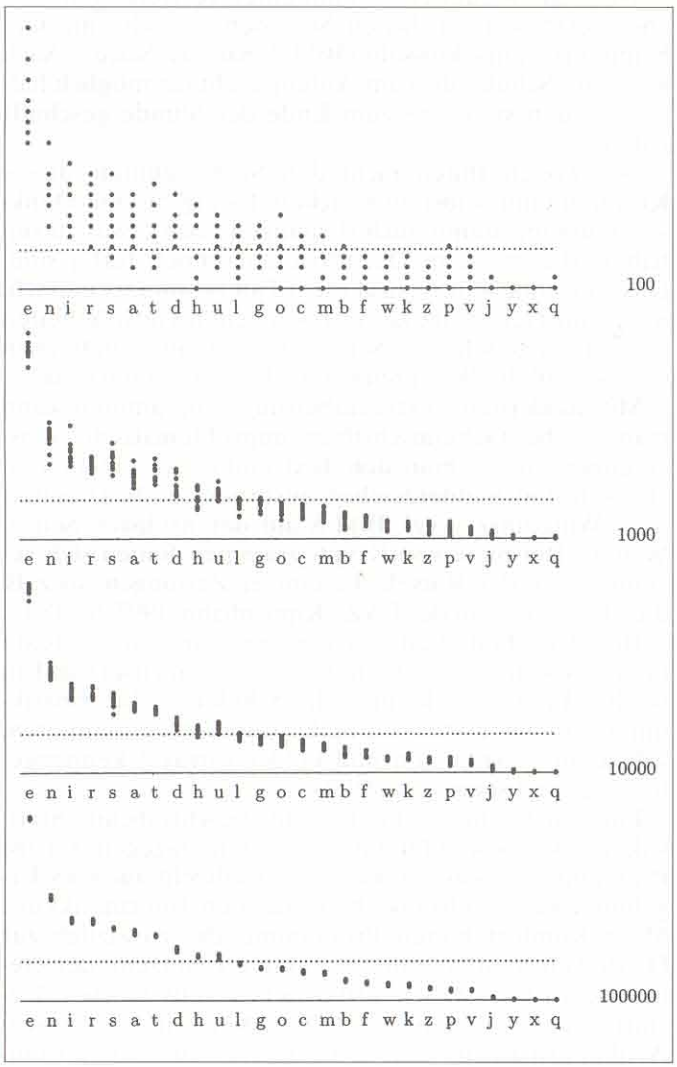
Legrand benutzt zum Dechiffrieren also zwei Merkmale natürlicher Sprachen: Zum einen die Häufigkeitsverteilung der einzelnen Buchstaben, die alles andere als gleichverteilt sind, zum anderen bestimmte charakteristische Muster (Doppel-„e", *the*). Darüber hinaus verwendet er wahrscheinliche Wörter, denn *e*n Te+t m*t Leer?tellen* ist wegen der Redundanz der Sprache relativ leicht zu entziffern.

Natürliche Sprachen und Statistik

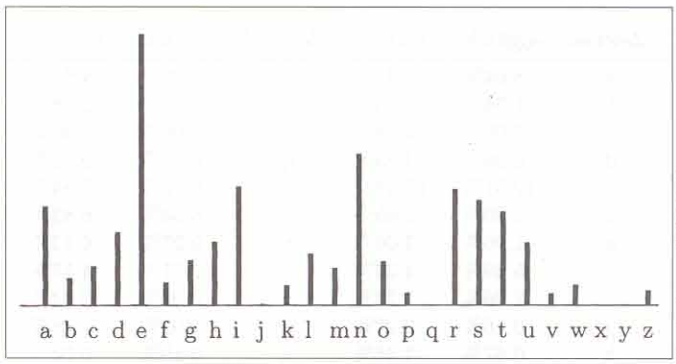
An Texten kann man (möglichst mit Computerhilfe) sehr schön das Gesetz der großen Zahl studieren. Bei einer umfangreichen Untersuchung an Artikeln aus der Süddeutschen Zeitung wurden jeweils 10 Stichproben im Umfang von 100, 1000, 10000 und 100000 Zeichen entnommen und die Häufigkeitsverteilung ermittelt (vgl. Bild 3).

Wenn man die Zeichen nicht nach der Häufigkeit, sondern nach dem Alphabet sortiert, bekommt man „Häufigkeitsgebirge", die für die jeweiligen Sprachen charakteristisch sind: „Im Deutschen sind besonders auffällig die e-Spitze und der n-Gipfel, die f-g-h-i-Flanke mit anschließender j-k-Senke, die o-p-q-Senke mit anschließendem r-s-t-u-Kamm" (Bauer, 1995, S. 215; vgl. Bild 4).

Bild 3: Schwankungen der Häufigkeiten der Einzelzeichen im Deutschen.



Quelle: Bauer, 1995, S. 221



Quelle: Bauer, 1995, S. 215

Bild 4: Häufigkeitsgebirge im Deutschen.

„Demgegenüber bestehen im Englischen signifikante Unterschiede: Es ist ein a-Gipfel ausgeprägter, es besteht ein h-i-Kamm und ein l-m-n-o-Kamm, der r-s-t-u-Kamm hat einen t-Gipfel; jedoch finden sich b-c-d-Flanke, j-k-Senke und v-w-x-y-z-Niederung wieder" (Bauer, 1995, S. 216; vgl. Bild 5).

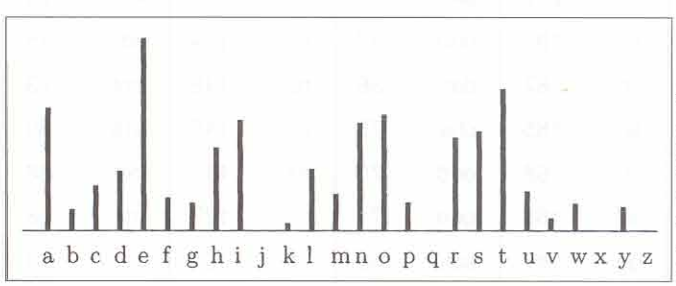
Während die „Häufigkeitsgebirge" einen intuitiven Eindruck von der Buchstabenverteilung natürlicher Sprachen liefern, gibt es inzwischen sehr leistungsfähige statistische Verfahren zur Sprachanalyse mit Computern. Ein eindrucksvolles Beispiel ist auf der Homepage des europäischen Forschungszentrums der Firma Xerox zu finden:

`http://www.xrce.xerox.com/cgi-bin/mltt/LanguageGuesser`

Das Programm erkennt mit hoher Treffsicherheit schon nach wenigen eingetippten Wörtern, um welche der 32 vorgebenen Sprachen es sich handelt, ohne daß der Computer den „Sinn" der eingegebenen Sätze erkennen muß. Solche statistischen Verfahren werden z. B. auch bei der auf deutsche Texte spezialsierten Suchmaschine „fireball" verwendet (vgl. den Artikel von Tolksdorf und Paulus in diesem Heft, S. 22 ff.).

Natürlich sind auch einfache Geheimschriften mit statistischen Methoden relativ leicht zu „knacken", wie E. A. Poe in der Geschichte vom Goldkäfer gezeigt hat. Denn im Fall von monoalphabetischen Chiffrierungen bleiben die statistischen Besonderheiten der natürlichen Sprache voll erhalten, egal welche möglicherweise exotischen Zeichen den einzelnen Buchstaben zuordnet wurden.

Bild 5: Häufigkeitsgebirge im Englischen.



Quelle: Bauer, 1995, S. 215

Zeichen	englisch	deutsch	Zeichen	englisch	deutsch
a	8.04%	6.47%	n	7.09%	9.84%
b	1.54%	1.93%	o	7.60%	2.98%
c	3.06%	2.68%	p	2.00%	0.96%
d	3.99%	4.83%	q	0.11%	0.02%
e	12.51%	17.48%	r	6.12%	7.54%
f	2.30%	1.65%	s	6.54%	6.83%
g	1.96%	3.06%	t	9.25%	6.13%
h	5.49%	4.23%	u	2.71%	4.17%
i	7.26%	7.73%	v	0.99%	0.94%
j	0.16%	0.27%	w	1.92%	1.48%
k	0.67%	1.46%	x	0.19%	0.04%
l	4.14%	3.49%	y	1.73%	0.08%
m	2.53%	2.58%	z	0.09%	1.14%

Quelle: Bauer, 1995, S. 223

Bild 6: Zeichenwahrscheinlichkeiten im Englischen und im Deutschen.

Für eine elementare Kryptoanalyse geht man zunächst von der Häufigkeitskeitsverteilung der Einzelzeichen aus (vgl. Bild 6). Zeichen, die etwa gleich oft auftreten, werden zu sog. *Cliquen* zusammengefaßt. Im Deutschen ergeben sich {e} {n} {irsat} {dhu} {lgocm} {bfwkz} {pv} {jyxq}, im Englischen dagegen {e} {t} {oani} {rsh} {dl} {ucwm} {fygpb} {vk} {xjqz} (Bauer, 1995, S. 223; Beutelspacher, ^{3}1993, S. 18; Niederdrenk-Felgner, 1988, S. 99).

Beim Ersetzen der Einzelzeichen findet man die ersten beiden Buchstaben relativ eindeutig – es sei denn, daß in dem Text häufig von „Ananaszüchtern in Alaska" die Rede ist. Auch der Roman „Anton Voyls Fortgang" dürfte Schwierigkeiten bereiten, da er gänzlich ohne „e" geschrieben wurde (Perec, ^{3}1998) – das ist aber sicherlich nicht der Normalfall. Bei den Buchstaben aus der nächsten Clique („irsat" im Deutschen) ist eine eindeutige Zuordnung i. allg. nicht von vornherein möglich; dort wird man verschiedene Ersetzungen ausprobieren müssen.

Wertvolle Anhaltspunkte für die Entzifferung bieten auch die sehr unterschiedlichen Häufigkeiten von Zeichenpaaren (*Bigrammen*) und -tripeln (*Trigrammen*). In Bild 7 sind jeweils die zehn häufigsten Bi- und Trigramme der deutschen und der englischen Sprache aufgeführt (vgl. Bauer, 1995, S. 233 ff; Beutelspacher, ^{3}1993, S. 25 – die Zahlenangaben weichen bei unterschiedlicher Textbasis geringfügig voneinander ab). Wenn eine Clique von Einzelzeichen durch die Häufigkeit nicht getrennt werden kann, zieht man die Bigrammhäufigkeiten heran. Wenn dies immer noch nicht zum Erfolg führt, untersucht man die Trigrammhäufigkeiten. Darüber hinaus können auch Mustererkennungstechniken (zur Erkennung „wahrscheinlicher Wörter") eingesetzt werden.

Bild 7: Tabelle der zehn häufigsten Bi- und Trigramme im Deutschen und im Englischen (Zahlenangaben jeweils bezogen auf 10 000).

Deutsch				Englisch			
er	409	ein	122	th	315	the	353
en	400	ich	111	he	251	ing	111
ch	242	nde	89	an	172	and	102
de	227	die	87	in	169	ion	75
ei	193	und	87	er	154	tio	75
nd	187	der	86	re	148	ent	73
te	185	che	75	on	145	ere	69
in	168	end	75	es	145	her	68
ie	163	gen	71	ti	128	ate	66
ge	147	sch	66	at	124	ver	64

nach: Bauer, 1995, S. 233 ff.

Häufigkeitsanalyse von Kryptogrammen im Unterricht

Wenn Sie wieder einmal mit einer Vertretungsstunde „beglückt" werden, lassen Sie doch einfach eine Geheimschrift entschlüsseln (Bild 8, nächste Seite)! Auch wenn die Schüler dies am Anfang nicht für möglich halten, werden sie es bis zum Ende der Stunde geschafft haben.

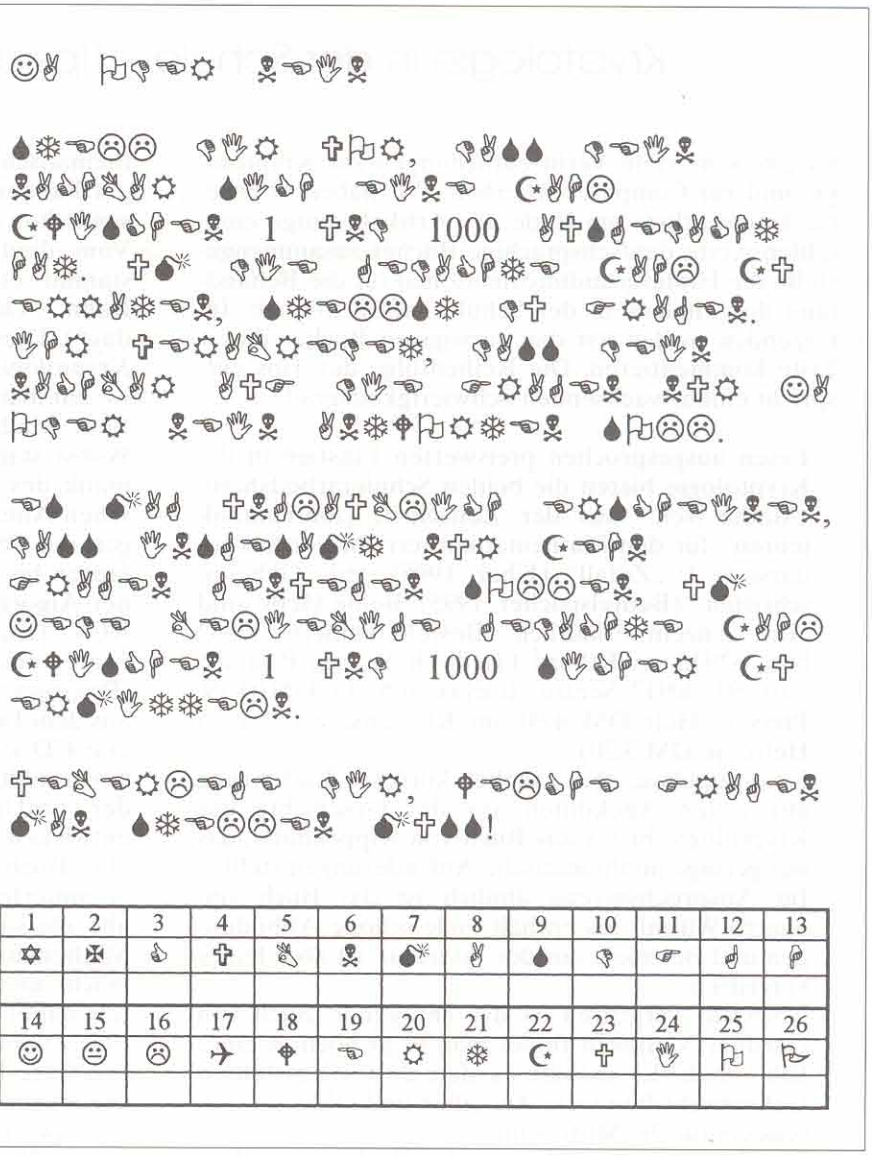
Wir wollen Ihnen nicht den Spaß nehmen, dieses Kryptogramm selbst zu knacken: Es ergibt eine Denksportaufgabe, damit auch diejenigen noch etwas zu tun haben, die mit dem Dechiffrieren schnell fertig sind. Das Entschlüsseln ist in diesem Fall besonders einfach, weil – im Gegensatz zu professionellen Gewohnheiten – die Leerstellen und Satzzeichen erhalten geblieben sind, so daß die Wortgrenzen noch zu erkennen sind.

Mit modernen Textverarbeitungsprogrammen kann man solche Geheimschriften unproblematisch selbst erzeugen, indem man den Text einfach in einem Zeichensatz mit Sonderzeichen ausdruckt (z. B. Dingbats oder Wingdings – vgl. Bild 8 auf der nächsten Seite). Weitere Beispiele von Kryptogrammen finden sich regelmäßig in der Rätselecke einiger Zeitungen, so z. B. die „Findlinge" in der FAZ (Kippenhahn, 1997, S. 118 f.).

Bei der Häufigkeitsanalyse umfangreicherer Texte bietet es sich natürlich an, Computer einzusetzen. Ein solches Programm können die Schüler im Informatikunterricht im ersten Lernjahr selbst programmieren, sobald sie Textdateien und Felder (arrays) kennengelernt haben (Sennholz, 1995).

Für den Mathematikunterricht (beschreibende Statistik in der Sekundarstufe I) sollten dagegen fertige Programme eingesetzt werden – vielleicht auch als Ergebnis einer „Auftragsarbeit" für den Informatikkurs. Mehr Komfort bieten Programme, die zusätzlich zur Häufigkeitsanalyse eine interaktive Ersetzung der Geheimschriftbuchstaben ermöglichen (ein solches Dechiffrierprogramm ist im LOG IN-Service zu diesem Artikel erhältlich, siehe S. 119). Aber auch ohne Com-

Bild 8:
Arbeitsbogen zur Entschlüsselung.



puter gibt es anregende Aufgaben für kleine „Kryptologen" – beispielsweise im Arbeitsheft „Zufall" der Mathe-Welt (Uher, 1995).

Ein anspruchsvolles Projekt zur Kryptoanalyse monoalphabetischer Chiffrate (Informatikunterricht der Sekundarstufe II) zielt auf ein Programm zur automatischen Erzeugung einer Rohentschlüsselung. Dazu reicht die Häufigkeitsanalyse von Einzelbuchstaben nicht aus; im Programm müssen auch noch (mindestens) die Bigrammhäufigkeiten untersucht werden (vgl. Künzell, 1998).

Wie geht es weiter?

Man darf nicht glauben, daß grundsätzlich jede monoalphabetische Chiffrierung unsicher ist – auch moderne Verfahren wie RSA sind monoalphabetisch in dem Sinne, daß jedem „Zeichen" eines „Klartextes" bei gegebenem Schlüssel genau ein „Geheimzeichen" zugeordnet wird. Nur besteht dieser „Klartext" nicht aus Buchstaben einer natürlichen Sprache, sondern aus riesigen Zahlen (ca. 200 Stellen)!

Zunächst wollen wir aber der Frage nachgehen, ob es auch sichere Chiffrierungen normaler Texte gibt. Und in der Tat: Nach vielen Versuchen, die sich alle als „knackbar" herausgestellt haben, ist es am Anfang dieses Jahrhunderts gelungen, eine Chiffriermethode zu erfinden, die beweisbar sicher ist und bei korrekter Anwendung nicht gebrochen werden kann.

(Fortsetzung folgt)

Danksagung

Wir möchten Herrn Dr. Franz von der Walther-Rathenau-Oberschule (Gymnasium) in Berlin-Grunewald danken, daß er sein Dechiffrierprogramm zur Häufigkeitsanalyse und interaktiven Zeichenersetzung (mit Turbo-PASCAL-Quellcode) für den LOG IN-Service zur Verfügung gestellt hat.

Im **LOG in-Service** (s. S. 119) erhalten Sie ein PASCAL-Quell-Programm zur Dechiffrierung monoalphabetisch verschlüsselter Texte sowie eine Datei mit der Kopiervorlage des Arbeitsbogens vom Bild 8.

StD Helmut Witten
Landesbildstelle Berlin
BICS
Wickingerufer 7
10555 Berlin

E-Mail: witten@bics.be.schule

Irmgard Letzner
Fritz-Karsen-Schule
Onkel-Bräsig-Straße 76-78
12359 Berlin

E-Mail: letzner@math.fu-berlin.de

Prof. Dr. Ralph-Hardo Schulz
Freie Universität Berlin
Fachbereich Mathematik und Informatik
Institut für Mathematik II
Arnimallee 3
14195 Berlin

E-Mail: schulz@math.fu-berlin.de

Kryptologie in der Schule – Tips zum Weiterlesen

Es gibt sehr viele Veröffentlichungen zur Kryptologie und zur Computersicherheit. Wir haben in unserer Literaturliste am Ende des Artikels einige empfehlenswerte deutschsprachige Bücher zusammengestellt, die Hintergrundinformationen für die Behandlung des Themas in der Schule liefern können. Im folgenden wollen wir die wichtigsten Bücher dieser Liste kommentieren. Die Reihenfolge der Tips entspricht einem wachsenden Schwierigkeitsgrad.

- ▷ Einen ausgesprochen preiswerten Einstieg in die Kryptologie bieten die beiden Schülerarbeitshefte „Mathe-Welt" aus der Zeitschrift „mathematik lehren" für den Mathematikunterricht der Sekundarstufe I: „Zufall" (Uher, 1995) und „Geheimschriften" (Beutelspacher, 1995). Beide Hefte sind auch einzeln erhältlich (Bestellnummern 32932 bzw. 32933 im Erhard Friedrich Verlag, Postfach 10 01 50, 30917 Seelze, Telefax: (05 11) 4 00 04 19, Preis je Heft DM 4,80, im Klassensatz (mind. 5 Hefte) je DM 3,20).
- ▷ Eine populäre, aber sachlich korrekte Einführung mit vielen Anekdoten aus der Geschichte der Kryptologie bietet das Buch von Kippenhahn, das nur geringe mathematische Anforderungen stellt.
- ▷ Im Anspruchsniveau ähnlich ist das Buch von Sgarro/Würmli. Es enthält viele schöne Abbildungen und Beispiele aus der Literatur, ist aber leider vergriffen.
- ▷ Ebenfalls vergriffen ist das anregende Buch von Gardner; vielleicht findet man es ja noch in einer Bibliothek. Es enthält zu den dort vorgestellten Geheimschriften viele Aufgaben und lohnt sich besonders für die Mittelstufe.
- ▷ Empfehlenswert sind auch die Bücher von Beutelspacher. Der „Klassiker" Kryptologie bietet eine gut lesbare Einführung, die allerdings schon etwas mehr Mathematik voraussetzt als die vorher genannten Bücher. Wer es kürzer und einfacher mag, ist mit seinem neuen, sehr preiswerten Buch Geheimsprachen gut bedient. Beutelspacher/Schwenk/Wolfenstetter setzen dagegen mehr mathematisches Verständnis voraus, geben dafür auf gut hundert Seiten einen Überblick über die wesentlichen modernen Verfahren.
- ▷ Vom deutschen Informatik-Pionier F. L. Bauer stammt eine hochgelobte Monographie (David Kahn: „The best single book on cryptology today"). Dieses Buch, das zunächst unter dem Titel *Kryptologie – Methoden und Maximen* erschienen ist, hat mathematisch einfaches Hochschulniveau.
- ▷ Vom Mathematiker und Informatiker Reinhard Wobst stammt *Abenteuer Kryptologie*. „Die Thematik des Buches reicht von spannenden historischen Anekdoten bis hin zu neuesten Entwicklungen der Politik, von den einfachsten Chiffrierverfahren bis hin zur genauen Untersuchung moderner Algorithmen wie IDEA, RC5, RSA, DES und RC4" (Klappentext). Die beigefügte CD enthält Programme zu den Algorithmen im Quelltext (Programmiersprache C) sowie Texte und „Links" aus dem Internet zur Krypto-Politik.
- ▷ Die CD zum Buch von Kaderali enthält dagegen einen multimedialen Kurs zur Kryptologie, der an der FernUniversität in Hagen für das Fernstudium entwickelt wurde.
- ▷ Das Buch von Stallings, einem international renommierten Sicherheitsexperten, richtet sich an alle, die sich für Netzwerksicherheit interessieren. Mathematische Grundlagen werden plausibel gemacht und in Anhängen zu den jeweiligen Kapiteln näher erläutert.
- ▷ Phil Zimmermanns Anleitung zu dem berühmten Freeware-Programm *Pretty Good Privacy* ist nicht nur wegen der originellen deutschen Übersetzung lesenswert; die beigefügte CD enthält das Programm PGP für alle gängigen Betriebssysteme.
- ▷ Last but not least soll das Standardwerk von Schneier über die moderne Kryptographie empfohlen werden. Es ist unentbehrlich für jeden Programmierer, der kryptographische Algorithmen implementieren will; für die Schule ist es eher zum Nachschlagen geeignet (800 Seiten und 1653 Literaturverweise!).

Literatur

Bauer, F. L.: Entzifferte Geheimnisse – Methoden und Maximen der Kryptologie. Berlin u. a.: Springer, 1995.

Baumann, R.: Didaktik der Informatik. Stuttgart u. a.: Klett-Verlag, ^{2}1996.

Baumann, R.: Digitales Geld – Bestellen und Bezahlen im Internet. In: LOG IN, 17 (1997), H. 2, S. 30-38.

Beutelspacher, A.: Kryptologie – Eine Einführung in die Wissenschaft vom Verschlüsseln, Verbergen und Verheimlichen. Braunschweig: Vieweg, ^{3}1993.

Beutelspacher, A.: Mathe-Welt – Geheimschriften. In: mathematik lehren, 13 (1995), Heft 72, S. 23-46.

Beutelspacher, A.; Schwenk, J.; Wolfenstetter, K.-D.: Moderne Verfahren der Kryptographie – Von RSA zu Zero-Knowledge. Braunschweig: Vieweg, 1995.

Beutelspacher, A.: Geheimsprachen – Geschichte und Techniken. München: Beck, 1997.

Gardner, M.: Das verhexte Alphabet – Tips und Tricks für Geheimschriften. Frankfurt/M. u. a.: Ullstein, 1981.

Gröhndal, B: Die Entdeckung der Public-Key-Kryptograhpie – Ehre, wem Ehre gebührt. In: TELPOLIS (http://www.heise.de/tp/deutsch/special/krypto/1381/1.html), 20.01.98.

Heibey, H.-W.; Pfitzmann, A.; Sandl, U.: Kryptographie – Herausforderung für Staat und Gesellschaft. In: LOG IN, 16 (1996), H. 5/6, S.37-43.

Kaderali, F.: Kryptologie: Technischer Datenschutz in Kommunikationsnetzen. Bonn u. a.: Addison-Wesley, 1997 (mit CD).

Kippenhahn, R.: Verschlüsselte Botschaften – Geheimschrift, Enigma und Chipkarte. Reinbek bei Hamburg: Rowohlt, 1997.

Klafki, W.: Neue Studien zur Bildungstheorie und Didaktik – Zeitgemäße Allgemeinbildung und kritisch-konstruktive Didaktik. Weinheim und Basel: Beltz, ^{2}1991.

Künzell, St.: Binnendifferenzierung im Informatikunterricht. In: LOG IN 18 (1998); H. 1, S. 51-55 (Teil 1: Voraussetzungen); H. 2, S. 51-55 (Teil 2: Unterrichtsplanung); dieses Heft: S. 75 ff. (Teil 3: Unterrichtsverlauf).

Lenning, W.: Edgar Allan Poe. Reinbek bei Hamburg: Rowohlt, 1959.

Niederdrenk-Felgner C.: Algorithmen der elementaren Zahlentheorie – Computer im Mathematikunterricht, Heft 1. Tübingen: Deutsches Institut für Fernstudien – DIFF, 1988.

Perec, G.: Anton Voyls Fortgang. Frankfurt: Zweitausendeins, ^{3}1998.

Poe, E. A.: Der Goldkäfer. In: Werke I. Olten: Walter, 1966, S. 859-914.

Ruhmann, I.; Schulzki-Haddouti, C.: Abhör-Dschungel – Geheimdienste lesen ungeniert mit. In: c't, 16 (1998), H. 5, S. 82-93.

Schneier, B.: Angewandte Kryptographie – Protokolle, Algorithmen und Sourcecode in C. Bonn u. a.: Addison-Wesley, 1996.

Schulz, R.-H.: Codierungstheorie – eine Einführung. Braunschweig; Wiesbaden: Vieweg, 1991.

Schulz, R.-H.: Primzahlen im öffentlichen Chiffrierverfahren. In: mathematik lehren, 11 (1993), Heft 61, S. 56-64.

Schulz, R.-H.: Primfaktorzerlegung – Experimente zum Zeitaufwand. In: LOG IN, 16 (1996), H. 5/6, S. 22-26.

Sennholz, K.: Verschlüsselte Botschaften. In: Informatik betrifft uns, 8 (1995), H. 2, S. 1-23.

Sgarro, A; Würmli, M.: Geheimschriften – Verschlüsseln und Enträtseln von Geheimtexten. Augsburg: Weltbild Verlag, 1991.

Stallings, W: Sicherheit im Datennetz. München u. a.: Prentice Hall, 1995.

Tolksdorf, R., Paulus, O. K.: Informationen im Web erschließen. In diesem Heft, S. 22 ff.

Uher, B.: Mathe-Welt: Zufall. In: mathematik lehren, 13 (1995), Heft 71, S. 23-46.

Weis, R.: Zwei Schlüssel einer Nachricht – Kryptographie mit öffentlichen Schlüsseln. In: PC Magazin Spezial (Kryptographie), 5'98 , S. 24-29.

Witten, H.: Codierungstheorie – Ein Überblick. In: LOG IN, 14 (1994), H. 5/6, S. 13-18.

Wobst, R.: Abenteuer Kryptologie – Methoden, Risiken und Nutzen der Datenverschlüsselung. Bonn u. a.: Addison-Wesley 1997 (mit CD).

Zimmermann, Ph.: PGP – Pretty Good Privacy, das Verschlüsselungsprogramm für Ihre private elektronische Post. Bielefeld: Art d'Ameublement, ^{3}1997 (mit CD).