Grundlagen der RSA-Verschlüsselung

Torsten Linnemann
Diplomergänzungsstudium «Didaktik der Mathematik»
Modul «Mathematik ohne Computer»
Leitung: Prof. A. Beutelspacher

9. Mai 2010

Inhaltsverzeichnis

1 Einleitung

Im Modul «Mathematik ohne Computer» des Studiengangs «Diploma of Advanced Studies in Fachdidaktik Mathematik» der Universität Bern im Frühlingssemester 2009 ist diese Arbeit als Leistungsnachweis entstanden.

Grundlage der Arbeit sind einige Erfahrungen, die ich mit einer Projektwoche einer 10. Klasse zur Kryptologie und eines Kurses «Anwendungen der Mathematik» im Maturjahr gemacht habe.

Verwendet habe ich in diesen Unterrichtseinheiten jeweils das Modul «Grundlagen der RSA-Verschlüsselung» von K. Schläpfer [2], das im Rahmen des Mathe-Prismas http://www.matheprisma.uni-wuppertal.de zur Verfügung gestellt wird.

Mit den Beispielen und den Ver- und Entschlüsselungen kamen die Lernenden gut zurecht. Der theoretische Teil war allerdings nur schwer zu erschliessen und ich habe jeweils mit einem Lehrervortrag die Theorie der RSA-Verschlüsselung liefern müssen.

Diese Arbeit ist die Ausformulierung des Lehrervortrags für Schülerinnen und Schüler der Sekundarstufe II. Sie liefert einige zahlentheoretische Sätze und deren Anwendung für die RSA-Verschlüsselung.

2 Einführung

Eine gute Darstellung der RSA-Verschlüsselung findet sich im Internet im Mathe-Prisma der Uni Wuppertal (http://www.matheprisma.uni-wuppertal.de). Hier sollen nur einige theoretische Grundlagen gelegt werden, damit der Darstellung im Mathe-Prisma ohne Ablenkung gefolgt werden kann.

Den Namen hat das Verfahren von Ronald Rivest, Adi Shamir und Leonard Adleman, die das Verfahren 1977 beschrieben haben [3].

Es handelt sich um ein «asymmetrisches» Verfahren, bei dem alle benötigten Daten öffentlich ausgetauscht werden können - und zwar so, dass nur der Empfänger die gesendete Verschlüsselung in Klartext übersetzen kann. Eigentlich ja verwunderlich, dass eine praktisch nicht zu knackende Verschlüsselung existiert, bei der alle Daten öffentlich ausgetauscht werden können. Kein Wunder, dass es bis in die 1970er Jahre dauerte, bis diese Verschlüsselung entdeckt wurde. Sie beruht darauf, dass es praktisch unmöglich ist, aus dem Ergebnis einer Modul-Rechnung einer Potenz auf den Exponenten zu schliessen.

Es ist allerdings auch nicht bewiesen, dass es nicht vielleicht ein einfaches System gibt, dies doch zu tun. Damit wäre dann diese Verschlüsselung hinfällig. Aber, wenn es doch ein einfaches System gäbe, würde es dann verraten werden? Dieses System wäre zum Beispiel für Geheimdienste viel wert. Und bei diesen arbeiten viele MathematikerInnen. . .

Auch Hacker könnten sich mit einem solchen System eine goldene Nase verdienen. Es ist also nicht gesagt, dass es veröffentlich wird, wenn es gelingt, das RSA-System zu «knacken»

Das RSA-System beruht auf der Mathematik der Primzahlen: Es ist sehr schwierig, sehr grosse Zahlen in ihre Primfaktoren zu zerlegen. Wenn beispielsweise zwei 200-stellige Primzahlen miteinander multipliziert werden, ist es kaum möglich, aus dem Resultat wieder auf die Primzahlen zu schliessen.

Um genauer zu verstehen, wie verschlüsselt wird, müssen wir etwas weiter ausholen und einige zahlentheoretische Sätze darstellen. Dazu brauchen wir zwei Abschnitte: „Grösster gemeinsamer Teiler und das modulare Inverse“ und „Reste bei Potenzen“.

3 Grösster gemeinsamer Teiler und das modulare Inverse

Wenn eine Zahl kein Teiler einer anderen Zahl ist, so bleibt ein Rest. Damit umzugehen ist das Thema der Modulo-Rechnung. Zunächst die Definition:

3.1 Modulo-Rechnung

Dieses Kapitel folgt im Wesentlichen dem Buch «Kryptologie»1 von Albrecht Beutelspacher, [1].

Definition 3.1 *Es gilt* $a \mod m = r$*, falls* $a = km + r$ *mit* $r < m$*.* $(a, k, m, r \in \mathbb{N})$

Ein kleiner Satz:

Satz 3.1 *1.) Aus* $r = a$ *mod* n *und* $s = b$ *mod* n *folgt* $rs = ab \mod n$*.*
2.) Aus $a \mod m = r$ *folgt* $a^n \mod m = r^n \mod m$*.*

Beweis: Es gilt: Aus $a \mod m = r$ und $b \mod m = s$ folgt $a+b \mod m = r+s \mod m$. Das letzte mod benötigen wir, falls $r + s \geq m$. Wiederholtes Addieren ergibt eine Multiplikation. Also gilt der erste Teil des Satzes. Und wiederholtes Multiplizieren ist eine Potenzrechnung. Also gilt der entsprechende Satz auch hier. □

3.2 Der grösste gemeinsame Teiler

Den grössten gemeinsamen Teiler (ggT) zweier Zahlen per Primfaktorzerlegung zu finden kann mühsam sein, wenn die auftretenden Primzahlen gross sind. Beispiel $1517 = 41{\cdot}37$ und $1763 = 41 \cdot 43$. Hier den ggT zu finden ist, zumindest ohne Taschenrechner eher aufwändig.

Viel effizienter ist der sogenannte *Euklidische Algorithmus*. Er führt schnell zum Ziel. Es ist allerdings nicht ohne weiteres offensichtlich, dass er die richtige Lösung bietet. Der Algorithmus beruht auf Division mit Rest.

Beispiel 3.1 Der ggT von $a = 792$ und $b = 75$. Wir teilen 792 durch 75 mit Rest und führen das Verfahren dann weiter:

$$792 = 10 \cdot 75 + 42$$

$$75 = 1 \cdot 42 + 33$$

$$42 = 1 \cdot 33 + 9$$

$$33 = 3 \cdot 9 + 6$$

$$9 = 1 \cdot 6 + 3$$

$$6 = 2 \cdot 3$$

□

Wir behaupten nun, dass 3 der grösste gemeinsame Teiler von 792 und 75 ist. Um dies zu belegen, machen wir uns folgendes klar:

Satz 3.2 *Es gelte für natürliche Zahlen a, b, q und r:*

$$a = bq + r \text{ und } 0 \leq r \leq b - 1 \quad (1)$$

Dann ist ggT(a, b) =ggT(b, r).

Teilt eine Zahl a und b so muss sie auch r teilen. Ansonsten bliebe in Gleichung 1 nur auf der rechten Seite ein Rest bei Division durch diese Zahl. Also ist der ggT von a und b auch ein Teiler von r und damit ggT$(a, b) \leq$ ggT(b, r).

Genauso muss jede Zahl, die b und r teilt, auch a teilen. Also ist der ggT von b und r auch ein Teiler von a. Es gilt also ggT$(a, b) \geq$ ggT(b, r).Damit stimmen die beiden ggT überein. □

Im obigen Beispiel verändert sich also von Schritt zu Schritt der ggT nicht:

ggT$(792, 75)$ =ggT$(75, 42)$ =ggT$(42, 33)$ =ggT$(33, 9)$ =ggT$(9, 6)$ =ggT$(6, 3) = 3$.

Satz 3.3 Euklidischer Algorithmus *Der ggT zweier Zahlen a und b lässt sich bestimmen, indem natürliche Zahlen mit*

$$a = bq + r \text{ und } 0 \leq r \leq b - 1$$

bestimmt werden.
Anschliessend wird das Verfahren auf b und r an. Dies wird so lange durchgeführt, bis der ggT abzulesen ist.

Anschaulich wird das Verfahren schön von Gorski und Müller-Philipp [2] auf Seite 71-72 beschrieben.

Aufgabe 3.1 Bestimme den ggT von 120 und 84.

Aufgabe 3.2 Bestimme den ggT von 1764 und 105

3.3 Vielfachsummen

Satz 3.4 Vielfachsummendarstellung *Sei d der ggT der Zahlen a und b. Dann gibt es ganze Zahlen x und y mit der Eigenschaft*

$$d = xa + yb.$$

Eine solche Darstellung heisst Vielfachsummendarstellung. des ggT von a und b.

Beispiel 3.2 Wir wählen $a = 71$ und $b = 15$. Zunächst berechnen wir den ggT von a und b:

$$71 = 4 \cdot 15 + 11$$

$$15 = 1 \cdot 11 + 4$$

$$11 = 2 \cdot 4 + 3$$

$$4 = 1 \cdot 3 + 1$$

Nun lösen wir diese Gleichungskette von unten nach oben auf. Allerdings rechnen wir die Gleichungen nicht aus – sonst würde sich nur das nutzlose $1 = 1$ ergeben, sondern fassen die Terme nur geeignet zusammen.

$$\begin{aligned} 1 &= 4 - 1 \cdot 3 \\ &= 4 - 1 \cdot (11 - 2 \cdot 4) = 3 \cdot 4 - 1 \cdot 11 \\ &= 3 \cdot (15 - 1 \cdot 11) - 1 \cdot 11 = 3 \cdot 15 - 4 \cdot 11 \\ &= 3 \cdot 15 - 4 \cdot (71 - 4 \cdot 15) = 19 \cdot 15 - 4 \cdot 71 \end{aligned}$$

Also gilt

$$1 = 19 \cdot 15 + (-4) \cdot 71$$

□

Mit diesem Beispiel lässt sich der allgemeine Beweis des Satzes erklären. Sei $a = bq + ar$. Wenn wir eine Vielfachsummendarstellung für b und r haben, dann könnnen wir daraus eine für a und b ableiten. Dies ist genau das Prinzip des «Hinaufhangelns», welches wir am Beispiel durchgeführt haben:

Sei also $1 = xr + yb$ mit ganzen Zahlen x und y. Wegen $r = a - bq$ können wir diese Gleichung wie folgt umschreiben.

$$1 = x(a - bq) + yb.$$

Wir lösen die Klammer auf und erhalten mit

$$1 = xa + (y - xq)b = xa + yb$$

eine Vielfachsummendarstellung für a und b. □

Das Verfahren, mit dem wir die Vielfachsummendarstellung erhalten haben, nennt sich auch der «erweiterte euklidsche Algorithmus».

Aufgabe 3.3 Bestimme ganze Zahlen mit der Eigenschaft $3 = 792x + 75y$.

Aufgabe 3.4 Bestimme ganze Zahlen mit der Eigenschaft $1 = 17x + 55y$.

3.4 Modulare Inverse

Das Inverse einer Zahl x ist die Zahl y für die gilt $xy = 1$. Es gilt natürlich $y = 1/x$. Also lässt sich in den ganzen Zahlen, ausser für $x = 1$, nie ein Inverses finden. Interessanterweise ist das dann bei der Modulo-Rechnung wieder möglich.

Satz 3.5 Satz von der molularen Inversen *Sind a und n teilerfremde Zahlen, so gibt es eine ganze Zahl b mit der Eigenschaft*

$$ab \text{ mod } n = 1.$$

Also hat ab geteilt durch n den Rest 1.

Der Beweis ist mit unserem jetzigen Wissensstand ganz einfach: Da der ggT von a und n gleich 1 ist, gibt es ganze Zahlen x und y mit folgender Eigenschaft:

$$1 = ax + ny.$$

Da ny durch n teilbar ist, ergibt ax bei Division durch n den Rest 1, also ist

$$ax \text{ mod } n = 1.$$

Mit $b = x$ folgt die Behauptung. □

Algorithmus zur Bestimmung der modularen Inversen: Seien a und n ganze Zahlen mit $\text{ggT}(a, n) = 1$.

Zuerst wird die Vielfachsummendarstellung $1 = ax + yn$ bestimmt. Dann ist x die modulare Inverse von a modulo n.

Beispiel 3.3 Was ist die Inverse von 15 modulo 71?

Wegen $1 = 19 \cdot 15 + (-4) \cdot 71$ ist

$$19 \cdot 15 \text{ mod } = (1 + 4 \cdot 71) \text{ mod } 71 = 1.$$

Also ist 19 die modulare Inverse von 15 modulo 71. □

Bemerkung: Hier kann es passieren, dass $b < 0$. Für den RSA-Algorithmus werden wir eine positive Zahl benötigen. Wir wählen statt b eine Zahl d mit $d_k = (b + kn)$. Wird k gross genug gewählt, so ist d positiv. Für den RSA-Algorithmus werden wir das kleinste positive d_k wählen. Die entscheidende Eigenschaft $ad \text{ mod } n = 1$ gilt weiterhin, denn $a \cdot (b + kn) \text{ mod } n = (ab + akn) \text{ mod } n = ab \text{ mod } n = 1$.

Zusammenfassung:

- *Der euklidische Altorithmus kann von einem Rechner äusserst bequem ausgeführt werden. Der grösste gemeinsame Teiler zweier Zahlen kann einfach berechnet werden. Insbesondere kann leicht entschieden werden, ob zwei Zahlen den grössten gemeinsamen Teiler 1 haben.*
- *Wenn a und n teilerfremde Zahlen sind, so kann leicht eine positive Zahl d berechnet werden mit der Eigenschaft*

$$a \cdot d \text{ mod } n = 1$$

4 Reste bei Potenzen

Und jetzt kommt gleich der am schwierigsten zu beweisende Satz, den wir in diesem Kurs benötigen:

Satz 4.1 Kleiner Satz von Fermat *Für alle Primzahlen p und natürliche Zahlen a gilt*

$$(a^p - a) \mod p = 0 \tag{2}$$

Der hier aufgeführte Beweis benutzt einiges an Mathematik: Binomialkoeffizienten, das Summenzeichen und vollständige Induktion.

Schülerinnen und Schüler, die diese Bereiche der Mathematik noch nicht kennen, müssen den Satz einfach als Black Box verwenden.

Beweis: Induktionsverankerung: Für $a = 1$ gilt jedenfalls: $1^p - 1 \mod p = 0$.

Annahme: Der Satz gilt für ein n. Zu beweisen ist jetzt, dass der Satz auch für $n + 1$ gilt.

$$(a+1)^p - (a+1) \mod p = \left(\sum_{k=0}^{p} \left(\left(\begin{array}{c} p \\ k \end{array} \right) a^p \right) - a - 1 \right) \mod p$$

Für $1 \leq k \leq p-1$ ist $\left(\begin{array}{c} p \\ k \end{array} \right)$ durch p teilbar. (p ist prim). Die obige Gleichung lässt sich also weiter umformen:

$$= (a^p + 1 - a - 1) \mod p = (a^p - a) \mod p = 0,$$

was ja nach Voraussetzung gilt. □

Beispiel 4.1 $4^2 - 4 \mod 2 = 16 - 4 \mod 2 = 12 \mod 2 = 0$ □

Beispiel 4.2 $5^3 - 5 \mod 5 = 120 \mod 5 = 0$ □

Beispiel 4.3 $6^5 - 6 = 7770 \mod 6 = 0$

□

Dass dieser Satz gilt, ist nicht offensichtlich. Der Beweis ist tiefliegend. Und mit tiefsinnigen Tatsachen lässt sich oft viel anfangen. Zuerst müssen aber noch viele weitere Sätze aus dem kleinen Satz von Fermat gefolgert werden.

Folgerung 4.1 $a^{p-1} \mod p = 1$
Hier ist p eine Primzahl und a und p sind teilerfremd.

Wir folgern aus dem letzten Satz:

$$a^p - a \mod p = 0 \Rightarrow a^p \mod p = a \mod p \Rightarrow a^p = kp + a \Rightarrow a^{p-1} = \frac{kp}{a} + 1$$

Da $a^{p-1} \in \mathbb{N}$ folgt $\frac{kp}{a} \in \mathbb{N}$. Da p und a teilerfremd sind, gilt also $\frac{kp}{a} = cp$, also

$$a^{p-1} = cp + 1 \Rightarrow a^{p-1} \mod p = 1$$

□

Beispiel 4.4 $21^4 \mod 5 = 194481 \mod p = 1$

Das liess sich ja noch leicht nachrechnen, wir wissen jetzt aber auch: □

Beispiel 4.5 $1024^{16} \mod 17 = 1$ und das sprengt schon eher unsere Kopfrechenfertigkeiten. □

Die nächste Folgerung enthält jetzt bereits das Produkt zweier Primzahlen – und legt damit die Grundlage von RSA:

Satz 4.2 Euler *Hier sind p und q Primzahlen. Ausserdem sei a teilerfremd zu pq. Dann gilt:*

$$a^{(p-1)(q-1)} \mod pq = 1 \tag{3}$$

Irgendwie müssen wir den letzten Satz mit sich selbst multiplizieren...

Ist a teilerfremd zu q, so ist auch a^{p-1} teilerfremd zu q. Mit dem letzten Satz folgt also:

$a^{(p-1)(q-1)} \mod q = 1$, oder, anders formuliert $a^{(p-1)(q-1)} = kq + 1$

Ebenso folgt dann natürlich

$$a^{(p-1)(q-1)} = lp + 1$$

Hier sind k und l natürliche Zahlen, die uns aber nicht weiter interessieren. Es gilt jedenfalls $kq = lp$. Also ist $kq = lp$, sowohl durch q als auch durch p teilbar. Mithin $kq = lp = mqp$.

Wir erhalten $a^{(p-1)(q-1)} = kq + 1 = lp + 1 = mpq + 1$

Dies bedeutet $a^{(p-1)(q-1)} \mod pq = 1$ □

5 RSA-Verschlüsselung

Dies ist eine Kurzfassung der RSA-Schlüsselerzeugung. Sie dient als Ergänzung zur Durcharbeitung des Moduls des Mathe-Prismas [4].

5.1 Wahl des Schlüssels

- Es werden zwei Primzahlen p und q gewählt.
- $n = pq$.
- Es wird eine Zahl e gewählt, die teilerfremd zu $(p-1)(q-1)$ ist und d so berechnet, dass gilt

  $$ed \bmod (p-1)(q-1) = 1$$

  (modulare Inverse). Es gilt also

  $$ed = 1 + k(p-1)(q-1)$$

- Öffentlicher Schlüssel: e und n
- Privater Schlüssel: d (d kann grösser als Null gewählt werden)
- Geheim: p, q und $(p-1)(q-1)$

5.2 Verschlüsselung und Entschlüsselung

Die Zahl m, die kleiner als n sein muss, wird verschlüsselt als

$$c = m^e \bmod n$$

Die Zahl c wird entschlüsselt als

$$m' = c^d \bmod n$$

Satz 5.1 vom korrekten Dechiffrieren *Mit obiger Verschlüsselungsprozedur gilt $m' = m$. Der Algorithmus arbeitet also korrekt.*

Erinnerung an Satz 3.1: Aus $a \mod m = r$ folgt $a^n \mod m = r^n \mod m$.

Für $a = m^e$ ergibt das:

Aus $m^e \bmod n = c$ folgt $m^{ed} \bmod n = c^d \bmod n = m'$.

Wir müssen also noch zeigen, dass $m^{ed} \bmod n = m$ gilt.

Nach Definition von e und d gilt $ed = 1 + k(p-1)(q-1)$.

An dieser Stelle wird eine Fallunterscheidung vorgenommen[1]:

1. Fall: m ist teilerfremd zu n. Es ergibt sich

$$m^{ed} \bmod n = m^{1+k(p-1)(q-1)} \bmod n = m \cdot m^{k(p-1)(q-1)} \bmod n$$

[1] Herr Michael Hoppe hat mich auf die Notwendigkeit der Fallunterscheidung hingewiesen und mir auch gleich den Beweis im zweiten Fall geliefert. Vielen Dank

Nun wenden wir den ersten Teil von Satz 3.1 an und erhalten

$$= (m \bmod n) \left(m^{k(p-1)(q-1)} \bmod n \right)$$

Da $m < n$ ist $m \bmod n = m$.

Zu zeigen bleibt $\left(m^{k(p-1)(q-1)} \bmod n \right) = 1 \bmod n$.

Es wird der zentrale Satz von Euler, Satz 4.2 verwendet. Es ergibt sich

$$\left(m^{k(p-1)(q-1)} \bmod n \right) = 1 \bmod n$$

2. Fall: m ist Vielfaches von q, aber nicht von p (der umgekehrte Fall geht analog). Wir rechnen

$m^{k(p-1)(q-1)} \bmod p = (m^{(p-1)})^{(q-1)k} \bmod p = 1^{(q-1)k} \bmod p = 1 \bmod p$, denn nach Folgerung 3.1 ist $m^{p-1} = 1 \bmod p$.

Das heisst, p teilt $m^{k(p-1)(q-1)} - 1$ und da q ja m teilt folgt, dass $n = pq$ den Ausdruck $m(m^{k(q-1)(p-1)} - 1) = m^{k(q-1)(p-1)+1} - m = m^{ed} - m$ teilt , also $m^{ed} = m \bmod n$.

Somit ergibt sich nach dem Dechiffrieren mit dem privaten Schlüssel tatsächlich wieder m. □

Literatur

[1] Beutelspacher, A. (2002): Krytologie. Vieweg-Verlag: Braunschweig/Wiesbaden.

[2] Gorski, A.; Müller-Philipp, S. (2008): Leitfaden Arithmetik. Vieweg-Verlag: Braunschweig/Wiesbaden.

[3] Rivest, R.; Shamir, A.; Adleman, L. (1977): A Method for Obtaining Digital Signatures and Public-Key Cryptosystems. http://people.csail.mit.edu/rivest/Rsapaper.pdf, Zugriff Mai 2009

[4] Schäfer, K. (2002): RSA–Verschlüsselung. http://www.matheprisma.uni-wuppertal.de/Module/RSA/index.htm, Zugriff Mai 2009