Fachhochschule Wedel

Seminar IT-Sicherheit

Cryptanalytic Attacks on RSA

1.4 Intractable Number-Theoretic Problems

von: Sebastian Krumm
Matrikelnummer: 100037

Dozent: Prof. Dr. Gerd Beuster
Abgabetermin: 06.06.2016

Inhaltsverzeichnis

1. Einleitung

Grundlage dieser Seminararbeit ist das Kapitel „Intractable Number-Theoretic Problems“ aus dem Buch „Cryptanalytic Attacks on RSA“ von Song Y. Yan. [1]

In diesem Teil des Buches werden, nach einigen Beispielen von bis heute ungelösten Problemen der Zahlentheorie und einem Blick auf die Komplexitätsklassen, mehrere schwer lösbare zahlentheoretische Probleme vorgestellt.

Nach einem Exkurs über zahlentheoretische Probleme sowie, analog zum Kapitel im Buch, einer Betrachtung von bis heute ungelösten Problemen, liegt das Hauptaugenmerk dieser Arbeit auf dem „Integer Factorization Problem (IFP)“. Dabei wird allgemein auf das Faktorisierungsproblem ganzer Zahlen und dessen Bedeutung eingegangen und neben ausgewählten einfacheren Verfahren wird mit dem „Quadratischen Sieb“ ein schnelles, bedeutungsvolles Verfahren zur Lösung dieses Problems vorgestellt.

2. Zahlentheoretische Probleme

Neben der Geometrie ist die Zahlentheorie das älteste Gebiet der Mathematik.

Zu Problemstellungen in der Zahlentheorie gehören Aufgaben wie die Teilbarkeit von Zahlen, die Suche und Überprüfung von Primzahlen oder die (Prim-) Faktorzerlegung. Was viele Probleme gemeinsam haben, ist zugleich eine bedeutsame Eigenschaft in der Zahlentheorie. Die Probleme sind oftmals einfach anzugeben und meist für jedermann verständlich, wohingegen die Lösung häufig sehr komplex ist.

Forschungen in diesem Bereich begannen bereits einige Jahrhunderte v. Chr. und sind bis heute noch von großer Bedeutung. Euklid konnte bspw. um 300 v. Chr. beweisen, dass es keine größte Primzahl gibt. Ein weiterer berühmter Zahlentheoretiker ist Pierre de Fermat, der im 17. Jahrhundert eine Vermutung aufstellte, die als der „Große Fermatsche Satz“ bekannt ist.

$$a^n + b^n = c^n$$

Diese besagt, dass für die obige Gleichung keine ganzzahlige Lösung existiert, wenn n größer als 2 ist.

Erst vor einigen Jahren (1995) gelang es Andrew Willes diese Vermutung zu beweisen.

Dies ist ein eindrucksvolles Beispiel dafür, wie lang mitunter an zahlentheoretischen Problemen geforscht wird, bis eine Lösung oder ein Beweis gefunden werden kann.

2.1 Problemklassen

Um Aussagen über die Komplexität von zahlentheoretischen Problemen machen und verschiedene Probleme miteinander vergleichen zu können, wurden Komplexitätsklassen eingeführt.

Diese sind in Ihren Einzelheiten zwar nicht Gegenstand dieser Arbeit, die Wichtigsten sollen aber für das weitere Verständnis an dieser Stelle kurz aufgeführt werden.

Die Klasse P: Probleme dieser Klasse werden als praktisch oder effizient lösbar bezeichnet, da sie von einer deterministischen Maschine in polynomieller Zeit gelöst werden können. Aufgrund der Tatsache, dass sie von einer deterministischen Turing-Maschine gelöst werden können, sind sie auch wirklich realisierbar.

Die Klasse NP: Probleme dieser Klasse können von einer nichtdeterministischen Turing-Maschine in polynomieller Zeit gelöst werden, vermutlich aber nicht von einer deterministischen Turing-Maschine. Sie gelten daher als schwer lösbar. (intractable) [2]

2.2 Bedeutung

Neben der unumstrittenen Bedeutung innerhalb der Mathematik war die Zahlentheorie ursprünglich überwiegend für die Wirtschaft und den Handel von Bedeutung. Im Laufe der Zeit erstreckte sich das Anwendungsgebiet der Zahlentheorie aber auch zunehmend auf Bereiche außerhalb der Mathematik.

Hervorzuheben ist aktuell zweifellos das Anwendungsgebiet der Informatik und dort insbesondere der Bereich der Codierung und Kryptografie, denn viele Sicherheits- und Verschlüsselungsverfahren beruhen auf Erkenntnissen der Zahlentheorie.

Neue Forschungserfolge hätten unter Umständen direkte Auswirkungen auf solche Verschlüsselungsverfahren, was ein enormes Sicherheitsrisiko darstellen kann.

Auch das RSA-Kryptosystem ist ein Verfahren, dass auf der Zahlentheorie beruht, nämlich auf den „Intractable Number-Theoretic Problems“.

An dieser Stelle sei darauf hingewiesen, dass nicht alle Probleme, die aktuell als schwer lösbar eingestuft sind, dies auch wirklich sind. Solange kein Beweis vorliegt, bedeutet es lediglich, dass noch kein Algorithmus gefunden wurde, der in polynomieller Zeit zur Lösung kommt. Hier liegt das erwähnte Sicherheitsrisiko der Verschlüsselungsverfahren, die auf eben diesen Problemen basieren.

Im weiteren Verlauf von Interesse sind daher jene schwer lösbaren (intractable) Probleme, zu denen nicht bewiesen ist, dass kein effizienter Algorithmus existiert.

Das „Integer Factorization Problem“ ist solch ein Beispiel, bei dem nicht bewiesen ist, dass es schwer lösbar ist. Dementgegen ist das „Traveling Salesman Problem“ ein Fall, bei dem unter der allgemein geläufigen Annahme $P \neq NP$ dessen schwere Lösbarkeit bzw. „intractability“ gilt. [1]

3. Intractable Problems

3.1 Beispiele

Folgend sind zwei bis heute ungelöste Probleme der Zahlentheorie vorgestellt.

- Starke Goldbachsche Vermutung

  „Jede gerade Zahl größer als 2 ist Summe zweier Primzahlen.“

Diese Vermutung stammt aus einem Brief von Christian Goldbach an Leonhard Euler aus dem Jahr 1742. [3] [4] Die Mathematiker sind zwar davon überzeugt, dass diese Vermutung gültig ist, doch bewiesen wurde es bis heute nicht vollständig. Einzig Tomas Oliveira e Silva zeigte die Gültigkeit dieser Vermutung für alle Zahlen bis $4 \cdot 10^{18}$.

Selbst ein im Jahr 2000 vom britischen Verlag Faber & Faber ausgelobtes Preisgeld von einer Million Dollar konnte dies nicht ändern, da das Preisgeld nicht ausgezahlt wurde, weil bis 2002 kein Beweis eingegangen war. [3]

- Primzahlzwillings-Vermutung

  „Es gibt unendlich viele Primzahlzwillinge.“

Primzahlzwilling meint ein Paar aus zwei Primzahlen, deren Abstand 2 beträgt. Beispiele für Primzahlzwillinge sind (3,5), (5,7) und (17,19). Es ist ungewiss und bisher nicht bewiesen, ob diese Vermutung Gültigkeit besitzt, denn je größer die betrachteten Zahlen sind, desto weniger Primzahlen werden gefunden.

Der bisher größte bekannte Primzahlzwilling ist

$$3756801695685 \cdot 2^{666669} \pm 1$$

Diese beiden Zahlen bestehen aus 200.700 Ziffern. [1]

3.2 Integer Factorization Problem

Zwei Zahlen miteinander zu multiplizieren ist einfach und wird einem Schüler bereits sehr früh in der Schulzeit gelehrt. Mit etwas Training und Routine gelingt dies selbst bei größeren Zahlen recht schnell und häufig nur durch Kopfrechnen.

Will man beispielsweise die Zahlen 29 und 103 miteinander multiplizieren, kommt man ohne größere Schwierigkeiten auf die Lösung.

$$29 \cdot 103 = 2987$$

Stellt man diese Frage allerdings genau andersrum, also in welche Faktoren die Zahl 2987 zerlegt werden kann, ist die Lösung nicht mehr so einfach und wird besonders bei immer größer werdenden Zahlen sehr komplex. Dieses Problem wird als Faktorisierungsproblem ganzer Zahlen bezeichnet.

Unter dem Faktorisierungsproblem ganzer Zahlen versteht man ganz allgemein die Aufgabe, zu einer natürlichen, zusammengesetzten Zahl einen nichttrivialen Teiler zu ermitteln. Nichttrivial bedeutet in diesem Fall, dass weder die Zahl selbst, noch die Zahl 1 als Teiler gültig sind. Algorithmen, die zur Lösung dieses Problems dienen, nennt man Faktorisierungsverfahren.

Genau genommen unterscheidet man zwischen der Faktorzerlegung und der Primfaktorzerlegung einer zusammengesetzten Zahl. In der Regel interessiert man sich allerdings für die Primfaktorzerlegung, die durch rekursive Anwendung von Faktorisierungsverfahren in Kombination mit Primzahltests berechnet werden kann. Viele Verfahren setzen bereits das Bekanntsein von Primzahlen in einem bestimmten Bereich voraus. Aus diesem Grund sind im Folgenden die Primfaktoren gemeint, wenn von Faktoren die Rede ist.

Wie für zahlentheoretische Probleme typisch, ist auch hier die Aufgabenstellung sehr einfach formuliert und für jeden verständlich, wohingegen die Lösung nicht so simpel ist, wie beim Multiplizieren zweier Zahlen. Obwohl diese Problemstellung, dessen Ursprung in der Antike liegt und erstmals populär wurde, nachdem durch Euklid von Alexandria 300 v. Chr. bekannt wurde, dass jede natürliche Zahl eine eindeutige Primfaktorzerlegung besitzt, zu den ältesten und bedeutungsvollsten Problemen der Zahlentheorie zählt, existiert selbst mithilfe der besten heutigen Computer aktuell kein Verfahren, dass beliebig große Zahlen in effizienter Zeit in seine Faktoren zerlegt. [1] [2] [5]

Von Carl Friedrich Gauss stammt folgendes Zitat, welches auch die heutige Lage zum Problem der Faktorisierung natürlicher Zahlen gut beschreibt.

> *Dass die Aufgabe, die Primzahlen von den zusammengesetzten zu Unterscheiden und letztere in ihre Primfaktoren zu zerlegen, zu den wichtigsten und nützlichsten der gesamten Arithmetik gehört und die Bemühungen und den Scharfsinn sowohl der alten wie auch der neuen Geometer in Anspruch genommen hat, ist so bekannt, dass es überflüssig wäre, hierüber viele Worte zu verlieren. Trotzdem muß man gestehen, daß alle bisher angewendeten Methoden entweder auf spezielle Fälle beschränkt oder so mühsam und weitläufig sind, daß sie auf größere Zahlen meistenteils kaum angewendet werden können. Außerdem aber dürfte es die Würde der Wissenschaft erheischen, alle Hilfsmittel zur Lösung jenes berühmten Problem fleißig zu vervollkommen.* [6]

Weitere bedeutende Mathematiker, die sich auf diesem Gebiet und vor allem mit der Primfaktorzerlegung ganzer Zahlen beschäftigt haben: [7] [4]

- Eratosthenes von Kyrene (276 – 194 v. Chr.)
- Leonardo Pisano Fibonacci (1170 – 1250)
- Pierre de Fermat (1170 – 1250)
- Leonhard Euler (1707 – 1783)
- Adrien-Marie Legendre (1752 – 1833)

Von aktuellem Interesse ist das Faktorisiserungsproblem ganzer Zahlen besonders bei dem kryptografischen Verfahren RSA, da dessen Sicherheit eben auf dem Problem und der Schwierigkeit der Faktorisierung großer Zahlen beruht. Das Finden eines effizienten Algorithmus zur Zerlegung großer Zahlen wäre also gleichbedeutend mit dem Brechen des RSA-Kryptosystems. Dadurch wird die Relevanz der Entwicklung schneller Faktorisierungsalgorithmen und der Forschung in diesem Gebiet deutlich. [1]

3.2.1 Mathematische Grundlagen

3.2.1.1 Sieb des Eratosthenes

Das Sieb des Eratosthenes ist nach seinem Erfinder Eratosthenes von Kyrene benannt und ein Verfahren, um Primzahlen in einem gewählten Intervall von 2 bis n zu ermitteln. Die Zahlen werden in aufsteigender Reihenfolge aufgeschrieben und beginnend mit der 2 als erste bekannte Primzahl markiert. Daraufhin werden alle Vielfachen der 2 in der Liste gestrichen. Dies wird mit jeder weiteren Primzahl bis $\sqrt{n}$ wiederholt. Die übriggebliebenen Zahlen sind die Primzahlen. Natürlich sind auch Intervalle möglich, die mit Zahlen ≥ 2 beginnen. [8]

Außerdem kann mit diesem Verfahren auch die Primfaktorzerlegung ermittelt werden. Dafür muss zusätzlich notiert werden, wann eine Zahl als Vielfaches welcher Primzahl gestrichen wird. Die Zahl 45 wird beispielsweise als Vielfaches der 3 und 5 gestrichen. Zusätzlich muss dann aber noch überprüft werden, ob eine der Primzahlen mehrfach die zu zerlegende Zahl teilt. In diesem Beispiel ist das die 3 , die die Zahl 45 genau zweimal teilt.

Das Ergebnis lautet dann $45 = 3^2 \cdot 5$.

Speziell für große Zahlen ist die Verwendung dieses Verfahrens zur Ermittlung der Primfaktorzerlegung ungeeignet, weshalb es als solches so gut wie keine Verwendung findet.

3.2.1.2 Der größte gemeinsame Teiler

Der größte gemeinsame Teiler (ggT) zweier natürlicher Zahlen u und v ist die größte natürliche Zahl, die sowohl u als auch v ohne Rest teilt.

Für die Erweiterung auf ganze Zahlen gilt, dass nur die Beträge der Zahlen betrachtet werden. So ist beispielsweise $ggt(3,-12) = ggT(3,|-12|) = ggT(3,12) = 3$.

Der ggT hat zwei wichtige Eigenschaften:

Für $v = 0$ gilt: $ggT(u,0) = u$
und für $u, v \neq 0$ gilt: $ggT(u,v) = ggT(u, v \bmod u)$

Diese beiden Eigenschaften werden beim Euklidischen Algorithmus genutzt. [8]

3.2.1.3 Euklidischer Algorithmus

Mit dem euklidischen Algorithmus lässt sich relativ schnell der größte gemeinsame Teiler zweier natürlicher Zahlen u und v berechnen. Dazu werden die beiden Zahlen u und v bis zum Erreichen des Restes 0 wie folgt rekursiv mit Rest geteilt:

$$
\begin{aligned}
u &= q_0 \cdot v + r_1 \\
v &= q_1 \cdot r_1 + r_2 \\
r_1 &= q_2 \cdot r_2 + r_3 \\
r_2 &= q_3 \cdot r_3 + r_4 \\
&\vdots \\
r_{k-2} &= q_{k-1} \cdot r_{k-1} + r_k \\
r_{k-1} &= q_k \cdot r_k + 0\ .
\end{aligned}
$$

Die Zahl r_k ist dann das gesuchte Ergebnis $ggT(u, v) = r_k$.

Ein einfaches Beispiel: $ggT(1071{,}1029)$

$$
\begin{aligned}
1071 &= 1 \cdot 1029 + 42 \\
1029 &= 24 \cdot 42 + 21 \\
42 &= 2 \cdot 21 + 0
\end{aligned}
$$

Der größte gemeinsame Teiler von 1071 und 1029 ist somit 21. [2]

3.2.1.4 Kongruenz

Mit zahlentheoretischer Kongruenz meint man, dass zwei Zahlen bei Division durch ein Modul in ihrem Rest übereinstimmen. Dies ist genau dann der Fall, wenn die beiden Zahlen sich um ein Vielfaches des Moduls unterscheiden. Stimmen die Zahlen in ihrem Rest nicht überein, nennt man sie inkongruent bezüglich des Moduls. Ein Beispiel soll zugleich die in dieser Arbeit verwendete Schreibweise verdeutlichen:

$$7 \equiv 13\ (mod\ 3)$$

„Die Zahlen 7 *und* 13 *sind kongruent modulo* 3*.“*

$7 \div 3 = 2$ **Rest 1** und $13 \div 3 = 4$ **Rest 1**

bzw.

$$13 - 7 = 6 = 2 \cdot 3$$

Die Theorie der Kongruenzen wurde ebenfalls von Carl Friedrich Gauß entwickelt und stammt, wie das erwähnte Zitat auf Seite 6, aus dem Werk „Disquisitiones Arithmeticae“. [2] [8]

3.2.1.5 Quadratische Reste

Eine Zahl a ist „quadratischer Rest“ modulo m , wenn eine Lösung der Kongruenz

$$x^2 \equiv a\ (mod\ m)\text{ , mit } a\ \in\ \mathbb{Z},\ m\ \in \mathbb{N} \text{ und } ggT(a,m) = 1$$

existiert. Existiert diese Lösung nicht, heißt a „nicht-quadratischer Rest“ modulo m.
Ein Beispiel für $m = 5$ soll diese Definition verdeutlichen:

$$1^2 \equiv a\ (mod\ 5) \rightarrow a = 1$$
$$2^2 \equiv a\ (mod\ 5) \rightarrow a = 4$$
$$3^2 \equiv a\ (mod\ 5) \rightarrow a = 4$$
$$4^2 \equiv a\ (mod\ 5) \rightarrow a = 1$$

Es sind 1 und 4 quadratische Reste modulo 5. Die Zahlen 2 und 3 kommen auf der rechten Seite nicht vor und sind nicht-quadratische Reste modulo 5.

Euler entwarf hierzu ein Theorem, welches die Frage, ob a ein quadratischer Rest ist, effizient beantworten kann. Demnach ist a genau dann ein quadratischer Rest modulo p (p sei eine ungerade Primzahl), wenn

$$a^{\frac{p-1}{2}} \equiv 1\ (mod\ p)\ .$$

In Worten, wenn die Berechnung von „ $a^{\frac{p-1}{2}}\ modulo\ p$ " die Lösung 1 ergibt. [2] [4]

3.2.1.6 Legendre Symbol

Aufbauend auf Eulers Theorem zum quadratischen Rest wurde das Legendre Symbol für ungerade Primzahlen definiert als: [2]

$$\left(\frac{a}{p}\right) = \begin{cases} 0, & \text{wenn a} \equiv 0\ (\text{mod p}) \\ 1, & \text{wenn a quadratischer Rest modulo p} \\ -1, & \text{wenn a nicht} - \text{quadratischer Rest modulo p} \end{cases}$$

Am Beispiel $m = 5$:

$\text{a} = 1 \rightarrow 1^{\frac{5-1}{2}} \equiv 1\ (\text{mod } 5)$ → quad. Rest modulo 5 → $\left(\frac{1}{5}\right) = 1$

$\text{a} = 2 \rightarrow 2^{\frac{5-1}{2}} \equiv 4\ (\text{mod } 5)$ → nicht-quad. Rest modulo 5 → $\left(\frac{2}{5}\right) = -1$

$\text{a} = 3 \rightarrow 3^{\frac{5-1}{2}} \equiv 4\ (\text{mod } 5)$ → nicht-quad. Rest modulo 5 → $\left(\frac{3}{5}\right) = -1$

$\text{a} = 4 \rightarrow 4^{\frac{5-1}{2}} \equiv 1\ (\text{mod } 5)$ → quad. Rest modulo 5 → $\left(\frac{4}{5}\right) = 1$

3.2.2 Faktorisierungsverfahren

In diesem Abschnitt geht es um Verfahren, die zur Zerlegung zusammengesetzter Zahlen in ihre Primfaktoren Anwendung finden. Um zu Beginn kleine Faktoren finden bzw. ausschließen zu können, nutzt man in der Praxis meist die Probedivision. Je nach Größe der zu faktorisierenden Zahl kommen danach weitere Verfahren mit ihren spezifischen Stärken zum Einsatz. So eignet sich die Faktorisierungsmethode von Fermat zum Beispiel besonders gut, um Teiler in der Nähe von $\sqrt{n}$ mit n als die zusammengesetzte Zahl zu finden.

Dagegen sind bei sehr großen Zahlen (bis 100-120 Dezimalstellen) schnelle Verfahren wie das Quadratische Sieb anderen überlegen. Auf die drei genannten Verfahren wird im Weiteren eingangen.

Folgende Liste gibt einen Überblick über weitere bekannte Faktorisierungsverfahren: [5]

- Faktorisierungsmethode von Lehman
- Kettenbruchmethode
- Methode der eliptischen Kurven
- Zahlkörpersieb
- Pollard-Rho-Methode
- Pollard p-1 Methode
- Pollard p+1 Methode

Zudem im Bereich der Quantencomputer:

- Shor-Algorithmus

3.2.2.1 Die Probedivision

Ältestes und intuitivstes Verfahren ist die Probedivision. Bei der Probedivision einer Zahl n wird zuerst eine Grenze B festgelegt, bis zu der dividiert wird. Als B wird in der Regel $\sqrt{n}$ gewählt. Beginnend bei 2 wird n nun der Reihe nach durch jede Primzahl geteilt. Wenn das Ergebnis eine ganze Zahl ist, ist ein Teiler gefunden. Wenn kein Teiler gefunden wurde, ist n eine Primzahl. Voraussetzung für dieses Verfahren ist das Bekanntsein aller Primzahlen bis zur festgelegten Grenze B.

Man kann schnell feststellen, dass dieses Verfahren bei großen Zahlen mit einem hohen nötigen Berechnungsaufwand und im Fall der Speicherung der getesteten Zahlen auch mit einem hohen Speicherplatzbedarf verbunden ist, da alle Primzahlen von 2 bis $\sqrt{n}$ getestet werden. In der Praxis wird es aber durchaus in Kombination mit weiteren Verfahren verwendet. Mittels Probedivision können kleine Faktoren ermittelt bzw. ausgeschlossen werden, wonach die restliche Faktorisierung durch ein anderes Verfahren fortgesetzt werden kann.

Die Probedivision als Algorithmus eignet sich nur für Zahlen bis $\approx 10^7$ und wird für größere Zahlen zu langsam. Sie benötigt im schlimmsten Fall $2\frac{\sqrt{n}}{\ln n}$ Divisionen.

Beispiel: $n = 7597 \rightarrow B = \sqrt{7597} = 87$

Zu prüfen sind also die Primzahlen 2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47, 53, 59, 61, 67, 71, 73, 79, 83. Erst bei der Zahl 71 findet man die Lösung, wodurch sich der zweite Faktor mit $7597 \div 71 = 107$ ergibt. In diesem Fall benötigte es 19 Divisionen, bevor ein Teiler gefunden werden konnte. Dies ist also zugleich ein Beispiel für den schlimmsten Fall, denn

$$2\frac{\sqrt{7597}}{\ln 7597} \approx 19{,}51$$

Ein gar besonderes Beispiel ist dies jedoch nicht, sondern eher erwartungsgemäß, da bei der Probedivsion die Laufzeiten des durchschnittlichen und schlechtesten Fall in der gleichen Größenordnung liegen. [8]

3.2.2.2 Faktorisierungsmethode von Fermat

Das Faktorisiserungsverfahren von Fermat basiert auf dem Finden von quadratischen Kongruenzen und der Ermittlung der in diesen vorhandenen Faktoren. Es ist Grundlage vieler anderer Faktorisierungsverfahen und dient auch dem im nächsten Kapitel vorgestellten Quadratischen Sieb als Basis. Wie bereits erwähnt eignet sich diese Methode besonders gut um Teiler in der Nähe von $\sqrt{n}$ einer zusammengesetzten Zahl n zu finden. Ziel des Verfahrens ist es, zu einer zusammengesetzten Zahl n Werte für x und y zu finden, so dass folgende Kongruenzen erfüllt sind:

$$\mathrm{x}^2 \equiv \mathrm{y}^2 \ (\mathrm{mod\ n}) \qquad (1)$$

und

$$\mathrm{x} \not\equiv \pm \mathrm{y} \ (\mathrm{mod\ n}) \qquad (2)$$

Wenn für zwei ganze Zahlen (1) gilt, so gilt nach Definition von ≡

$$n \mid x^2 - y^2$$

und damit nach der dritten binomischen Formel

$$n \mid (x + y)(x - y)\ .$$

Die Bedingung (2) lässt sich auch folgendermaßen formulieren:

$$n\ teilt\ weder\ (x + y)\ noch\ (x - y)\ .$$

Ausgehend von $\left|\sqrt{n}\right| + 1$ als erster Schritt werden systematisch zwei Quadratzahlen gesucht, die zueinander kongruent sind, deren Wurzeln allerdings inkongruent zueinander sind. Für den speziellen Fall, dass es sich bei n um eine Quadratzahl handelt, ist man bereits nach dem ersten Schritt, also nach dem Ziehen der Wurzel von n, fertig und hat einen Teiler gefunden. Ansonsten ist x_0 mit $\left|\sqrt{n}\right| + 1$ die Zahl, mit der die Suche beginnt. Hier erkennt man, weshalb das Faktorisierungsverfahren von Fermat gut geeignet ist für das Finden von Teilern in der Nähe von $\sqrt{n}$, denn es beginnt schlichtweg in diesem Bereich.

Nun wird $x_0^2 - n$ berechnet. Danach wird solange x_i^2 um 1 erhöht und anschließend $x_i^2 - n$ berechnet, bis mit $y^2 = x_i^2 - n$ eine Quadratzahl gefunden wurde.

Ist auf diese Weise eine Quadratzahl y^2gefunden, so gilt $x_i^2 - n = y^2 \rightarrow x_i^2 \equiv y^2$, womit die Kongruenz aus (1) erfüllt ist. [5]

Die Teiler von n ergeben sich dann aus $ggT(n, x + y)\ und\ ggT(n, x - y)$.

Als Beispiel sei die Zahl n = 2041 zu faktorisieren. Ausgehend von $x_0^2 = \left|\sqrt{2041}\right| + 1 = 46^2$ wird die Folge der Quadratzahlen x_i^2 und die Folge der Differenzen $x_i^2 - n$ gebildet.

x_i^2	46^2	47^2	48^2	49^2	50^2	51^2	…	85^2	…
$x_i^2 - n$	75	168	263	360	459	560	…	5184	…

In diesem Beispiel wird erst bei $85^2 - n = 5184 = 72^2$ eine Quadratzahl gefunden.

Nun können mittels $ggT(2041, 85 + 72) = 13$ und $ggT(2041{,}85 - 72) = 157$ die Teiler von $n = 2041$ ermittelt werden.

Bei sehr großen zusammengesetzten Zahlen n dauert dieses Verfahren möglicherweise sehr lang, weshalb andere Faktorisierungsverfahren, wie das Quadratische Sieb, die Methode von Fermat weiterentwickelt haben, um eine schnellere Lösungsfindung herbeizuführen.

3.2.2.3 Quadratisches Sieb

Das quadratische Sieb, abgekürzt als QS, ist einer der schnellsten Faktorisierungsalgorithmen für große Zahlen. Wie bereits erwähnt, bildet die Faktorisierungsmethode von Fermat die Grundlage für dieses Verfahren. Zudem sei erwähnt, dass das QS eine Verbesserung des Algorithmus von Dixon ist, welcher in dieser Arbeit nicht behandelt wird. [9] Für Zahlen mit 100-120 Dezimalstellen existiert kein schnelleres Verfahren zur Faktorisierung, als das des quadratischen Siebs. Bei noch größeren Zahlen ist das Zahlkörpersieb dem QS überlegen. [10] [11]

In diesem Kapitel wird das obige Zahlenbeispiel zur Faktorisierungsmethode von Fermat fortgesetzt, um dessen Verbesserung bzw. Weiterentwicklung zu verdeutlichen.

Ausgangspunkt ist folgende Tabelle:

x_i^2	46^2	47^2	48^2	49^2	50^2	51^2	…
$x_i^2 - n$	75	168	263	360	459	560	…

Zu diesem Zeitpunkt sind bereits fünf Quadratzahlen x_i^2 berechnet, aber noch keine Quadratzahl $y^2 = x_i^2 - n$ gefunden worden.

Da aber auch ein Produkt mehrerer Zahlen $x_i^2 - n$ eine Quadratzahl y^2 ergibt bzw. ergeben kann, gilt

$$\mathrm{y}^2 \equiv \left(\mathrm{x}_{\mathrm{i}_1}^2 - \mathrm{n}\right) \cdot \ldots \cdot \left(\mathrm{x}_{\mathrm{i}_\mathrm{k}}^2 - \mathrm{n}\right) \equiv \mathrm{x}_{\mathrm{i}_1}^2 \cdot \ldots \cdot \mathrm{x}_{\mathrm{i}_\mathrm{k}}^2 \equiv \left(\mathrm{x}_{\mathrm{i}_1} \cdot \ldots \cdot \mathrm{x}_{\mathrm{i}_\mathrm{k}}\right)^2 \equiv \mathrm{x}^2 \ (\mathrm{mod}\ \mathrm{n})$$

Sofern hierbei analog zu (2)

$$x \not\equiv \pm \mathrm{y} \ (\mathrm{mod}\ \mathrm{n})$$

gilt, ist eine Zerlegung von n in Faktoren gefunden.

Das bedeutet, dass eine Lösung bereits existieren kann, bevor eine einzelne Berechnung $x_i^2 - n$ eine Quadratzahl ergibt. Dies ist, wie zuvor angedeutet, genau das Ziel der Verfahren, die die Faktorisierungsmethode von Fermat verbessern.

Im Beispiel ist das Produkt folgender $x_{i_k}^2 - n$ eine Quadratzahl:

$$75 \cdot 168 \cdot 360 \cdot 560 = 50400^2 = y^2$$

Modulo $n = 2041$ ist diese Quadratzahl y^2 kongruent zum Produkt der entsprechenden $x_{i_k}^2$:

$$46^2 \cdot 47^2 \cdot 49^2 \cdot 51^2 = (46 \cdot 47 \cdot 49 \cdot 51)^2 = 5402838^2 = x^2$$

Somit ist die Kongruenzbedingung (1) erfüllt:

$$5402838^2 \equiv 50400^2 \ (mod\ 2041)$$

da

$$5402838^2 \equiv 794 \ (mod\ 2041)\ \textit{und}\ 50400^2 \equiv 794 \ (mod\ 2041)$$

Außerdem gilt die Nichtkongruenz (2):

$$5402838 \not\equiv \pm 50400 \ (mod\ 2041)$$

da

$$5402838 \ \equiv 311 \ (mod\ 2041)\ und\ 50400 \ \equiv 1416 \ (mod\ 2041)$$

Mit diesen Werten für x und y können nun wie zuvor mit $ggT(5402838 + 50400, 2041) = 157\ und\ ggT(5402838 - 50400, 2041) = 13$ die Faktoren von 2041 ermittelt werden.
[6] [7] [9]
Doch wie findet man ein Produkt mehrerer Zahlen $x_i^2 - n$, das eine Quadratzahl y^2 ergibt?

Dieses Problem ist es, bei dem das Verfahren des quadratischen Siebs ansetzt.

In der Primfaktorzerlegung einer Quadratzahl muss jeder Primfaktor einen geraden Exponenten haben. Gesucht ist daher, sofern eine Lösung existiert, eine Auswahl der Zahlen $x_i^2 - n$, bei der die jeweilige Primfaktorzerlegung zu einem Produkt mit ausschließlich geraden Exponenten kombiniert werden kann.

Daher wird jede der Zahlen $x_i^2 - n$ zunächst in ihre Primfaktoren zerlegt.

Im Beispiel sieht das wie folgt aus:

$$75 = 3 \cdot 5^2$$
$$168 = 2^3 \cdot 3 \cdot 7$$
$$360 = 2^3 \cdot 3^2 \cdot 5$$
$$560 = 2^4 \cdot 5 \cdot 7$$

Die auf der vorherigen Seite bereits gezeigte, geeignete Auswahl ist:

$$75 \cdot 168 \cdot 360 \cdot 560 = 2^{10} \cdot 3^4 \cdot 5^4 \cdot 7^4 = y^2$$

In der Primfaktorzerlegung des Produkts dieser vier Zahlen hat jeder Primfaktor einen geraden Exponenten.

Die Suche nach einer geeigneten Auswahl lässt sich wie folgt durchführen. Zunächst wird eine Faktorbasis a festgelegt. Dies sind diejenigen Primzahlen bis zu einer festgelegten Grenze B, die bei der Untersuchung der möglichen Kombinationen eine Rolle spielen sollen. Sinnvollerweise orientiert man sich dabei an dem Vorkommen der Primzahlen in der bereits durchgeführten Primfaktorzerlegung der Zahlen $x_i^2 - n$. Im obigen Beispiel tauchen Primzahlen bis zur 7 auf, weshalb es unsinnig ist, die Grenze bspw. mit $B = 100$ festzulegen. Für das Beispiel sei daher $B = 12$. So ist die Faktorbasis

$$a = 2, 3, 5, 7, 11.$$

Nun wird jedes Element der Faktorbasis auf quadratische Reste überprüft. Nach Eulers Theorem kommen nur jene Primzahlen p in Betracht, bei denen n quadratischer Rest modulo p ist. Dies wird mittels Legendre-Symbol berechnet:

$$„\ a^{\frac{p-1}{2}}\ modulo\ p\ "$$

Nur wenn das Ergebnis dieser Berechnung 1 ergibt, wird p Element des Siebs.

Die Berechnung ergibt:

$$p = 3 \rightarrow 2041^{\frac{3-1}{2}} \equiv 1\ (mod\ 3) \quad \rightarrow \quad \left(\frac{2041}{3}\right) = 1$$
$$p = 5 \rightarrow 2041^{\frac{5-1}{2}} \equiv 1\ (mod\ 5) \quad \rightarrow \quad \left(\frac{2041}{5}\right) = 1$$
$$p = 7 \rightarrow 2041^{\frac{7-1}{2}} \equiv 1\ (mod\ 7) \quad \rightarrow \quad \left(\frac{2041}{7}\right) = 1$$
$$p = 11 \rightarrow 2041^{\frac{11-1}{2}} \equiv 10\ (mod\ 11) \quad \rightarrow \quad \left(\frac{2041}{11}\right) = -1$$

Somit hat die Zahl $p = 11$ den Legendre-Wert -1 und wird aus der Faktorbasis gestrichen. Alle anderen Zahlen aus a werden in das Sieb S übernommen.
Die Zahl $p = 2$ geht per Definition mit in S ein.

$$S = 2, 3, 5, 7$$

Nun wird mit jedem Element S_i des Siebs die Folge der $x_i^2 - n$ Zahlen durchlaufen. Dabei wird, so oft es ganzzahlig geht, durch S_i geteilt.

Der Siebvorgang wird wie folgt durchgeführt. Ist unter den ersten p Zahlen $x_i^2 - n$ eine durch p teilbare Zahl, so sind auch alle in p-Schritten folgenden Zahlen durch p teilbar.Ebenso gilt, wenn eine Zahl $x_i^2 - n$ nicht durch p teilbar ist, so sind auch alle in p-Schritten folgenden Zahlen nicht durch p teilbar.

Formal lässt sich dieser Sachverhalt beschreiben mit

$$\text{Für alle } k \in \mathbb{Z} \text{ gilt:} \quad x_i^2 - n \equiv (x_i + k \cdot p)^2 - n\ (mod\ p).$$

Unter den ersten p Zahlen $x_i^2 - n$ können höchstens zwei durch p teilbare Zahlen sein, denn $x_i^2 - n$ ist ein Polynom zweiten Grades, welches höchsten zwei Nullstellen hat. Diese treten bei den ersten p aufeinander goldenen Werten x_i irgendwo auf.

Im Beispiel sind von den ersten 5 Zahlen die erste (75) und die vierte (360) durch 5 teilbar. Daher sind in der Folge auch die $6.$, $11.$, $16.$ usw. sowie die $9.$, $14.$, $19.$ usw. durch 5 teilbar.

	75	168	263	360	459	560
2	2^0	2^3	2^0	2^3	2^0	2^4
	75	21	263	45	459	35
3	3^1	3^1	3^0	3^2	3^3	3^0
	25	7	263	5	17	35
5	5^2	5^0	5^0	5^1	5^0	5^1
	1	7	263	1	17	7
7	7^0	7^1	7^0	7^0	7^0	7^1
	1	1	263	1	17	1

Diejenigen Elemente $x_i^2 - n$, die zum Schluss zu 1 geworden sind, lassen sich durch das Sieb S mit einem geeigneten Exponenten Vektor darstellen. Die Zahl 75 ergibt beispielsweise den Exponentenvektor

$\begin{bmatrix}0\\1\\2\\0\end{bmatrix}$, denn anhand der Tabelle kann man ablesen: 75 = $2^0 \cdot 3^1 \cdot 5^2 \cdot 7^0$.

Da es nicht auf die tatsächlichen Exponenten ankommt, sondern nur ob diese gerade oder ungerade sind, lässt sich mit den modulo 2 reduzierten Exponentenvektoren rechnen.

$\begin{bmatrix}0\\1\\0\\0\end{bmatrix}$ entspricht 75 $\begin{bmatrix}1\\1\\0\\1\end{bmatrix}$ entspricht 168 $\begin{bmatrix}1\\0\\1\\0\end{bmatrix}$ entspricht 360 $\begin{bmatrix}0\\0\\1\\1\end{bmatrix}$ entspricht 560

Unter diesen Vektoren wird nun eine Menge von linear abhängigen Vektoren gesucht, also eine Auswahl von gewissen Vektoren, deren Summe modulo 2 gerechnet den Nullvektor ergibt. Dieser Schritt entspricht der Lösung eines linearen Gleichungssystems und wird nicht nicht explizit vorgeführt.

$$\lambda_0 \begin{bmatrix}0\\1\\0\\0\end{bmatrix} + \lambda_1 \begin{bmatrix}1\\1\\0\\1\end{bmatrix} + \lambda_2 \begin{bmatrix}1\\0\\1\\0\end{bmatrix} + \lambda_3 \begin{bmatrix}0\\0\\1\\1\end{bmatrix} = \begin{bmatrix}0\\0\\0\\0\end{bmatrix}$$

Für dieses Beispiel ergibt sich $\lambda_0 = \lambda_1 = \lambda_2 = \lambda_3 = 1$ als Lösung. Es bilden also alle vier aufgeführten modulo 2 reduzierten Exponentenvektoren eine linear abhängige Menge. Wie bereits gezeigt, ergibt das zugehörige Produkt $75 \cdot 168 \cdot 360 \cdot 560$ eine Quadratzahl.

Das quadratische Sieb hat eine Laufzeit von $e^{\sqrt{\ln n \cdot \ln \ln n}}$. [10] [11] [12]

3.2.3 Zusammenfassung und Ausblick

Die Sicherheit des RSA-Kryptosystems basiert primär auf der Schwierigkeit, große zusammengesetzte Zahlen zu faktorisieren. Bei der Suche nach Faktoren einer zusammengesetzten Zahl ist einerseits die Zahl an sich und andererseits, meist von der Zahl abhängig, das eingesetzte Faktorisierungsverfahren entscheidend. Verschiedene Algorithmen haben in Bezug auf die zu faktorisierende Zahl ihre jeweiligen Stärken, weshalb sie fallspezifisch eingesetzt werden.

Bei einer zu untersuchenden Zahl sollte zunächst ein Primzahltest erfolgen, da solch ein Test mit den existierenden Algorithmen recht schnell erledigt ist und der Versuch, eine Zahl zu faktorisieren, unter Umständen eine exponentielle Laufzeit erfordert. [5]

Bei zusammengesetzten Zahlen bis zu einer maximalen Größe von $\approx 10^7$ sollte die Probedivision durchgeführt werden. Für Zahlen in der Größenordnung von 100-120 Dezimalstellen ist das quadratische Sieb das bisher schnellste Faktorisiserungsverfahren. Ist die untersuchende Zahl noch größer, hat sich das Zahlkörpersieb als bestes Verfahren durchgesetzt. Dazwischen, also zwischen dem Anwendungsbereich der Probedivision und den beiden Siebverfahren, nutzt man Verfahren wie den Pollard-Rho-Algorithmus ($\approx 10^4 - \approx 10^7$) oder die elliptischen Kurven (bis 40 Dezimalstellen).

Die Faktorisierung von Zahlen ist ein aktives Forschungsgebiet, bei dem davon auszugehen ist, dass es zukünftig Neues geben wird. Neben Verbesserungen an Algorithmen und der Technik ist speziell das Gebiet der Quantencomputer als spannendes Thema zu nennen. Mit diesem neuartigen Konzept und dem bereits erwähnten Shor-Algorithmus ist es theoretisch möglich, Zahlen in polynomieller Zeit zu faktorisieren. Praktisch ist der Aufwand für diese Technik jedoch sehr hoch und zurzeit ist es nicht möglich, einen Quantenrechner in geeigneter Größe zu realisieren. [13]

Literaturverzeichnis

[1] Yan, Song Y.**:** *Cryptanalytic Attacks on RSA,* S. 41-54, Springer, 2007.

[2] Bundschuh, Peter**:** *Einführung in die Zahlentheorie,* Springer, 2013.

[3] Fuss, Paul Heinrich (Hrsg.): *Correspondance mathématique et physique de quelques célèbres géomètres du XVIIIème siècle.* (Band 1), S. 125–129, 1843.

[4] Stewart, Ian: *Professor Stewarts mathematisches Sammelsurium,* Rowohlt Taschenbuch Verlag, 2011.

[5] Riesel, Hans**:** *Prime Numbers and Computer Methods for Factorization,* Birkhäuser, 1994.

[6] Gauss, Carl Friedrich**:** *Disquisitiones Arithmeticae,* Kessel, Norbert, 2009.

[7] Robertson, Edmund**:** *The MacTutor History of Mathematics*, http://www-history.mcs.st-andrews.ac.uk/index.html, 07.05.2016.

[8] Reiss, Kristina: *Basiswissen Zahlentheorie,* Springer, 2007.

[9] Dixon, John D: *Asymptotically Fast Factorization of Integers,* Mathematics of Computation, Volume 36, S. 255-260, American Mathematical Society, 1981.

[10] Pomerance, Carl: *The quadratic sieve factoring algorithm*, Advances in Cryptology, Volume 209, Lecture Notes in Computer Science, S. 169-182, Springer, 1985.

[11] Pomerance, Carl: *A Tale of Two Sieves,* Volume 43, Notices of the AMS, S. 1473-1485, 1996, http://www.ams.org/notices/199612, 21.05.2016.

[12] Lang, Hans Werner: *Quadratisches Sieb*, http://www.inf.fh-flensburg.de/lang/krypto/algo/quadraticsieve.htm, 21.05.2016.

[13] Shor, Peter W.: *Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer,* SIAM Journal on Computing, S. 1484-1509, 1997.