素因数分解問題評価報告書

２００４年1月２３日

日本電信電話株式会社

青木 和麻呂
植田　広樹
内山　成憲

素因数分解問題評価報告書

日本電信電話株式会社

2004年1月23日

1 前書き

RSA 暗号の安全性は素因数分解の困難性に基づいている。素因数分解の実験はいろいろと行われているが、単発的であり系統だったものがほとんどない。そのため実際に RSA 暗号に使われるような大きさの数を分解するために必要な時間やハードウェアの規模について予測することが困難であった。そこで本報告書ではひとつのハードウェア環境及び同一のソフトウェアで一連の数を同一の方法で素因数分解することによって各種のパラメータを決定することを目的とする。

なお本稿では、数体ふるい法の詳細については文献 [LL93, IK03] に譲ることとし、詳細には立ち入らない。

2 分解する対象

RSA Data Security 社が提供していた旧 RSA factoring challenge の問題から 100, 110, 120, 130, 140, 150 桁の数を扱うことする[1]。これらに 90 桁の数を一つ加えたものは次の通りである。

```
GNFS-90 =
  157902323883677749535156526668194527838287440076 61\
  084895594640602042317271452517048270676 7
RSA-100 =
  152260502792253336053561837813263742971806811496 13\
  806886579084945801229632589528976540003506920061 39
RSA-110 =
  357942341797258687749918078325684554030037780242 28\
  226193532908190484670252364677411513516111204504 06\
  0317568667
RSA-120 =
  227010481295437363334259960947493668895875336466 08\
  478003817325824700916267577973538979115157404916 67\
  47880487470296548479
RSA-130 =
  180708208868740480595165616440590556627810251676 94\
  013491701270214500566625402440483873411275908123 03\
  371781887966563182013214880557
RSA-140 =
  212902463182587575474978820162715174978067039632 77\
  216278233383215381949984056495911366573853021918 31\
  678310738799531723088956923087344193647 1
RSA-150 =
  155089812478348440509606754370011861770654545830 99\
  543065546694577431263270346346595436333502757772 90\
  253914539967874140270035016317721868408907959646 83
```

[1]ただし現 RSA 社は現在では別の系統の数を challenge 問題としている。

以下では、上記 90 桁から 150 桁の数を実際に分解することにより適切なパラメータを選択することを試みる。

3 使用するハードウェア

今回の実験では、全て次の構成の PC を用いた。

> Pentium 4 (Northwood), 2.53GHz, FSB 533MHz, Intel Desktop Boards D850EMV2, i850e chip set, 1GB RDRAM, PC800, FreeBSD 4.7-RELEASE-p13, Intel Compiler 7.1 (Build 20030617Z)

以下のデータで、ふるい処理は単純に台数による並列化が可能なので、時間は 1 台に積算した値を提示した。線形代数処理は 16 台の PC を 100baseT 全二重の switching HUB で接続を行なった場合の実測を実行時間とした。その他の処理については、全体に占める割合が少ないので、特に算出していない。

4 記号

n: 分解する数。`RSA-100` など。

$f(x)$: 数体を定義する多項式 ($\in \mathbf{Z}[x]$)

d: $\deg f$

M: $f(M) \equiv 0 \pmod{n} \quad (M \in \mathbf{Z})$

s: 傾斜 (skew)

`rp`: rational 側の factor base の上限

`rlp`: rational 側の large prime の上限

`ap`: algebraic 側の factor base の上限

`alp`: algebraic 側の large prime の上限

n_q: special-q の個数

`qs`: 最小の special-q (およそ $\pi(\texttt{ap}) = \pi(\texttt{qs}) + n_q$)

(a, b): line sieve におけるふるい領域 ($|a| \leq h_a$, $0 < b \leq h_b$)

(c, d): lattice sieve におけるふるい領域 ($|c| \leq \texttt{hc}$, $0 < d \leq \texttt{hd}$)

5 GNFS の多項式選択

ここでは `RSA-100` を例として多項式の作り方を具体的に述べる。他の数については、ほぼ同様の処理を行なった。多項式選択には、PC 1 台で分解対象が小さいものから大きいものにかけて数分から 1 日程度の時間をかけた。

今回用いた方法は、基本的に Montgomery-Murphy の方法 [M99] を用いるが細部についてはかなり省略している。

5.1 ノルムの大きさの判定基準

$$f(M) \equiv 0 \pmod{n}$$

となる d 次多項式 f と自然数 M が与えられたときこの多項式に対して領域

$$S = \{(a,b) | -h_a \leq a \leq h_a,\ 1 \leq b \leq h_b\}$$

における f の大きさ $\mathrm{SIZE}(f,S)$ を

$$\mathrm{SIZE}(f,S) = \log(h_b M) + \frac{1}{2}\log\left(\int_S |f(-a/b)(-b)^d|^2 dadb\right)$$

と定義する。これが小さいほど S において $|a+bM|$ も $|f(-a/b)(-b)^d|$ も小さい傾向があることを示すと考えられる。

5.2 最高次係数の動く範囲

$d = 5$ とし

$$f(x) = a_5x^5 + a_4x^4 + a_3x^3 + a_2x^2 + a_1x + a_0$$

とする。

$d = 5$ の場合は a_5 は 1 から $n^{1/6}$ までを動きうる。S をこれまでの実験結果から妥当と思われる値に定め、SIZE 評価によるいくつかの実験の結果 $n^{0.090}$ 付近に最適な値が多い傾向があることがわかった[2]。そこでこれを中心として前後を調べることにする。RSA-100 では $n^{0.085}$ から $n^{0.095}$ までを調べた。つまり a_5 の範囲は次の通りである。

$$(a_5^{\min} =)\ 269475891 \leq a_5 \leq 2644513508\ (= a_5^{\max})$$

さらに射影根が多くなるようにするため a_5 が小さな素数のべき乗で割れるようにしておく。RSA-100 では、多項式選択の時間[3] を考え

$$2^3 \cdot 3^2 \cdot 5 \cdot 7 = 2520\ (= a_5^{\mathrm{step}})$$

の倍数とする。したがって a_5 は

$$(a_5^{\max} - a_5^{\min})/a_5^{\mathrm{step}} = 942475$$

個を動くことになる。

5.3 ノルムの小さな多項式の選択

a_5 を決めても M には自由度があるので a_4 以下は変化する。その中で algebraic 側が最小になるものを探す。後に 1 次係数が M 程度の大きさの 2 次式を加える

[2] 4 次式では 0.127 近辺が良く、6 次式では 0.085 近辺が良いようである。ただし、理論的な裏づけがあるわけではないので一般的なことは何も言えない。

[3] 小さくとると、より多くの多項式を試せ、より良い多項式を生成できる可能性があるが、時間がかかる。

ことを考慮すると a_1 は M 程度の大きさであることが適当である。係数の傾斜 (skew) を s とすると

$$a_1 \approx a_5 s^4 \approx M$$

である。よって s は

$$s \approx (M/a_5)^{1/4}$$

を満たす。この s によっておおよそ

$$|a_4| \leq a_5 s$$

となることが要請される。

ここで a_4 の調整は M の平行移動によって行われる。M を $M+k$ に平行移動すると a_4 は $a_4 + 5ka_5$ に変わる。最初の a_4 は 0 に近いから $-s \leq 5k \leq s$ である。すなわち $-s/5 \leq k \leq s/5$ となるが余裕をみて

$$-s \leq k \leq s$$

とする。まとめると

$$M_0 = \mathrm{round}((n/a_5)^{1/5})$$

とし

$$M = M_0 + k, \quad -s \leq k \leq s$$

で M を動かすことにする。

そして各 M に対し n を M-進展開して a_4 以下を決める。この多項式 f に対して $\mathrm{SIZE}(f, S)$ を計算する。ただし領域 S は面積 $2h_a h_b$ を一定とし h_a/h_b を $\mathrm{SIZE}(f, S)$ が最小になるようにとるものとする。`RSA-100` の場合 S の面積は過去の結果から 2^{42} とした。

このように a_5, M を動かして $\mathrm{SIZE}(f, S)$ が小さなものをいくつか (10 個から 100 個) 探す。

`RSA-100` の場合 SIZE の一番小さな多項式は次のものであった。

$$\begin{aligned} f(x) = \quad & 2218671000 \quad x^5 \\ & +722975049906 \quad x^4 \\ & -113835927498852 \quad x^3 \\ & -13938430864062173 \quad x^2 \\ & +248440408936881498 \quad x \\ & +419147805253061235 \end{aligned}$$

5.4 Murphy の α 関数

α 関数は $(-b)^5 f(-a/b)$ が同じ大きさの平均的な数よりもどのくらい smooth になっているかを判定するものである。値が 0 ならば普通であり、小さい (つまり負の値) ほど smooth であることを示す。

次のようにして計算される。まず $(-b)^5 f(-a/b)$ が p の何乗で割れるかの平均 $v_p(f)$ は

計算法 1) p が g の判別式を割らなければ射影型を含めて $f(x) \equiv 0 \pmod{p}$ の異なる解の個数を q_p とすれば平均しておよそ

$$v_p(f) = q_p \frac{p}{p^2 - 1}$$

である。

計算法 2) p が g の判別式を割る場合は領域内のランダムな点 (a, b) を選んで実際に $(-b)^5 f(-a/b)$ の p べき指数を計算して平均を求める。RSA-100 の場合は 20000 個の点 (a, b) について計算した。

普通のランダムな数は素数 p で平均して

$$\mathrm{std}_p = \frac{1}{p - 1}$$

回割れる。この 2 つのべき指数の差の分、つまり $p^{\mathrm{std}_p - v_p(f)}$ だけ大きい約数を $(-b)^d f(-a/b)$ は普通のランダムな数よりも持つことになる。そこで log をとって和を作ったものを

$$\alpha(f) = \sum_p \log p \, (\mathrm{std}_p - v_p(f))$$

と定義する。

RSA-100 の計算では最初の 20 個の素数に対しては計算法 2 を用い、それ以上では計算法 1 を 1000 番目の素数まで用いて和をとった。

5.5 一次式を加えて α 関数が最小となる多項式を選択する

f, M が与えられたとき $f(x)$ に 2 次式 $(j_1 x + j_0)(x - M)$ を加えた多項式

$$g(x) = f(x) + (j_1 x + j_0)(x - M)$$

を作る。$g(M) \equiv 0 \pmod{n}$ でもあるので g, M も n に対する GNFS 用の多項式として使える。

ここで $(-b)^5 g(-a/b)$ が小さな素数 p で多くの約数を持つようにする。その判定基準は上記の α 関数である。

RSA-100 では $-20 \le j_1 \le 20$, $-2^{12} \le j_0 < 2^{12}$ の範囲で調べた。

$t \bmod p$ を固定するとき

$$f(t) + (j_1 t + j_0)(t - M) \equiv 0 \pmod{p}$$

となる j_1, j_0 は $\bmod p$ で定まるので j_1 を固定して j_0 の最小の値を決めればそこから step p での “ふるい” で調べることができる。具体的には $\log p$ を加えて行く。

RSA-100 では最初の 50 個の素数について “ふるい” を行った後で値が一番大きいところから j_1, j_0 を決定した。この g に対して $\alpha(g)$ を計算する。

5.6 ノルムを最小にする

g, M に対して M を微小変化 (± 200) させて SIZE(g, S) が最小となるようにする。変化があればこの g を新たな f として前節から繰り返す。変化がなければ g, M を出力して終了する。

RSA-100 の場合は 2 回の繰り返しで次の多項式を得た。

$$\begin{aligned} g(x) = \quad & 476148960 \quad x^5 \\ & +33466236556 \quad x^4 \\ & -95242541476020 \quad x^3 \\ & +24540020572973215 \quad x^2 \\ & -3475579183967599680 \quad x \\ & -2599927782355220688836 \end{aligned}$$

これは SIZE のみによる順位で第 5 位であった多項式に

$$-(x + 2112)(x - M)$$

を加え、M を 1261737131078349311 から 1261737131078349405 へ 94 だけ平行移動したものである。

6 GNFS の factor base およびそれに関連するパラメータの選択

今回のパラメータ選択において、ap については、qs = rp とし、およそ $\pi(\mathtt{ap}) = \pi(\mathtt{qs}) + n_q$ とした。また、ある special-q に対する algebraic 側の factor base の上限は、小さい方から i 番目の special-q を q_i とおいたとき、$q_{0.95i}$ とした。

6.1 方針

ここではすべてを lattice sieve で行うことにする。RSA-100, RSA-110 等では line sieve を併用した方が短時間で終えることができるがより大きな数に対しては line sieve はほとんど役に立たず、より大きな数に対する実行時間を類推するには小さな数に対しても lattice sieve のみでの実行時間を計測することが適切な結論を出すものと考えられる。以下、この節では lattice sieve のみを用いることとする。

6.2 パラメータのとり方

信頼できるデータを下にして、未知の部分は単純な補間による。

RSA-100 については先行して、かなり大規模にパラメータを変化させる実験から最適と思われるパラメータが得られているので、その結果

$\mathtt{rp} = 0.25\mathrm{e}6$, $\mathtt{ap} = 3.75\mathrm{e}6$, $n_q = 0.15\mathrm{e}6$, $\mathtt{rlp} = 17\mathrm{e}6$, $\mathtt{alp} = 38\mathrm{e}6$, $\mathtt{hc} = \mathtt{hd} = 2^{11}$

を基礎にする。もちろん今回生成の多項式とは異なるものを利用しており、ほかのパラメータも違うのでこれをそのまま使うことは必ずしも適当ではない。

RSA-140 CWI のレポート [CDL+99] によると

> rp = 3.5e6, ap = 12.17e6, n_q = (0.51e6?), rlp = 500e6, alp = 1e9, hc = hd = 2^{13}

である。ただし、lattice sieve で 55%の relation を、line sieve で 45%の relation を得ているので、lattice sieve のみにするには n_q を 1.0e6 にする必要がある。それには

> rp = 3.5e6, ap = 20.08e6, n_q = 1.0e6

とすれば良い。必要な relation の個数は 67e6 個である。

これを補間する。ただし安全を見てやや多めにする。

	rp
RSA-100	0.3e6
RSA-140	4.0e6

この間を単純に対数比例配分する (=1.91 倍/10 桁)

	rp
RSA-100	0.3e6
RSA-110	0.57e6
RSA-120	1.01e6
RSA-130	2.09e6
RSA-140	4.0e6

同じく rlp は 2.33 倍し、alp はその 2 倍。

	rp	rlp	alp
RSA-100	0.3e6	17e6	34e6
RSA-110	0.57e6	40e6	80e6
RSA-120	1.01e6	92e6	184e6
RSA-130	2.09e6	215e6	430e6
RSA-140	4.0e6	500e6	1e9

special-q の個数について

$$n_q = \text{必要な relation の個数}/\text{1 つの } q \text{ で得られる relation の個数}$$

であるが分母は hc, hd のとり方によっても変化するのでここまでのところでは決められない。

必要な relation の個数は RSA-140 では 67e6 個である。RSA-100 では記述がないが実験してみると 1 つの q で平均 23 個の relation がえられるので $23 \times 150\text{e}3 = 3.5\text{e}6$ 個とする。少し多めにとって対数比例配分する。

	rp	rlp	alp	relation	$\log_2$ hc	$\log_2$ hd	relation/1q	n_q
RSA-100	0.3e6	17e6	34e6	4.0e6	11	11	23	174e3
RSA-110	0.57e6	40e6	80e6	8.2e6	12?	11?		
RSA-120	1.01e6	92e6	184e6	16.7e6	12?	12?		
RSA-130	2.09e6	215e6	430e6	34.2e6	13?	12?		
RSA-140	4.0e6	500e6	1e9	70.0e6	13	13	67	1.04e6

6.3 RSA-150 の戦略

このまま RSA-150 の類推をしたいところだが 158 桁 (以下 GNFS-158 と書く) の分解 [BFK02] のパラメータがこれまでのものと著しく異なるので両者の折衷にならざるを得ない。実験で補正するしかないであろう。

	rp	rlp	alp	relation	$\log_2$ hc	$\log_2$ hd	relation/1q	n_q
RSA-150	12.0e6	2e9	2e9	140e6?	13	13	67	2.0e6?
GNFS-158	30.0e6	4e9	4e9	309e6	13	13	59	5.2e6(a)
GNFS-164	40.0e6	4e9	4e9	458e6	14	13	50	9.1e6(a)

注 (a) line sieve の分も lattice sieve で行った場合の補正を加えた。

なお、GNFS-164 の結果は [AUK+04] にある。

実験の結果 rp を補正する。RSA-140 の値はあくまで line sieve と lattice sieve を併用することを前提に決定されているものなので最適ではなくても不思議はない。RSA-130 までの実験結果によっては後で RSA-140 を修正する。

	rp	rlp	alp	relation	$\log_2$ hc	$\log_2$ hd	relation/1q	n_q
RSA-100	0.3e6	18e6	40e6	4.0e6	11	11	23	174e3
RSA-110	0.8e6	40e6	80e6	8.2e6	12?	11?		
RSA-120	1.6e6	92e6	184e6	16.7e6	12	11	25	668e3
RSA-130	3.0e6	215e6	430e6	34.2e6	13?	12?		
RSA-140	4.0e6	500e6	1e9	70.0e6	13	13	67	1.04e6

6.4 RSA-130 の hc, hd の決定

hc, hd を大きくするとひとつの special-q においてふるいの時間は増えるが、得られる relation の個数も増える。この最適値を求めたい。なお、この節での計算機実験は Pentium 4 [3.0GHz] での実行時間であることに注意せよ。

6.4.1 hc $= 2^{12}$, hd $= 2^{11}$ の場合

最初の 100 個の q での効率は

$$2202 \text{ relation}/96 \text{ 秒} = 22.94$$

と高い。しかし得られた個数が少ないので 34.2e6/22 = 1.55e6 個の q を用いることになる。

1550001 番目の q から 100 個の q に対する効率は $2233/171 = 13.06$ とかなり下がる。

中間点 (750001 ∼ 750100 番目の q) では 133 秒で 2285 個の relation が得られているので、これらを平均して

$$(2202 + 2285 + 2233)/(96 + 133 + 171) = 16.8$$

が全体の効率である。

6.4.2 $\texttt{hc} = 2^{12}$, $\texttt{hd} = 2^{12}$ の場合

それぞれ最初、560001 番目、1120001 番目から 100 個の q では

	生成 relation 数	実行時間 [秒]
1	3046	159
560001	3492	195
1120001	3471	223

となるので効率は $(3046 + 3492 + 3471)/(159 + 195 + 223) = 17.35$ となる。

6.4.3 $\texttt{hc} = 2^{13}$, $\texttt{hd} = 2^{12}$ の場合

	生成 relation 数	実行時間 [秒]
1	4617	281
371701	5136	317
743401	5343	344

となるので効率は $(4617 + 5136 + 5343)/(281 + 317 + 344) = 16.03$ である。

結論として $\texttt{hc} = \texttt{hd} = 2^{12}$ を採用する。

	rp	rlp	alp	relation	$\log_2 \texttt{hc}$	$\log_2 \texttt{hd}$	relation/1q	n_q
RSA-100	0.3e6	18e6	40e6	4.0e6	11	11	23	174e3
RSA-110	0.8e6	40e6	80e6	8.2e6	12	11	25	330e3
RSA-120	1.6e6	92e6	184e6	16.7e6	12	11	25	668e3
RSA-130	3.0e6	215e6	430e6	34.2e6	12	12	33	1.03e6
RSA-140	4.0e6	500e6	1e9	70.0e6	13	13	67	1.04e6

7 実験データの報告

対象となる数に対しては 5 次式がもっとも有効であると考えられる。この実験で最適な値を絞り込むことにする。また、対象とした数に対しては 4 次式はまだ有効であるが大きくなるにつれて 5 次式との差が大きくなるはずである。実験によってこのことを確認する。さらに、対象とした数に対しては 6 次式は有効ではない。しかし数が大きくなるにつれてその効率は 5 次式に近づいて行くはずである。実験によってこのことを確認する。但し、評価期間の制約から 4 次式と 6 次式は 90, 100, 110, 120 桁のみの実験となった。

7.1 多項式

以下に、実験に用いた多項式 gnfs90 から rsa150 およびそれに対応する M を示す。5 次多項式で RSA-100 及び RSA-110 については、最初に生成した多項式での実験結果から、明らかに性能が悪そうなので生成し直した。生成し直した多項式を、それぞれ rsa100f、rsa110g と表記する。なお、rsa140 は今回新たに生成することはなく文献 [CDL+99] の多項式をそのまま使った。

7.1.1 GNFS-90

$$
\begin{aligned}
\texttt{gnfs90d4} = & \\
& 230700960 \quad x^5 \\
& -1126041076 \quad x^4 \\
& -2800530116008 \quad x^3 \\
& +361291922417541 \quad x^2 \\
& -19404936952585372 \quad x \\
& -8090459983574955158 \\
M = & \ 14691556985141025
\end{aligned}
$$

$$
\begin{aligned}
\texttt{gnfs90d4} = & \\
& 384095471760 \quad x^4 \\
& -241081261647426 \quad x^3 \\
& -4687866717940381 \quad x^2 \\
& +268017730716567141252 \quad x \\
& -9349131240830077277886 5 \\
M = & \ 25321377834100183152
\end{aligned}
\qquad
\begin{aligned}
\texttt{gnfs90d6} = & \\
& 34739040 \quad x^6 \\
& -700184704 \quad x^5 \\
& -40151123010 \quad x^4 \\
& -333940064067 \quad x^3 \\
& +8190990388272 \quad x^2 \\
& +241490187216714 \quad x \\
& -581408893208572 \\
M = & \ 40700066437843
\end{aligned}
$$

7.1.2 RSA-100

$$
\begin{aligned}
\texttt{rsa100} = & \\
& 162177120 \quad x^5 \\
& +28143053314 \quad x^4 \\
& +23324146463349 \quad x^3 \\
& -37091958614999369 \quad x^2 \\
& -38913224302195502499 \quad x \\
& +6320108185403991362425 \\
M = & \ 1565018714315923398
\end{aligned}
\qquad
\begin{aligned}
\texttt{rsa100f} = & \\
& 476148960 \quad x^5 \\
& +33466236556 \quad x^4 \\
& -95242541476020 \quad x^3 \\
& +24540020572973215 \quad x^2 \\
& -34755791839675996 80 \quad x \\
& -2599927782355220688836 \\
M = & \ 1261737131078349405
\end{aligned}
$$

$$\texttt{rsa100d4} = \begin{aligned} &11280637368000 && x^4 \\ &+ 1297331842676166 && x^3 \\ &- 6398231616999999973 && x^2 \\ &- 3867985592992020706164 && x \\ &- 502980915481497659791952 9 \\ &M = 3408500839386006478662 \end{aligned}$$

$$\texttt{rsa100d6} = \begin{aligned} &54885600 && x^6 \\ &- 6171978062 && x^5 \\ &- 279021830192 && x^4 \\ &+ 7107344769942 && x^3 \\ &+ 230170783897105 && x^2 \\ &+ 12995121899819102 && x \\ &- 1211347205408698220 \\ &M = 1739888725534017 \end{aligned}$$

7.1.3 RSA-110

$$\texttt{rsa110} = \begin{aligned} &8806117320 && x^5 \\ &- 6602302439733 && x^4 \\ &- 10418783160743424 && x^3 \\ &+ 2509093800839550332 && x^2 \\ &+ 2444993039496642448884 && x \\ &- 484899833751348621692744 \\ &M = 83522930105029420917 \end{aligned}$$

$$\texttt{rsa110g} = \begin{aligned} &2186636760 && x^5 \\ &+ 7090231275050 && x^4 \\ &- 7779420006796361 && x^3 \\ &- 7338302252559380692 && x^2 \\ &+ 2945060505193947891936 && x \\ &+ 5283706951828717562159 92 \\ &M = 110358880444076439675 \end{aligned}$$

$$\texttt{rsa110d4} = \begin{aligned} &63063320720400 && x^4 \\ &- 108858128876245362 && x^3 \\ &- 333714480816166136741 && x^2 \\ &+ 2151401060734002095936690 && x \\ &- 1855021161851883623239525160 \\ &M = 867978684105357920487813 \end{aligned}$$

$$\texttt{rsa110d6} = \begin{aligned} &3356406900 && x^6 \\ &- 129410945870 && x^5 \\ &+ 8416221526418 && x^4 \\ &+ 87105425158528 && x^3 \\ &- 5730778027746978 && x^2 \\ &+ 19623529153740031 && x \\ &+ 1160691112322070568 \\ &M = 46916226934871871 \end{aligned}$$

7.1.4 RSA-120

$$\texttt{rsa120} = \begin{aligned} &1554355923120 && x^5 \\ &+ 279300728665360 && x^4 \\ &+ 235068853764040024 && x^3 \\ &- 366715788836593094 && x^2 \\ &- 19671660878164914170531 && x \\ &- 69422097616325345011725 99 \\ &M = 2709561965307563001574 \end{aligned}$$

$$\texttt{rsa120d4} =$$
$$\begin{aligned}
&895498088284800\,x^4\\
&+3413868460618755900\,x^3\\
&-24034697186813319912433\,x^2\\
&-28465301846766247738030655 2\,x\\
&-37422215960189606744506740 1260\\
M &= 126181391268425470300909327
\end{aligned}$$

$$\texttt{rsa120d6} =$$
$$\begin{aligned}
&6201111840\,x^6\\
&+65538566248\,x^5\\
&+69773963135376\,x^4\\
&-578627837775246\,x^3\\
&-345217686788156398\,x^2\\
&+1982291106588495261\,x\\
&-73006857823750108920\\
M &= 1822200063999191027
\end{aligned}$$

7.1.5 RSA-130

$$\begin{aligned}
\texttt{rsa130} = &\\
&147039132240\,x^5\\
&+871623037904469\,x^4\\
&+3086117472198489675\,x^3\\
&-7719799519497061434782\,x^2\\
&-9743342795049257456352467\,x\\
&+11536315812200841021988194190\\
M &= 414867094746941900457767
\end{aligned}$$

7.1.6 RSA-140

$$\begin{aligned}
\texttt{rsa140} = &\\
&439682082840\,x^5\\
&+390315678538960\,x^4\\
&-7387325293892994572\,x^3\\
&-1902715324374298871 4824\,x^2\\
&-6344102569446461791 3930613\,x\\
&+31855391707147435039 2223507494\\
M &= 34435657809242536951779007
\end{aligned}$$

7.1.7 RSA-150

$$\begin{aligned}
\texttt{rsa150} = &\\
&39579179880240\,x^5\\
&+118091572936301268\,x^4\\
&+99882037492763164770\,x^3\\
&-56447112339915941335654 51\,x^2\\
&-17749003945730989474189 029270\,x\\
&+14495606942348552079748 145328451\\
M &= 13140845099321381544918 13836
\end{aligned}$$

7.2 factor base のパラメータ

今回利用した factor base に関するパラメータを表 1 に示す。RSA-140, RSA-150 は line sieve を併用した実験も行なったので、それを表 2 に示す。

表 1: factor base のパラメータ

	rp	rlp	ap	alp	s	qs	備考
gnfs90	300e3	9e6	833669	20e6	192	206951	
gnfs90d4	800e3	10e6	823003	19e6	716	330859	
gnfs90d6	150e3	6e6	2142353	80e6	20	836161	
rsa100	300e3	18e6	2746739	40e6	591	468109	
rsa100f	300e3	18e6	1981711	40e6	369	424247	
rsa100d4	1200e3	18e6	12474757	40e6	1200	282797	
rsa100d6	300e3	12e6	4992019	160e6	53	1422221	
rsa110	800e3	40e6	6243553	80e6	600	1133857	
rsa110g	800e3	40e6	3995743	80e6	848	1069603	
rsa110d4	2400e3	40e6	3773773	80e6	2812	589759	
rsa110d6	600e3	20e6	8072513	320e6	23	1496549	
rsa120	1600e3	92e6	12474757	184e6	321	2264707	
rsa120d4	4800e3	100e6	10322033	140e6	5886	1363513	
rsa120d6	1200e3	48e6	13379207	640e6	58	2060059	
rsa130	3e6	215e6	20319361	430e6	1882	4107643	
rsa140	4e6	1e9	20761397	1e9	3992	4505773	line sieve 併用
	6e6	500e6	22960397	1e9	3992	4505773	
			19153333				空欄は上行と同一
rsa150	12e6	2e9	49312987	2e9	4049	14830997	line sieve 併用
			52864993				空欄は上行と同一

表 2: line sieve の factor base のパラメータ

	rp	rlp	ap	alp	h_a	h_b
rsa140	8e6	1e9	16777216	1e9	335544320	32768
rsa150	12e6	2e9	60e6	2e9	335544320	80000

さらに得られた様々なデータを次に示す。

	hc	hd	q_end	line で得た rel 数	lattice で得た rel 数	total rel 数
gnfs90	2 048	1 024	48 000	0	1 422 465	1 422 465
gnfs90d4	2 048	1 024	37 000	0	1 422 612	1 422 612
gnfs90d6	2 048	2 048	92 000	0	2 551 165	2 551 165
rsa100	2 048	2 048	160 000	0	3 879 533	3 879 533
rsa100f	2 048	2 048	112 000	0	3 053 728	3 053 728
rsa100d4	2 048	2 048	95 000	0	2 736 929	2 736 929
rsa100d6	2 048	2 048	240 000	0	5 329 204	5 329 204
rsa110	2 048	2 048	340 000	0	7 425 275	7 425 275
rsa110g	2 048	2 048	200 000	0	6 752 059	6 752 059
rsa110d4	4 096	2 048	220 000	0	6 185 022	6 185 022
rsa110d6	4 096	2 048	430 000	0	10 228 570	10 228 570
rsa120	4 096	2 048	650 000	0	14 225 875	14 225 875
rsa120d4	4 096	4 096	580 000	0	13 145 521	13 145 521
rsa120d6	4 096	4 096	720 000	0	19 304 631	19 304 631
rsa130	4 096	4 096	1 000 000	0	26 975 303	26 975 303
rsa140	8 192	8 192	1 000 000	10 168 552	69 296 581	79 465 133
rsa140g	8 192	8 192	1 130 000	0	65 393 900	65 393 900
rsa140g	8 192	8 192	904 000	0	51 340 137	51 340 137
rsa150	8 192	8 192	2 000 000	20 159 340	113 240 317	133 399 657
rsa150	8 192	8 192	2 200 000	0	124 804 557	124 804 557

	dup rel 数	残 rel 数	lost rel 数	line 時間 (s)	lattice 時間 (s)	total 時間 (s)
gnfs90	186 588	1 235 877	0	0	14 350	14 350
gnfs90d4	136 101	1 286 507	0	0	13 570	13 570
gnfs90d6	228 736	2 322 429	0	0	46 621	46 621
rsa100	587 798	3 291 735	0	0	79 631	79 631
rsa100f	391 235	2 662 493	0	0	54 271	54 271
rsa100d4	357 714	2 379 215	0	0	55 155	55 155
rsa100d6	584 401	4 744 803	0	0	145 247	145 247
rsa110	1 271 941	6 153 328	0	0	328 949	328 949
rsa110g	948 123	5 803 935	0	0	193 676	193 676
rsa110d4	956 470	5 228 552	0	0	250 945	250 945
rsa110d6	1 273 191	8 955 379	0	0	461 685	461 685
rsa120	2 389 912	11 835 963	0	0	790 138	790 138
rsa120d4	2 088 586	11 056 935	0	0	1 320 933	1 320 933
rsa120d6	2 370 817	16 933 814	0	0	1 517 888	1 517 888
rsa130	3 407 897	23 567 406	0	0	2 315 347	2 315 347
rsa140	12 378 307	67 086 826	0	838 452	7 889 757	8 728 209
rsa140g	7 186 499	58 207 401	0	0	8 577 408	8 577 408
rsa140g	4 982 485	46 357 652	0	0	6 726 258	6 726 258
rsa150	19 986 659	113 412 693	305	2 128 089	18 642 834	20 770 923
rsa150	12 666 924	112 137 633	305	0	20 597 260	20 597 260

	free rel 数	sc 前 rel 数	sc 前 FB 数	alprime0 要素数	rlprime0 要素数
gnfs90	2 363	1 238 240	1 325 906	844 320	481 593
gnfs90d4	13 750	1 300 257	1 368 708	818 391	550 316
gnfs90d6	540	2 322 969	2 545 100	2 146 731	398 905
rsa100	6 411	3 298 146	2 846 258	1 839 545	1 006 709
rsa100f	5 273	2 667 766	2 664 811	1 713 274	951 535
rsa100d4	24 960	2 404 175	2 557 942	1 580 867	977 074
rsa100d6	954	4 745 757	5 233 152	4 468 920	765 319
rsa110	12 286	6 165 614	5 518 625	3 451 808	2 066 813
rsa110g	11 442	5 815 377	5 443 581	3 407 556	2 036 022
rsa110d4	55 263	5 283 815	5 365 601	3 281 038	2 084 611
rsa110d6	1 574	8 956 953	9 784 707	8 534 027	1 251 490
rsa120	24 169	11 860 132	11 738 886	7 463 338	4 275 555
rsa120d4	126 435	11 183 370	10 928 533	6 178 137	4 750 423
rsa120d6	3 524	16 937 338	18 512 122	15 707 493	2 805 427
rsa130	47 433	23 614 839	24 521 904	15 300 885	9 221 041
rsa140	123 809	67 210 635	68 974 717	36 705 191	32 269 518
rsa140g	107 386	58 314 787	55 664 003	34 607 162	21 056 835
rsa140g	85 613	46 443 265	50 556 075	31 063 425	19 492 644
rsa150	210 243	113 622 936	126 224 606	67 309 955	58 914 643
rsa150	207 286	112 344 614	125 001 514	66 430 003	58 571 503

	1-pass 後 rel 数	1-pass 後 FB 数	sc 後 rel 数	sc 後 FB 数	sc 後 total w	sc 後 ave. w
gnfs90	798 835	791 694	331 876	330 872	5 623 433	16.944
gnfs90d4	875 758	859 333	374 083	373 082	6 482 139	17.328
gnfs90d6	1 243 916	1 232 301	515 742	514 740	8 800 410	17.064
rsa100	2 519 116	1 971 985	424 478	423 480	7 356 036	17.330
rsa100f	1 850 290	1 704 727	574 926	573 925	10 087 790	17.546
rsa100d4	1 594 376	1 585 057	724 636	723 634	12 848 598	17.731
rsa100d6	2 552 375	2 545 446	1 092 722	1 091 719	18 881 757	17.280
rsa110	4 606 507	3 731 140	893 009	892 011	16 142 280	18.076
rsa110g	4 207 271	3 581 036	891 802	890 801	16 518 559	18.527
rsa110d4	3 688 857	3 481 925	1 271 853	1 270 846	23 381 040	18.383
rsa110d6	4 776 194	4 729 413	1 884 183	1 883 182	34 029 045	18.060
rsa120	8 302 289	7 540 430	2 279 188	2 278 185	43 910 886	19.266
rsa120d4	8 207 377	7 471 071	2 455 106	2 454 102	45 839 280	18.671
rsa120d6	9 089 129	9 053 515	3 530 459	3 529 458	64 431 664	18.250
rsa130	15 596 081	14 838 636	4 563 119	4 562 121	87 722 757	19.224
rsa140	43 191 800	39 443 255	6 116 250	6 115 247	125 097 698	20.453
rsa140g	40 267 856	34 396 120	5 941 114	5 940 116	121 776 953	20.497
rsa140g	28 512 843	28 193 867	7 789 491	7 788 488	158 573 057	20.357
rsa150	67 846 880	67 813 102	16 593 213	16 592 212	335 877 843	20.241
rsa150	66 795 889	66 782 788	16 361 279	16 360 278	332 343 491	20.312

	lprime	alprime	rlprime	cprime	20way merge 後 rel 数	20way merge 後 FB 数	20way merge 後 total w
	要素数	要素数	要素数	要素数			
gnfs90	5	203 034	127 835	64	105 898	104 878	7 872 377
gnfs90d4	11	207 051	166 022	64	119 970	118 956	9 246 510
gnfs90d6	5	367 820	146 917	64	174 649	173 642	13 696 162
rsa100	6	272 978	150 496	64	163 437	162 435	13 487 097
rsa100f	5	365 142	208 778	64	190 616	189 614	16 664 599
rsa100d4	8	403 691	319 937	64	215 271	214 257	20 079 506
rsa100d6	5	789 496	302 222	64	336 137	335 114	33 891 248
rsa110	7	571 425	320 579	64	318 428	317 427	34 100 158
rsa110g	7	565 076	325 720	64	304 917	303 916	31 889 185
rsa110d4	9	714 237	556 602	64	384 185	383 169	42 320 324
rsa110d6	7	1 366 085	517 090	64	528 397	527 389	67 209 440
rsa120	9	1 460 541	817 635	64	687 247	686 239	101 496 903
rsa120d4	8	1 375 986	1 078 108	64	741 068	740 055	103 481 516
rsa120d6	6	2 509 741	1 019 711	64	904 269	903 253	139 815 927
rsa130	8	2 868 381	1 693 732	64	1 249 886	1 248 880	209 107 002
rsa140	8	3 782 325	2 332 914	64	1 611 769	1 610 763	294 459 963
rsa140g	8	3 693 767	2 246 341	64	1 651 577	1 650 572	293 205 775
rsa140g	8	4 734 617	3 053 863	64	1 842 573	1 841 554	348 263 555
rsa150	8	10 062 946	6 529 261	64	4 009 552	4 008 542	761 332 214
rsa150	8	9 898 427	6 461 843	64	4 061 097	4 060 083	749 694 248

	20way merge 後 ave. w	cut96 後 rel 数	cut96 後 FB 数	cut96 後 total w	cut96 後 ave. w
gnfs90	74.339	105 898	104 782	6 063 318	57.26
gnfs90d4	77.074	119 970	118 860	7 110 371	59.27
gnfs90d6	78.421	174 649	173 546	10 716 688	61.36
rsa100	82.522	163 437	162 339	10 648 753	65.16
rsa100f	87.425	190 616	189 518	13 181 907	69.15
rsa100d4	93.275	215 271	214 161	16 089 914	74.74
rsa100d6	100.826	336 137	335 018	27 513 589	81.85
rsa110	107.089	318 428	317 331	27 768 070	87.20
rsa110g	104.583	304 917	303 820	25 682 013	84.23
rsa110d4	110.156	384 185	383 073	34 490 578	89.78
rsa110d6	127.195	528 397	527 293	55 904 354	105.80
rsa120	147.686	687 247	686 143	84 825 348	123.43
rsa120d4	139.638	741 068	739 959	87 282 307	117.78
rsa120d6	154.6168	904 269	903 157	120 238 399	132.97
rsa130	167.301	1 249 886	1 248 784	178 678 686	142.96
rsa140	182.694	1 611 769	1 610 667	253 592 913	157.34
rsa140g	177.630	1 650 654	1 649 553	251 695 742	152.48
rsa140g	189.108	1 841 611	1 840 496	302 854 189	164.45
rsa150	189.880	4 009 552	4 008 446	665 517 429	165.98
rsa150	184.604	4 061 097	4 059 987	653 770 042	160.98

block Lanczos（block 長 128）の処理時間（16 台での並列処理）

```
rsa100      00:07:19
rsa110      00:28:49
rsa120      02:24:15
rsa130      08:47:28
rsa140      15:07:39
rsa150     101:31:17
```

表中の記号は次の通りである。

q_end n_q

dup rel 数 一意性処理で捨てた relation 数

rel 残数 total rel 数から dup rel 数を減じたもの

lost rel 数 本来あってはならないものだが、作業中に disk full などの理由で失ってしまった relation 数

sc 前 rel 数 一意性処理後の relation に、free relation を加えた relation の数

sc 前 FB 数 一意性処理後の relation で使われている large prime を含む factor base の数

alprime0 要素数 sc 前 FB 数のうち、algebraic 側で使われている factor base の数

rlprime0 要素数 sc 前 FB 数のうち、rational 側で使われている factor base の数

1-pass 後 rel 数 singleton 処理で、singleton を 1 回落したあとの relation 数

1-pass 後 FB 数 singleton 処理で、singleton を 1 回落したあとの relation で使われている factor base 数

sc 後 rel 数 singleton 及び clique 処理により、残った relation 数

sc 後 FB 数 singleton 及び clique 処理後に残った relation で使われている factor base 数

sc 後 total w singleton 及び clique 処理後に残った relation で行列を構成した際の weight

sc 後 ave. w singleton 及び clique 処理後に残った relation で構成された行列の 1 relation 辺りの weight

lprime 要素数 数体定義多項式の最高次係数の約数の個数 (factor base に含まれる)

alprime 要素数 sc 後 FB 数のうち、algebraic 側で使われている factor base の数

rlprime 要素数 sc 後 FB 数のうち、rational 側で使われている factor base の数

cprime 要素数 線形代数処理後の二次指標の調整で使うために algebraic 側の factor base から抜き出した素イデアルの個数で、常に 64 固定

`20way merge` 後 `rel` 数 20-way merge を実施した後に残っている relation-set 数

`20way merge` 後 `FB` 数 20-way merge を実施した後に残っている relation-set で使われている factor base の個数

`20way merge` 後 `total w` 20-way merge を実施した後に構成した行列の weight

`20way merge` 後 `ave. w` 20-way merge を実施した後に構成される行列の 1 relation-set 辺りの weight

`cut96` 後 `rel` 数 20-way merge を実施した後に構成される行列から weight の重い factor base を 96 個除いた後の relation 数であり、`20way merge` 後 `rel` 数に一致

`cut96` 後 `total w` 20-way merge を実施した後に構成される行列から weight の重い factor base を 96 個除いた後の合計 weight 数

`cut96` 後 `ave. w` 20-way merge を実施した後に構成される行列から weight の重い factor base を 96 個除いた後の 1 relation-set 辺りの weight

参考文献

[AUK+04] K. Aoki, H. Ueda, Y. Kida, T. Shimoyama, and Y. Sonoda. A trial of GNFS implementation (Part I) — Summary. In *2004 Symposium on Cryptography and Information Security*, number 2B1-3 in SCIS'04. Technical Group on Information Security (IEICE), 2004. (in Japanese).

[BFK02] F. Bahr, J. Franke, and T. Kleinjung. Factorization of 158-digit cofactor of $2^{953} + 1$. (available at `http://www.crypto-world.com/announcements/c158.txt`), 2002.

[CDL+99] S. Cavallar, B. Dodson, A. K. Lenstra, P. Leyland, W. Lioen, P. L. Montgomery, B. Murphy, H. te Riele, and P. Zimmermann. Factorization of RSA-140 Using the Number Field Sieve. In K. Y. Lam, E. Okamoto, and C. Xing, editors, *Advances in Cryptology — ASIACRYPT'99*, Volume 1716 of *Lecture Notes in Computer Science*, pp. 195–207. Springer-Verlag, Berlin, Heidelberg, New York, 1999.

[IK03] T. Izu and Y. Kida. A Status Report of Integer Factorization. *Transactions of the Japan Society for Industrial and Applied Mathematics*, Vol. 13, No. 2, pp. 151–165, 2003. (in Japanese).

[LL93] A. K. Lenstra and H. W. Lenstra, Jr., editors. *The development of the number field sieve*, Volume 1554 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, Heidelberg, 1993.

[M99] B. A. Murphy. *Polynomial Selection for the Number Field Sieve Integer Factorisation Algorithm*. PhD thesis, Australian National University, 1999. (`http://web.comlab.ox.ac.uk/oucl/work/richard.brent/ftp/Murphy-thesis.p%s.gz`).