Cryptanalysis Course
Part IV – Factorization

Tanja Lange
Technische Universiteit Eindhoven

30 November 2016

with some slides by
Daniel J. Bernstein

Q sieve

Sieving small integers $i > 0$ using primes $2, 3, 5, 7$:

1								
2	2							
3					3			
4	2	2						
5							5	
6	2				3			
7								7
8	2	2	2					
9					3	3		
10	2						5	
11								
12	2	2			3			
13								
14	2							7
15					3		5	
16	2	2	2	2				
17								
18	2				3	3		
19								
20	2	2					5	

etc.

Q sieve

Sieving i and $611 + i$ for small i using primes $2, 3, 5, 7$:

1				
2	2			
3		3		
4	2 2			
5			5	
6	2	3		
7				7
8	2 2 2			
9		3 3		
10	2		5	
11				
12	2 2	3		
13				
14	2			7
15		3	5	
16	2 2 2 2			
17				
18	2	3 3		
19				
20	2 2		5	

612	2 2	3 3		
613				
614	2			
615		3	5	
616	2 2 2			7
617				
618	2	3		
619				
620	2 2		5	
621		3 3 3		
622	2			
623				7
624	2 2 2 2	3		
625			5 5 5 5	
626	2			
627		3		
628	2 2			
629				
630	2	3 3	5	7
631				

etc.

Have complete factorization of the “congruences” $i(611 + i)$ for some i’s.

$14 \cdot 625 = 2^1 3^0 5^4 7^1.$
$64 \cdot 675 = 2^6 3^3 5^2 7^0.$
$75 \cdot 686 = 2^1 3^1 5^2 7^3.$

$14 \cdot 64 \cdot 75 \cdot 625 \cdot 675 \cdot 686$
$= 2^8 3^4 5^8 7^4 = (2^4 3^2 5^4 7^2)^2.$

$\gcd\{611, 14 \cdot 64 \cdot 75 - 2^4 3^2 5^4 7^2\}$
$= 47.$

$611 = 47 \cdot 13.$

Why did this find a factor of 611?
Was it just blind luck:
$\gcd\{611, \text{random}\} = 47$?

No.

By construction 611 divides $s^2 - t^2$
where $s = 14 \cdot 64 \cdot 75$
and $t = 2^4 3^2 5^4 7^2$.
So each prime > 7 dividing 611
divides either $s - t$ or $s + t$.

Not terribly surprising
(but not guaranteed in advance!)
that one prime divided $s - t$
and the other divided $s + t$.

Why did the first three completely factored congruences have square product?
Was it just blind luck?

Yes. The exponent vectors $(1, 0, 4, 1), (6, 3, 2, 0), (1, 1, 2, 3)$ happened to have sum 0 mod 2.

But we didn't need this luck!
Given long sequence of vectors, easily find nonempty subsequence with sum 0 mod 2.

This is linear algebra over $\mathbf{F}_2$.
Guaranteed to find subsequence
if number of vectors
exceeds length of each vector.

e.g. for $n = 671$:

$$1(n+\ 1) = 2^5 3^1 5^0 7^1;$$
$$4(n+\ 4) = 2^2 3^3 5^2 7^0;$$
$$15(n+15) = 2^1 3^1 5^1 7^3;$$
$$49(n+49) = 2^4 3^2 5^1 7^2;$$
$$64(n+64) = 2^6 3^1 5^1 7^2.$$

$\mathbf{F}_2$-kernel of exponent matrix is
gen by (0 1 0 1 1) and (1 0 1 1 0);
e.g., $1(n+1)15(n+15)49(n+49)$
is a square.

Plausible conjecture: $\mathbf{Q}$ sieve can separate the odd prime divisors of any n, not just 611.

Given n and parameter y:

Try to completely factor $i(n+i)$ for $i \in \{1, 2, 3, \ldots, y^2\}$ into products of primes $\leq y$.

Look for nonempty set I of i's with $i(n+i)$ completely factored and with $\prod_{i \in I} i(n+i)$ square.

Compute $\gcd\{n, s-t\}$ where $s = \prod_{i \in I} i$ and $t = \sqrt{\prod_{i \in I} i(n+i)}$.

How large does y have to be for this to find a square?

Uniform random integer in $[1, n]$ has $n^{1/u}$-smoothness chance roughly u^{-u}.

Plausible conjecture:
$\mathbf{Q}$ sieve succeeds
with $y = \lfloor n^{1/u} \rfloor$
for all $n \geq u^{(1+o(1))u^2}$;
here $o(1)$ is as $u \to \infty$.

More generally, if $y \in \exp\sqrt{\left(\frac{1}{2c}+o(1)\right)\log n \log\log n}$, conjectured y-smoothness chance is $1/y^{c+o(1)}$.

Find enough smooth congruences by changing the range of i's: replace y^2 with $y^{c+1+o(1)} = \exp\sqrt{\left(\frac{(c+1)^2+o(1)}{2c}\right)\log n \log\log n}$.

Increasing c past 1 increases number of i's but reduces linear-algebra cost. So linear algebra never dominates when y is chosen properly.

Improving smoothness chances

Smoothness chance of $i(n+i)$ degrades as i grows. Smaller for $i \approx y^2$ than for $i \approx y$.

Crude analysis: $i(n+i)$ grows.
$\approx yn$ if $i \approx y$;
$\approx y^2 n$ if $i \approx y^2$.

More careful analysis:
$n+i$ doesn't degrade, but
i is always smooth for $i \leq y$,
only 30% chance for $i \approx y^2$.

Can we select congruences to avoid this degradation?

Choose q, square of large prime.
Choose a "q-sublattice" of i's:
arithmetic progression of i's
where q divides each $i(n+i)$.
e.g. progression $q-(n \bmod q)$,
$2q-(n \bmod q)$, $3q-(n \bmod q)$,
etc.

Check smoothness of
generalized congruence $i(n+i)/q$
for i's in this sublattice.
e.g. check whether $i, (n+i)/q$ are
smooth for $i = q-(n \bmod q)$ etc.

Try many large q's.
Rare for i's to overlap.

e.g. $n = 314159265358979323$:

Original $\mathbf{Q}$ sieve:

i	$n+i$
1	314159265358979324
2	314159265358979325
3	314159265358979326

Use 997^2-sublattice,
$i \in 802458 + 994009\mathbf{Z}$:

i	$(n+i)/997^2$
802458	316052737309
1796467	316052737310
2790476	316052737311

Crude analysis: Sublattices eliminate the growth problem. Have practically unlimited supply of generalized congruences

$$(q-(n \bmod q))\frac{n+q-(n \bmod q)}{q}$$

between 0 and n.

More careful analysis: Sublattices are even better than that! For $q \approx n^{1/2}$ have $i \approx (n+i)/q \approx n^{1/2} \approx y^{u/2}$ so smoothness chance is roughly $(u/2)^{-u/2}(u/2)^{-u/2} = 2^u/u^u$, 2^u times larger than before.

Even larger improvements
from changing polynomial $i(n+i)$.

"Quadratic sieve" (QS) uses
$i^2 - n$ with $i \approx \sqrt{n}$;
have $i^2 - n \approx n^{1/2+o(1)}$,
much smaller than n.

"MPQS" improves $o(1)$
using sublattices: $(i^2 - n)/q$.
But still $\approx n^{1/2}$.

"Number-field sieve" (NFS)
achieves $n^{o(1)}$.

Generalizing beyond $\mathbf{Q}$

The $\mathbf{Q}$ sieve is a special case of the number-field sieve.

Recall how the $\mathbf{Q}$ sieve factors 611:

Form a square as product of $i(i + 611j)$ for several pairs (i, j):
$14(625) \cdot 64(675) \cdot 75(686)$
$= 4410000^2$.

$\gcd\{611, 14 \cdot 64 \cdot 75 - 4410000\}$
$= 47$.

The $\mathbf{Q}(\sqrt{14})$ sieve
factors 611 as follows:

Form a square
as product of $(i + 25j)(i + \sqrt{14}j)$
for several pairs (i, j):

$$(-11 + 3 \cdot 25)(-11 + 3\sqrt{14}) \cdot (3 + 25)(3 + \sqrt{14}) = (112 - 16\sqrt{14})^2.$$

Compute

$s = (-11 + 3 \cdot 25) \cdot (3 + 25)$,

$t = 112 - 16 \cdot 25$,

$\gcd\{611, s - t\} = 13$.

Why does this work?

Answer: Have ring morphism
$\mathbf{Z}[\sqrt{14}] \to \mathbf{Z}/611$, $\sqrt{14} \mapsto 25$,
since $25^2 = 14$ in $\mathbf{Z}/611$.

Apply ring morphism to square:

$$(-11 + 3 \cdot 25)(-11 + 3 \cdot 25)$$
$$\cdot (3 + 25)(3 + 25)$$
$$= (112 - 16 \cdot 25)^2 \text{ in } \mathbf{Z}/611.$$

i.e. $s^2 = t^2$ in $\mathbf{Z}/611$.

Unsurprising to find factor.

Generalize from $(x^2 - 14, 25)$
to (f, m) with irred $f \in \mathbf{Z}[x]$,
$m \in \mathbf{Z}$, $f(m) \in n\mathbf{Z}$.

Write $d = \deg f$,
$f = f_d x^d + \cdots + f_1 x^1 + f_0 x^0$.

Can take $f_d = 1$ for simplicity,
but larger f_d allows
better parameter selection.

Pick $r \in \mathbf{C}$, root of f.
Then $f_d r$ is a root of
monic $g = f_d^{d-1} f(x/f_d) \in \mathbf{Z}[x]$.

$$\mathbf{Q}(r) \leftarrow \mathcal{O} \leftarrow \mathbf{Z}[f_d r] \xrightarrow{f_d r \mapsto f_d m} \mathbf{Z}/n$$

Build square in $\mathbf{Q}(r)$ from congruences $(i - jm)(i - jr)$ with $i\mathbf{Z} + j\mathbf{Z} = \mathbf{Z}$ and $j > 0$.

Could replace $i - jx$ by higher-deg irred in $\mathbf{Z}[x]$; quadratics seem fairly small for some number fields. But let's not bother.

Say we have a square $\prod_{(i,j)\in S}(i - jm)(i - jr)$ in $\mathbf{Q}(r)$; now what?

$\prod(i - jm)(i - jr)f_d^2$
is a square in $\mathcal{O}$,
ring of integers of $\mathbf{Q}(r)$.

Multiply by $g'(f_d r)^2$,
putting square root into $\mathbf{Z}[f_d r]$:
compute r with $r^2 = g'(f_d r)^2 \cdot$
$\prod(i - jm)(i - jr)f_d^2$.

Then apply the ring morphism
$\varphi : \mathbf{Z}[f_d r] \to \mathbf{Z}/n$ taking
$f_d r$ to $f_d m$. Compute gcd$\{n,$
$\varphi(r) - g'(f_d m)\prod(i - jm)f_d\}$.
In $\mathbf{Z}/n$ have $\varphi(r)^2 =$
$g'(f_d m)^2 \prod(i - jm)^2 f_d^2$.

How to find square product of congruences $(i - jm)(i - jr)$?

Start with congruences for, e.g., y^2 pairs (i, j).

Look for y-smooth congruences: y-smooth $i - jm$ and y-smooth $f_d \operatorname{norm}(i - jr) = f_d i^d + \cdots + f_0 j^d = j^d f(i/j)$. Here "$y$-smooth" means "has no prime divisor $> y$."

Find enough smooth congruences. Perform linear algebra on exponent vectors mod 2.