A NEW METHOD FOR FACTORING LARGE SEMI-PRIMES

One of the more important problems of number theory remaining unsolved is how to effectively factor large semi-primes \( N = pq \) into their prime components \( p \) and \( q \). Although considerable efforts have been devoted to this subject over the last several hundred years, even the latest approaches using generalized sieving techniques and elliptic curve factorization still cannot easily break up semi-primes of several hundred digit length in any reasonable amount of time. It is our purpose here to introduce an alternate method for factoring such semi-primes based on our recently observation that all primes above \( N = 3 \) must have the form \( N = 6n \pm 1 \).

Our starting point, based on this observation, is that all semi-primes \( N \) (provided its components \( p \) and \( q \) each exceed a value of 3) must have one of the following two forms –

\[ N = pq = (6n+1)(6m+1) \text{ or } (6n+1)(6m-1) \]

The positive integers \( n \) and \( m \) are interchangeable. In terms of modular arithmetic, these two cases have \( N \mod (6) = 1 \) and 5, respectively. We will look at these two cases separately.

CASE 1: \( N \mod(6)=1 \)

Here we write-

\[ N = pq = (6n+1)(6m+1) = 36mn + 6(m+n) + 1 \]

This is equivalent to the linear Diophantine Equation-

\[ 6U + V = (N-1)/6 \]

where \( U = nm \) and \( V = n+m \). This equation has the integer solutions \( U = k = nm \) and \( V = (N-1)/6 = n+m \). Eliminating \( m \) from these two equalities, we get the quadratic-

\[ n^2 - n[(N-1) - 36k]/6 + k = 0 \]

which solves as-

\[ n = \frac{1}{12} \left\{ (N-1) - 36k \pm \sqrt{[(N-1) - 36k]^2 - 144k} \right\} = \frac{(p-1)}{6} \]

Since \( n \) is a positive integer we know that \( (p-1)/6 \) also equals the same integer. Furthermore the radical must be a perfect square. That is-
must be such that \( \sqrt{F} \) is an integer. For the root of \( F \) to be a real integer it is necessary that \( F > 0 \). This means that \( k \) must be restricted to the two intervals:

\[
k < \frac{(N + 1 - 2\sqrt{N})}{36} \quad \text{or} \quad k > \frac{(N + 1 + 2\sqrt{N})}{36}
\]

The value of \( k \) for which \( \sqrt{F} \) is an integer will typically be found near one of the zero points of \( F \). Once such an integer pair \([k, \sqrt{F}]\) has been found, it is then simple bookkeeping to determine the factors \( p \) and \( q \). We have:

\[
p, q = \frac{1}{2} \left\{ \sqrt{F} + 4N \pm \sqrt{F} \right\}
\]

**CASE 2: \( \text{N mod}(6) = 5 \)**

For this case we have:

\[N = pq = (6n+1)(6m-1)=36nm-6(n-m)-1\]

On setting \( nm=U \) and \( n-m=V \), we get \( 6U-V=(N+1)/6 \) with the solution \( U=k \) and \( V=[36k-(N+1)]/6 \). Eliminating \( m \) from these equations, we obtain the quadratic:

\[n^2 - n[36k - (N + 1)]/6 - k = 0\]

This has the solution:

\[n = \frac{1}{12} \left\{ 36k - (N + 1) \pm \sqrt{(36k - (N + 1))^2 + 144k} \right\}\]

Next we introduce the function:

\[G = [36k - (N + 1)]^2 + 144k = [36k - (N - 1)]^2 + 4N\]

We see that \( \sqrt{G} \) must be an integer for \( n \) and \( p \) to be such. Hence the problem reduces to finding the \( k \) which makes \( \sqrt{G} \) an integer. When looking at the function \( G=G(k) \) for fixed \( N \), we see that it is a parabola with zero slope at \( k=(N-1)/36 \), suggesting one start a search for integer \( \sqrt{G} \) there. This time there are no restricted regions on \( k \) and so the search would appear to be more difficult. We have, however, found that the search need not go far away from the zero slope point to get the desired \( k \). Here is a schematic of the \( G \) parabola:
Once the solution pair \([k, \sqrt{G}]\) has been found it will then be a simple matter to determine \(p\) and \(q\). We find for \(N=6n-1\) semi-primes that:

\[
\{N \pm \sqrt{G - 4N}\}
\]

It is pointed out that a better initial estimate for \(k\) is \(k = \{N + (\alpha^2 - 1)/\alpha \sqrt{N} - 1\}/36\), where \(\alpha\) is an unknown constant lying in the range \(0 < \alpha < 1\). This follows from the definition \(N=(6n+1)(6m-1)\). The problem with this initial guess is that \(\alpha\) is unknown beforehand. So we usually just make the guess \(\alpha = 1\) which produces the initial value \(k=(N-1)/36\). However a guess such as \(\alpha = 1/2\) or \(\alpha = 1/4\) will sometimes work better. Note that \(N=pq=\{\alpha \sqrt{N}\}\{1/\alpha \sqrt{N}\}\).

We can summarize our factoring procedure for \(N>>1\) via the following flow chart:
FACTORING OF SEVERAL SPECIFIC SEMI-PRIMES:

We next verify the above generic results by looking at five different semi-primes. We start with the five digit semi-prime $N=34417$ which has $N \mod(6)=1$. Here we should carry out a search in the neighborhoods of $k=945.748$ and $k=966.362$ according to the above inequalities for $k$. We begin our computer search near $k=945$ using the following one line program

$$\text{for } k \text{ from 942 to 946 do } \{k, \text{sqrt}(F)\} \text{od;}$$

It yields the output –

<table>
<thead>
<tr>
<th>$k$</th>
<th>sqrt(F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>943</td>
<td>204sqrt(2)</td>
</tr>
<tr>
<td>944</td>
<td>48sqrt(22)</td>
</tr>
<tr>
<td>945</td>
<td>144 ←Answer</td>
</tr>
<tr>
<td>946</td>
<td>12isqrt(46)</td>
</tr>
</tbody>
</table>

Hence $k=945$ produces the integer value $\text{sqrt}(F)=144$. We thus obtain-

$$p,q=0.5\{\text{sqrt}[144^2 +4(34417)]±144}\} = 127 \text{ or } 271$$

It is amazing how little effort it took to factor this semi-prime requiring only three evaluations for determining the integer value of $\text{sqrt}(F)$.

FLOW CHART FOR FACTORING THE SEMI-PRIME $N=pq$

- $N \mod(6)=1$
  - $F=[36k-(N+1)]^2-4N$
  - $36k>N+1-2\text{sqrt}(N)$ or $36k>N+1+2\text{sqrt}(N)$
  - $p,q=\{\text{sqrt}(F+4N)±\text{sqrt}(F)\}/2$
- $N \mod(6)=5$
  - $G=[36k-(N-1)]^2+4N$
  - $k_{\text{approx}}=(N-1)/36$
  - $p,q=\{\text{sqrt}(G)±\text{sqrt}(G-4N)\}/2$

both sqrt(F) and sqrt(G)

must be integers
To prove that this was not a fluke, consider the second and larger 6n+1 semi-prime $N=21428053$. There-

$$F = (36k - 21428054)^2 - 85712212$$

and one must search in the ranges-

$$k<594966 \quad \text{and} \quad k>595480$$

Trying the upper range first, we find $k=595483$ produces the integer value $\sqrt{F}=1188$. It took just 4 trials to find this pair. The values of $p$ and $q$ follow from-

$$p,q = 0.5\{\sqrt{1188^2+4(21428053)}\pm1188\} = 4073 \text{ and } 5261$$

The factoring was again accomplished with very little effort. Had we not found an integer solution in the right $k$ range then one would have needed to try the left lower range.

We next examine a $N \mod(6)=5$ case using the semi-prime $N=106577$. This time we require the use of the $G$ function which for this $N$ reads-

$$G = (36k - 106576)^2 + 426308$$

To find the value of $k$ which makes $\sqrt{G}$ equal to an integer we search about $k=106576/36 \approx 2960$. Letting $k=2960+b$, the Function $G$ can be rewritten as-

$$G = (36b-16)^2 + 4N$$

Carrying out a search over the range $-20<b<20$ we quickly find an integer answer $\sqrt{G}=738$ for $b=10$ using the MAPLE search program-

```maple
for b from -20 to 20 do {b, sqrt(G)} od;
```

From this follows that-

$$p,q = \frac{1}{2} \left\{ 738 \pm \sqrt{738^2 - 4(106577)} \right\} = 197 \text{ and } 541$$

As second example of factoring a $N \mod(6)=5$ semi-prime, we looked at the number $N=732010841$. This number is one of the largest semi-primes we have been able to factor on our PC using the present method. One expects an integer solution for $\sqrt{G}$ in the neighborhood of $k=N/36 \approx 20333634$. Carrying out a search about this point took a total of 614 trials to find desired integer values $\sqrt{G}=58458$ at $k=20333020$. Simple bookkeeping then produced $p=18169$ and
q=40289.

As a final demonstration, consider factoring the number \( N=455839 \) which has \( N \mod(6)=1 \). This semi-prime is often used in the literature to demonstrate Lenstra’s Elliptic Curve Factorization Method. We have that \( F=0 \) at \( k=12624.71 \) and \( k=12699.73 \). Carrying out a search near the two \( F=0 \) points shows that with just a single trial one finds that \( k=12700 \) produces the integer value \( \sqrt{F}=162 \). Thus we have

\[
p, q = 0.5\left\{\sqrt{162^2 + 4(455939)} - 162\right\} = 599 \text{ and } 761
\]

Thus we have factored 455839 as the product of the primes 599 and 761 with almost no effort compared to what is required via the elliptic curve approach.

It still remains to see if the present method will still work in a reasonable time (and for a reasonable number of trials) when \( N \) approaches 100 digit length numbers as used in public key cryptography. My home PC is limited to factoring \( N \)s no larger than 15 digit length. What is clear is that larger \( N \)s will require ever larger departures of the \( k \)s from their \( k=N/36 \) value despite of the fact that the actual departure in percent from \( N/36 \) drops with increasing \( N \). These departures become especially noticeable when the ratio of \( q/p \) becomes large such as when looking at the Fermat semi-prime \( 2^{32}+1 \). Perhaps an improved version of the existing sieve methods can be used in such a search. What is clear from the above examples is that a search for integer \( k \) near \( N/36 \) will often quickly produce a value for which \( \sqrt{F} \) or \( \sqrt{G} \) are equal to an integer and so \( N \) is factored.

U.H. Kurzweg
Gainesville, Florida
August 2013