FINDING LARGE SEMI-PRIMES

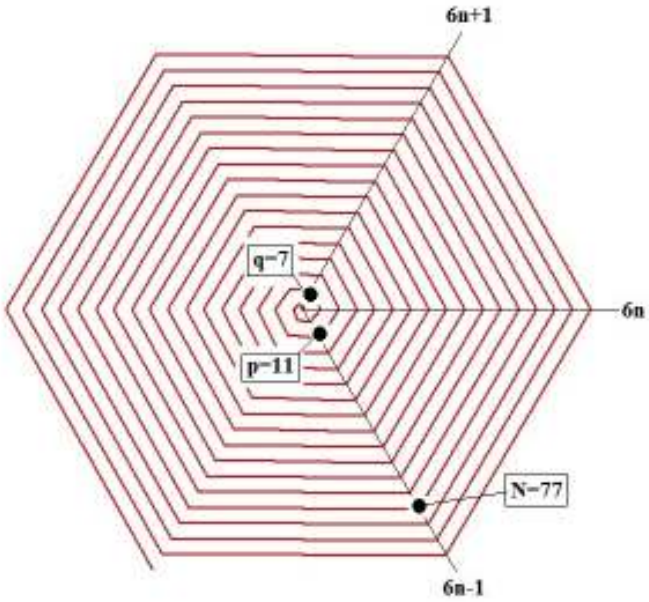
It is well known that any positive integer N may be decomposed into the product of one or more prime numbers. For example 111=3x37 and 6479=11x19x31. When the number of factors is just one term(1 and N excluded) it is a prime number and when it factors in the product of two primes it is referred to as a semi-prime. An example the number 1237 is a prime while 2501=41x61 is a semi-prime. A number with multiple factors such as 85085=5x7x11x13x17 is a composite. We want here to identify large semi-primes N=p·q since these play a critical role in public key cryptography.

Our starting point is our earlier defined number fraction –

$$f=[\sigma(N)-(N+1)]/N$$

, where σ(N) is the sigma function of number theory. The number fraction has zero value when N is a prime and a value greater than one when N is a composite number with multiple factors. A semi-prime will have a value near zero, but not zero. Typically if N is a k digit long semi-prime one can expect the value of f to equal about $10^{-k/2}$. Furthermore, a semi-prime N must have its mod(6) operation yield a value of 1 or -1 since we know that all primes above 3 have the form 6n±1. In terms of our earlier defined hexagonal integer spiral we have the following diagram for the semi-prime N=77 showing its components p=11 and q=7-

SEMI-PRIME N=77 AND ITS COMPONENTS p=11 AND q=7 AS LOCATED ON THE HEXAGONAL INTEGER SPIRAL



Performing a mod(6) operation on N=p·q produces (-1)=(-1)(+1) . This shows why p and q lie along different diagonals in the graph. The f value for N=77 is 18/77=0.2337, and so is small but not zero.

We have run through a set of numbers N which are semi-primes and find the results shown in the following table-

N=pq	N mod(6)	f	p	p mod(6)	q	q mod(6)
77	-1	0.2337	11	-1	7	+1
391	+1	0.10230	23	-1	17	-1
12091	+1	0.018195	113	-1	107	-1
765469	+1	0.0026780	1559	-1	491	-1
8932479	+1	0.00080837	70949	-1	1259	-1
123456763	+1	0.000281264	30703	+1	4021	+1
3000000089	-1	0.0000471499	115469	-1	25981	+1

We picked these semi-primes by finding the lowest (non-zero) value of f for a given range of N. Once such a value for f had been found, we next applied an ifactor(N) operation to get the components. Note that N, p and q always lie on either the 6n+1 or 6n-1 diagonal in the hexagonal integer plane. The mod(6) values for N show that N mod(6)=+1 implies that that both p and q lie on the same diagonal while a N mod(6)=-1 result says that p and q must lie on different diagonals. The values of f for these semi-primes decrease rapidly in value as the number of digits in N increases. We can estimate the value of f for a semi-prime by noting that-

$$f(N)=[p+q]/N\approx 2\,\mathrm{sqrt}(N)/N=2/\mathrm{sqrt}(N) \quad \text{since we know that } q<\mathrm{sqrt}(N)<p$$

That is, the value of f for a semi-prime is approximately 2/sqrt(N). For the ten digit number 3000000089 we have a number fraction estimaste f≈0.0000365. This is close to the exact value of f given in the above table. If f lies much above the approximation 2/sqrt(N) then we know we are dealing not with a semi-prime but rather with one having three or more prime factors. For example , N=453583 factors into four prime components 13 x 23 x 37 x 41 and has a value f=0.182264. If this where a semi-prime then the value of f should be near f=0.002969 and not 61 times larger.

We can summarize the above observations by noting that a semi-prime N=p·q has the following properties-

(1) It must be of the form 6n+1 or 6n-1, where n=1,2,3,..

(2) The value of its number fraction should lie near f=2/sqrt(N).

(3) If N mod(6)=1 than both p and q must lie along the same diagonal. If N mod(6)=-1 then p and q lie along different diagonals in the hexagonal spiral plane.

To test things out using these criteria for a large semi-prime consider the 12 digit long number N=460969682477. It has N mod(6)=-1 and an f estimate of 2/sqrt(N)=0.0000029457. Evaluating the actual f value yields f=[sigma(N)-(N+1)]/N =0.00000554281 and so we are close and one can conclude that N is a semi-prime. We can then proceed, using the technique discussed in earlier notes, to factor this number. It says essentially that –

$$p=Nf/2+sqrt\{(Nf/2)^2-N\}$$

Using the above values of N and f then yields p=2359603 and q=N/p=195359, so that-

460969682477=2359603 x 195359

The secret to a precise evaluation is to be able to find an accurate value of f=(p+q)/N quickly. This will no longer be possible when N has digit lengths above 50 or so. Note here that p=6(393267)+1 while q=6(32560)-1, so that p and q do indeed lie along different diagonals as predicted by (3).

June 1, 2013