Semiprimi e fattorizzazione col modulo

ing. R. Turco, prof. Maria Colonnese

Sommario
Nel seguito viene esaminato un Teorema ed una tecnica di fattorizzazione per numeri semiprimi di qualsiasi dimensione; inoltre viene valutato il criterio di arresto algoritmico e la bontà della tecnica descritta.

I semiprimi
Sono definiti "semiprimi" quei numeri N scomponibili direttamente nel prodotto di due fattori primi. I numeri RSA, ad esempio, sono da considerarsi tali.

Teorema
Sia la (1)

$$x^2 = a^2 \bmod N \qquad \textbf{(1)}$$

un' equazione modulo di secondo grado per semiprimi N non quadratici. Se *N è primo,* le due soluzioni x_1 e x_2 della (1) si dicono *soluzioni associate*; altrimenti *se N non è primo,* ma da scomporre in fattori primi, allora le soluzioni si dicono *soluzioni non associate*.

Se N non è primo e non è quadratico, indicando le "due soluzioni non associate" con x_1 e x_2 e con $x_2 > x_1$, allora N è scomponibile in due fattori non banali, ovvero diversi sia da 1 che da N stesso, tali che:

$$N = (x_2-x_1)(x_2+x_1) \quad \textbf{(2)}$$

Dimostrazione
Nella (1) se N non è primo e non quadratico allora è:

$$x_1^2 - x_2^2 = 0 \bmod N$$

Questo perché entrambe le soluzioni soddisfano a^2 mod N.

il che significa;

$$x_1^2 - x_2^2 = k * N$$

ovvero che N divide $x_1^2 - x_2^2$ o equivalentemente N divide $(x_2+x_1)(x_2-x_1)$. Con k=1 è maggiormente evidente.

Caso soluzioni associate
Ad esempio se a^2=4 ed N =7 (numero primo), le possibili soluzioni associate sono $x_1 = 2$ e $x_2 = -2$ ed in entrambi i casi si ottiene 4 nella (1).

In tale caso N è primo e non è scomponibile; per cui il caso di soluzioni associate non ha interesse ai fini della fattorizzazione dei semiprimi.

Caso di interesse: soluzioni non associate – fattorizzazione di semiprimi
La (1) si può interpretare anche in un altro modo; difatti la (1) afferma che x^2 ed a^2 differiscono per un multiplo di N o che x^2 è congruo ad a^2 modulo N:

$$x^2 = k * N + a^2 \quad \textbf{(3)}$$

Per cui se è noto N, fissato a^2 e variando k, a partire da zero, la (3) ha *due soluzioni non associate* se si trovano valori di x che sono interi positivi o negativi.

Anche qui è necessario, però, trovare un criterio di arresto algoritmico della (3). E′ anche chiaro che x^2 deve essere un "quadrato perfetto", lo stesso dicasi per a^2 e k * N, affinché x sia intero. Questo fa sì che la (3) deve essere una terna pitagorica, con tutte le proprietà relative.

Il metodo del modulo, con l′equazione (1), ci riporta ad un metodo analogo del quadrato perfetto generalizzato di fattorizzazione già esaminato in altri due lavori (vedi [Rif. 1][Rif. 2]).

Nel caso di N numero primo o di quadrato la (3) ci porta alle *soluzioni banali 1 ed N*.

E′ anche evidente che l′equazione (3) ha diverse soluzioni e che essendo un′equazione di secondo grado ce ne aspettiamo almeno due.

Le domande che ci si pone sono adesso essenzialmente:

- Quali soluzioni prendere tra le m disponibili?
- Quando arrestarsi nella ricerca delle soluzioni?

Soluzioni da scegliere e criterio d′arresto algoritmico.
Per questo metodo si possono calcolare le soluzioni x, finché x < N, al variare di valori successivi di k (da 0 a 1) e di a^2 (1,4,9,16, 25, etc) e ci si arresta se si ottengono due soluzioni consecutive con valori x interi. Le soluzioni si ritengono consecutive nell′ambito di uno stesso a^2, ma con due valori di k consecutivi. Solitamente le soluzioni si trovano consecutive per due valori di k iniziali. Gli esempi successivi, attraverso un excel, mostreranno quanto affermato.

Parallelizzazione algoritmica
Una semplificazione algoritmica rispetto ad altri algoritmi noti (ad esempio il "Quadratic sieve") è di integrare i due step che, di solito, sono considerati separati:

- Produzione di soluzioni o *'solutions collection'*
- Segnalazione delle soluzioni consecutive o *'solutions detection'*.

Per il primo step, su un computer con possibilità di calcolo parallelo e parecchie CPU, è possibile avviare parallelamente processi che calcolano archi di valori disgiunti di a^2. Ad esempio il processo A calcola, per valori di k da 0 a N, tutte le soluzioni che si possono ottenere per un determinato arco di valori di a^2, disgiunto dall′arco di valori assegnato ad un processo B etc.

Lo step 2 è attivabile all'interno dello step 1, cioè il programma può mantenere conto della "consecutività delle soluzioni" ed arrestarsi appena le trova o se ha superato il range di valori da considerare (x<N), scrivendone i valori su file o a video ad uso dell'utente. Nel seguito vari esempi.

Esempi di fattorizzazione con "soluzioni non associate"

N=35

Se N=35 e a^2=1, allora le possibili soluzioni di x^2=1 mod 35 sono x_1=1, x_2=6 oppure se consideriamo anche i numeri negativi anche x_3 = -1, x_4 = -6 (ma otteniamo lo stesso risultato). Per cui la scomposizione in fattori è:

$$N = (6-1)(6+1)=5*7$$

Con l'aiuto di un semplice excel si possono facilmente calcolare i valori. Nell'excel le parti in giallo della figura sono i "dati di input" che inseriamo durante l'analisi '*what-if*'.

Il valore della colonna **k** è preimpostato, con valori da 0 a 100: potevamo considerare per k solo i valori 0 e 1; ma nell'excel si è voluto evidenziare che possono uscire anche soluzioni per valori di k>1, che scartiamo.

Nell'excel **x^2** è il valore calcolato con la (3); mentre con **xQP** è indicata la radice quadrata di x^2 che se rappresenta una soluzione deve essere un valore intero.

Quando si individuano sull'excel nella colonna xQP due valori interi consecutivi, allora si sono trovate le due soluzioni non associate consecutive x_1 e x_2.

Primo tentativo. Dalla (3) con a^2=1 otteniamo due soluzioni consecutive, sia con k=0 che k=1.

Input N	Input a^2
35	1

K	X^2	xQP
0	1	1
1	36	6

$$N = (6-1)(6+1)=5*7$$

N=21 fattorizzazione con soluzioni non associate

N=21 e a^2=4, allora le possibili soluzioni di x^2=1 mod 21 sono x_1=2, x_2=5. Per cui la scomposizione in fattori è:

$$N = (5-2)(5+2)=3*7$$

Con l'aiuto dell'excel si possono vedere due situazioni. Dalla (3) con a^2=1 **non otteniamo** delle soluzioni consecutive; difatti nell'excel con tale valore di a^2, xQP presenta valori con virgola mobile tali da non presentare due soluzioni intere consecutive per k=0 e k=1.

Input N	Input a^2
21	1

K	x^2	xQP
0	1	1
1	22	4,690416
2	43	6,557439
3	64	8
4	85	9,219544
5	106	10,29563
6	127	11,26943
7	148	12,16553
8	169	13

Con $a^2=4$ otteniamo soluzioni consecutive con k=0 e k=1.

Input N	Input a^2
21	4

K	x^2	xQP
0	4	2
1	25	5

N=91

Con l'aiuto di un semplice excel si può osservare che fissato N le soluzioni consecutive, solitamente per k=0 e k=1, sono $x_1=3$ e $x_2=10$.

Input N	Input a^2
91	9

K	x^2	xQP
0	9	3
1	100	10

Per cui N = (10-3) *(10+3) = 7*13 = 91

N=713

Con l'excel si può osservare che con a=4 N=23*31.

N=1271

Con l'excel si può osservare che con a=5 N=31*41.

N=659161

Con l'excel si può osservare che con a=6192 N=53*12437.

N=758657

Con l'excel si può osservare che con a=6188 N=61*12437.

N=1118069
Con l'excel si può osservare che con a=9134 N=61*18329.

N=49 quadratico
N è un quadrato perfetto. Non usiamo il metodo perché genera le soluzioni banali 1 ed N.

N=7 numero primo
N è un primo. Non usiamo il metodo perché genera le soluzioni banali 1 ed N.

Considerazioni
Questo modo di procedere è una "tecnica di trasformazione della natura del problema": da un problema di prodotto (o produttorio) ad un problema di somme o differenze (sommatorie), cercando cioè di percorrere soprattutto strade dove si possa aggirare la complessità di calcolo.

Il metodo qui presentato, rispetto al "Quadratic Sieve" che è di validità generale per la fattorizzazione di un numero con molti fattori primi, è una semplificazione dedicata ai semiprimi N=p*q. La semplificazione qui è specialmente nella parte di collezione dei dati e di rilevazione delle soluzioni.

L'inconveniente che nasce basandosi sulla (3) è che se N è a parecchie cifre, il valore di a^2 da cercare è elevato ed aumenta il tempo di ricerca delle soluzioni.

Riferimenti
[1] **Fattorizzazione con algoritmo generalizzato con quadrati perfetti in ambito delle forme 6k± 1** - ing. Rosario Turco, dott. Michele Nardelli, prof. Giovanni Di Maria, Francesco Di Noto, prof. Annarita Tulumello, prof. Maria Colonnese
[2] **TEST DI PRIMALITA', FATTORIZZAZIONE E π(N) CON FORME 6k±1** - ing. Rosario Turco, dott. Michele Nardelli, prof. Giovanni Di Maria, Francesco Di Noto, prof. Annarita Tulumello
[3] **Le basi della crittografia** – ing. Rosario Turco (www.geocities.com/SiliconValley/Port/3264 sezione MISC)