UNIVERSITÀ DI TRENTO

Quadratic Sieve

introduzione e primo impatto sugli standard crittografici degli algoritmi di fattorizzazione per interi

Leonardo Errati

Università di Trento

22 ottobre 2021

Contenuti

Il crittosistema RSA

Crittosistema

Insieme di algoritmi crittografici in grado di implementare una particolare forma di sicurezza, ad esempio la confidenzialità o l'integrità dei dati.

In particolare, siamo interessati ad RSA.

RSA: Rivest–Shamir–Adleman

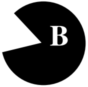
Crittosistema a chiave pubblica, ideato per garantire la confidenzialità dei dati. Opera tramite proprietà di natura algebrica.

RSA: key generation

Supponiamo Alice (A) e Bob (B) vogliano comunicare tra loro.

1. A e B si incontrano, scegliendo due primi p e q

$$
\begin{aligned}
p &= 11 \\
q &= 17
\end{aligned}
$$

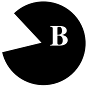


RSA: key generation

Supponiamo Alice (A) e Bob (B) vogliano comunicare tra loro.

1. A e B si incontrano, scegliendo due primi p e q

$$\begin{aligned} p &= 11 \\ q &= 17 \end{aligned}$$



2. A e B calcolano $N = pq$ e $\varphi(N) = (p-1)(q-1)$, quindi scelgono
   - un intero $e \leq \varphi(N)$ tale che $\gcd(e, \varphi(n)) = 1$
   - un intero d tale che $ed \equiv 1 \mod \varphi(n)$

RSA: encryption, decryption

3. chiave pubblica: (N, e), chiave privata: (N, d)

RSA: encryption, decryption

3. chiave pubblica: (N, e), chiave privata: (N, d)
4. criptare:

$$E(k) = k^e \mod N$$

RSA: encryption, decryption

3. chiave pubblica: (N, e), chiave privata: (N, d)
4. criptare:
$$E(k) = k^e \mod N$$
5. decriptare:
$$D(k) = k^d \mod N$$

RSA: encryption, decryption

3. chiave pubblica: (N, e), chiave privata: (N, d)
4. criptare:

$$E(k) = k^e \mod N$$

5. decriptare:

$$D(k) = k^d \mod N$$

Funziona perché...

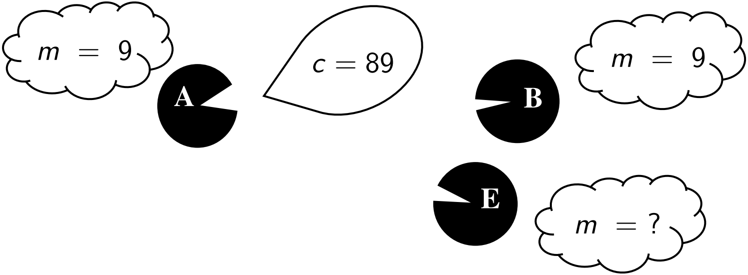
$$D(E(k)) = (k^e)^d = k^{ed} = k$$

dove abbiamo usato (in ordine) il piccolo teorema di Fermat ed il teorema cinese del resto.

Ora A e B possono separarsi e comunicare a distanza.

Un esempio di applicazione:

1. A e B scelgono $p = 11$, $q = 17$, risultando in $N = 187$ e $\varphi(N) = 160$; poi costruiscono le chiavi con $e = 7$ e $d = 23$
2. chiave pubblica: $(187, 7)$, chiave privata: $(187, 23)$
3. A vuole inviare 9, deve inviare $9^7 \mod 187 = 89$
4. B riceve 89, decodifica con $89^{23} \mod 187 = 9$



Take aways

Key management I

La chiave pubblica può essere divulgata, la chiave privata **deve essere protetta**. Questo è un concetto generale.

Take aways

Key management I

La chiave pubblica può essere divulgata, la chiave privata **deve essere protetta**. Questo è un concetto generale.

Key management II

La chiave pubblica non fornisce informazioni sulla funzione di decriptazione, ma scomporre $N = pq$ permette di calcolare $\varphi(N) = (p-1)(q-1)$ e trovare d! Il sistema è violato. Questo riguarda molti algoritmi simili ad RSA.

nome	anno	numero cifre	metodo fattorizzazione
RSA-100	1991	100	multiple-polynomial quadratic sieve
RSA-129	1994	129	quadratic sieve
RSA-130	1996	130	number field sieve
RSA-150	2004	150	general number field sieve
RSA-140	2020	140	quadratic sieve

Fattorizzare numeri interi

Fattorizzare grandi numeri

Come possiamo procedere nella fattorizzazione di un intero?

Come possiamo procedere nella fattorizzazione di un intero?

N=21

Fattorizzare grandi numeri

Come possiamo procedere nella fattorizzazione di un intero?

N=21
N=81
N=148
N=625
N=626

Fattorizzare grandi numeri

Come possiamo procedere nella fattorizzazione di un intero?

N=21
N=81
N=148
N=625
N=626
N=627

Come possiamo procedere nella fattorizzazione di un intero?

N=21

N=81

N=148

N=625

N=626

N=627

N=629

N=19 627

Come possiamo procedere nella fattorizzazione di un intero?

N=21

N=81

N=148

N=625

N=626

N=627

N=629

N=19 627

N=3 209 209 209 209 209 209 209

Trial Division

Trial Division

Dato un intero N, proviamo a trovare un suo divisore “alla cieca”, tramite tentativi: $2, 3, 5, 7, 11, \ldots$

Trial Division

Dato un intero N, proviamo a trovare un suo divisore “alla cieca”, tramite tentativi: $2, 3, 5, 7, 11, \ldots$

. . . buona fortuna!

$$N = 3\,209\,209\,209\,209\,209\,209\,209$$

Trial Division

Dato un intero N, proviamo a trovare un suo divisore “alla cieca”, tramite tentativi: $2, 3, 5, 7, 11, \ldots$

. . . buona fortuna!

$$N = 3\,209\,209\,209\,209\,209\,209\,209$$

$$\sqrt{N} = 56\,649\,882\,693.69327\ldots$$

Metodo di Fermat

Cerco x, y tali che $x^2 - y^2 = N$.

Metodo di Fermat

Cerco x, y tali che $x^2 - y^2 = N$.

$$x^2 - y^2 = (x - y)(x + y) = N$$

Ogni intero ha una fattorizzazione di questo tipo.

$$N = ab \implies N = \left(\frac{a+b}{2}\right)^2 - \left(\frac{a-b}{2}\right)^2$$

1. partiamo da $a_0 = \lceil \sqrt{N} \rceil$
2. se a_i non è tale che $b_i = \sqrt{a_i^2 - N}$ sia intero, $a_{i+1} = a_i + 1$
3. trovato tale a_i, $a = a_i$ e $b = b_i$
4. finalmente $N = (a - b)(a + b)$

1. partiamo da $a_0 = \lceil \sqrt{N} \rceil$
2. se a_i non è tale che $b_i = \sqrt{a_i^2 - N}$ sia intero, $a_{i+1} = a_i + 1$
3. trovato tale a_i, $a = a_i$ e $b = b_i$
4. finalmente $N = (a - b)(a + b)$

$$N = 119$$

$$\left\lceil \sqrt{N} \right\rceil = 11 \qquad \sqrt{a^2 - N} = 1.4142\ldots$$

$$12 \qquad \sqrt{a^2 - N} = 5$$

$$N = (12 - 5)(12 + 5)$$

Possiamo modificare leggermente la procedura.

Fermat

Cerchiamo a, b tali che $a^2 - b^2 = N$, ovvero $N = (a - b)(a + b)$; trovo i fattori come

$$(a - b) \quad (a + b)$$

Possiamo modificare leggermente la procedura.

Fermat

Cerchiamo a, b tali che $a^2 - b^2 = N$, ovvero $N = (a-b)(a+b)$; trovo i fattori come

$$(a-b) \qquad (a+b)$$

Fermat "modificato"

Cerchiamo a, b tali che $a^2 - b^2 \equiv 0 \mod N$, ovvero $kN = (a-b)(a+b)$ per un certo k; trovo i fattori come

$$\gcd(N, a-b) \qquad \gcd(N, a+b)$$

Non sempre funzionano al primo tentativo...

Fattorizzare interi

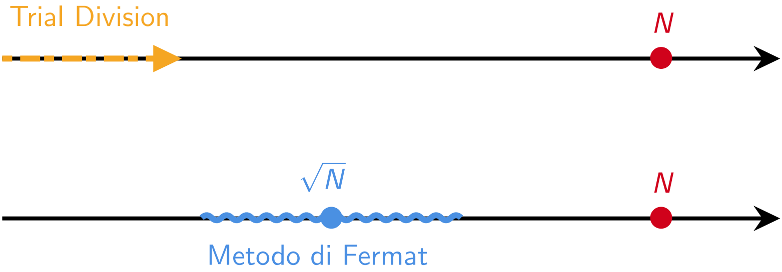
Non è un compito facile, su questo si basano crittosistemi come quelli simili ad RSA.

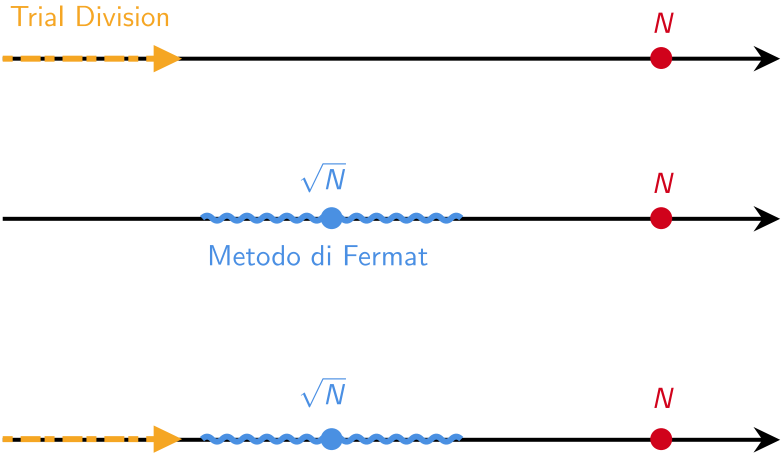
Fattorizzare interi

Non è un compito facile, su questo si basano crittosistemi come quelli simili ad RSA.

Costruire algoritmi migliori

La trial division ha complessità $O(n)$, è troppo dispendiosa.
Il metodo di Fermat ha ancora complessità $O(n)$.
Tuttavia, possono essere combinati...


Trial Division
N
$\sqrt{N}$
N
Metodo di Fermat


Trial Division
N
√N
N
Metodo di Fermat
√N
N

Quadratic Sieve: introduzione

Abbiamo visto il metodo di Fermat. Modificato o meno, esistono molti modi per trovare tali a, b.

Abbiamo visto il metodo di Fermat. Modificato o meno, esistono molti modi per trovare tali a, b.

"It occurred to me early in 1981 that one might use something akin to the sieve of Eratosthenes to quickly recognize the smooth values of Kraitchik's quadratic polynomial $Q(x) = x^2 - n$."

(Carl Pomerance, *A Tale of Two Sieves*)

Possiamo compiere due osservazioni da questa frase.

"[...] the smooth values of Kraitchik's quadratic polynomial $Q(x) = x^2 - n$."

Tale polinomio quadratico è esattamente quello usato nell'algoritmo di Fermat.

Ricordiamo l'algoritmo...

1. partiamo da $a_0 = \lceil\sqrt{N}\rceil$
2. se a_i non è tale che $b_i = \sqrt{a_i^2 - N}$ sia intero, $a_{i+1} = a_i + 1$
3. trovato tale a_i, $a = a_i$ e $b = b_i$
4. finalmente $N = (a - b)(a + b)$

“[...] the smooth values of Kraitchik’s quadratic polynomial
$$Q(x) = x^2 - n."$$

"[. . .] the smooth values of Kraitchik's quadratic polynomial
$$Q(x) = x^2 - n."$$

Intero B-regolare

Un intero N si dice essere B-regolare (B-smooth) se tutti i suoi divisori sono minori o uguali a B.

$270 = 2 \times 3^3 \times 5$ 19-regolare

$290 = 2 \times 5 \times 29$ non 19-regolare

$290 = 2 \times 5 \times 29$ 29-regolare

Proviamo a svolgere i calcoli necessari per $N = 1771$:

$$\begin{aligned}
43^2 &\equiv 78 = 2^1 \times 3^1 \times 13^1 && \mod N \\
44^2 &\equiv 165 = 3^1 \times 5^1 \times 11^1 && \mod N \\
45^2 &\equiv 254 = 2^1 \times 127^1 && \mod N \\
46^2 &\equiv 345 = 3^1 \times 5^1 \times 23^1 && \mod N \\
47^2 &\equiv 438 = 2^1 \times 3^1 \times 73^1 && \mod N \\
48^2 &\equiv 533 = 13^1 \times 41^1 \times 73^1 && \mod N \\
&\vdots \\
74^2 &\equiv 163 = 163^1 && \mod N \\
75^2 &\equiv 312 = 2^3 \times 3^1 \times 13^1 && \mod N \\
76^2 &\equiv 463 = 463^1 && \mod N
\end{aligned}$$

Proviamo a svolgere i calcoli necessari per $N = 1771$:

$$
\begin{aligned}
43^2 &\equiv 78 = 2^1 \times 3^1 \times 13^1 && \mod N \\
44^2 &\equiv 165 = 3^1 \times 5^1 \times 11^1 && \mod N \\
45^2 &\equiv 254 = 2^1 \times 127^1 && \mod N \\
46^2 &\equiv 345 = 3^1 \times 5^1 \times 23^1 && \mod N \\
47^2 &\equiv 438 = 2^1 \times 3^1 \times 73^1 && \mod N \\
48^2 &\equiv 533 = 13^1 \times 41^1 \times 73^1 && \mod N \\
&\vdots \\
74^2 &\equiv 163 = 163^1 && \mod N \\
75^2 &\equiv 312 = 2^3 \times 3^1 \times 13^1 && \mod N \\
76^2 &\equiv 463 = 463^1 && \mod N
\end{aligned}
$$

Per trovare una corrispondenza seguendo l'algoritmo di Fermat sono necessari al più $(N-1)-(\sqrt{N})$ calcoli.
Ma se studiamo questi valori...

$$43^2 \equiv 78 = 2^1 \times 3^1 \times 13^1 \mod N$$
$$75^2 \equiv 312 = 2^3 \times 3^1 \times 13^1 \mod N$$

...notiamo che $(43 \times 75)^2$ è un quadrato modulo N.

La B-regolarità ha limitato il numero di casi da considerare. In questo caso abbiamo considerato i valori 19-regolari.

Un paio di formalismi

Notazione: exponent vector

Sia $m = p_1^{e_1} \dots p_k^{e_k}$, diciamo **exponent vector di** m il vettore

$$m = \left(p_1^{e_1}, \dots, p_k^{e_k}\right)$$

per indicare in modo compatto i fattori primi di m, indicando con esponente nullo anche quelli che non appaiono.

$$\begin{aligned} 4788 &= 2^2 \times 3^2 \times 5^0 \times 7^1 \times 11^0 \times 13^0 \times 17^0 \times 19^1 \\ &= \left(2^2, 3^2, 5^0, 7^1, 11^0, 13^0, 17^0, 19^1\right) \end{aligned}$$

Prime counting function

Diciamo tale la funzione $\pi : \mathbb{Z} \longrightarrow \mathbb{N}$ che associa ad N il numero di primi minori o uguali ad N.

Teorema dei numeri primi

La funzione $p(N) = {}^{N}/_{\log(N)}$ approssima asintoticamente $\pi(N)$, nel senso che $\lim_{N \to +\infty} {}^{\pi(N)}/_{p(N)} = 1$.

Take aways

Il metodo di Pomerance

Invece di cercare a, b con $a^2 - b^2 = N$, cerco diversi a_i^2 tali che il prodotto dei rispettivi b_i sia un quadrato modulo N.
Questo richiede meno passaggi.

Exponent vectors

Descrivono tutti gli interi B-regolari.
Ovviamente $\pi(B)$, per B soglia di regolarità fissata, è la lunghezza dell'exponent vector.

$$N = 1771 \qquad (B = 13)$$

$$43^2 \equiv 78 = (2^1, 3^1, 5^0, 7^0, 11^0, 13^1) \mod N$$

$$75^2 \equiv 312 = (2^3, 3^1, 5^0, 7^0, 11^0, 13^1) \mod N$$

$$(43 \times 75)^2 \equiv (78 \times 312) = (2^4, 3^2, 5^0, 7^0, 11^0, 13^2) \mod N$$

Take away

Per ottenere un quadrato, gli esponenti nell'exponent vector devono essere pari. Possiamo lavorare in modulo 2.

Quadratic Sieve

Spazio vettoriale di riferimento

Per una soglia B fissata, detto $k = \pi(B)$, possiamo vedere l'exponent vector $\left(p_1^{e_1}, \dots, p_k^{e_k}\right)$ come vettore di soli esponenti e basi prime sottintese: $(e_1, \dots, e_k)$.
In questo modo lavoriamo sullo $\mathbb{Z}_2$-spazio vettoriale di dimensione $k = \pi(B)$ costituito dagli interi B-regolari.

Possiamo sfruttare tutte le strutture dell'algebra lineare per risolvere l'ultima delle questioni: quanti a_i devo calcolare?

Teorema

Fissato B e dati k interi B-regolari $m_1, \ldots, m_k$ distinti, dove $k > \pi(B)$, il prodotto di una loro sottosequenza è un quadrato.

Teorema

Fissato B e dati k interi B-regolari $m_1, \dots, m_k$ distinti, dove $k > \pi(B)$, il prodotto di una loro sottosequenza è un quadrato.

Dalla teoria vista, questo equivale a scrivere $m_i = \left(e_1^{(i)}, \dots, e_n^{(i)}\right)$ e notare che il seguente sistema lineare ammette soluzione.

$$\begin{cases} x_1 e_1^{(1)} + \cdots + x_k e_1^{(k)} \equiv 0 \mod 2 \\ x_1 e_2^{(1)} + \cdots + x_k e_2^{(k)} \equiv 0 \mod 2 \\ \quad \vdots \\ x_1 e_{\pi(B)}^{(1)} + \cdots + x_k e_{\pi(B)}^{(k)} \equiv 0 \mod 2 \end{cases}$$

Infatti stiamo lavorando in uno spazio di dimensione $\pi(B)$ ed abbiamo una combinazione lineare di $k > \pi(B)$ elementi.

Algoritmo per QS

Dato N intero positivo da valutare:

1. scegliere una soglia B, che limiterà la quantità di calcoli e costruirà la struttura lineare del problema
2. partendo ad esempio da $x = \lceil\sqrt{N}\rceil$, calcolare diversi a_i in $\{x, x \pm 1, x \pm 2, \ldots\}$ in quantità maggiore di $\pi(B)$
3. calcolare b_i come $a_i^2 \mod N$ (**non sono necessariamente quadrati**) ed i loro exponent vectors v_i modulo N
4. risolvere il sistema lineare visto per individuare l'insieme J tale che $a^2 = b^2 \mod N$, dove $b^2 = \prod_{i \in J} b_i$ ed $a^2 = \prod_{i \in J} a_i^2$
5. uno tra $\gcd(N, a - b)$ ed $\gcd(N, a + b)$ potrebbe essere un fattore (**non necessariamente proprio**)

1. $N = 1771$; scegliere una soglia B, che limiterà la quantità di calcoli e costruirà la struttura lineare del problema

$$B = 13 \qquad \pi(B) = 6$$

2. partendo ad esempio da $x = \lceil \sqrt{N} \rceil$, calcolare diversi a_i in $\{x, x \pm 1, x \pm 2, \dots\}$ in quantità maggiore di $\pi(B)$

$$
\begin{aligned}
(a_1)^2 = 43^2 &\equiv 78 &= (2^1, 3^1, 5^0, 7^0, 11^0, 13^1) &\mod N \\
(a_2)^2 = 44^2 &\equiv 165 &= (2^0, 3^1, 5^1, 7^0, 11^1, 13^0) &\mod N \\
(a_3)^2 = 49^2 &\equiv 630 &= (2^1, 3^2, 5^1, 7^1, 11^0, 13^0) &\mod N \\
(a_4)^2 = 50^2 &\equiv 729 &= (2^0, 3^6, 5^0, 7^0, 11^0, 13^0) &\mod N \\
(a_5)^2 = 56^2 &\equiv 1365 &= (2^0, 3^1, 5^1, 7^1, 11^0, 13^1) &\mod N \\
(a_6)^2 = 73^2 &\equiv 16 &= (2^4, 3^0, 5^0, 7^0, 11^0, 13^0) &\mod N \\
(a_7)^2 = 75^2 &\equiv 312 &= (2^3, 3^1, 5^0, 7^0, 11^0, 13^1) &\mod N
\end{aligned}
$$

3. calcolare b_i come $a_i^2 \mod N$ (**non sono necessariamente quadrati**) ed i loro exponent vectors v_i in modulo N

4. risolvere il sistema lineare visto per individuare l'insieme J tale che $a^2 = b^2 \mod N$, dove $b^2 = \prod_{i \in J} b_i$ ed $a^2 = \prod_{i \in J} a_i^2$

$$\begin{cases} x_1 + x_3 + x_7 \equiv 0 \mod 2 \\ x_1 + x_2 + x_5 + x_7 \equiv 0 \mod 2 \\ x_2 + x_3 + x_5 \equiv 0 \mod 2 \\ x_3 + x_5 \equiv 0 \mod 2 \\ x_2 \equiv 0 \mod 2 \\ x_1 + x_5 + x_7 \equiv 0 \mod 2 \end{cases}$$

$$\begin{cases} x_1 + x_3 + x_7 \equiv 0 \mod 2 \\ x_1 + x_5 + x_7 \equiv 0 \mod 2 \\ x_3 + x_5 \equiv 0 \mod 2 \\ x_2 \equiv 0 \mod 2 \end{cases}$$

Ad esempio la soluzione trovata in precedenza corrisponde ad $\underline{x} = (1, 0, 0, 0, 0, 0, 1)$, ovvero a_1, a_7, che rispetta il sistema.

$$\begin{aligned} (a_1)^2 = 43^2 \equiv \ \ 78 \ &= \left(2^1, 3^1, 5^0, 7^0, 11^0, 13^1\right) \mod N \\ (a_7)^2 = 75^2 \equiv 312 \ &= \left(2^3, 3^1, 5^0, 7^0, 11^0, 13^1\right) \mod N \end{aligned}$$

5. uno tra $\gcd(N, a-b)$ ed $\gcd(N, a+b)$ potrebbe essere un fattore (**non necessariamente proprio**)

$$a = 43 \times 75, \qquad b = \left(2^4, 3^2, 5^0, 7^0, 11^0, 13^2\right)$$

$$\gcd(N, a-b) = 11$$

$$\gcd(N, a+b) = 161$$

$$1771 = 7 \times 11 \times 23$$

Conclusioni

Nascita di QS

L'idea del trovare sottosequenze degli a_i è del matematico inglese Dixon: Pomerance discute l'esistenza di un metodo per migliorare l'efficacia della ricerca.

"The time to factor n is now about $\exp\sqrt{(p \log n \log\log n)}$*; namely, the factor* $\sqrt{2}$ *in the exponent is missing. Is this a big deal? You bet. [...] And so was born the quadratic sieve method as a complexity argument and with no numerical experiments."*

(Carl Pomerance, *A Tale of Two Sieves*)

Rompere RSA?

nome	anno	numero cifre	metodo fattorizzazione
RSA-100	1991	100	multiple-polynomial quadratic sieve
RSA-129	1994	129	quadratic sieve
RSA-130	1996	130	number field sieve
RSA-150	2004	150	general number field sieve
RSA-140	2020	140	quadratic sieve

nome	anno	numero cifre	metodo fattorizzazione
RSA-100	1991	100	multiple-polynomial quadratic sieve
RSA-129	1994	129	quadratic sieve
RSA-130	1996	130	number field sieve
RSA-150	2004	150	general number field sieve
RSA-140	2020	140	quadratic sieve

Lunghezza delle chiavi RSA

RSA attualmente usa interi da 1024 a 4096 bits.
Il tempo necessario per fattorizzare RSA-140 è stato 2000 MIPS-years.

Take away

Rompere RSA

Dopo le sue prime implementazioni ha effettivamente reso non affidabili i sistemi a 100 bit, ma è sufficiente aumentare la lunghezza delle chiavi.

Take aways

Rompere RSA

Dopo le sue prime implementazioni ha effettivamente reso non affidabili i sistemi a 100 bit, ma è sufficiente aumentare la lunghezza delle chiavi.

Multiple Polynomial Quadratic Sieve

Un singolo polinomio $x^2 - N$ non fornisce facilmente sufficienti interi B-regolari, ma è possibile utilizzare diversi polinomi nella forma $(ax + b)^2 - N$ per a, b parametri. Questo permette di distribuire il calcolo in parallelo.