Malleability of RSA moduli

Luis Dieulefait and Jorge J. Urroz, UB and UPC, Barcelona

Santander, February, 2019

Problem. (Malleability of Factoring) Given and RSA modulus n find another integer n' so that the factorization of n' will help to factorize n.

Problem. (Malleability of Factoring) Given and RSA modulus n find another integer n' coprime to n :) so that the factorization of n' will help to factorize n.

Conjecture. Factoring is not malleable.

Conjecture. Factoring is not malleable.

Theorem. Given any $n = pq$ RSA modulus there exist another integer n' so that factoring n' allow us to factor n in polynomial time.

Conjecture. Factoring is not malleable.

Theorem. Given any $n = pq$ RSA modulus there exist another integer n' so that factoring n' allow us to factor n in polynomial time.

$n' = 2^n - 1$

A particular case

Theorem (L. Dieulefait and J. Jiménez Urroz, 2009)

Let $n = pq$ $z < p, q < 2z$, be and RSA modulus such that either we have $2^{p-1} \not\equiv 1 \pmod{q}$ or $2^{q-1} \not\equiv 1 \pmod{p}$ and let $n' = 2^n - 1$. Then, with the factorization of n' we can find a prime divisor of n in polynomial time.

A particular case

Theorem (L. Dieulefait and J. Jiménez Urroz, 2009)

Let $n = pq$ $z < p, q < 2z$, be and RSA modulus such that either we have $2^{p-1} \not\equiv 1 \pmod q$ or $2^{q-1} \not\equiv 1 \pmod p$ and let $n' = 2^n - 1$. Then, with the factorization of n' we can find a prime divisor of n in polynomial time.

Proof To factor n we use an oracle $\mathcal{O}$ that allow us to factor any given n‘ coprime to n. Let $S = \{r \pmod n \neq 1, r | n', \text{ prime}\}$

A particular case

Theorem (L. Dieulefait and J. Jiménez Urroz, 2009)

Let $n = pq$ $z < p, q < 2z$, be and RSA modulus such that either we have $2^{p-1} \not\equiv 1 \pmod q$ or $2^{q-1} \not\equiv 1 \pmod p$ and let $n' = 2^n - 1$. Then, with the factorization of n' we can find a prime divisor of n in polynomial time.

Proof To factor n we use an oracle $\mathcal{O}$ that allow us to factor any given n‘ coprime to n. Let $S = \{r \pmod n \neq 1, r | n', \text{ prime}\}$

Algorithm.

- Send n' in binary form to $\mathcal{O}$.
- Take $r \in S$ and compute $(r - 1, n) = p$.

Step 1. There exist such r. Indeed if every prime of $2^n - 1$ is 1 modulo n then $2^n - 1 \equiv 1 \pmod{n}$ or $2^{n-1} \equiv 1 \pmod{n}$

$$2^{n-1} \equiv 1 \pmod{p}, \text{ and } 2^{n-1} \equiv 1 \pmod{q}$$

But

$$2^{n-1} = 2^{(p-1)q+q-1} \equiv 2^{q-1} \pmod{p}$$

So,

$$2^{q-1} \equiv 1 \pmod{p}, \text{ and } 2^{p-1} \equiv 1 \pmod{q}$$

Step 1. There exist such r. Indeed if every prime of $2^n - 1$ is 1 modulo n then $2^n - 1 \equiv 1 \pmod{n}$ or $2^{n-1} \equiv 1 \pmod{n}$

$$2^{n-1} \equiv 1 \pmod{p}, \text{ and } 2^{n-1} \equiv 1 \pmod{q}$$

But

$$2^{n-1} = 2^{(p-1)q+q-1} \equiv 2^{q-1} \pmod{p}$$

So,

$$2^{q-1} \equiv 1 \pmod{p}, \text{ and } 2^{p-1} \equiv 1 \pmod{q}$$

Step 2. $2^n \equiv 1 \pmod{r}$ and $2^{r-1} \equiv 1 \pmod{r}$ Hence

$$2^{(n,r-1)} \equiv 1 \pmod{r}$$

and $(n, r-1) \neq 1, n$. Note that $(n, r-1) = (n, r \pmod{n} - 1)$

Pseudoprimes

Definition. An integer n so that $2^n - 1 \equiv 1 \pmod{n}$ is called a pseudoprime.

Pseudoprimes

Definition. An integer n so that $2^n - 1 \equiv 1 \pmod{n}$ is called a pseudoprime.

The previous algorithm does not work for pseudoprime modulus.

Pseudoprimes

Definition. An integer n so that $2^n - 1 \equiv 1 \pmod{n}$ is called a pseudoprime.

The previous algorithm does not work for pseudoprime modulus.

Are there any?...

Pseudoprimes

Definition. An integer n so that $2^n - 1 \equiv 1 \pmod{n}$ is called a pseudoprime.

The previous algorithm does not work for pseudoprime modulus.

Are there any?...well... yes 341 is the smallest

Pseudoprimes

Definition. An integer n so that $2^n - 1 \equiv 1 \pmod{n}$ is called a pseudoprime.

The previous algorithm does not work for pseudoprime modulus.

Are there any?...well... yes 341 is the smallest

Are there infinitely many pseudoprimes?

Pseudoprimes

Definition. An integer n so that $2^n - 1 \equiv 1 \pmod{n}$ is called a pseudoprime.

The previous algorithm does not work for pseudoprime modulus.

Are there any?...well... yes 341 is the smallest

Are there infinitely many pseudoprimes?

Theorem. (Alford, Granville, Pomerance, 1994) There are infinitely many Charmichael numbers.

A Charmichael number is a composite number n such that $b^{n-1} \equiv 1 \pmod{n}$ for all $(b, n)) = 1$. Example $561 = 3 \cdot 11 \cdot 17$.

Theorem. (Pomerance, 1981) Given $x > 0$, the number of pseudoprimes up to x is less than

$$x \exp(-\frac{1}{2} \log x \log \log \log x / \log \log x)$$

Theorem. (Pomerance, 1981) Given $x > 0$, the number of pseudoprimes up to x is less than

$$x \exp(-\frac{1}{2} \log x \log\log\log x / \log\log x)$$

Proposition For large z, the number of RSA moduli $n = pq$, $z < p, q < 2z$ pseudoprimes are less than

$$\left(\frac{z}{\log z}\right)^2 \frac{(\log\log\log z)^2}{\log z}$$

Proof.
$2^{(p-1,q-1)} \equiv 1 \pmod{n}$ not possible if $(p-1, q-1) < \log z$. Let $\pi(d,z) = |\{p \equiv 1 \pmod{d}, z < p < 2z \text{prime }\}|$.

$$\sum_{\substack{z<p,q<2z \\ (p-1,q-1)>\log z}} 1 = \sum_{\log z<d<z} \pi(d,z)^2 \sim \sum_{\log z<d<z} \left(\frac{z}{\varphi(d)\log z}\right)^2$$

Since

$$\varphi(d) = d\prod_{p|d}\left(1-\frac{1}{p}\right) > d\prod_{p<\log d}\left(1-\frac{1}{p}\right) > \frac{Cd}{\log\log d}$$

$$\sum_{\log z<d<z} \frac{1}{\varphi(d)^2} < c\sum_{\log z<d<z} \frac{\log\log d}{d^2} < \frac{c(\log\log\log z)^2}{\log z}.$$

Theorem. (Barban-Davenport-Halberstam, 1963-1966)

$$\sum_{d \leq z^{1-\varepsilon}} \left| \psi(d, z) - \frac{z}{\varphi(d)} \right|^2 \ll \frac{z^2}{(\log z)^A},$$

with a constant depending only in ε and A.

Primitive roots and the general case.

To avoid the pseudoprime moduli, we will choose another integer m and $n' = m^n - 1$ with a prime factor not 1 modulo n.

Primitive roots and the general case.

To avoid the pseudoprime moduli, we will choose another integer m and $n' = m^n - 1$ with a prime factor not 1 modulo n.

Definition. Given a prime p, a primitive root modulo p is an integer so that $< m >= \mathbb{F}_p^*$. $m^d \not\equiv 1 \pmod{p}$ for any $d < p - 1$.

Primitive roots and the general case.

To avoid the pseudoprime moduli, we will choose another integer m and $n' = m^n - 1$ with a prime factor not 1 modulo n.

Definition. Given a prime p, a primitive root modulo p is an integer so that $< m >= \mathbb{F}_p^*$. $m^d \not\equiv 1 \pmod{p}$ for any $d < p - 1$.

If $n = pq$, $q < p$ and m is a primitive root modulo p, $m^{n-1} \not\equiv 1 \pmod{n}$, since $m^{q-1} \not\equiv 1 \pmod{p}$.

Question. How difficult is to find a primitive root modulo p without knowing p?.

There are $\varphi(p-1)$ primitive roots modulo p. Hence the probability to find one is

$$\frac{\varphi(p-1)}{p-1} = \prod_{q|p-1}\left(1-\frac{1}{q}\right) > \prod_{q<\log p}\left(1-\frac{1}{q}\right) > \frac{c}{\log\log p}.$$

Question. How difficult is to find a primitive root modulo p without knowing p?.

There are $\varphi(p-1)$ primitive roots modulo p. Hence the probability to find one is

$$\frac{\varphi(p-1)}{p-1} = \prod_{q|p-1}\left(1-\frac{1}{q}\right) > \prod_{q<\log p}\left(1-\frac{1}{q}\right) > \frac{c}{\log\log p}.$$

In particular a random set of size $C\log\log p$ should have positive probability to contain a primitive root modulo p. Since $p < n$ a set of size $C\log\log n$ should have positive probability to contain a primitive root modulo p. The probability for a set of this size to contain no primitive roots is

$$\left(1-\frac{c}{\log\log p}\right)^{C\log\log p} \sim e^{-Cc}.$$

Results.

(E. Bach, 1997) Let $g(p)$ the least primitive root modulo p.

$$g(p) \leq e^{\gamma} \log p (\log \log p)^2 (1 + \varepsilon).$$

Results.

(E. Bach, 1997) Let $g(p)$ the least primitive root modulo p.

$$g(p) \leq e^{\gamma} \log p (\log\log p)^2 (1+\varepsilon).$$

Theorem (V. Shoup, 1992) Under GRH, $g(p) \ll (\log p)^6$

Results.

(E. Bach, 1997) Let $g(p)$ the least primitive root modulo p.

$$g(p) \leq e^{\gamma} \log p (\log\log p)^2 (1+\varepsilon).$$

Theorem (V. Shoup, 1992) Under GRH, $g(p) \ll (\log p)^6$

Conjecture (Artin, 1927) Any given integer a not $1, -1$ or a perfect square is a primitive root for a positive proportion of primes $\prod_q \left(1 - \frac{1}{q(q-1)}\right) \sim 0.37395.$

Results.

(E. Bach, 1997) Let $g(p)$ the least primitive root modulo p.

$$g(p) \leq e^{\gamma} \log p (\log \log p)^2 (1 + \varepsilon).$$

Theorem (V. Shoup, 1992) Under GRH, $g(p) \ll (\log p)^6$

Conjecture (Artin, 1927) Any given integer a not $1, -1$ or a perfect square is a primitive root for a positive proportion of primes $\prod_q \left(1 - \frac{1}{q(q-1)}\right) \sim 0.37395$.

Theorem. (Heath-Brown, 1986) Among $3, 5, 7$ there is a primitive root for infinitely many p

For each integer m set $n'_m = (m^n - 1)/(m-1)$, and $S_m = \{r \pmod{n} \neq 1 : r \text{ prime } r | n'_m\}$.

For each integer m set $n'_m = (m^n - 1)/(m-1)$, and
$S_m = \{r \pmod{n} \neq 1 : r \text{ prime } r | n'_m\}$.

Algorithm The m-ary representation of n is c independent of m

- m=2
- Send (c, m) to $\mathcal{O}$
- $S =$, $m = m + 1$. Return
- take $r \in S$ and compute $d = (r - 1, n)$.

For each integer m set $n'_m = (m^n - 1)/(m-1)$, and
$S_m = \{r \pmod{n} \neq 1 : r \text{ prime } r | n'_m\}$.

Algorithm The m-ary representation of n is c independent of m

- m=2
- Send (c, m) to $\mathcal{O}$
- $S =$, $m = m + 1$. Return
- take $r \in S$ and compute $d = (r-1, n)$.

Theorem (L. Dieulefait and J. Jiménez Urroz)

Let $n = pq$ $z < p, q < 2z$, be and RSA modulus . Then, under GRH the previous algorithm gives a prime divisor of n in polynomial time.

Proof

Lemma Let $n = pq$ and RSA modulus and m such that $(m-1, n) = 1$. Then $(n'_m, m-1) = 1$. If $r|(n'_m, m-1)$, then $n'_m = \sum_{j=0}^{n-1} m^j \equiv n \pmod{r}$.

Step 1. There exist such r. Indeed if every prime of n'_m is 1 modulo n then $m'_n \equiv 1 \pmod{n}$ or $m^{n-1} \equiv 1 \pmod{n}$

$$m^{n-1} \equiv 1 \pmod{p}, \text{ and } m^{n-1} \equiv 1 \pmod{q}$$

But

$$m^{n-1} = 2^{(p-1)q+q-1} \equiv m^{q-1} \pmod{p}$$

which is not possible.

Proof

Lemma Let $n = pq$ and RSA modulus and m such that $(m-1, n) = 1$. Then $(n'_m, m-1) = 1$. If $r | (n'_m, m-1)$, then $n'_m = \sum_{j=0}^{n-1} m^j \equiv n \pmod r$.

Step 1. There exist such r. Indeed if every prime of n'_m is 1 modulo n then $m'_n \equiv 1 \pmod n$ or $m^{n-1} \equiv 1 \pmod n$

$$m^{n-1} \equiv 1 \pmod p, \text{ and } m^{n-1} \equiv 1 \pmod q$$

But

$$m^{n-1} = 2^{(p-1)q+q-1} \equiv m^{q-1} \pmod p$$

which is not possible.

Step 2. $m^n \equiv 1 \pmod r$ and $m^{r-1} \equiv 1 \pmod r$. Hence

$$m^{(n,r-1)} \equiv 1 \pmod r$$

and $(n, r-1) \neq 1, n$. Note that $(n, r-1) = (n, r \pmod n) - 1)$

And... ¿Without cheating?

And... ¿Without cheating? We are looking for a number n' which helps to factorize n.

And... ¿Without cheating? We are looking for a number n' which helps to factorize n.

Defintion. Given a field K. An elliptic curve over K is the set

$$E/K := \left\{(x, y) \in K \times K \,:\, y^2 = x^3 + ax + b, a, b \in K\right\} \cup \{O\}$$
$$4a^3 + 27b^2 \neq 0.$$

And... ¿Without cheating? We are looking for a number n' which helps to factorize n.

Defintion. Given a field K. An elliptic curve over K is the set

$$E/K := \left\{(x,y) \in K \times K \,:\, y^2 = x^3 + ax + b, a, b \in K\right\} \cup \{O\}$$
$$4a^3 + 27b^2 \neq 0.$$

Theorem. $E/\mathbb{F}_q$ is an abelian group of size

$$|E/\mathbb{F}_q| = q + 1 - a_q$$

where

$$|a_q| \leq 2\sqrt{q}.$$

Defintion. Given an integer $n = pq$ an elliptic curve modulo n is the set

$$E_n := E/\mathbb{F}_p \times E/\mathbb{F}_q.$$

Defintion. Given an integer $n = pq$ an elliptic curve modulo n is the set

$$E_n := E/\mathbb{F}_p \times E/\mathbb{F}_q.$$

$$|E_n| = |E/\mathbb{F}_p| \times |E/\mathbb{F}_q|.$$

Defintion. Given an integer $n = pq$ an elliptic curve modulo n is the set

$$E_n := E/\mathbb{F}_p \times E/\mathbb{F}_q.$$

$$|E_n| = |E/\mathbb{F}_p| \times |E/\mathbb{F}_q|.$$

Lemma. Let $n = pq$ with $p \approx q$. Then,

$$||E_n| - n| \leq cn^{3/4}.$$

Theorem (L. Dieulefait and J. Jiménez Urroz, 2019)

Let $n = pq$, and E_n and elliptic curve modulo n. Then knowing $|E_n|$ we can factor n in polynomial time.

Theorem (L. Dieulefait and J. Jiménez Urroz, 2019)

Let $n = pq$, and E_n and elliptic curve modulo n. Then knowing $|E_n|$ we can factor n in polynomial time.

Proof.

Theorem (J. Cilleruelo-J. Jiménez Urroz)

In an arc of lenght $cn^{1/4}$ of the hyperbola $xy = n$ with $x, y \geq n^{1/2}$ there are at most 4 points of integer coordinates.

Theorem (L. Dieulefait and J. Jiménez Urroz, 2019)

Let $n = pq$, and E_n and elliptic curve modulo n. Then knowing $|E_n|$ we can factor n in polynomial time.

Proof.

Theorem (J. Cilleruelo-J. Jiménez Urroz)

In an arc of lenght $cn^{1/4}$ of the hyperbola $xy = n$ with $x, y \geq n^{1/2}$ there are at most 4 points of integer coordinates.

Theorem (L. Dieulefait and J. Jiménez Urroz, 2019)

Let $n = pq$, and E_n and elliptic curve modulo n. Then knowing $|E_n|$ we can factor n in polynomial time.

Proof.

Theorem (J. Cilleruelo-J. Jiménez Urroz)

In an arc of lenght $cn^{1/4}$ of the hyperbola $xy = n$ with $x, y \geq n^{1/2}$ there are at most 4 points of integer coordinates.

So, we ask the oracle for the factor of E_n of size $n^{1/2}$. Note that $p + 1 - a_p$ and $q + 1 - a_q$ are two of those points.

Theorem (L. Dieulefait and J. Jiménez Urroz, 2019)

Let $n = pq$, and E_n and elliptic curve modulo n. Then knowing $|E_n|$ we can factor n in polynomial time.

Proof.

Theorem (J. Cilleruelo-J. Jiménez Urroz)

In an arc of lenght $cn^{1/4}$ of the hyperbola $xy = n$ with $x, y \geq n^{1/2}$ there are at most 4 points of integer coordinates.

So, we ask the oracle for the factor of E_n of size $n^{1/2}$. Note that $p + 1 - a_p$ and $q + 1 - a_q$ are two of those points. Use Coppersmith algorithm to find p.

Theorem (Coppersmith)

If we know an integer $n = pq$ and we know the high order $log_2 N$ bits of p, then in polynomanl time in $\log(n)$ we can recover p and q.

Theorem

Finding the number of points of elliptic curves modulo n is equivalent to factoring n.

Theorem

Finding the number of points of elliptic curves modulo n is equivalent to factoring n.

Let $\hat{E}, \tilde{E}, \bar{E}$ the three possible twists of E. Then

$$
\begin{aligned}
E &= (p - a_p)(q - a_q) = n - pa_q - qa_p + a_p a_q \\
\hat{E} &= (p + a_p)(q + a_q) = n + pa_q + qa_p + a_p a_q, \\
\tilde{E} &= (p - a_p)(q + a_q) = n + qa_q - qa_p - a_p a_q, \\
\bar{E} &= (p + a_p)(q - a_q) = n - pa_q + qa_p - a_p a_q.
\end{aligned}
$$

Lemma

$$|E| + |\hat{E}| + |\tilde{E}| + |\bar{E}| = 4n$$
$$E\hat{E} = \tilde{E}\bar{E}.$$

Then, knowing E and $\hat{E}$, we compute its product, $M = E\hat{E}$ and its sum $L = E + \hat{E}$, and we have

$$\tilde{E}\bar{E} = M$$
$$\tilde{E} + \bar{E} = 4n - L$$

so $\tilde{E}$ and $\bar{E}$ are the solutions of the quadratic polynomial $X^2 - (4n - L)X + M$.

Lemma

$$|E| + |\hat{E}| + |\tilde{E}| + |\bar{E}| = 4n$$
$$E\hat{E} = \tilde{E}\bar{E}.$$

Then, knowing E and $\hat{E}$, we compute its product, $M = E\hat{E}$ and its sum $L = E + \hat{E}$, and we have

$$\tilde{E}\bar{E} = M$$
$$\tilde{E} + \bar{E} = 4n - L$$

so $\tilde{E}$ and $\bar{E}$ are the solutions of the quadratic polynomial $X^2 - (4n - L)X + M$.

$\gcd(E + \bar{E}, n) = p$