2.8 Small Exponents

Question: Is RSA in danger if someone chooses a small public exponent e?

The exponent $e = 1$ is nonsensical since it leaves plaintexts unencrypted.

The exponent $e = 2$ doesn't work for RSA since it is even and thus not coprime with $\lambda(n)$. Nevertheless the related RABIN cipher uses $e = 2$. Here the receiver of the message must be able to take square roots $\bmod\, n$, and this works since he knows the prime factors of n (see later). (By the way he must also be able to recognize the true plaintext among several different square roots.)

Same Message for Several Receivers

For RSA the smallest potentially suited exponent is $e = 3$. However it enables an attack that applies as soon as someone sends the same message a to *three* different receivers A, B, and C. Let their public keys be $(n_\text{A}, 3)$, $(n_\text{B}, 3)$, and $(n_\text{C}, 3)$. Assume the modules n_A, n_B, and n_C are pairwise coprime, otherwise the attacker factorizes at least two of them and reads a. Then (with some luck) she intercepts three ciphertexts

$$c_\text{A} = a^3 \bmod n_\text{A}, \quad c_\text{B} = a^3 \bmod n_\text{B}, \quad c_\text{C} = a^3 \bmod n_\text{C},$$

with $0 \le a < n_\text{A}, n_\text{B}, n_\text{C}$, thus $a^3 < n_\text{A} n_\text{B} n_\text{C}$. Using the chinese remainder algorithm she constructs an integer $\tilde{c} \in \mathbb{Z}$ with

$$0 \le \tilde{c} < n_\text{A} n_\text{B} n_\text{C}$$

such that

$$\tilde{c} \equiv c_\text{X} \bmod n_\text{X} \quad \text{for X} = \text{A}, \text{B}, \text{C}.$$

By uniqueness $\tilde{c} = a^3$ in $\mathbb{Z}$. So she computes $a = \sqrt[3]{\tilde{c}}$ by taking the integer root in $\mathbb{Z}$. This is an efficient procedure. (In this situation she doesn't succeed with computing the private exponents.)

This attack obviously generalizes to other "small" shared public exponents e: If the same message is sent to e different people, then everybody can read it. This attack is not completely unrealistic: Think for example of fixed "protocol information" at the beginning of a larger message. Even in classical cryptography an important maxim was: *Never encrypt the same plaintext with different keys.*

In practice the exponent $e = 2^{16} + 1 = 65537$ is considered as sufficiently secure for "normal" situations.

Stereotypical Message Parts

Consider the key parameters (n, e, d). Imagine an attack with known plaintext that reads:

`Der heutige Tagesschluessel ist:********`

("The master key for today is: ...", example by Julia Dietrichs) with known (stereotypical) 32 byte part "`Der heutige Tagesschluessel ist:`", and unknown 8 byte part "`********`".

This message is encoded by the 8-bit character code ISO-8859-1 (used for German texts) as a sequence of 40 bytes or 320 bits, and for encryption by RSA interpreted as an integer $a \in [0 \ldots n-1]$ (assume n has more then 320 bits, and $e = 3$). Decompose a as $a = b + x$ where b corresponds to the known, and x, to the unknown part. Since the latter forms the end of the message and consists of 64 bits we know $x < 2^{64}$. Encryption yields the ciphertext

$$c = a^e \bmod n = (b+x)^e \bmod n.$$

Hence the secret x is a root of the polynomial

$$(T+b)^e - c \in (\mathbb{Z}/n\mathbb{Z})[T].$$

At first sight this observation doesn't seem alarming since we know of no general efficient algorithms that compute roots. However algorithms for certain special cases exist, for instance:

> COPPERSMITH's **algorithm**
> Let $f \in (\mathbb{Z}/n\mathbb{Z})[T]$ be a polynomial of degree r. The algorithm finds all roots x of f with $0 \le x < \sqrt[r]{n}$ (or proves that there are none).
> The execution time is polynomial in $\log n$ and r.
> (The algorithm uses the "LLL algorithm" for reduction of lattice bases.)

In our example n has at least 321 bits, and $e = 3$. Thus the algorithm outputs x since $x^3 < 2^{192} < 2^{320} < n$.

This is only a simple example of a larger class of attacks for special situations that amount to a computation of roots $\bmod n$.

Exercise. Modify the attack for a situation where the unknown part of the plaintaxt consists of some contiguous letters surrounded by known plaintext sequences.