On Squaring Modulo Mersenne Numbers

David Naccache[1] and Ofer Yifrach-Stav[1]

DIÉNS, ÉNS, CNRS, PSL University, Paris, France
45 rue d'Ulm, 75230, Paris CEDEX 05, France
ofer.friedman@ens.fr, david.naccache@ens.fr

Abstract. During the design of a new primitive inspired by Squash we accidentally stumbled on the observation described in this note.
Let n be a k-bit Mersenne number whose factors are unknown. Consider an ℓ-bit secret number $x = 2^{k/2}a+b$. We observe that there are parameter configurations where a chunk of the value b^2 is leaked even if $k < 2\ell$.
This observation does not endanger any known scheme and in particular not Squash.

1 The observation

During the design of a new lightweight primitive inspired by Squash [3] we accidentally stumbled on the observation described in this short note.

The initial intent was to get an arbitrary input m of k bits whose entropy is $k' \leq k$ and hash m into a k-bit output c where each bit has entropy k'/k. A good candidate for doing so is modular squaring. In particular, working modulo a Mersenne number has notable computational advantages. This note shows that squaring modulo a Mersenne number does not provide this desirable entropy spreading property, even when $m^2 > n$ for some parameter configurations as some of the bits of c may depend *only* on specific bits of m.

The way in which this was accidentally discovered is interesting by its own right. The designed hash function was used as a building-block in an IoT malware analysis prototype. Packets including metadata and data were fed into a GAN that had to learn normal protocol behavior. Because part of the hashed data (the LSBs) consisted of constant system commands while the other part was a random nonce (the MSBs) to our surprised the GAN declared that part of the hash was... part of the protocol's semantics. This happened during the covariance detection phase where data is input into a filter reacting to the repeated appearance of sufficiently large patterns in the dataset. Looking into the reason for which this happened, we discovered the arithmetic phenomenon described in this note.

Let n be a k-bit Mersenne number $n = 2^k - 1$ whose factors are unknown. Consider an ℓ-bit secret number $x = 2^{k/2}a + b$. Although k is prime (and hence odd) we simplify it here as an even number to avoid managing unbalanced halves.

An attacker is given $c = x^2 \bmod n$. What can s/he learn about a and b individually?

We have:

$$c = x^2 = (2^{k/2}a + b)^2 = 2^k a^2 + 2^{1+k/2}ab + b^2 = 2^{1+k/2}ab + a^2 + b^2 \bmod n$$

Denoting $\Delta = ab$ and $\Gamma = a^2 + b^2$ we get

$$c = 2^{1+k/2}\Delta + \Gamma \bmod n$$

Γ is the modular sum of a k-bit number (b^2) and a $2\ell - k$ bit number (a^2). We observe that if $2\ell - k < k$, i.e. $\ell < k$, then Γ has good chances to be in $\mathbb{Z}$. Note that even if Γ exceeds n by one or two bits, those will wrap around and blur only a few LSBs of Γ leaving the remaining bits of $\Gamma \bmod n$ identical to those of Γ in $\mathbb{Z}$.

We now turn to analyzing the effect of adding to Γ the quantity $2^{1+k/2}\Delta$. We start by observing that Δ is of size size ℓ. We distinguish in Δ two parts Δ_H (of size $k/2$) and Δ_L (of size $\ell - k/2$), i.e. $\Delta = 2^{\ell-k/2}\Delta_H + \Delta_L$.

We see that the addition of $2^{1+k/2}\Delta$ to Γ will blur the $\ell - k/2$ MSBs (because of Δ_L) and the $k/2$ LSBs (because of the wrapping of Δ_H). This will leave $k - \ell$ bits of Γ exposed.

Γ is essentially of size $2\log_2 b$ and is nothing but b^2 with its $2\ell - k$ (size of a^2) LSBs blurred. All in all it appears that c features the bits of b^2 between positions $\max(k/2, 2\ell - k)$ and $3k/2 - \ell$ which therefore depend only on b which, in our application, was a constant assortment of commands sent to the device.

It goes without saying that the home-made hash function was replaced by a standard Squash. This note confirms that the use of moduli of the form $2^k \pm r$ where r is small should always be analyzed carefully as episodically weaknesses due to this choice arise, e.g. [1] or [2][1].

2 Example

Let $n = 2^{1009} - 1$ and fix randomly:

```
a = 00000004 b6b610e6 4e2d3680 139cca0b
b = 13fceaff 599d4f4e fa14b5c7 82d2f55c
    05c2c3ee 108fdd03 3f161099 237cb257
    24ac47a7 be03b21d d293d5e5 43e83374
    47dd3589 960fc891 669477c6 b7498278
```

Indeed, the quantities $c = (b + 2^{509}a)^2 \bmod n$ and b^2 coincide in their central bits as shown in red:

[1] Note that while very different, the observation in this note is somewhat reminiscent of [2] page 10, section 4.2.

$$
\begin{aligned}
c = {} & \texttt{1887fa50 303e3d1a c6c9b433 0e0087f4} \\
& \texttt{256fbc49 1d4628c7 7c45ca72 bbb65a96} \\
& \texttt{47c964b4 23ff555e 22cbea2f 5e8eaaca} \\
& \texttt{16eeabeb 7e988c3a cb3289e3 3136b061} \\
& \texttt{602e98ff dbd6560e e2d43566 aa9ef7b5} \\
& \texttt{6207638c 656dd780 5110d904 bfc4a799} \\
& \texttt{fe09cce3 01ba1cc 7bc61ac93 ec41c55b} \\
& \texttt{882cad79 cd602f49 ec00aa8f 3a06b} \\
b^2 = {} & \texttt{184c165b d9601185 a8e14d91 ab8e0cfa} \\
& \texttt{0cac609f 8800030f 0327a865 e25c1d21} \\
& \texttt{957e2e15 cf5a290e 1fdaa07f bb68064c} \\
& \texttt{b5942217 ba885076 f8a3f8ba 440a1061} \\
& \texttt{602e98ff dbd6560e e2d43566 aa9ef7b5} \\
& \texttt{6207638c 656dd780 5110d904 bfc4a799} \\
& \texttt{fe09cce2 fef319ca a5a387bc 1473eb06} \\
& \texttt{7c6fd770 e258cdaf b8f433ae e1907}
\end{aligned}
$$

References

1. N. Borisov, M. Chew, R. Johnson, and D. Wagner. Multiplicative differentials. In *Fast Software Encryption, 9th International Workshop, FSE 2002, Leuven, Belgium, February 4-6, 2002, Revised Papers*, volume 2365 of *Lecture Notes in Computer Science*, pages 17–33. Springer, 2002.
2. K. Ouafi and S. Vaudenay. Smashing SQUASH-0. In A. Joux, editor, *Advances in Cryptology - EUROCRYPT 2009*, pages 300–312, Berlin, Heidelberg, 2009. Springer Berlin Heidelberg.
3. A. Shamir. SQUASH – a new MAC with provable security properties for highly constrained devices such as RFID tags. In K. Nyberg, editor, *Fast Software Encryption*, pages 144–157, Berlin, Heidelberg, 2008. Springer Berlin Heidelberg.