Resistance of Randomized Projective Coordinates Against Power Analysis

W. Dupuy, S. Kunz-Jacques

DCSSI Crypto Lab
Paris, France

September 7, 2005

Outline

1. Background
   - Elliptic Curves
   - Randomized Projective Coordinates
2. Attack on Optimized Curves
   - Optimized Parameters
   - Target of the Attack
   - Attack Methodology

Overview

- New Goubin-style side-channel attack
- Target: scalar multiplication on elliptic curves
- Chosen-ciphertext
- Defeats randomized projective coordinates countermeasure

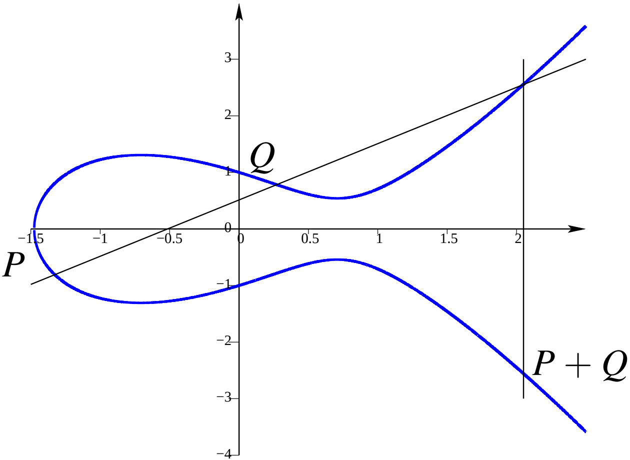
Outline

Elliptic Curves on Finite Fields

- set $\mathcal{C}$ of solutions of a non-singular cubic equation
- Choices for the ground field $\mathbb{K}$:
  $\mathbb{K} = \mathbb{F}_p$ with p a large prime $(y^2 = x^3 + a_4x + a_6)$
  or $\mathbb{K} = \mathbb{F}_{2^n}$ $(y^2 + xy = x^3 + a_2x^2 + a_6)$
- Group law on the points of the curve together with a "point at infinity" (neutral element)
- Costly operation used in crypto: $u \rightarrow uP = P + \ldots + P$, $u \in \mathbb{N}$, $P \in \mathcal{C}$

Elliptic Curve: example

$$y^2 = x^3 - 1.5x + 1$$



Randomized Projective Coordinates

- $P = (x, y) \in \mathcal{C}$ is represented by $(X, Y, Z) = (xZ, yZ, Z)$, for any $Z \in \mathbb{K} - \{0\}$
- avoids computing inverses in computations
- if Z is randomized, is a DPA countermeasure

Goubin Observation (PKC' 03)

- Despite projective randomization, if (X, Y, Z) represents (x, y), $x = 0 \Longrightarrow X = 0$ (and $y = 0 \Longrightarrow Y = 0$)

$\Longrightarrow$ points with $x = 0$ are distinguished points:

- If Hamming weights can be observed, distinguished points can be detected

Why do Distinguished Points Matter?

- Their appearance can be detected in the course of a computation

$\Longrightarrow$ Can be used to build tests of the form:

$$\{\text{ secret bit } b = 0\}$$
$$\Longleftrightarrow$$
$$\{\text{distinguished point appears}\}$$

- We build a class of distinguished points for optimized curves

Outline

Optimized Parameters used in EC Cryptography

- group law on $\mathcal{C}$: rational expressions must be computed
- point adding $(P+Q)$ or doubling $(P+P)$ cost measured in number of elementary operations in the ground field $\mathbb{K}$:$+$, $\times$, square, inverse

$\Longrightarrow$ fast operations in the ground field are needed

- one common technique: use sparse polynomials P $(\mathbb{K}=\mathbb{F}_{2^n}=\mathbb{F}_2[X]/P)$ or sparse primes $(\mathbb{K}=\mathbb{F}_p)$: modular reduction easier

Example of Optimized Parameters: fields for NIST Curves

Binary fields:

- $P_{233}(x) = x^{233} + x^{74} + 1$
- $P_{283}(x) = x^{283} + x^{12} + x^{7} + x^{5} + 1$
- ...

Prime fields:

- $p_{192} = 2^{192} - 2^{64} - 1$
- $p_{224} = 2^{224} - 2^{96} + 1$
- ...

Sparsity

- $P = X^n + 1 + \sum_{i=0}^{I} X^{a_i}$
- $p = 2^n - 1 + \sum_{i=0}^{I} \varepsilon_i 2^{a_i}, \ \varepsilon_i = \pm 1$
- sparsity: I small

Multiplication by the Generator in an Optimized Field

- $\mathbb{K} = \mathbb{F}_p$: if $p = 2^n - 1$ (Mersenne prime), multiplication by 2 = left circular shift ($2^n = 1 \bmod p$)
- $\mathbb{K} = \mathbb{F}_{2^n}$: same with $P = X^n + 1$, multiplication by X
- if $p = 2^n - 1+$ few terms, multiplication by 2 $\simeq$ left circular shift

Multiplication by the Generator in an Optimized Field

- $z = \sum \alpha_i u^i$, $\alpha_i =$ bit
- generator $u = X$ ($\mathbb{K} = \mathbb{F}_{2^n}$) or $u = 2$ ($\mathbb{K} = \mathbb{F}_p$)
- $\mathrm{H}(z)$: hamming weight of z

$$\boxed{\begin{array}{c} u \times z \simeq z <<< 1 \\ \mathrm{H}(u^\lambda \times z) \simeq \mathrm{H}(z) \text{ if } \lambda \text{ small} \end{array}}$$

Observable Point in Projective Coordinates

Suppose $P = (u^\lambda, y) \in \mathcal{C}$, λ small ($u = 2$ if $\mathbb{K} = \mathbb{F}_p$, $u = X$ if $\mathbb{K} = \mathbb{F}_{2^n}$)

For any projective representation (X, Y, Z) of P, $\mathrm{H}(X) \simeq \mathrm{H}(Z)$

Indeed, $X = u^\lambda Z$.
Like the distinguished points of Goubin, these points can be observed through hamming weights.

Target of the Attack

- A black-box performing $P \to k \times P$ on a known optimized curve;
  - k secret
  - P controlled by the attacker
- uses a standard anti-SPA scalar multiplication algorithm (eg double-and-add-always)
- randomized projective coordinates are used
- no exponent randomization

Double-and-add always algorithm

Input: $P \in \mathcal{C}, k = \sum_{i=0}^{n} k_i 2^i$ an integer
Output: $R = nP$

$R_0 \leftarrow 0$
for $i = n$ downto 0 **do**
$R_0 \leftarrow 2R_0$
$R_1 \leftarrow R_0 + P$
$R_0 \leftarrow R_{k_i}$
end for
return R_0

Attack Initiation

- Build a distinguished point P_0: find the smallest λ s.t. there exists $P_0 = (u^\lambda, y)$ on the curve (NIST recommended curves: $\lambda \leq 5$)
- input $\frac{1}{2}P_0$ to the black box
- If the MSB k_n of k is 0, P_0 is observed in the first step of double-and-add
- Knowing k_n and assuming $k_{n-1} = b$, P_0 is observed in second pass on input $\frac{1}{2k_n+b}P_0$
- ...

Attack Methodology

- Once $k_n, \ldots k_{i+1}$ are known, some μ_i can be computed s.t. P_0 is observed during pass $n-i$ on input
  - $\frac{1}{\mu_i} P_0$ if $k_i = 0$
  - $\frac{1}{\mu_i + 1} P_0$ if $k_i = 1$
- μ_i might not be coprime with the order of P_0, in that case $\mu_i + 1$ is

Detecting the point P_0

- Assume that when $P = (X, Y, Z)$ is manipulated, H(X) and H(Z) can be observed (possibly up to some noise)
- Statistical test on $U =$ H$(X) -$ H(Z)
- If $P = P_0$, H$(X) \simeq$ H(Z)
- If $P \neq P_0$ it is reasonable to expect that H(Z) and H(X) are uncorrelated

Point Detection: Basic Statistical Test

- Estimate through several measures the standard deviation of $U = \mathrm{H}(X) - \mathrm{H}(Z)$
- For a threshold t, we decide $P = P_0$ if $\sigma(U) < t$ and $P \neq P_0$ otherwise
- With probability $1/2$, $P = P_0$: compute a threshold s.t.

$$P(\text{deciding } P = P_0 | P \neq P_0) = P(\text{deciding } P \neq P_0 | P = P_0)$$

Point Detection: Better Statistical Test

- Compute the distribution of U under the hypotheses $P = P_0$, $P \neq P_0$
- Perform several experiments and choose the hypothese for which the observed values are the most likely (Neyman-Pearson test)
- Case $P \neq P_0$: computation easy (uncorrelated Hamming weights); $P = P_0$: harder, esp. in the prime field case, because of carries
- Theoretically, Neyman Pearson test is optimal
- Because of approximations made, basic test better

Simulated Results on NIST Curves

Curve	Experiments per bit	$I\lambda$
p_{192}	6	2
p_{224}	10	6
p_{256}	11	12
p_{384}	7	3
p_{521}	3	0
B_{233}	2	1
B_{283}	7	15
B_{409}	2	1
B_{571}	4	15

Table: Experiments Required for a 90% Confidence Level (no added noise)

Simulated Results and Curve Properties

- Two key parameters:
  - Number I of parasitic terms in field definition
  - λ for the distinguished point (u^λ, y) found on the curve
- Number of measures per bit required roughly $\propto I\lambda$
- In the prime (resp. binary case), $V(U) = I\lambda/2$ (resp $\simeq (I+1)\lambda/2$)

Conclusion

- Favour cryptosystems where the secret is not used for scalar multiplication
- Use exponent randomization (or more specific countermeasures)
- Need to better understand the effect of optimizations on security

BRIP (Mamiya et Al, CHES' 04)

Input: $P \in \mathcal{C}, k = \sum_{i=0}^{n} k_i 2^i$ an integer
Output: nP

$R_0 \leftarrow$ random point R
$R_1 \leftarrow -R_0$
$R_2 \leftarrow P - R_0$
for $i = n$ downto 0 **do**
  $R_0 \leftarrow 2R_0$
  $R_0 \leftarrow R_0 + R_{k_i+1}$
end for
return $R_0 + R_1$

works because R is added $2^n - \sum_{i=0}^{n-1} 2^i - 1 = 0$ times