The art of cryptography, summer 2013

Lattices and cryptography

Prof. Dr. Joachim von zur Gathen

A further cryptanalytic use of basis reduction is to break certain pseudo-random number generators.

The most popular pseudorandom generators are the *linear congruential pseudorandom generators*. We have a modulus $m \in \mathbb{N}$, two integers s, t, a *seed* $x_0 \in \mathbb{N}$, and define

$$x_i = s x_{i-1} + t \text{ in } \mathbb{Z}_m \tag{1}$$

for $i \geq 1$.

In the generator (1), we have

$$\begin{aligned} x_i &= s x_{i-1} + t \text{ in } \mathbb{Z}_m, \\ x_{i+1} &= s x_i + t \text{ in } \mathbb{Z}_m. \end{aligned}$$

In order to eliminate s and t, we subtract and find

$$x_i - x_{i+1} = s(x_{i-1} - x_i) \text{ in } \mathbb{Z}_m.$$

Similarly we get

$$x_{i+1} - x_{i+2} = s(x_i - x_{i+1}) \text{ in } \mathbb{Z}_m.$$

Multiplying by appropriate quantities, we obtain

$$\begin{aligned} (x_i - x_{i+1})^2 &= s(x_i - x_{i+1})(x_{i-1} - x_i) \\ &= (x_{i+1} - x_{i+2})(x_{i-1} - x_i) \text{ in } \mathbb{Z}_m. \end{aligned}$$

Thus from four consecutive values x_{i-1}, x_i, x_{i+1}, x_{i+2} we get a multiple

$$m' = (x_i - x_{i+1})^2 - (x_{i+1} - x_{i+2})(x_{i-1} - x_i)$$

of m.
If the required $\gcd$s are 1, then we can also compute guesses s' and t' for s and t, respectively. We can then compute the next values $x_{i+3}, x_{i+4}, \ldots$ with these guesses and also observe the generator. Whenever a discrepancy occurs, we refine our guesses.

Instead of outputting all of x_i, we only use part of it, say the top half of its bits. More generally, we take an integer approximation parameter $\alpha \geq 1$ and for $i \geq 1$ output an α-approximation y_i to x_i with

$$|x_i - y_i| \leq \alpha. \tag{2}$$

There are many such y_i, and we need a deterministic way of determining one of them. A natural choice is

$$y_i = \left\lfloor \frac{x_i}{\alpha} \right\rfloor \cdot \alpha; \tag{3}$$

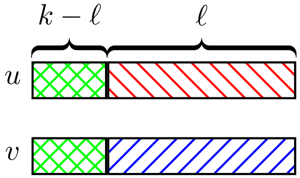
We use the symmetric system of representatives modulo m

$$R_m = \{-\lfloor m/2 \rfloor, \ldots, \lfloor (m-1)/2 \rfloor\},$$

where $u \operatorname{srem} m \in R_m$ is the representative of $u \in \mathbb{Z}$ and $u = (u \operatorname{srem} m)$ in $\mathbb{Z}_m$. For an approximation parameter α and $u \in \mathbb{R}$, the α-*vicinity* of u is the set of integers whose distance from u is at most α:

$$V_\alpha(u) = \{v \in \mathbb{Z} \colon |u - v| \leq \alpha\}. \tag{4}$$

If u and $v \in \mathbb{Z}$ are positive k-bit integers and their first $k - \ell$ bits agree, then $|u - v| < 2^{\ell+1}$ and $v \in V_{2^{\ell+1}}(u)$. But due to carries, the reverse may be false. As an example, we take $k = 6$, $0 \leq \ell \leq 4$, $47 = (101111)_2 \in V_1(48) \subseteq V_{2^\ell}(48)$, and $48 = (110000)_2 \in V_1(47) \subseteq V_{2^\ell}(47)$. But the two (or more) leading bits of the 6-bit integers 47 and 48 do not agree.


$k-\ell$
ℓ
u
v

We first show that key recovery from $y_1, \ldots, y_n$ is usually possible when $t = 0$ in (1), which we now assume. Later, we reduce the general case to this one. The unknown integers $x_1, \ldots, x_n$ satisfy

$$\begin{aligned} x_{i+1} &= s x_i \text{ in } \mathbb{Z}_m, \\ x_i &= s^{i-1} x_1 \text{ in } \mathbb{Z}_m, \text{ for } 1 \leq i \leq n. \end{aligned} \tag{5}$$

We consider the lattice $L = L_{s,m}$ spanned by the rows $a_1, \ldots, a_n \in \mathbb{Z}^n$ of the following $n \times n$ matrix:

$$A = \begin{pmatrix} m & 0 & 0 & \cdots & 0 \\ -s & 1 & 0 & \cdots & 0 \\ -s^2 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -s^{n-1} & 0 & 0 & \cdots & 1 \end{pmatrix}. \tag{6}$$

As above, we write

$$z_i = x_i - y_i \text{ with } |z_i| \leq \alpha \tag{7}$$

for each i. The z_i are unknown, and our task is to find them. (5) implies that

$$\begin{aligned} z_i = x_i - y_i &= s^{i-1}(y_1 + z_1) - y_i \\ &= s^{i-1} z_1 + (s^{i-1} y_1 - y_i) \text{ in } \mathbb{Z}_m. \end{aligned}$$

This is a set of linear congruences, but in contrast to the homogeneous congruences (5), they are inhomogeneous with (known) constants

$$c_i = s^{i-1} y_1 - y_i. \tag{8}$$

The lattice basis reduction works on n linearly independent vectors in $\mathbb{Z}^n$, and the first element b_1 of the reduced basis that it produces satisfies $\|b_1\| \leq 2^{(n-1)/2}\lambda_1(L)$. We now need a generalization which gives a bound on each $\|b_i\|$ in terms of the successive minima $\lambda_i(L)$.

THEOREM 9. *Let $L \subseteq \mathbb{R}^n$ be the lattice generated by its reduced basis $b_1, \ldots, b_\ell \in \mathbb{R}^{\ell \times n}$. Then*
$\|b_i\| \leq 2^{(\ell-1)/2} \cdot \lambda_i(L) \leq 2^{(\ell-1)/2}\lambda_\ell(L)$ *for all $i \leq \ell$.*

LEMMA 10. *There is at most one $z \in \mathbb{Z}^n$ with $Az = c$ in $\mathbb{Z}_m^n$ and*

$$\|z\| \leq \frac{m}{\lambda_n(L) \cdot (2^{(n+1)/2} + 1)}. \tag{11}$$

Given A, c, and m, one can determine in polynomial time whether such a z exists, and if so, compute it.

Lemma 10 with c as in (8) and (7) imply that if

$$\alpha \leq \frac{m}{\lambda_n(L) \cdot (2^{(n+1)/2} + 1)}, \tag{12}$$

then the approximated generator with $t = 0$ can be broken. In (12), we have to analyze $\lambda_n(L)$. More specifically, we show an upper bound on $\lambda_n(L)$ for almost all $s \in \mathbb{Z}_m$.

To this end, we need a new tool, namely the *dual lattice* L^* of a lattice $L \subseteq \mathbb{R}^n$, which is defined as

$$L^* = \{v \in \mathbb{R}^n \colon x \star v \in \mathbb{Z} \text{ for all } x \in L\}.$$

LEMMA 13. *If* $A = (a_1, \ldots, a_n) \in \mathbb{R}^{n \times n}$ *is nonsingular and* L *the lattice generated by the rows of* A*, then* $B = (A^T)^{-1} \in \mathbb{R}^{n \times n}$ *is a basis of the dual lattice* L^**.*

We use the following fact without proof.

THEOREM 14. *If* λ_1^* *is the length of a shortest nonzero vector in* L^**, then* $\lambda_1^* \cdot \lambda_n(L) \le n^2$*.*

Recall:

$$A = \begin{pmatrix} m & 0 & 0 & \cdots & 0 \\ -s & 1 & 0 & \cdots & 0 \\ -s^2 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -s^{n-1} & 0 & 0 & \cdots & 1 \end{pmatrix}.$$

We next derive such a lower bound for most $s \in \mathbb{Z}_m$. For notational simplicity, we study the lattice $M = mL^*$ generated by the rows of

$$\begin{pmatrix} 1 & s & s^2 & \cdots & s^{n-1} \\ 0 & m & 0 & \cdots & 0 \\ 0 & 0 & m & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & m \end{pmatrix}.$$

Consider, for a positive bound $C < m$, the set

$$E_C = \left\{ s \in \mathbb{Z} \colon \begin{array}{c} |s^i t \text{ srem } m| < C \text{ for } 0 \le i < n \text{ and} \\ \text{some } t \in \mathbb{Z} \text{ with } \gcd(t, m) = 1 \end{array} \right\}$$

of exceptional multipliers s. We will later assume m to be prime, so that the gcd condition corresponds to $t \text{ srem } m \neq 0$. We have $\lambda_1(M) \ge C$ for all $s \in \mathbb{Z}_m \setminus E_C$.

LEMMA 15. *Let $n \ge 2$ and $s \in E_C$. Then there exist $d_1, \ldots, d_n \in \mathbb{Z}$, not all divisible by m, with*

$$\begin{aligned} &\sum_{1 \le i \le n} d_i s^{i-1} = 0 \text{ in } \mathbb{Z}_m, \\ &|d_i| < (nC)^{1/(n-1)} + 2 \text{ for all } i \le n. \end{aligned} \tag{16}$$

Theorem 17. *Let m be a k-bit prime, $n \geq 19$, $\epsilon > 0$, $2^{5n} \leq m^{1-\epsilon}$,*

$$\ell \leq (1-\epsilon)(1-\frac{1}{n})(k-1) - 4n,$$

and $\alpha = 2^{\ell}$. Given s and m and α-approximations $y_1, \ldots, y_n$ to the output of the generator (1) with $t = 0$, the generator can be broken in polynomial time for all but at most $m^{1-\epsilon}$ values $s \in \mathbb{Z}_m$.

This result is almost optimal in the following sense. We think of k as being large and of ϵ as small. Then the upper bound on $\ell \approx \log_2 \alpha$ is roughly $(1 - 1/n)k$, so that the approximations y_i only have about k/n bits of information about x_i.

We have broken the generator when $t = 0$, and now reduce the general case of (1) with arbitrary t to this one. Let $x'_i = x_{i+1} - x_i$ for $i \geq 0$. Then

$$x'_{i+1} = x_{i+2} - x_{i+1} = (sx_{i+1} + t) - (sx_i + t) = s(x_{i+1} - x_i) = sx'_i \text{ in } \mathbb{Z}_m,$$

so that the sequence $x'_1, x'_2, \ldots$ satisfies (1) with $t = 0$. Their approximations can be recovered from the original ones, as described below, with a loss of two bits.

We have to cope with the following issue. In the standard formulation (1), we take $\{0, 1, \ldots, m-1\}$ as representatives of $\mathbb{Z}_m$, and these integers are approximated in the generator. Thus instead of x'_i, we have to use

$$x_i^* = \begin{cases} x'_i = x_{i+1} - x_i & \text{if } x_{i+1} - x_i \geq 0, \\ x'_i + m = x_{i+1} - x_i + m & \text{otherwise.} \end{cases} \tag{18}$$

Then $x_0^*, x_1^*, \ldots$ satisfy (1) with $t = 0$. From approximations y_i to x_i, as observed for the attack, we have to determine approximations to the x_i^*

According to the case distinction in (18), we set

$$y_i^* = \begin{cases} y_{i+1} - y_i & \text{if } x_{i+1} - x_i \geq 0, \\ y_{i+1} - y_i + m & \text{otherwise.} \end{cases} \tag{19}$$

In both cases we have $|x_i^* - y_i^*| \leq 2\alpha$.

In our attack, we are only given the y_i and do not know the sign of $x_{i+1} - x_i$. But we can (almost) deduce it. Namely, if y_i and y_{i+1} differ by at least 2α, say $y_i \geq y_{i+1} + 2\alpha$, then $x_i \geq y_i - \alpha \geq y_{i+1} + \alpha \geq x_{i+1}$ and we have the sign. If $|y_i - y_{i+1}| < 2\alpha$, we do not know this sign and pursue both possibilities. Hopefully the y_i are sufficiently random so that this undesirable branching occurs only rarely.

Finally take

$$y'_i = \begin{cases} y_{i+1} - y_i & \text{if } y_{i+1} \geq y_i + 2\alpha, \\ y_{i+1} - y_i + m & \text{if } y_{i+1} \leq y_i - 2\alpha, \\ \text{both } y_{i+1} - y_i \text{ and } y_{i+1} - y_i + m & \text{if } |y_{i+1} - y_i| < 2\alpha, \end{cases}$$

and call the algorithm for Theorem 17 with s, m, $t = 0$, and 2α for α and the 2α-approximations $y'_1, \ldots, y'_n$.