Better random walks

for

Pollard's rho method

Edlyn Teske

University of Waterloo

September 16, 1998

http://www.cacr.math.uwaterloo.ca/~eteske

The discrete logarithm problem (DLP):

$$G \text{ finite cyclic group },$$

$$g \ , h \in G \ ,$$

$$\langle g \rangle = G \ .$$

Task: Find x such that $h = g^x$.

Pollard's Rho Method:

$$F : G \to G \ ,$$
$$Y_0 \in G \ , \qquad Y_{k+1} = F(Y_k) \ , \qquad k \in \mathbb{N}_0 \ .$$

Y_0 , F randomly chosen $\Longrightarrow$

$$E(\lambda + \mu) \approx \sqrt{\frac{\pi}{2}|G|} = 1.253 \ldots \cdot \sqrt{|G|} \ .$$

Finding a match:

Task: find Y_i, Y_j such that
$$Y_i = Y_j \quad \text{and} \quad i < j \ .$$

(cycle-finding methods of
Floyd, Brent, Schnorr-Lenstra, ...)

Schnorr-Lenstra, with optimized parameters:
Experimentally:
$$j \approx 1.13 \ (\lambda + \mu) \ .$$
(in average)

1.13 = "Delay factor"

Therefore:

If the iterating function $F : G \to G$ behaves as a random mapping,

we expect to find a match $Y_i = Y_j$ after

$$\approx 1.41\sqrt{|G|}$$

steps.

Question:
How can we efficiently simulate a random mapping?

Pollard's original application:

$G = (\mathbb{Z}/p\mathbb{Z})^*$, p prime ,
g = primitive root of p , $h \in (\mathbb{Z}/p\mathbb{Z})^*$.

$$Y_0 = h$$

or

$$Y_0 = h * g^\alpha , \qquad \alpha \in_R \{1, \ldots, p\} .$$

$$Y_{k+1} = \begin{cases} Y_k * g , & 0 < Y_k < p/3 , \\ {Y_k}^2 , & p/3 < Y_k < 2p/3 , \\ Y_k * h , & 2p/3 < Y_k < p . \end{cases}$$

Experimental results:

#digits of group order	average $\frac{\#\text{steps}}{\sqrt{\lvert G\rvert}}$
3	1.570
4	1.591
5	1.553
6	1.477
7	1.539
8	1.605
9	1.511
10	1.493
11	1.690
12	1.713
13	1.667
average	1.553

DL-computation in $(Z/pZ)^*$

Pollard's original application, generalized:

Divide G into three pairwise disjoint sets

$$T_1 , \quad T_2 , \quad T_3 .$$

$$Y_{k+1} = \begin{cases} Y_k * g , & Y_k \in T_1 , \\ {Y_k}^2 , & Y_k \in T_2 , \\ Y_k * h , & Y_k \in T_3 . \end{cases}$$

#digits of group order	average #steps $\frac{\text{\#steps}}{\sqrt{|G|}}$
3	1.786
4	1.743
5	1.809
6	1.767
7	1.670
8	1.758
9	1.731
10	1.943
11	1.834
12	1.772
13	1.858
average	1.759

DL-computation in subgroups of $(Z/pZ)^*$ of prime order

Computation in $(Z/pZ)^*$ versus computation in $G \subset (Z/pZ)^*$, $|G|$ prime

Consider 3 mappings:

$$F_1(Y) = Y * g \ ,$$
$$F_2(Y) = Y^2 \ ,$$
$$F_3(Y) = Y * h \ .$$

When is F_i a bijection?

	in $(Z/pZ)^*$	in $G \subset (Z/pZ)^*$, $|G|$ prime
F_1	always	always
F_2	never	always
F_3	always	always

Elliptic curve analogue:

Elliptic curve E,
over finite field $\mathbb{F}_q$,

point $P \in E$ of prime order,
point $Q \in \langle P \rangle$.

ECDLP:

Find x such that

$$Q = xP \ .$$

Define sequence (Y_k):

$Y_0 = Q$.

Divide $\langle P \rangle$ into three pairwise disjoint sets T_1, T_2, T_3.

$$Y_{k+1} = \begin{cases} Y_k + P \ , & Y_k \in T_1 \ , \\ 2Y_k \ , & Y_k \in T_2 \ , \\ Y_k + Q \ , & Y_k \in T_3 \ . \end{cases}$$

I.e., *Adding* and *Doubling* steps.

Experimental results:

#digits of subgroup order p	average $\frac{\#\text{steps}}{\sqrt{p}}$
3	1.891
4	1.776
5	1.773
6	1.800
7	1.825
8	1.703
9	1.773
10	1.804
11	1.948
12	1.856
13	1.924
average	1.807

DL-computation in elliptic curve subgroups of prime order (Pollard's walk, generalized)

Find better walks:

1.) *Additive walks*
(= walks with adding steps only)

2.) *Mixed walks*
(= walks with adding and doubling steps)

Partitioning of G:

r = small positive integer ($r \in [3, 100]$) :

We want

$$G = T_1 \cup \cdots \cup T_r \ ,$$

with roughly equally large sets T_s.

Take hash function

$$v : G \to \{1, \ldots, r\}$$

and define

$$T_s := \{y \in G \ : \ v(y) = s\} \ , \qquad s = 1, \ldots, r \ .$$

1.) Additive walks

Example: $r = 20$

20 pairwise disjoint sets $T_1, \ldots, T_r$.

$Y_0 = mP + nQ$ with $m, n \in_R \{1, 2, \ldots, |G|\}$

For $1 \le s \le 20$ let

$$A_s = m_s P + n_s Q \ ,$$

with m_s , $n_s \in_R \{1, 2, \ldots, |G|\}$.

Then let for $k \in \mathbb{N}_0$:

$$Y_{k+1} = Y_k + A_s \ ,$$

where s such that $Y_k \in T_s$, $1 \le s \le 20$.

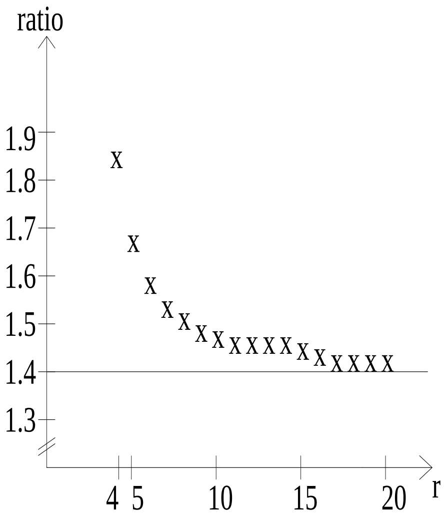
Experimental results:

#digits of subgroup order p	average $\frac{\#\text{steps}}{\sqrt{p}}$
3	1.454
4	1.453
5	1.453
6	1.462
7	1.445
8	1.443
9	1.440
10	1.441
11	1.428
12	1.431
13	1.405
average	1.452

DL-computation in elliptic curve subgroups of prime order (additive walk, with $r = 20$)

Now vary r:

r	$\#\text{steps}/\sqrt{p}$ (average values) (11000 computations, $100 < p < 10^9$)
4	1.842
5	1.660
6	1.568
7	1.524
8	1.508
9	1.488
10	1.475
11	1.456
12	1.455
13	1.451
14	1.454
15	1.447
16	1.430
17	1.420
18	1.420
19	1.417
20	1.419
30	1.413
40	1.388
50	1.408
100	1.384



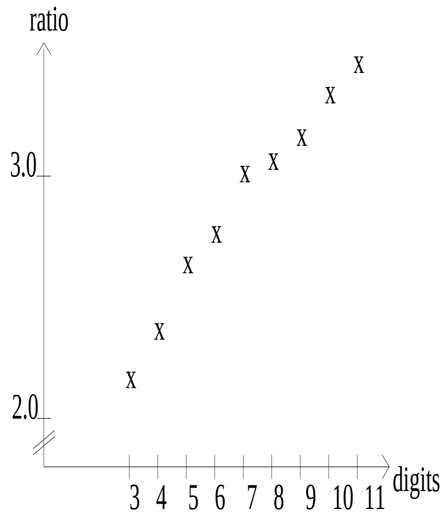
$$\left(\text{ratio} = \frac{\#\text{steps}}{\sqrt{p}}, \quad r = \#\text{addings} \right)$$

...while for $r = 3$: no constant ratio !

#digits of p	$\#\text{steps}/\sqrt{p}$ (average values)
3	2.156
4	2.365
5	2.651
6	2.780
7	3.007
8	3.079
9	3.166
10	3.327
11	3.484

average values for $r = 3$ addings

Compare: with Pollard's "original" function: ratio ≈ 1.8 and constant !



Performance of additive walks $r = 3$ addings.

$\left(\text{ratio} \;=\; \frac{\#\text{steps}}{\sqrt{p}} \right)$

Summing up:

Pollard's walk (2 adding, 1 doubling):
ratio ≈ 1.8

Walk with 3 addings:
no constant ratio

Walk with r addings,
$15 \le r \le 100$: $1.45 \ge \text{ratio} \ge 1.38$

Random walk:
expected ratio ≈ 1.4

Conclusions:

Increasing the number of partitions improves performance.

Doubling step improves performance.

2.) Consider mixed walks:

Two parameters:

r and q .

$r + q$ pairwise disjoint sets of roughly equal size

$$T_1 , \ldots , T_r \; , \qquad\qquad V_1 , \ldots , V_q \; .$$

For $1 \le s \le r$ let

$$A_s = m_s P + n_s Q \; .$$

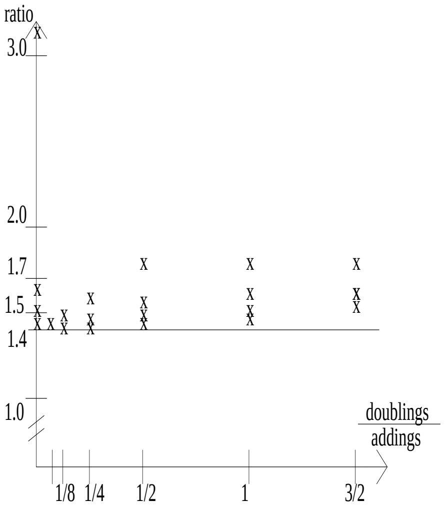
Put

$Y_0 = mP + nQ$ with $m, n \in_R \{1, 2, \ldots , |G|\}$.

$$Y_{k+1} = \begin{cases} Y_k + A_s \; , & Y_k \in T_s \; , \; 1 \le s \le r \; , \\ 2Y_k \; , & Y_k \in V_u \; , \; 1 \le u \le q \; . \end{cases}$$

r (#addings)	q (#doublings)	$\frac{\#\text{steps}}{\sqrt{p}}$
3	0	3.122
2	1	1.776
2	2	1.770
2	3	1.769
5	0	1.644
4	1	1.589
4	2	1.569
4	4	1.627
4	6	1.618
9	0	1.513
8	1	1.473
8	2	1.496
8	4	1.487
8	8	1.576
8	12	1.623
17	0	1.451
16	1	1.453
16	2	1.427
16	4	1.435
16	8	1.455
16	16	1.502
16	24	1.546

Mixed walks for elliptic curve subgroups of 9-digit prime group orders p



$$\left(\frac{\text{doublings}}{\text{addings}} = \frac{|V_1 \cup \cdots \cup V_q|}{|T_1 \cup \cdots \cup T_r|} \right)$$

So far: only experimental results

- Limited range:
  only group orders $\leq 10^{13}$ were used for our statistics.
  Because:
- Variance for #steps requires computation of many examples.
  For example, average ratios taken over 2000 examples may vary up to 7%.

Any theoretical results?

Yes!

Additive walk (Y_k)**:**
Put $p = |\langle P \rangle|$.

$$Y_k = \alpha_k P + \beta_k Q \ ,$$

where $\alpha_0, \beta_0 \in_R \{1, \dots p\}$ and

$$\alpha_{k+1} \equiv \alpha_k + m_s \ (\mathrm{mod} p) \ , \qquad \beta_{k+1} \equiv \beta_k + n_s \ (\mathrm{mod} p) \ , \qquad s \in \{1, \dots, r\} \ .$$

Put $\delta_k = \alpha_k + \beta_k x \ (\mathrm{mod} p)$: $\qquad (Q = xP)$

$$\Longrightarrow \qquad Y_k = \delta_k P \ .$$

Thus, $Y_{k+1} = F(Y_k)$ corresponds to

$$\delta_{k+1} \ \equiv \ \delta_k + (m_s + n_s x) \ (\mathrm{mod} p) \ ,$$

$$s = s(Y_k) \in \{1, \dots, r\} \ ,$$

determined e.g. via hash function.

(For doubling steps of mixed walks:
$$\delta_{k+1} \ \equiv \ 2\delta_k \ (\mathrm{mod} p).)$$

One-to-one correspondence:

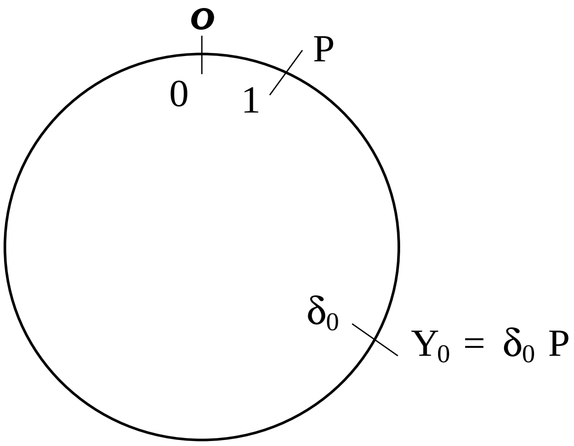
walks (Y_k) in elliptic curve subgroup of order p

$\longleftrightarrow$

walks (δ_k) in $\mathbb{Z}/p\mathbb{Z}$

$A_1, \dots, A_r \;\longleftrightarrow\; r$ distances of form $m_s + n_s x$

Circumference of circle $= p$
(= order of elliptic curve subgroup $\langle P \rangle$)



Random walks on the integers mod n:

$r \in \mathbb{N}$, $p_1, \dots, p_r > 0$ such that

$$\sum_{j=1}^{r} p_j = 1.$$

$n \in \mathbb{N}$, $a_1, \dots, a_r \in \{1, \dots, n\}$,

$$\tilde{a} := (a_1, \dots, a_r) \;.$$

Define probability distribution on $\{1, \dots, n\}$:

$$P_{\tilde{a}}(a) := \begin{cases} p_j & \text{if } a = a_j \\ 0 & \text{otherwise .} \end{cases}$$

Define: *Random walk on the integers mod n supported by r elements:*

$$e_0 = 0, \qquad e_{i+1} = e_i + a, \qquad i = 0, 1, 2, \dots$$

where $a \in_R \{a_1, \dots, a_r\}$ according to $P_{\tilde{a}}$.

Distribution of random walk after m steps:

$$P_{\tilde{a}}^{*m} \;.$$

(= $m-$fold convolution of $P_{\tilde{a}}$ with itself).

U = uniform distribution:

$$U(a) = \frac{1}{n}\;, \quad a \in \{1\;, \dots\;, n\}\;.$$

Variation distance between P and U:

$$\|P - U\| := \frac{1}{2} \sum_{a} |P(a) - \frac{1}{n}|\;.$$

Theorem (Hildebrand, '94):
Let $r \geq 2$ and $\varepsilon > 0$.
For sufficiently large primes n
there exists some constant

$$\gamma > 0\;,$$

which may depend on r, ε, $p_1, \dots, p_r$,
but not on n,
such that for

$$m = \lfloor \gamma n^{2/(r-1)} \rfloor$$

we have

$$E(\|P_{\tilde{a}}^{*m} - U\|) < \varepsilon\;.$$

The expectation is taken over a uniform choice of the $\tilde{a} = (a_1, \dots, a_r)$.

Hildebrand "$\Longrightarrow$"
Our random walk (δ_k) on the integers $\bmod p$ is expected to get close to uniformly distributed after

$$const \cdot p^{2/(r-1)} \quad \text{steps.}$$

Back to initial problem
(Random walks for Pollard's rho method):
Expected #steps in random case: const $\cdot \sqrt{p}$.

If $r \geq 6$ then

$$\lim_{p \to \infty} \frac{p^{2/r}}{\sqrt{p}} = 0 \ .$$

Therefore:
Fraction of steps where walk supported by r points behaves differently from a random walk becomes arbitrarily small for $r \geq 6$.

Can we do better than random?

Idea:
Enforce non-bijectivity of the partial mappings.

Example:
in the case of $(\mathbb{Z}/p\mathbb{Z})^*$: see Pollard.

for other groups: ???

generic solution: ???

#digits of subgroup order p	Number of examples computed
3	$100 \cdot 100$
4	$100 \cdot 100$
5	$100 \cdot 100$
6	$100 \cdot 100$
7	$100 \cdot 100$
8	$80 \cdot 40$
9	$40 \cdot 30$
10	$30 \cdot 30$
11	$25 \cdot 20$
12	$30 \cdot 5$
13	$30 \cdot 1$

Rho method:

Uses a pseudorandom function where all elements of a space are accessible on each iteration.

Lambda method:

Requires an ordering of elements with each iteration taking a step whose distance is small relative to the total size of the space.

Both methods can be parallized!
(van Oorschot/Wiener, 1996)