On some properties of PRNGs based on block ciphers in counter mode

Alexey Urivskiy, Andrey Rybkin, Mikhail Borodin
JSC InfoTeCS, Moscow, Russia
alexey.urivskiy@mail.ru

2016

Pseudo Random Number Generators

G: $\{0,1\}^m \to \{0,1\}^M$ for $M \gg m$

Typical assumptions for a cryptographic PRNG:

- **G** is efficiently computable
- the seed is uniformly distributed on $\{0,1\}^m$
- **G** is '**random-like**': no polynomial statistical test can distinguish **G** from a truly random generator with uniform distribution (informally)

Pseudo Random = Unpredictable

Predictability problem: predict the next output bit for **G** with probability better than ½ if all previously output bits are known

Next-bit test: G passes the test if the next bit cannot be predicted by any polynomial predictor.

Theorem [Yao'82] : if **G** passes the next-bit test it will pass any polynomial statistical test.

PRNGs based on block ciphers – G1

$E(K, T)$ – block cipher

$K \in V_k$ – key

$T \in V_n$ – message

G1:

for *i*=0 to *M* do

$$count \coloneqq (IV + i) \; mod \; 2^n$$

$$a_i \coloneqq E(K, count)$$

Consider the case $\boldsymbol{M < N = 2^n}$.

PRNGs based on block ciphers – G1

G1 is highly appreciated and widely used – ISO/IEC 18031 **`CTR_DRBG`**.

However, if **G1** has output a symbol, it will never output it again → For $\boldsymbol{M \sim \sqrt{N}}$ due to the **birthday paradox** becomes distinguishable from a truly random uniform generator.

PRNGs based on block ciphers – G2

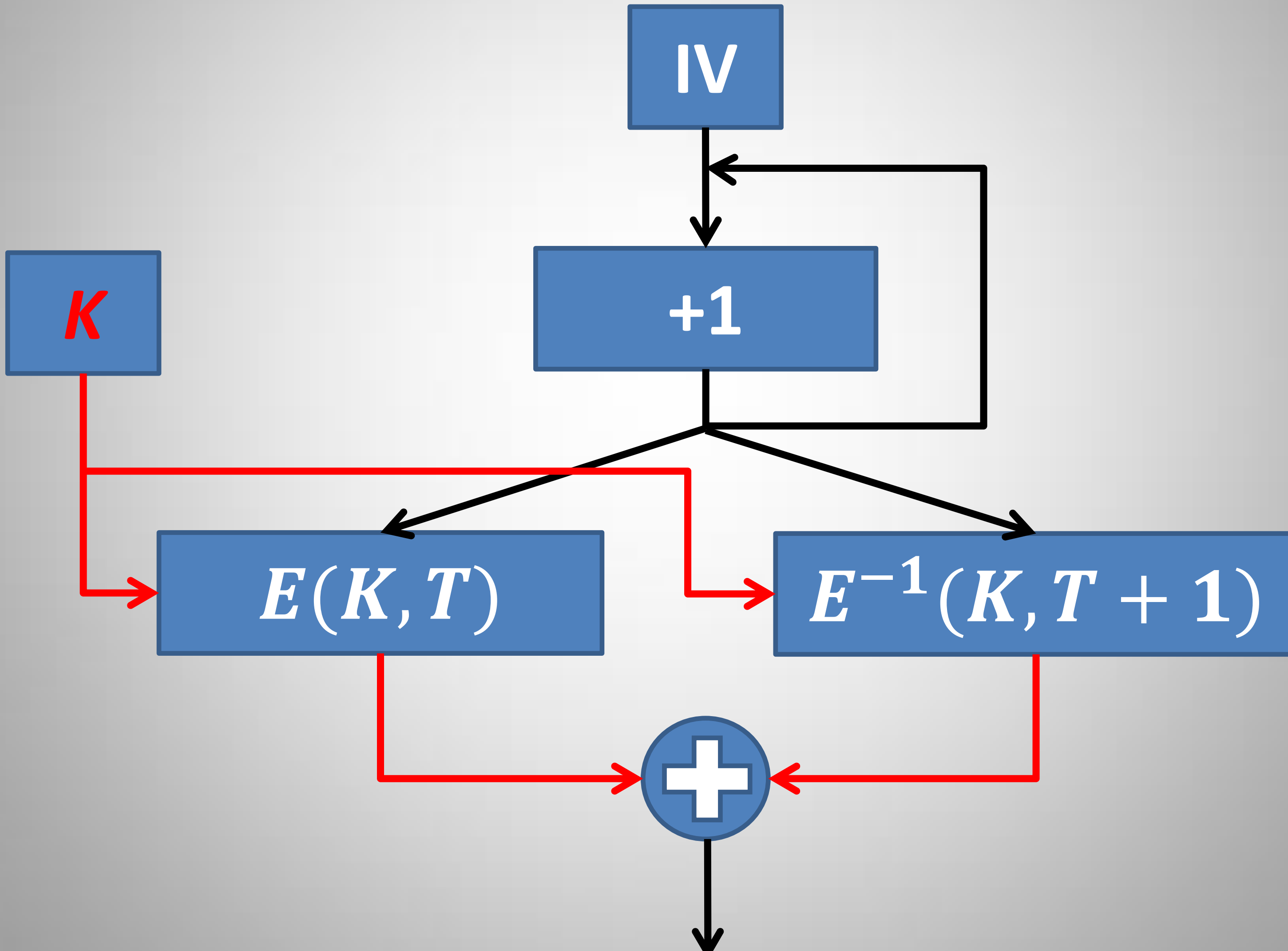
G2:

for i = 0 to M do

$$count \coloneqq (IV + i)\ mod\ 2^n$$

$$a_i \coloneqq E(K, count) \oplus E^{-1}(K, count + 1)$$

PRNGs based on block ciphers – G2



PRNGs based on block ciphers – G2

G2:

for i = 0 to M do

$$count := (IV + i) \bmod 2^n$$

$$a_i := E(K, count) \oplus E^{-1}(K, count + 1)$$

Will a second cipher help? and how?

Idealized model for PRNGs

Assumption 1: encryption (decryption) procedure of an *n*-bit **block cipher** with a **random key** is a **random permutation** on V_n

0	1	2	…	2^n-2	2^n-1
$\sigma(0)$	$\sigma(1)$	$\sigma(2)$	…	$\sigma(2^n-2)$	$\sigma(2^n-1)$

Typical to cryptanalysis: a ‘good’ block cipher with a random key must be indistinguishable from a random permutation.

Idealized model for PRNGs

Assumption 2: **encryption and decryption procedures** of a block cipher with the same random key are independent so they are the two **random and independent permutations** on V_n.

G1I: $E(K,T) \to \sigma(T)$

G2I: $E(K,T) \oplus E^{-1}(K,T+1) \to \sigma_1(T) \oplus \sigma_2(T+1)$

Does IV matter ?

$$\sigma^{'}(i) = \sigma(\,(i + \boldsymbol{IV}\,)\;mod\;2^{n}\,)$$

$$\Downarrow$$

Another choice of IV leads to a different $\sigma^{'}$ given σ, however from the same set: $\boldsymbol{IV = 0}$

G1I: $E(K,T) \rightarrow \sigma(T)$

G2I: $E(K,T) \oplus E^{-1}(K,T+1) \rightarrow \sigma_1(T) \oplus \sigma_2(T)$

Output sequences

G1I: $a_0, \quad a_1, \quad a_2, \quad \ldots, \; a_{N-2}, \; a_{N-1}$

$$N \cdot (N-1) \cdot (N-2) \cdot \ldots \cdot 2 \cdot 1 = \boldsymbol{N!}$$

G2I: $\bigoplus_{i \in V_n} \sigma(i) = 0 \;\Rightarrow\; \bigoplus_{i=0}^{N-1} a_i = \bigoplus_{i \in V_n} \left(\sigma_1(i) \oplus \sigma_2(i)\right) = 0$

$a_0, \quad a_1, \quad a_2, \; \ldots, \; a_{N-2}, \; a_{N-1}$

$$N \cdot N \cdot N \cdot \ldots \cdot N \cdot 1 \le \boldsymbol{N^{N-1}}$$

Output sequences

Theorem (***Hall'52***). For **any** sequence

$a_0, a_1, \dots, a_{N-1}, \quad a_i \in V_n, \quad i = 0,1,\dots,2^n - 1,$

satisfying the condition

$$\bigoplus_{i=0}^{N-1} a_i = 0$$

there exists at least one pair of permutations σ_1, σ_2 on V_n such that $a_i = \sigma_1(i) \oplus \sigma_2(i)$.

$$a_0, \quad a_1, \quad a_2, \ \dots, \ a_{N-2}, \ a_{N-1}$$

$$N \quad \cdot \quad N \quad \cdot \quad N \cdot \dots \quad \cdot \quad N \quad \cdot \quad 1 = \boldsymbol{N^{N-1}}$$

Output sequences: summary

	G1I $\sigma(count)$	**G2I** $\sigma_1(count) \oplus \sigma_2(count)$
Type	all elements are different	any fixed $N-1$ elements are arbitrary
Number (of length N)	$N!$	N^{N-1}

$$\frac{N!}{N^{N-1}} = \frac{\sqrt{2\pi N} \cdot N^N}{e^N} \frac{N}{N^N} = e^{-\left(N - \ln N\sqrt{2\pi N}\right)}$$

Equivalent representation for G2I

$$
\begin{array}{c|c}
\oplus & \begin{array}{cccccc} 0 & 1 & 2 & 3 & \ldots & N-1 \end{array} \\
\hline
\begin{array}{c} 0 \\ 1 \\ 2 \\ 3 \\ \vdots \\ N-1 \end{array} &
\mathbf{M} = \begin{bmatrix}
0 & 1 & 2 & 3 & \ldots & N-1 \\
1 & 0 & 3 & 2 & \ldots & N-2 \\
2 & 3 & 0 & 1 & \ldots & N-3 \\
3 & 2 & 1 & 0 & \ldots & N-4 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
N-1 & N-2 & N-3 & N-4 & \ldots & 0
\end{bmatrix}
\end{array}
$$

Equivalent representation for G2I

Definition. A sequence of pairs of indices $(i_0, j_0), (i_1, j_1), \dots, (i_{N-1}, j_{N-1})$, $i_l, j_l \in \{0,1,\dots,N-1\}$, $i_k \neq i_t, j_k \neq j_t$ for any $t \neq k$ is called **a trajectory on $\mathbf{M}$.**

Definition. The sequence $\mathbf{M}(i_0, j_0), \mathbf{M}(i_1, j_1), \dots, \mathbf{M}(i_{N-1}, j_{N-1})$ is called **the output of the trajectory** $(i_0, j_0), (i_1, j_1), \dots, (i_{N-1}, j_{N-1})$

Equivalent representation for G2I

Proposition. Between
the set of ordered pairs of permutations on V_n and
the set of trajectories on matrix $\mathbf{M}$
a **one-to-one correspondence** can be defined so that
the sum of the pair of permutations will coincide with
the output of the corresponding trajectory.

Corollary. The generation process is: the s-th output symbol a_s is **chosen randomly** from $\mathbf{M}$.
After that **the row and the column** containing a_s **are struck out from** $\mathbf{M}$.

Equivalent representation for G2I

$$\mathbf{M} = \begin{bmatrix} 0 & 1 & 2 & 3 & \dots & N-1 \\ 1 & 0 & 3 & 2 & \dots & N-2 \\ 2 & 3 & 0 & 1 & \dots & N-3 \\ 3 & 2 & 1 & 0 & \dots & N-4 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ N-1 & N-2 & N-3 & N-4 & \dots & 0 \end{bmatrix}$$

Equivalent representation for G2I

$$\mathbf{M} = \begin{bmatrix} 0 & 1 & 2 & 3 & \dots & N-1 \\ 1 & 0 & 3 & 2 & \dots & N-2 \\ 2 & 3 & 0 & 1 & \dots & N-3 \\ 3 & 2 & 1 & 0 & \dots & N-4 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ N-1 & N-2 & N-3 & N-4 & \dots & 0 \end{bmatrix}$$

$a_0 = 3$

$(2,1)_{,}$

Equivalent representation for G2I

$$\mathbf{M} = \begin{bmatrix} 0 & 1 & 2 & 3 & \dots & N-1 \\ 1 & 0 & 3 & 2 & \dots & N-2 \\ 2 & 3 & 0 & 1 & \dots & N-3 \\ 3 & 2 & 1 & 0 & \dots & N-4 \\ \vdots & & \vdots & \vdots & \ddots & \vdots \\ N-1 & N-2 & N-3 & N-4 & \dots & 0 \end{bmatrix}$$

$$a_0 = 3$$

$$(2,1)_,$$

Equivalent representation for G2I

$$\mathbf{M} = \begin{bmatrix} 0 & 1 & 2 & 3 & \dots & N-1 \\ 1 & 0 & 3 & 2 & \dots & N-2 \\ 2 & 3 & 0 & 1 & \dots & N-3 \\ 3 & 2 & 1 & 0 & \dots & N-4 \\ \vdots & & \vdots & \vdots & \ddots & \vdots \\ N-1 & N-2 & N-3 & N-4 & \dots & 0 \end{bmatrix}$$

$a_0 = 3, a_1 = 2$

$(2,1)_,$ $(1,3)$

Equivalent representation for G2I

$$\mathbf{M} = \begin{bmatrix} 0 & 1 & 2 & 3 & \dots & N-1 \\ 1 & 0 & 3 & 2 & \dots & N-2 \\ 2 & 3 & 0 & 1 & \dots & N-3 \\ 3 & 2 & 1 & 0 & \dots & N-4 \\ \vdots & & \vdots & & \ddots & \vdots \\ N-1 & N-2 & N-3 & N-4 & \dots & 0 \end{bmatrix}$$

$$a_0 = 3, a_1 = 2$$

$(2,1)_{,}$ $(1,3)$

Equivalent representation for G2I

$$\mathbf{M} = \begin{bmatrix} 0 & 1 & 2 & 3 & \dots & N-1 \\ 1 & 0 & 3 & 2 & \dots & N-2 \\ 2 & 3 & 0 & 1 & \dots & N-3 \\ 3 & 2 & 1 & 0 & \dots & N-4 \\ \vdots & & \vdots & & \ddots & \vdots \\ N-1 & N-2 & N-3 & N-4 & \dots & 0 \end{bmatrix}$$

$$a_0 = 3, a_1 = 2, a_2 = N-3$$

(2,1), (1,3), (N-1,2)

Equivalent representation for G2I

$$\mathbf{M} = \begin{bmatrix} 0 & 1 & 2 & 3 & \dots & N-1 \\ 1 & 0 & 3 & 2 & \dots & N-2 \\ 2 & 3 & 0 & 1 & \dots & N-3 \\ 3 & 2 & 1 & 0 & \dots & N-4 \\ \vdots & & & & \ddots & \vdots \\ N-1 & N-2 & N-3 & N-4 & \dots & 0 \end{bmatrix}$$

$$a_0 = 3, a_1 = 2, a_2 = N-3$$

(2,1), (1,3), (N-1,2)

Conditional probability

Conditional probability $P(a_s|a_{s-1}, a_{s-2}, \dots, a_0)$ is the probability for a generator to output a_s provided $a_{s-1}, a_{s-2}, \dots, a_0$ were output before.

To estimate $P(a_s|a_{s-1}, a_{s-2}, \dots, a_0)$ for **G2I** it suffices to estimate how many a_s are left in $\mathbf{M}$ after ***s* rows and columns were struck out**.

Conditional probability

G1I

$$P(a_s|a_{s-1}, a_{s-2}, \dots, a_0) = \begin{cases} 0, & \text{if } a_s \in \{a_{s-1}, a_{s-2}, \dots, a_0\}; \\ \dfrac{1}{N-s}, & \text{otherwise.} \end{cases}$$

G2I

$$\boldsymbol{P_1 = \frac{N-2s}{(N-s)^2} \le P(a_s|a_{s-1}, a_{s-2}, \dots, a_0) \le \frac{N-s}{(N-s)^2} = P_2}$$

$$\boldsymbol{P_1 < \frac{1}{N} < P_2}$$

Conditional probability: summary

	G1I $\sigma(count)$	G2I $\sigma_1(count) \oplus \sigma_2(count)$	Ideal
Lower bound	$\mathbf{0}$	$\frac{\boldsymbol{N-2s}}{\boldsymbol{(N-s)^2}}$	$\frac{\mathbf{1}}{\boldsymbol{N}}$
Number of symbols	$\boldsymbol{s}$	$\leq \boldsymbol{N-s}$	$\boldsymbol{N}$
Upper bound	$\frac{\mathbf{1}}{\boldsymbol{N-s}}$	$\frac{\mathbf{1}}{\boldsymbol{N-s}}$	$\frac{\mathbf{1}}{\boldsymbol{N}}$
Number of symbols	$\boldsymbol{N-s}$	$\leq \boldsymbol{s}$	$\boldsymbol{N}$

Thank you!
Questions?