COMS30048 lecture: week #20

Agenda: explore (pseudo-)random bit generation, via

- 1. an "in theory", i.e., design-oriented perspective, and
- 2. an "in practice", i.e., implementation-oriented perspective.

Caveat!

~ 2 hours \Rightarrow introductory, and (very) selective (versus definitive) coverage.



COMS30048 lecture: week #20

Bad news: in *theory*, we need to consider each of

1. random bit, i.e., an

 $x \in \{0, 1\}$

which is random,

2. random bit sequence, i.e., an

 $x \in \{0, 1\}^n$

which is random (e.g., for an AES cipher key *k*),

3. random *number*, i.e., an

 $x\in\{0,1,\ldots,n-1\}$

which is random (e.g., for an RSA modulus $N = p \cdot q$).



COMS30048 lecture: week #20

Good news: in *practice*, we don't because 1. ⇒ 2.

• concatenate n random bits together, i.e.,

$$x = x_0 \parallel x_1 \parallel \cdots \parallel x_{n-1},$$

- produce *x* as output.
- ▶ 2. ⇒ 3.
 - if $n = 2^{n'}$ for some integer n', then
 - generate an *n*'-bit sequence *x*' per the above,
 - interpret x' as the integer

$$x = \sum_{i=0}^{i < n'} x'_{i'}$$

produce x as output.

• if $n \neq 2^{n'}$ for any integer n', then

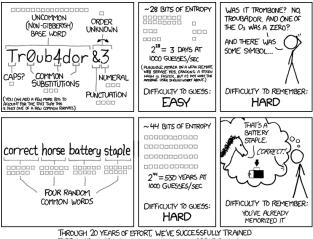
- let n' be the smallest integer such that $2^{n'} > n$,
- generate an n'-bit sequence x' per the above,
- interpret x' as the integer

$$x = \sum_{i=0}^{i < n'} x'_i,$$

- if $x \ge n$, reject (or discard) it and try again; otherwise, if x < n, produce x as output.
- \therefore we can focus on random bits (and ignore numbers).



Part 1: in theory (1) Entropy



EVERYONE TO USE PASSWORDS THAT ARE HARD FOR HUMANS TO REMEMBER, BUT EASY FOR COMPUTERS TO GUESS.

③ Daniel Page (csdsp@bristol.ac.uk) Applied Cryptology



Part 1: in theory (2) Entropy

Definition

The concept of **entropy** is a measure of uncertainty with respect to a random variable. Less formally, the entropy of some x relates to how much you know (resp. do not know) about x: if some x could be one of 2^n possible values, it is said to have n bits of entropy. In addition, we say

- 1. an *x* with n > 0 bits of entropy is termed **entropic**, and
- 2. if an entropic *x* has negligible probability of having been generated before, it is deemed **fresh entropy**.



Part 1: in theory (2) Entropy

Definition

The concept of **entropy** is a measure of uncertainty with respect to a random variable. Less formally, the entropy of some x relates to how much you know (resp. do not know) about x: if some x could be one of 2^n possible values, it is said to have n bits of entropy. In addition, we say

- 1. an *x* with n > 0 bits of entropy is termed **entropic**, and
- 2. if an entropic *x* has negligible probability of having been generated before, it is deemed **fresh entropy**.
- **Example**: given a 32-bit sequence *x*,
 - ▶ if *x* is random, then it has 32 bits of entropy,
 - if $x_0 = 0$ and $x_1 = 1$ (i.e., the two LSBs of x are known), then it has 30 bits of entropy,
 - if HW(x) = 14 (i.e., x has Hamming weight 14), then it has ~ 29 bits of entropy.



Part 1: in theory (3) Entropy

Definition

A **noise source** is a non-deterministic, physical process which provides a means of generating an *unconditioned* (or raw) entropic output.



Part 1: in theory (3) Entropy

Definition

A noise source is a non-deterministic, physical process which provides a means of generating an *unconditioned* (or raw) entropic output.

• Example (see [8, Section 5.2], or [14, Section 3]):

- 1. hardware-based:
 - time between emission of (e.g., α or β) particles during radioactive decay,
 - thermal (or Johnson-Nyquist) noise stemming from a resistor or capacitor,
 - frequency instability (or "jitter") of a ring oscillator,
 - fluctuation of hard disk seek-time and access latency,
 - noise resulting from a disconnected audio input (or ADC),
 - ► ..
- 2. software-based:
 - a high resolution system clock or cycle counter,
 - elapsed time between user input (e.g., key-presses or mouse movement),
 - content of input/output buffers (e.g., disk caches),
 - operating system state (e.g., load) or events (e.g., network activity),

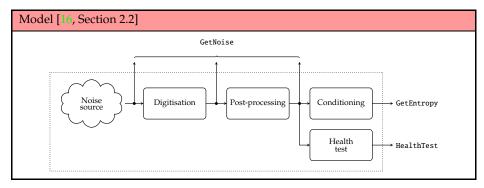
<u>ا ا</u>



Part 1: in theory (4) Entropy

Definition

An entropy source is a construction, based on a noise source, which provides a means of generating a *conditioned* entropic output.





Part 1: in theory (5) Randomness

Definition

Per [15, Section 4], an ideal random bit-sequence

 $x = \langle x_0, x_1, \dots x_{n-1} \rangle$

will exhibit the following properties

1.	unpredictable	$\uparrow \uparrow \uparrow$	the probability of guessing x_i is close to $\frac{1}{2}$
2.	unbiased		$x_i = 0$ and $x_i = 1$ occur with equal probability
3.	uncorrelated		x_i and x_j are statistically independent
its of entro	nnv.		,

and contain *n* bits of entropy.



Part 1: in theory (5) Randomness

Definition

Per [15, Section 4], a pseudo-random bit-sequence

 $x=\langle x_0,x_1,\ldots x_{n-1}\rangle$

"looks random", i.e., exhibits the same properties as an ideal random sequence, *but* is generated algorithmically and thus likely contains less than *n* bits of entropy.



Part 1: in theory (6) (Pseudo-)random bit generators

Definition

A Random Bit Generator (RBG) can be used to generates a sequence of random bits. There are two more specific cases, namely

True Random Bit Generator (TRBG)	≡
Pseudo-Random Bit Generator (PRBG)	Ξ

Non-deterministic Random Bit Generator (NRBG) Deterministic Random Bit Generator (DRBG)

with the right-hand terms preferred by [15]. Based on this, it is reasonable to say that

TRBG \equiv NRBG \simeq entropy source.



Part 1: in theory (6) (Pseudo-)random bit generators

Definition

A Random Bit Generator (RBG) can be used to generates a sequence of random bits. There are two more specific cases, namely

True Random Bit Generator (TRBG)	Ξ
Pseudo-Random Bit Generator (PRBG)	≡

Non-deterministic Random Bit Generator (NRBG) Deterministic Random Bit Generator (DRBG)

with the right-hand terms preferred by [15]. Based on this, it is reasonable to say that

TRBG \equiv NRBG \simeq entropy source.

Idea: informally at least,



... we'll consider a *hybrid* construction.



Part 1: in theory (7) (Pseudo-)random bit generators

Definition

Consider a deterministic, polynomial-time algorithm *G*. Given a **seed** $\varsigma \in \{0, 1\}^{n_{\varsigma}}$ as input, it produces $G(\varsigma) \in \{0, 1\}^{n_{r}}$ as output where $n_{r} = f(n_{\varsigma})$ for some polynomial function *f*. As such, we call *G* a **Pseudo-Random Generator (PRG)** if

- 1. for every n_{ς} it holds that $n_r > n_{\varsigma}$, and
- 2. for all polynomial-time destinguishers D, there exists a negligible function negl such that

 $|\Pr[D(G(\varsigma)) = 1] - \Pr[D(r) = 1]| \le \operatorname{negl}(n_{\varsigma})$

where ς and r are chosen uniformly at random from $\{0, 1\}^{n_{\varsigma}}$ and $\{0, 1\}^{n_{\tau}}$ respectively.



Syntax

Having fixed the (finite) space S of states, a concrete Pseudo-Random Generator (PRG) is defined by

- 1. an algorithm SEED : $\mathbb{Z} \times \{0, 1\}^{n_c} \rightarrow S$ that
 - accepts a security parameter and an n_c-bit seed as input, and
 - produces an initial state as output
- 2. an algorithm Update : $S \rightarrow S \times \{0, 1\}^{n_b}$ that
 - accepts a current state as input, and
 - produces a next state and an n_b-bit block of pseudo-random bits as output.



Part 1: in theory (8) (Pseudo-)random bit generators

• Translation: assuming $n_r = l \cdot n_b$ for some *l*, then

1. use TRBG
$$\rightsquigarrow$$

$$\begin{cases}
\text{generate a sufficiently large,} \\
\text{high-entropy seed } \zeta
\end{cases}$$
2. use PRBG \rightsquigarrow

$$\begin{cases}
\theta[0] \leftarrow \text{Seed}(\lambda, \zeta) \\
\theta[1] , b[0] \leftarrow \text{Update}(\theta[0]) \\
\theta[2] , b[1] \leftarrow \text{Update}(\theta[1]) \\
\vdots \\
\theta[i+1], b[i] \leftarrow \text{Update}(\theta[i]) \\
\vdots
\end{cases}$$

meaning that

$$b = \underbrace{b[0]}_{n_b \text{-bits}} \parallel \underbrace{b[1]}_{n_b \text{-bits}} \parallel \cdots \parallel \underbrace{b[l-1]}_{n_b \text{-bits}} \equiv G(\varsigma)$$

 $l \cdot n_b = n_r$ -bits

provides the output required per the PRG definition.

Daniel Page (csdsp@bristol.a Applied Cryptology

```
int getRandomNumber()
{
return 4; // chosen by fair dice roll.
// guaranteed to be random.
}
```

http://xkcd.com/221

② Daniel Page (csdsp@bristol.ac.uk) Applied Cryptology



Part 1: in theory (10) (Pseudo-)random bit generators

▶ Problem: we need to assess the quality of our construction (and output from it).

Solution:

- 1. for some instanciations, we can develop a proof,
- 2. for some instanciations, we must apply
 - online (e.g., continuously or periodically *during* use), and/or
 - offline (i.e., once before use)

statistical tests (see, e.g., [8, Section 5.4]) to sample outputs; note that

- the intention is to detect weakness (meaning a PRBG can only be rejected by a test),
- the conclusion is itself probabilistic, meaning use of multiple tests amplifies confidence.



Definition

A PRBG is said to pass all statistical tests iff. no polynomial-time algorithm can, with probability greater than $\frac{1}{2}$, distinguish the output from a ideal random bit-sequence of the same length.

Definition

A PRBG is said to pass the **next-bit test** iff. no polynomial-time algorithm can, with probability greater than $\frac{1}{2}$, predict the (*n* + 1)-th bit of output given the previous *n* bits.

Theorem (Yao [11])

If a PRBG passes the next-bit test, it will pass all statistical tests.



Definition

Per [15, Section 4], imagine an attacker compromises the PRBG state at time *t*: we term a PRBG **back-tracking resistant** (resp. **prediction resistant**) if said attacker cannot distinguish between an (unseen) PRBG output at time t' < t (resp. t' > t) and an ideal random bit-sequence of the same length.

Definition

A Cryptographically Secure Pseudo-Random Bit Generator (CS-PRBG) is simple a PRBG whose properties make it suitable for use within a cryptographic use-case. A CS-PRBG should (at least)

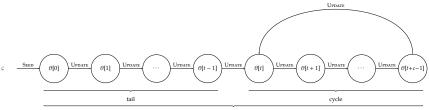
- 1. be a PRBG of sufficient quality, i.e., pass the next-bit test, and
- 2. resist state compromise attacks, i.e., be back-tracking and prediction resistant.



Part 1: in theory (13) (Pseudo-)random bit generators

Problem: our construction is deterministic, so

- the same ς will yield the same $\theta[0]$ and hence any $\theta[j]$ for j > 0,
- recovery of ζ allows computation of any $\theta[j]$ for $j \ge 0$,
- recovery of $\theta[i]$ allows computation of any $\theta[j]$ for j > i,
- the set \hat{S} is finite, so per



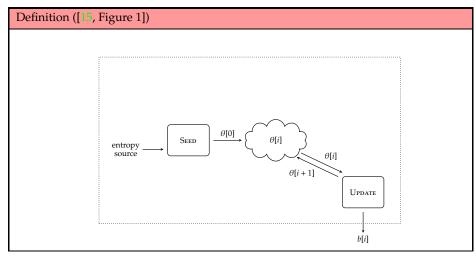
period

the state, and thus also the output, will eventually cycle.

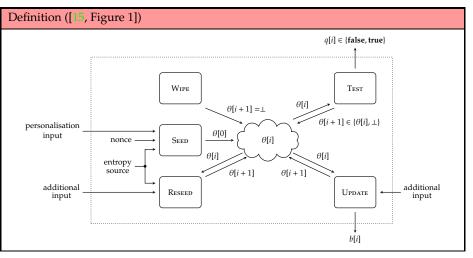
► Solution:

- 1. select parameters that mitigate such issues, and
- 2. introduce selected *non*-determinism.











Part 2: in practice (1)

(Sub-)agenda: explain selected, example designs, organised into 4 classes, i.e.,

- 1. "classic",
- 2. software-oriented,
- 3. hardware-oriented,
- 4. system-oriented,

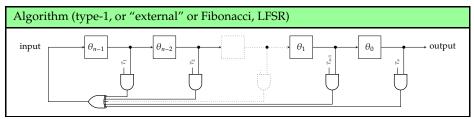
with a focus on design properties and trade-offs between them, e.g.,

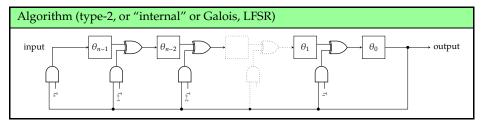
- efficiency,
- security, i.e., quality of (pseudo-)random output,
- interface,
- assumptions,
- ► ...



Part 2: in practice (2) Class #1: "classic"

Design: Linear-Feedback Shift Registers (LFSR) [5, 6].







Design: Blum-Blum-Shub (BBS) [10].

Algorithm (BBS.SEED)

Input: A security parameter λ , and a seed ς **Output:** An initial state $\theta[0]$

Use entropy provided by ς to perform the following steps:

- 1. Select two random ($\lambda/2$)-bit primes *p* and *q* such that $p \equiv q \equiv 3 \pmod{4}$, and compute $N = p \cdot q$.
- 2. Select a random $s \in \{0, 1, \dots, N-1\}$ such that gcd(s, N) = 1.
- 3. Compute $s[0] = s^2 \pmod{N}$.
- 4. Return $\theta[0] = (N, s[0])$.



Part 2: in practice (3) Class #2: software-oriented

Design: Blum-Blum-Shub (BBS) [10].

Algorithm (BBS.UPDATE)

Input: A current state $\theta[i] = (N, s[i])$ **Output:** A next state $\theta[i + 1]$, and $n_b = 1$ bit pseudo-random output b[i]

- 1. Compute $s[i + 1] = s[i]^2 \pmod{N}$.
- 2. Let $b[i] = s[i + 1] \pmod{2}$, i.e., b[i] = LSB(s[i + 1]).
- 3. Return $\theta[i + 1] = (N, s[i + 1])$ and b[i].



Part 2: in practice (4) Class #2: software-oriented

Design: ANSI X9.31 [13, Appendix A.2.4].

Algorithm (X9.31.SEED)

Input: A security parameter λ , and a seed ς **Output:** An initial state $\theta[0]$

1. Use λ to select a block cipher with an n_k -bit key size and n_b -bit block size, e.g.,

3DES	\sim	$n_b = 64$,	$n_k = 192$
AES-128	\sim	$n_b = 128$,	$n_k = 128$
AES-192	\sim	$n_b = 128$,	$n_k = 192$
AES-256	\sim	$n_b = 128$,	$n_k = 256$

- 2. Use entropy provided by ς to derive an n_k -bit cipher key k (or pre-select a k for the PRBG).
- 3. Use entropy provided by *c* to derive an *n*_b-bit block *s*[0].
- 4. Return $\theta[0] = (k, s[0])$.



Part 2: in practice (4) Class #2: software-oriented

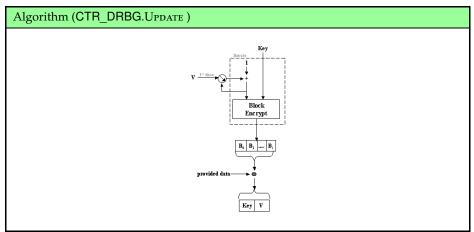
Design: ANSI X9.31 [13, Appendix A.2.4].

Algorithm (X9.31.UPDATE)

Input: A current state $\theta[i] = (k, s[i])$ **Output:** A next state $\theta[i + 1]$, and n_b -bit pseudo-random output b[i]

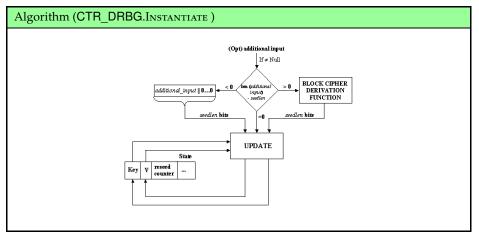
- 1. Compute t' = Enc(k, t), where *t* is a n_b -bit representation of the current time.
- 2. Compute $b[i] = \text{Enc}(k, t' \oplus s[i])$.
- 3. Compute $s[i + 1] = \text{Enc}(k, t' \oplus b[i])$.
- 4. Return $\theta[i + 1] = (k, s[i + 1])$ and b[i].





http://csrc.nist.gov/publications/nistpubs/800-90A/SP800-90A.pdf

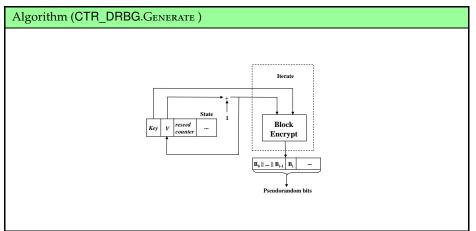
Daniel Page (csdsp0bristol.ac.u Applied Cryptology University of BRISTOL



http://csrc.nist.gov/publications/nistpubs/800-90A/SP800-90A.pdf

© Daniel Page (csdsp@bristol.ac.u

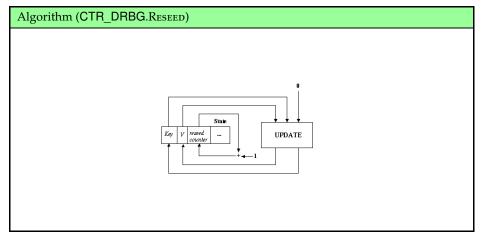
University of BRISTOL



http://csrc.nist.gov/publications/nistpubs/800-90A/SP800-90A.pdf

© Daniel Page (csdsp@bristol.ac.u Applied Cryptology





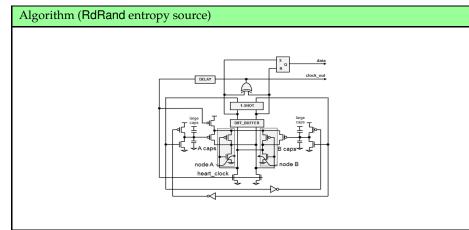
http://csrc.nist.gov/publications/nistpubs/800-90A/SP800-90A.pdf

© Daniel Page (csdsp0bristol.ac.uk Applied Cryptology

Market University of

Part 2: in practice (6) Class #3: hardware-oriented

Design: Intel Secure Key [12].



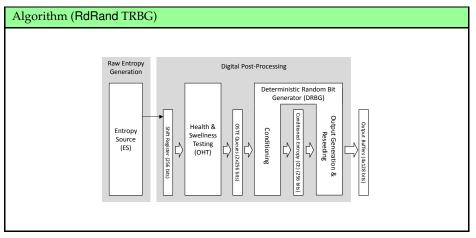
http://www.cryptography.com/public/pdf/Intel_TRNG_Report_20120312.pdf

© Daniel Page (csdsp0bristol.ac.)

Market University of

Part 2: in practice (6) Class #3: hardware-oriented

Design: Intel Secure Key [12].



http://www.cryptography.com/public/pdf/Intel_TRNG_Report_20120312.pdf

© Daniel Page (csdsp0bristol.ac

Applied Cryptology

Main University of の間 BRISTOL

Part 2: in practice (7) Class #3: hardware-oriented

Design: Intel Secure Key [12].

Listing (RdRand interface)

```
1 bool rdrand64( uint64_t* r ) {
2   bool success;
3
4   asm( "rdrand %0 ; setc %1"
5      : "=r" (*r), "=qm" (success) );
6
7   return success;
8 }
```

Listing (RdRand interface)

```
1 bool rdrand64_retry( uint64_t* r, int l ) {
2     int i = 0;
3
4     do {
5        if( rdrand64( r ) ) {
6            return true;
7        }
8     } while( i++ < l );
9
10     return false;
11 }</pre>
```



Part 2: in practice (8) Class #4: system-oriented

Design: Linux.

- circa 1994(ish):
 - maintain entropy pool θ[i], injecting entropy, e.g., from system-related events,
 - define a predicate

$$P(\theta[i]) = \begin{cases} \text{ false } & \text{if estimated entropy in } \theta[i] \text{ is deemed insufficient} \\ \text{true } & \text{if estimated entropy in } \theta[i] \text{ is deemed sufficient} \end{cases}$$

based on the concept of entropy estimation,

expose θ[i] to user-space via the (pseudo) files

write to $/\text{dev}/\text{random} \simeq \text{inject entropy into } \theta[i]$

read from /dev/random $\simeq \begin{cases} \text{ if } P(\theta[i]) = \text{false, block then sample from PRNG (re)seeded from } \theta[i] \\ \text{ if } P(\theta[i]) = \text{true,} \\ \text{ then sample from PRNG (re)seeded from } \theta[i] \end{cases}$

read from /dev/urandom \simeq sample from PRNG (re)seeded from $\theta[i]$



Part 2: in practice (8) Class #4: system-oriented

Design: Linux.

- circa 2014(ish):
 - update re. additional system call

```
ssize_t getrandom( void* x, size_t n, unsigned int flags )
```

where

```
\texttt{getrandom} \simeq \left\{ \begin{array}{ll} \text{if PRNG has not been initialised, then do} & \texttt{block} \\ \text{if PRNG has} & \texttt{been initialised, then do not block} \end{array} \right.
```

this yields clear(er) semantics, and avoids need for file handle.



Design: Linux.

- circa 2016(ish):
 - update re. PRNG, which is changed from being based on SHA-1 to ChaCha20,
 - this yields, e.g., lower latency with respect to sampling output.



Part 2: in practice (8) Class #4: system-oriented

Design: Linux.

- circa 2020(ish):
 - update re. file-based semantics

 $/dev/urandom \simeq do not block$

 $/dev/random \simeq \begin{cases} if PRNG has not been initialised, then do block if PRNG has been initialised, then do not block do not block been initialised. The second second$



Conclusions

Quote

Any one who considers arithmetical methods of producing random digits is, of course, in a state of sin.

- von Neumann (http://en.wikiquote.org/wiki/Randomness)

Quote

The generation of random numbers is too important to be left to chance.

- Coveyou (http://en.wikiquote.org/wiki/Randomness)

Quote

The design of such pseudo-random number generation algorithms, like the design of symmetric encryption algorithms, is not a task for amateurs.

- Eastlake, Schiller, and Crocker [14]



Conclusions

Take away points:

- 1. A high-quality source of randomness is fundamental to more or less *every* security proof: it might be an assumption in in theory, but in practice this issue requires care.
- 2. Iff. you need to develop your own PRBG implementation, use a standard (e.g., NIST SP800-90A [15]) design or framework ...
- 3. ... often such a design can leverage a primitive (e.g., a block cipher) you need anyway, thus reducing effort, attack surface, etc.
- 4. Some golden rules:
 - use a large, high-entropy seed,
 - avoid reliance on a single entropy source where possible,
 - opt for a cryptographically secure design and ensure it is parameterised correctly,
 - hedge against failure via robust pre- and post-processing where need be,
 - include quality tests on pseudo-randomness generation (e.g., alongside functional unit testing),
 - don't compromise security for efficiency,
 - ► .



Additional Reading

- Wikipedia: Randomness. URL: https://en.wikipedia.org/wiki/Randomness.
- Wikipedia: Pseudorandomness. URL: https://en.wikipedia.org/wiki/Pseudorandomness.
- Wikipedia: /dev/random. URL: https://en.wikipedia.org/wiki/dev/random.
- Wikipedia: RDRAND. uRL: https://en.wikipedia.org/wiki/RDRAND.
- K.H. Rosen. "Chapter 7: Discrete probability". In: Discrete Mathematics and Its Applications. 7th ed. McGraw Hill, 2013.
- A.J. Menezes, P.C. van Oorschot, and S.A. Vanstone. "Chapter 5: Pseudorandom bits and sequences". In: Handbook of Applied Cryptography. CRC, 1996. URL: http://cacr.uwaterloo.ca/hac/about/chap5.pdf.
- D. Johnston. Random Number Generators Principles and Practices: A Guide for Engineers and Programmers. 1st ed. De|G Press, 2018.
- D. Eastlake, J. Schiller, and S. Crocker. Randomness Requirements for Security. Internet Engineering Task Force (IETF) Request for Comments (RFC) 4086. 2005. URL: http://tools.ietf.org/html/rfc4086.



References

- [1] Wikipedia: /dev/random. uRL: https://en.wikipedia.org/wiki/dev/random (see p. 43).
- [2] Wikipedia: Pseudorandomness. URL: https://en.wikipedia.org/wiki/Pseudorandomness (see p. 43).
- [3] Wikipedia: Randomness. URL: https://en.wikipedia.org/wiki/Randomness (see p. 43).
- [4] Wikipedia: RDRAND. uRL: https://en.wikipedia.org/wiki/RDRAND (see p. 43).
- [5] S.W. Golomb. Shift Register Sequences. 3rd ed. https://doi.org/10.1142/9361. Aegean Park Press, 2017 (see p. 25).
- M. Goresky and A. Klapper. Algebraic Shift Register Sequences. 1st ed. https://doi.org/10.1017/CB09781139057448. Cambridge University Press, 2012 (see p. 25).
- [7] D. Johnston. Random Number Generators Principles and Practices: A Guide for Engineers and Programmers. 1st ed. De/G Press, 2018 (see p. 43).
- [8] A.J. Menezes, P.C. van Oorschot, and S.A. Vanstone. "Chapter 5: Pseudorandom bits and sequences". In: Handbook of Applied Cryptography. CRC, 1996. URL: http://cacr.uwaterloo.ca/hac/about/chap5.pdf (see pp. 7, 8, 18, 43).
- K.H. Rosen. "Chapter 7: Discrete probability". In: Discrete Mathematics and Its Applications. 7th ed. McGraw Hill, 2013 (see p. 43).
- [10] L. Blum, M. Blum, and M. Shub. "A Simple Unpredictable Pseudo-Random Number Generator". In: SIAM Journal on Computing 15.2 (1986), pp. 364–383 (see pp. 26, 27).
- [11] A.C. Yao. "Theory and application of trapdoor functions". In: Symposium on Foundations of Computer Science (SFCS). 1982, pp. 80–91 (see p. 19).
- [12] Intel Digital Random Number Generator (DRNG) Software Implementation Guide. Tech. rep. Intel Corp., 2012. URL: http://software.intel.com/sites/default/files/m/d/4/1/d/8/441_Intel_R_DRNG_Software_Implementation_ Guide_final_Aug7.pdf (see pp. 34-36).
- [13] Digital Signatures Using Reversible Public Key Cryptography for the Financial Services Industry. American National Standards Institute (ANSI) Standard X9.31. 1993 (see pp. 28, 29).
- [14] D. Eastlake, J. Schiller, and S. Crocker. *Randomness Requirements for Security*. Internet Engineering Task Force (IETF) Request for Comments (RFC) 4086. 2005. URL: http://tools.ietf.org/html/rfc4086 (see pp. 7, 8, 41, 43).



References

- [15] Recommendation for Random Number Generation Using Deterministic Random Bit Generators. National Institute of Standards and Technology (NIST) Special Publication 800-90A. 2012. URL: http://csrc.nist.gov (see pp. 10–13, 20, 22, 23, 30–33, 42).
- [16] Recommendation for the Entropy Sources Used for Random Bit Generation. National Institute of Standards and Technology (NIST) Special Publication 800-90B. 2012. URL: http://csrc.nist.gov (see p. 9).

