Key extraction from general non-discrete signals

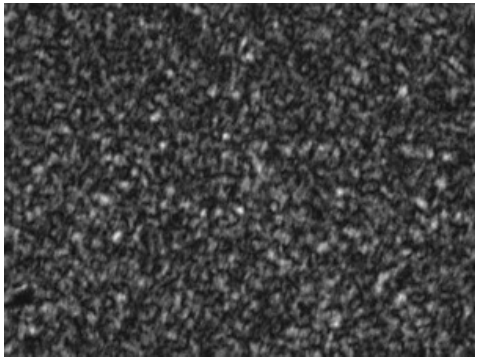
Evgeny Verbitskiy, Pim Tuyls, Chibuzo Obi, Berry Schoenmakers, Boris Škorić

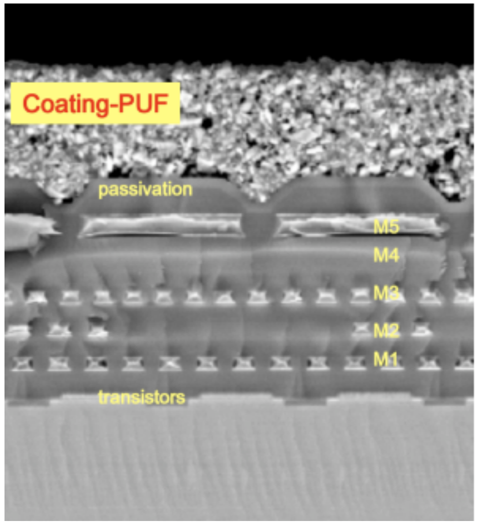
WISSEC 2008
13&14 November, Eindhoven

Outline

- Key extraction from continuous sources
- Defining properties of a Continuous-Source Fuzzy Extractor
- Partitioning scheme
- What if attacker has better knowledge of the source?

Almost all real-life sources generate *real* numbers, not discrete.



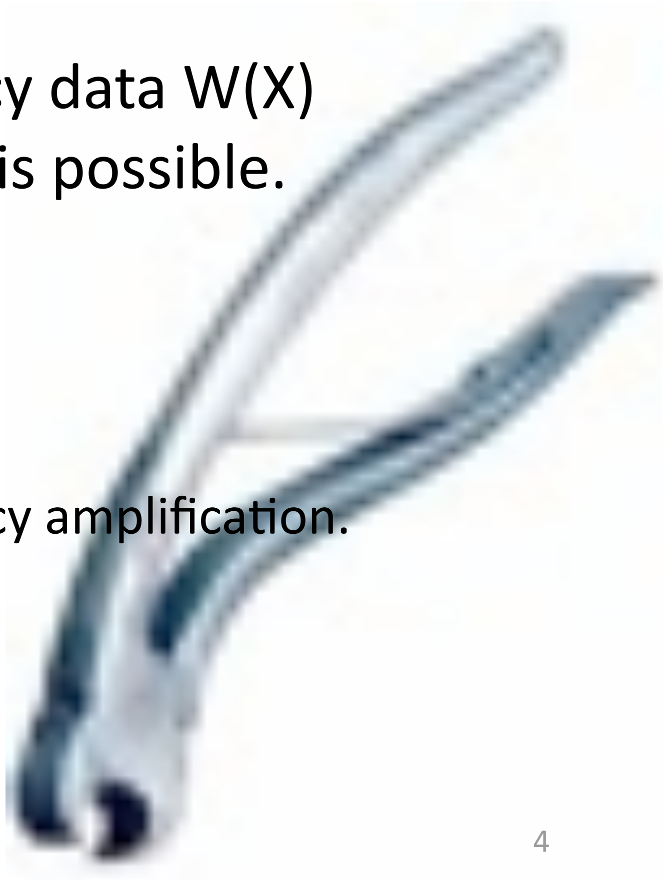


Key extraction

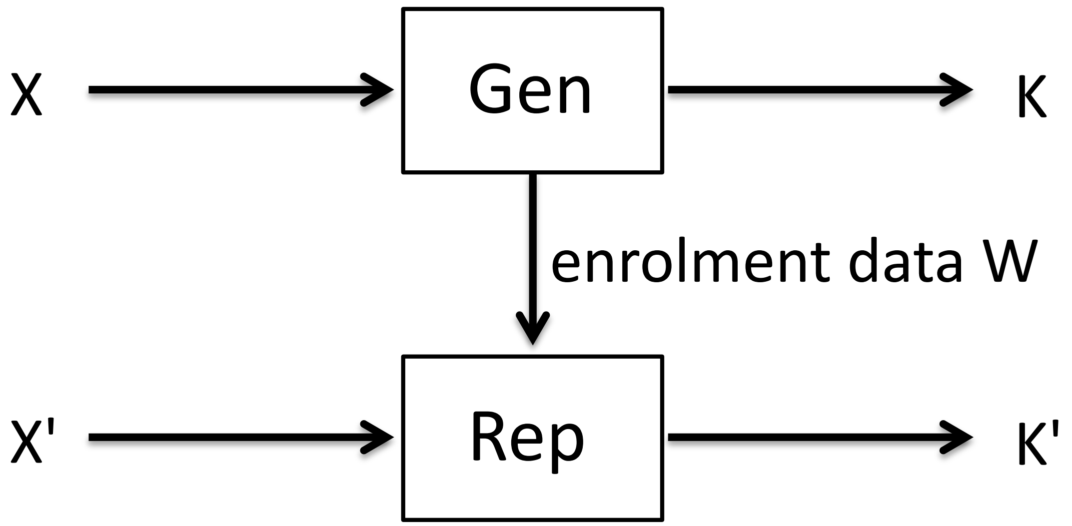
- Privacy amplification:
  Given a non-uniform source X, derive an L-bit string f(X) as uniformly distributed as possible on $\{0,1\}^L$.

- Information reconciliation:
  If the source is noisy, then some redundancy data W(X) must be given before privacy amplification is possible.
  - biometrics
  - PUFs

- Fuzzy extractor:
  - Does both information reconciliation and privacy amplification.
  - Extracts secret key K from noisy source.
  - Aims for high entropy H(K | W).



Fuzzy Extractor



Traditionally defined for discrete source X. **But most sources are continuous!**

- extra step: discretization of X
- degree of freedom that can be exploited

We extend the definition [Buhan et al. 2007] of

Continuous-Space Fuzzy Extractor

- Correctness
- Security

[Dodis et al. 2003]

3.2 Fuzzy Extractors

Definition 5. An $(\mathcal{M}, m, \ell, t, \epsilon)$-*fuzzy extractor*

avoid

Correctness definitions

Old way
- Requires distance measure
- Hard to see failure prob.

1. t-correct:
   If d(x, x')<t then K'=K.

2. Worst case ϵ-stochastically noise resilient:

$$\forall \mathrm{x} \quad \mathrm{Prob}[\mathrm{Rep}(X', w_x) = k_x] \geq 1 - \varepsilon$$

3. On average ϵ-stochastically noise resilient:
   For (k_x, w_x)=Gen(x):

$$\int \mathrm{Prob}[\mathrm{Rep}(X', w_x) = k_x] \mathrm{d}x \geq 1 - \varepsilon$$

Security definitions

1. (m,δ)-secure.
   $H_\infty(X) \geq m \Rightarrow \Delta(KW, U_L W) \leq \delta$.

$H_\infty(X)$ not defined for cont. distribution

2. Worst case m-secure:
   $\forall w \quad H_\infty(K|W=w) \geq m$.

3. On average m-secure:

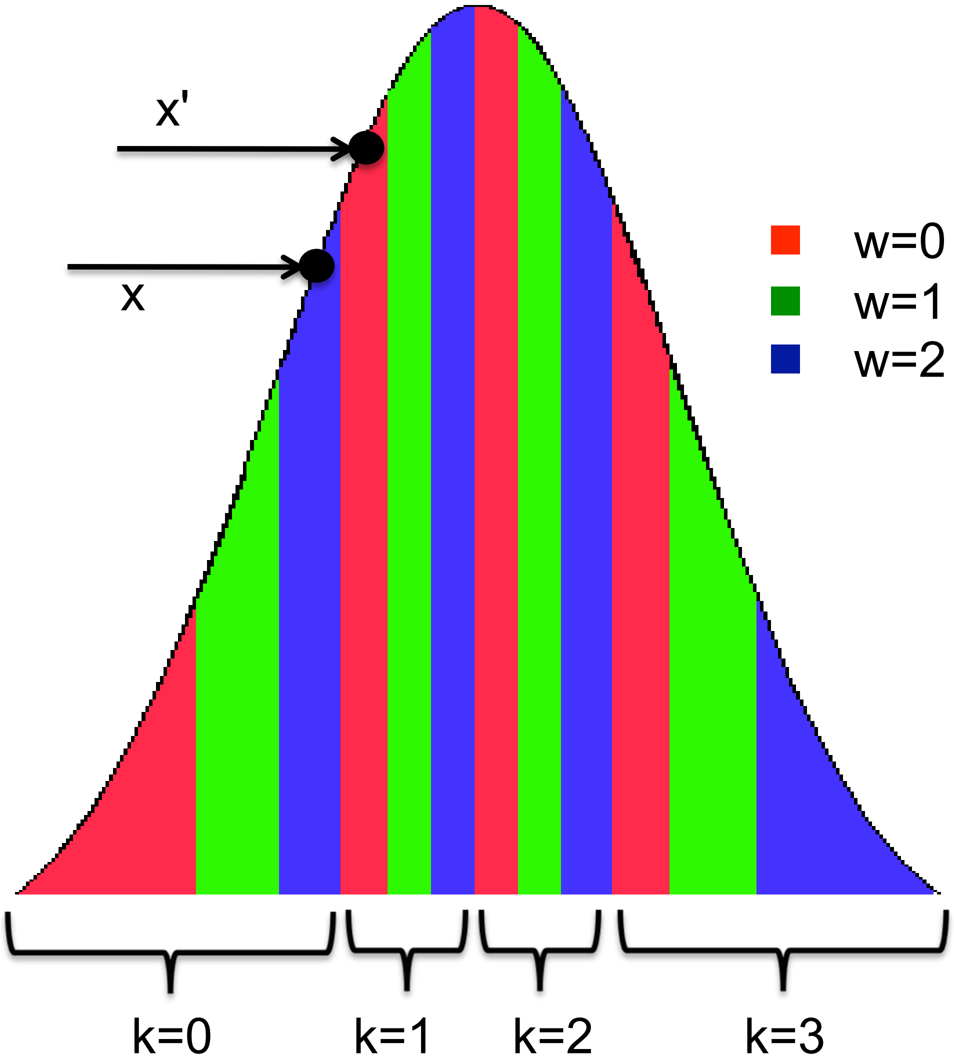
$$\tilde{H}_\infty(K \mid W) \geq m$$

average conditioning

Continuous-Space Fuzzy Extractor: Partitioning scheme

Two nested equiprobable partitions

- secret K = outer index
- helper W = inner index



Enrollment

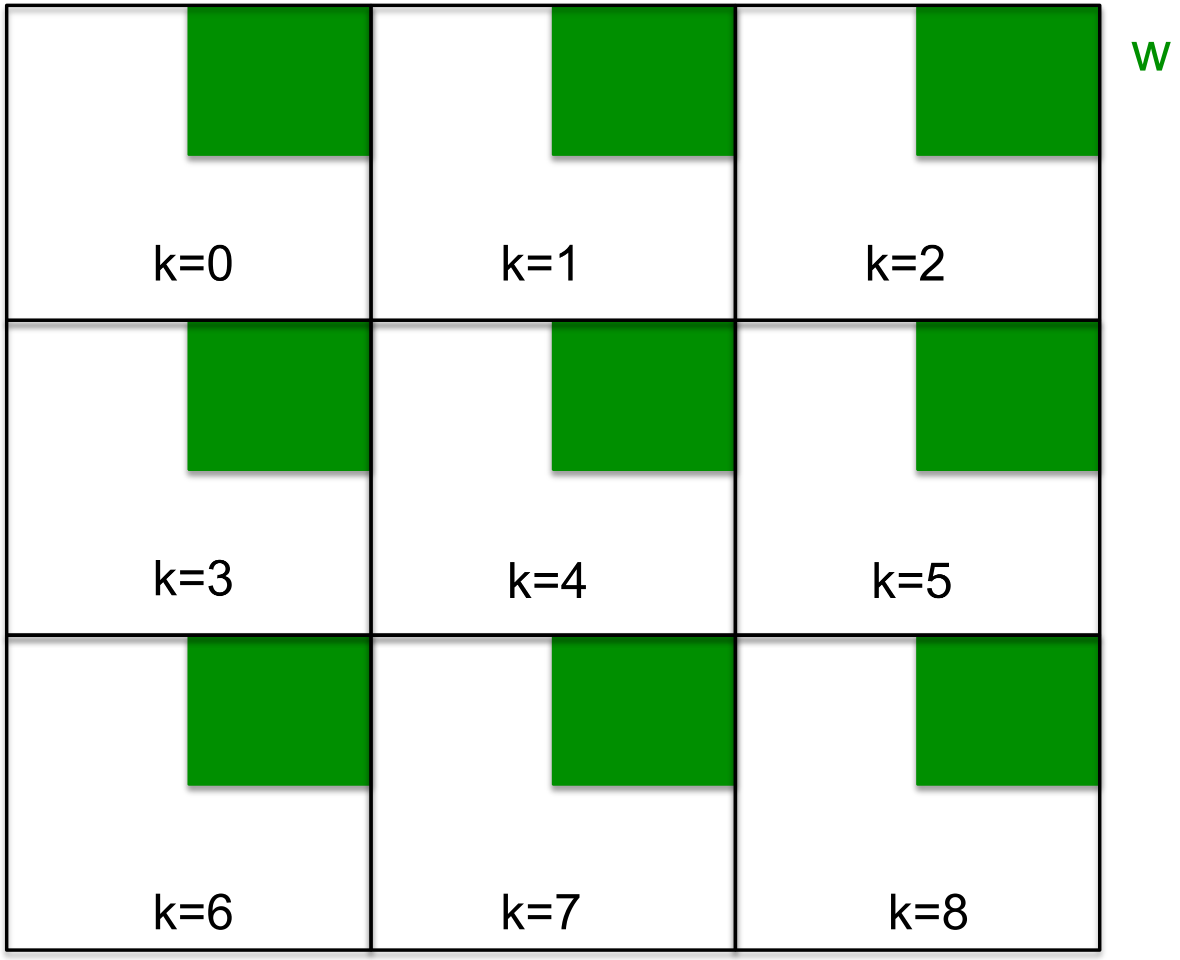
- Measure x.
- k=0.
- Store w=2.

Reconstruction

- Measure x'.
- Read w.
- Go to nearest blue interval.
- Read off k=0.

Gap between (k,w) and (k±1,w) reduces noise.

Partitioning scheme: 2D toy example



Gaps between (k,w) and (k+Δk,w) reduce noise.

Properties of the partitioning scheme

$K \in \{0,1\}^L$. $W \in \{0,1\}^b$.

- Security of the extracted key K: **$H(K|W) = H(K) = L$**.
  - *Helper data reveals nothing about key*.
  - *Key is uniform*.
  - *"Worst-case L-secure"*.
- Leakage about the source X: **$I(X; W) = H(W) = b$**.
  - *Helper data leaks b bits about raw measurement*.
  - *Inevitable!*
- Correctness properties:
  - *depends on specific noise distribution*.

What if source distribution is not known exactly?

Partitioning scheme based on best guess

- Key not exactly uniform
- **Attacker may have better knowledge of X and exploit it!**

Lemma: $$\tilde{H}_\infty(K \mid W) \geq L - \log(1 + \delta\, 2^{L+b})$$

with $$\delta = \frac{1}{2} \sum_{k,w} \left| \Pr[K = k \wedge W = w] - \frac{1}{2^{L+b}} \right|$$

Gaussian case: $$\delta \leq \frac{\sqrt{(\tilde{\sigma} - \sigma)^2 + (\tilde{\mu} - \mu)^2}}{\min(\sigma, \tilde{\sigma})}$$

Why <u>average</u> conditioning on W?

Attacker does not control the helper data.

Conclusions

- **Adapted Fuzzy Extractor definition for non-discrete source**
  - correctness and security properties
  - generalization of [Buhan et al.]
- **Explicit construction for known prob. densities**
  - discretization: exploitable extra degree of freedom
  - nested equiprobable intervals
  - perfectly uniform key
  - noise reduced by gaps between intervals (k,w) and (k+Δk,w)
- **Effect of incomplete knowledge about source**
  - worst case assumption: attacker has full knowledge
  - average-case conditioning on W
  - derived bound on min-entropy of extracted key