Secrecy of High-Entropy Sources

Adam Smith, MIT (visiting HU)

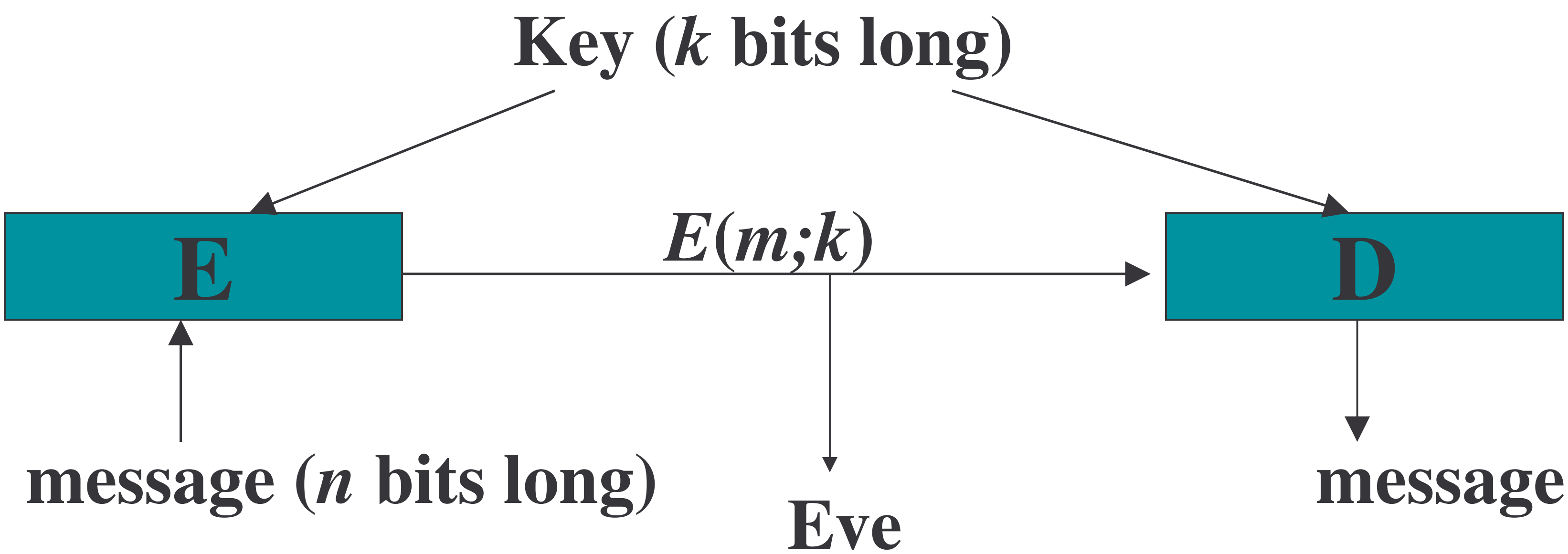
Joint work with Yevgeniy Dodis, NYU

Unconditional Secrecy When Information Leakage is Unavoidable

Adam Smith, MIT (visiting HU)

Joint work with Yevgeniy Dodis, NYU

Symmetric Encryption



- Shannon: Symmetric Encryption without computational assumptions requires $k \geq n$ (achieved by one-time pad)
- Russell and Wang 2002 [RW02]: What can be said when the message is guaranteed to have high entropy?

Russell-Wang: Entropic Security

Entropic security for symmetric encryption [RW02]:

1. No computational assumptions (statistical secrecy)
2. Assume message distribution has **high entropy**
3. Constructions with short key (not possible without #2)

Motivation:

- Systematic study, simplification of [RW02] definition
- Understand “high-entropy secrets” in simple setting
- Develop tools for settings other than encryption

Russell-Wang: Entropic Security

Entropic security for symmetric encryption [RW02]:

1. No computational assumptions (statistical secrecy)
2. Assume message distribution has **high entropy**
3. Constructions with short key (not possible without #2)

This talk:

- Definitions & Background
- Equivalent characterizations
- Simpler constructions
- Lower bounds
- Application to other settings

Definitions: Symmetric Encryption

- (No security requirements yet)
- Encryption Scheme: Pair of functions (E,D) :
  - E takes message $m \in \{0,1\}^n$
    key $s \in \{0,1\}^k$
    randomness $i \in \{0,1\}^r$

    Not shared
  - Ciphertext is $E(m,s;i)$ (write $E(M)$ for random i,s)
  - Decryption: $D(E(m,s;i)\ ,s) = m$ (with probability 1)
- Parameters: $n = |m|$, $k=|s|$
- $s \leftarrow U_k$ (= uniform distribution on $\{0,1\}^k$)

Min-Entropy of Random Variables

- There are various ways to measure entropy…
- Min-entropy: For random variable M on $\{0,1\}^n$:

$$H_\infty(M) = -\log\left(\max_m \Pr[M=m]\right)$$

- Uniform on $\{0,1\}^n$: $H_\infty(U_n) = n$
- " Message has min-entropy t " means that
  - No message arises with probability $\geq 2^{-t}$
  - Adversary's probability of guessing the message is $\leq 2^{-t}$

Entropic Security [RW02]

Definition: (E,D) is (λ,ε)-entropically secure if

$\forall$ distributions M on $\{0,1\}^n$ with $H_\infty(M) \geq n-\lambda$

$\forall$ (adversaries) $A:\{0,1\}^* \rightarrow \{0,1\}$

$\forall$ predicates $g:\{0,1\}^n \rightarrow \{0,1\}$

$\exists$ random variable A' (independent of M)

$$\left| \Pr[A(E(M)) = g(M)] - \Pr[A' = g(M)] \right| \leq \varepsilon$$

Entropic Security [RW02]

Definition: (E,D) is (λ,ε)-entropically secure if

$\forall$ distributions M on $\{0,1\}^n$ with $H_\infty(M) \geq n\text{-}\lambda$

$\forall$ (adversaries) $A:\{0,1\}^* \to \{0,1\}$

$\forall$ predicates $g:\{0,1\}^n \to \{0,1\}$

$\exists$ random variable A' (independent of M)

$$\left| \Pr[A(E(M)) = g(M)] - \Pr[A' = g(M)] \right| \leq \varepsilon$$

Caveats:

- Assumes that message has high entropy!
  What if the adversary knows more than you think he knows?
- Computational "issues": what happens when such a scheme gets plugged into more complex situations?

Entropic Security [RW02]

Definition: (E,D) is (λ,ε)-entropically secure if

$\forall$ distributions M on $\{0,1\}^n$ with $H_\infty(M) \geq n-\lambda$

$\forall$ (adversaries) $A:\{0,1\}^* \to \{0,1\}$

$\forall$ predicates $g:\{0,1\}^n \to \{0,1\}$

$\exists$ random variable A' (independent of M)

$$\left|\ \Pr[A(E(M)) = g(M)] - \Pr[A' = g(M)]\ \right| \leq \varepsilon$$

[RW02] There exist (λ,ε)-ES schemes with

$$k \approx \lambda + 3 \log(1/\varepsilon)$$

This work: equivalent definition, new constructions, lower bounds.

Context: Perfect Security [Shannon]

- Shannon: Perfect Security $\Leftrightarrow$ message independent of ciphertext

  $\forall$ distrib's M on $\{0,1\}^n$: $\quad M$ independent of $E(M)$

- Equivalently $\quad \forall\, m,m' \in \{0,1\}^n$: $E(m) \equiv E(m') \equiv E(U_n)$

  (sufficient to require independence only for $M=U_n$)

- Theorem: Perfect security requires $k \geq n$.

- "Proof": Take any possible ciphertext $\boldsymbol{c}$

  Perfect Secrecy $\Rightarrow$ $\boldsymbol{c}$ can be decrypted to any $m \in \{0,1\}^n$

  Each key decrypts $\boldsymbol{c}$ to at most one message

  $\geq 2^n$ different keys

Context: Computational Security [GM84]

Definition: (E,D) is semantically-secure if

$\forall$ distributions M on $\{0,1\}^n$

$\forall$ PPT (prob. poly. time) circuits (adversaries) A

$\forall$ functions $g:\{0,1\}^n \to \{0,1\}^*$

$\exists$ random variable A' (independent of M)

$$\left| \Pr[A(E(M)) = g(M)] - \Pr[A' = g(M)] \right| \leq \text{negligible}$$

Definition: (E,D) is message-indistinguishable if

$$\forall\, m,m' \in \{0,1\}^n \quad E(m) \approx_{\mathrm{PPT}} E(m')$$

Theorem [GM84]: Definitions above are equivalent.

Statistical Security?

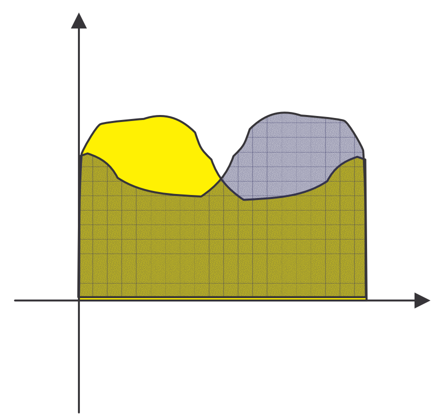
- Natural Generalizations: replace computational indistinguishability with statistical indistinguishability:
- Statistical Difference (L_1): For distributions $p_0(x)$, $p_1(x)$:

$$SD(p_0,p_1) = \tfrac{1}{2} \sum_x | p_0(x) - p_1(x) |$$

- *SD* measures distinguishability:

  If $b \leftarrow \{0,1\}$, $x \leftarrow p_b$ then

$$\max_A | \Pr[A(x)=b] - \tfrac{1}{2} | = \tfrac{1}{2} SD(p_0,p_1)$$



- (Notation: $X_1 \approx_\varepsilon X_2$ if $SD(X_1,X_2) \le \varepsilon$)

Statistical Security?

- Natural generalizations: replace computational indistinguishability with statistical indistinguishability

Definition: (E,D) is statistically ε-semantically-secure if
$\forall$ distrib's M, $\forall A$, $\forall g:\{0,1\}^n \to \{0,1\}^*$, $\exists A'$:

$$\left| \Pr[A(E(M)) = g(M)] - \Pr[A' = g(M)] \right| \leq \varepsilon$$

Definition: (E,D) is statistically ε-message-indistinguishable if
$\forall m,m' \in \{0,1\}^n$: $E(m) \approx_\varepsilon E(m')$

- Def's are equivalent, imply $k \geq n$ (as in perfect secrecy) but proofs go through 2-point distributions $M \leftarrow \{m,m'\}$

Entropic Security [RW02]

Definition: (E,D) is (λ,ε)-entropically secure if

$\forall$ distributions M on $\{0,1\}^n$ with $H_\infty(M) \geq n\text{-}\lambda$

$\forall$ (adversaries) $A{:}\{0,1\}^* \to \{0,1\}$

$\forall$ predicates $g{:}\{0,1\}^n \to \{0,1\}$

$\exists$ random variable A' (independent of M)

$$\left|\ \Pr[A(E(M)) = g(M)] - \Pr[A' = g(M)]\ \right| \leq \varepsilon$$

[RW02] There exist (λ,ε)-ES schemes with

$$k \approx \lambda + 3\log(1/\varepsilon)$$

Two constructions: twists on the one-time pad.

[RW02]: Two constructions

1. $E(m,s) = m \oplus b(s)$, with $b : \{0,1\}^k \to \{0,1\}^n$.
   - $b(\cdot)$ is carefully chosen: range is "δ-biased set"
   - Fourier-based proof works only for uniform message
   - $k \approx 2 \log n + 3 \log (1/\varepsilon)$ (here $\lambda = 0$)
2. $E(m,s;\ i) = (\phi_i\ ,\ \phi_i(m) + s)$
   - $\{\phi_i : \{0,1\}^n \to \{0,1\}^n\}$ are 3-wise independent permutations
   - $k \approx \lambda + 3 \log (1/\varepsilon)$ (works for all λ)
   - $3n$ bits of additional randomness, difficult proof

Outline

- Equiv. Def: Indistinguishability for high-entropy sources

  Intuition: Indistinguishable schemes ≈ extractors
- Two Simple, General Constructions:
  - Step in an expander graph
  - Random hash functions (less high-tech)
- Lower bounds: $k \geq \lambda$, (special case: $k \geq \lambda + \log(1/\varepsilon)$)
- "Stronger" Equiv. Def.: all functions hard to predict

  (not only predicates)

Indistinguishability for High Entropy

Def: (λ,ε)-entropically secure if $\forall\, M$, $H_\infty(M) \geq n\text{-}\lambda$, $\forall\, A$ $\forall$ pred. g

$\exists\, A'$: $\left|\ \Pr[A(E(M)) = g(M)] - \Pr[A' = g(M)]\ \right| \leq \varepsilon$

Recall: (Ordinary) semantic security $\Rightarrow$

$\forall$ distributions M,M': $E(M) \approx_{PPT} E(M')$

Definition: (E,D) is (t,ε)-indistinguishable (IND) if

$\forall$ distributions M,M' with $H_\infty(M)$, $H_\infty(M') \geq t$:

$$SD(E(M),E(M')) \leq \varepsilon$$

Proposition: **(λ,ε)-ES** equiv. to **(t , ε')-IND** for $t = n\text{-}\lambda\text{-}1$

Proof: (λ,ε)-ES $\Rightarrow$ $(n\text{-}\lambda\text{-}1, 4\varepsilon)$-IND

Fact: $H_\infty(M) \geq t \Rightarrow$ M is mixture of flat distrib's on 2^t pts.

- Take any M_0, M_1 of min-entropy $\geq t = n\text{-}\lambda\text{-}1$

  (Sufficient to prove lemma for flat distrib's on 2^t points)

- Suppose M_0, M_1 have disjoint support:

  Use $g(x) = b$ if $x \in \mathrm{supp}(M_b)$ and $M^* = M_b$ for $b \leftarrow \{0,1\}$

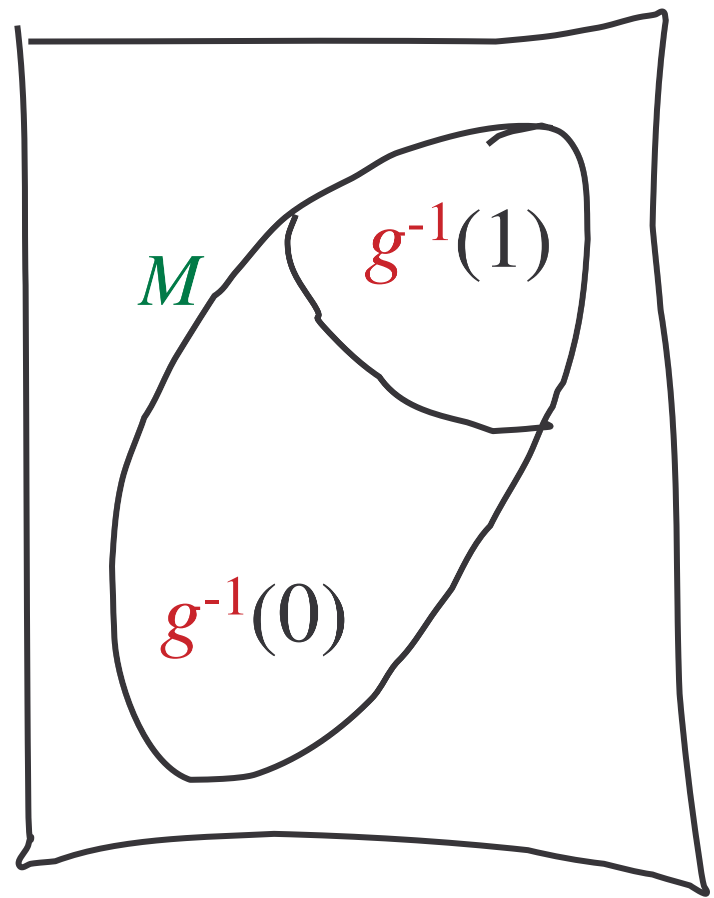
- $H_\infty(M^*) = t+1 = n\text{-}\lambda$ $\Rightarrow$ No A predicts g better than $½+\varepsilon$

  $\Rightarrow SD(\, E(M_0), E(M_1)\,) \leq 2\varepsilon$

- If M_0, M_1 not disjoint, find M_2 disjoint to both.

Proof: (n-λ-1,ε)-IND ⇒ (λ,ε)-ES

- Say $\Pr[A(E(M))=g(M)] \geq (1\text{-}p)+\varepsilon$ where $p = \Pr[g(M)=1] \leq ½$
- We want: M_0,M_1 disting'd by $A(E(\cdot))$
- **Try #1**: $M_b = g^{-1}(b)$
- **Problem**: $g^{-1}(1)$ may be too small (Min-entropy of M_1 too low – get weaker reduction)



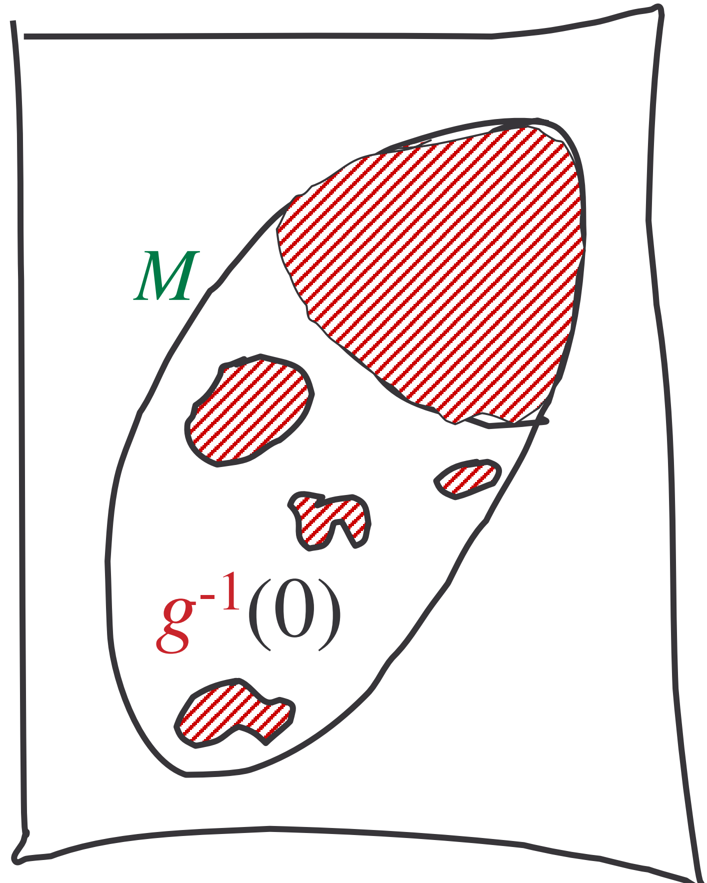
Proof: (n-λ-1,ε)-IND ⇒ (λ,ε)-ES

- Say $\Pr[A(E(M))=g(M)] \geq (1-p)+\varepsilon$ where $p = \Pr[g(M)=1] \leq ½$
- We want: M_0,M_1 disting'd by $A(E(\cdot))$
- **Try #2**: add random points from $g^{-1}(0)$

$q_m = \Pr[A(E(m))=1]$

$r_b = \Pr[A(E(M))=1 \mid g(M)=b]$

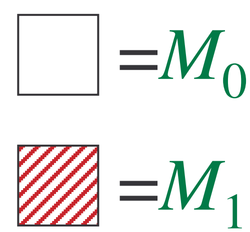
$= \mathbf{E}[\ q_M \mid g(M)=b\ ]$



In expectation: $\Pr[A(E(M_0))] = r_0$

$\Pr[A(E(M_1))] = 2p\ r_1 + (1-2p)r_0$

$\ldots \Rightarrow \Pr[A(E(M_1))] - \Pr[A(E(M_0))] \geq 2\varepsilon$



Recall: Indistinguishability

Def: (λ,ε)-entropically secure if $\forall M$, $H_\infty(M) \geq n\text{-}\lambda$, $\forall A$ $\forall$ pred. g $\exists A'$: $\left| \Pr[A(E(M)) = g(M)] - \Pr[A' = g(M)] \right| \leq \varepsilon$

Def: (t,ε)-indistinguishable (IND) if $\forall M_0, M_1$, $H_\infty(M_b) \geq t$:

$$E(M_0) \approx_\varepsilon E(M_1)$$

Proposition: **(λ,ε)-ES** equiv. to **(t , ε')-IND** for $t = n\text{-}\lambda\text{-}1$

- How can we use this?
- **Intuition**:

Indistinguishability $\approx$ extractor with "invertibility"

Two General Constructions

#1 : Steps on an expander graph

#2: Random Hashing

Expander Graphs

- Important tool in … everything.
- Expander = regular, undirected graph w
  - Let A = adjacency matrix of d-regular (
  - Vector (1,…,1) has eigenvalue d
  - Other eigenvalues $\in [-d,d]$
- *G* is a β-expander if other
- Random walks converge quickly:

When β is very small, walk converges in 1 step

Fact: If $H_\infty(p) \geq t$, then walk is ε-far from uniform after at most

$$\frac{n - t + 2\log(1/\varepsilon)}{2\log(1/\beta)}$$ steps, where $|G| = 2^n$.

Using Graphs for Encryption

- Encryption of m = random step from m
- Take regular G with $V=\{0,1\}^n$ and $d=2^k$
- Consider $\mathbf{E(m,s) = N(m,s)}$

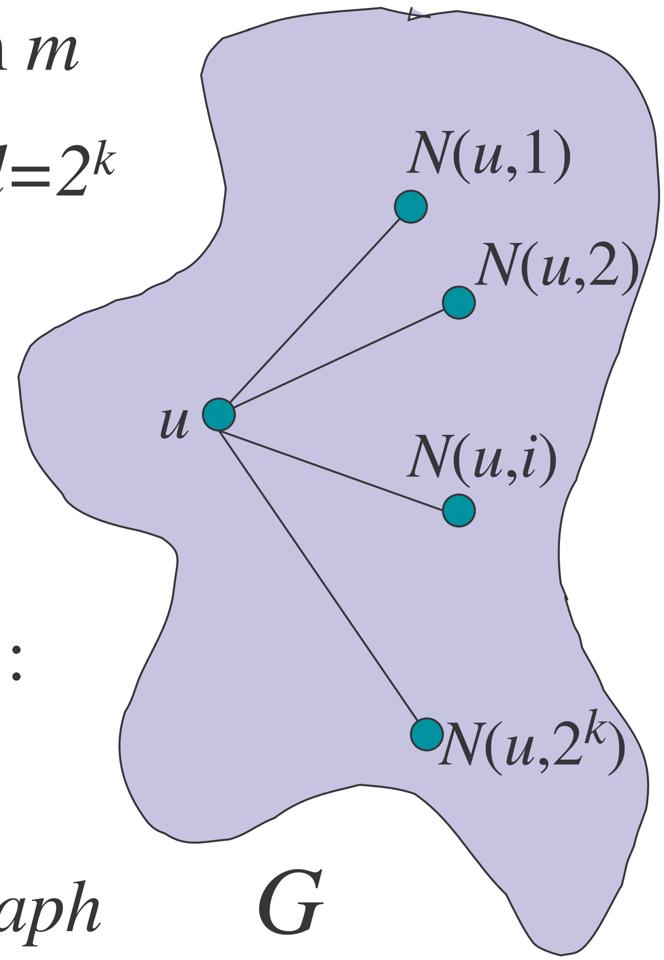
  ($N(u,i) = i^{\text{th}}$ neighbour of node u)

Q: When can you decrypt?

A: Need labeling N with an inverter N':

$$N'(\,N(u,i)\,,\,i) = u$$

Exercise: *Every regular undirected graph has an invertible labeling.*



Using Graphs for Encryption

- Encryption of m = random step from m
- Take regular G with $V=\{0,1\}^n$ and $d=2^k$
- Consider $\boldsymbol{E(m,s) = N(m,s)}$

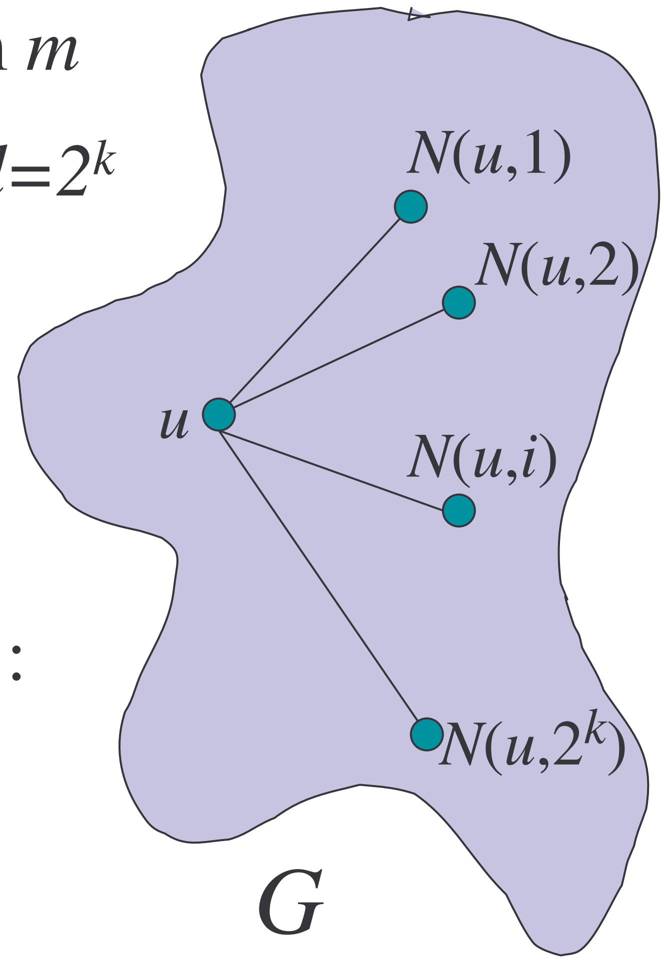
  ($N(u,i) = i^{\text{th}}$ neighbour of node u)

Q: When can you decrypt?

A: Need labeling N with an inverter N':

$$N'(\,N(u,i)\,,\,i) = u$$

Easier exercise: *Cayley graphs are invertible.*



Tangent: Cayley Graphs

- Let (V,*) be a group, $B=\{g_1,\ldots,g_d\}$ a set of generators.

  Cayley graph for $(V,*,B)$ has vertex set V and edges:

  $$E = \{\ (u,\ g*u) \mid u \in V,\ g \in B\ \}.$$

- Graph is undirected if B contains its inverses.
  - E.g. hypercube $\{0,1\}^n$ with B ={vectors of weight 1}
- Natural labeling is $N(u,i) = g_i*u$
- Invertible since $N'(w,i) = g_i^{-1}*w$
- Graphs in this talk are Cayley graphs

Using Graphs for Encryption

- Take regular G with $V=\{0,1\}^n$ and $d=2^k$
- Consider $\boldsymbol{E(m,s) = N(m,s)}$
  ($N(u,i) = i$ th neighbour of node u)

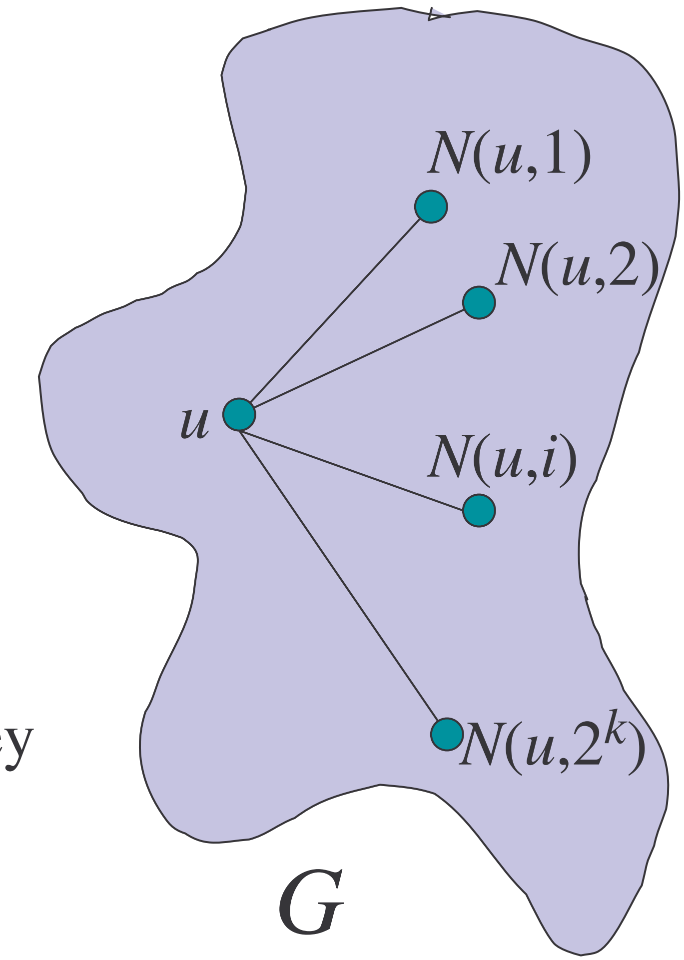
Q: When is E (t,ε)-indistinguishable?

A: When walk converges in 1 step.

Sufficient: G is β-expander with $\beta^2 \leq \varepsilon^2\, 2^{t-n}$

Theorem[LPS]: There exist (explicit) Cayley graphs with $\beta^2 \approx 1/d = 2^{-k}$

Corollary: *There exist (λ,ε)-ES encryption schemes with* $\boldsymbol{k \approx \lambda + 2\log(1/\varepsilon)}$



[RW02]: Two constructions

1. $E(m,s) = m \oplus b(s)$, with $b : \{0,1\}^k \rightarrow \{0,1\}^n$.
   - $b(\cdot)$ is carefully chosen: range is "δ-biased set"
   - Fourier-based proof works only for uniform message
   - $k \approx 2 \log n + 3 \log (1/\varepsilon)$ (here $\lambda = 0$)
2. $E(m,s;\, i) = (\phi_i\, ,\, \phi_i(m) + s)$
   - $\{\phi_i : \{0,1\}^n \rightarrow \{0,1\}^n\}$ are 3-wise independent permutations
   - $k \approx \lambda + 3 \log (1/\varepsilon)$ (works for all λ)
   - $3n$ bits of additional randomness, difficult proof

[RW02]: First construction

1. $E(m,s) = m \oplus b(s)$, with $b : \{0,1\}^k \rightarrow \{0,1\}^n$.
   - $b(\cdot)$ is carefully chosen: range is "δ-biased set"
   - Fourier-based proof works only for uniform message
   - $k \approx 2 \log n + 3 \log (1/\varepsilon)$ (here $\lambda = 0$)

Same scheme, new analysis:

- $G =$ Cayley graph for $\{0,1\}^n$ with generators $\{b(s) \mid s \in \{0,1\}^k\}$
- [BSVW] observe that G is a δ-expander (degree $= n^2/\delta^2$)
- Previous slide $\Rightarrow$ $\boldsymbol{k = \lambda + 2 \log n + 2 \log (1/\varepsilon)}$

  (Same proof works for all λ)

Two General Constructions

#1 : Steps on an expander graph

#2: Random Hashing

Hashing Construction

Goals:

- Schemes with simple combinatorial proofs
- Generalize second construction of Russell and Wang

Outline:

- Modify “Left-over Hash Lemma”
  (a.k.a. “Privacy Amplification”)
- One proof for simplified scheme and Russell-Wang construction

Pairwise Independent Hash Functions

- A collection of functions $\mathcal{H}=\{h_i\}$, h_i: $\mathcal{X}\rightarrow\mathcal{Y}$ is 2-wise independent if $\forall\, x,x' \in \mathcal{X}$, $x \neq x'$, and $\forall\, y,y' \in \mathcal{Y}$:

$$\Pr_{H\leftarrow\mathcal{H}}[\ H(x)=y \text{ and } H(x')=y'\ ] = 1/|Y|^2$$

- Equivalently: $\forall\, x,x' \in \mathcal{X}$, $x \neq x'$, whe

$H(x)$, $H(x')$ are independent a

- Typical construction: If $\mathcal{X}=\{0,1\}^n$, $\mathcal{Y}=\{$ $\}^p$, $p \leq n$,

View $\mathcal{X}=\{0,1\}^n$ as GF(2^n), use

$$\mathcal{H}=\left\{\ x \mapsto \textbf{last-}p\textbf{-bits}(ax+b)\ \middle|\ a,b \in \mathrm{GF}(2^n)\ \right\}$$

Requires $\approx 2n$ bits of randomness

Left-over Hash Lemma / Privacy Amplification [BBR,IZ,…]

LOHL [IZ89]: Let $\mathcal{H}=\{h_i\}$ be 2-wise : (n bits) $\rightarrow$ (p bits)

If $H_\infty(\mathbf{M}) \geq t$ and $t \geq p + 2\log(1/\varepsilon)$ then

$(H , H(\mathbf{M})) \approx_\varepsilon (H , U_p)$, when $H \leftarrow \mathcal{H}$.

- Good for extractors, but not encryption…

LOHL': Let $\mathcal{H}=\{h_i\}$ be 2-wise : (n' bits) $\rightarrow$ (n bits)

If $\mathbf{A},\mathbf{B}$ indep., and $H_\infty(\mathbf{A}) + H_\infty(\mathbf{B}) \geq n + 2\log(1/\varepsilon)$ then

$(H , \mathbf{A} \oplus H(\mathbf{B})) \approx_\varepsilon (H , U_n)$, when $H \leftarrow \mathcal{H}$

Modified Left-over Hash Lemma

LOHL': Let $\mathcal{H}=\{h_i\}$ be 2-wise : (n' bits) $\to$ (n bits)

If $\mathbf{A},\mathbf{B}$ indep., and $H_\infty(\mathbf{A}) + H_\infty(\mathbf{B}) \geq n + 2\log(1/\varepsilon)$ then

$(H , \mathbf{A} \oplus H(\mathbf{B})) \approx_\varepsilon (H , U_n)$, when $H \leftarrow \mathcal{H}$

Proof idea: As with LOHL, compute **collision probability**

- $\mathrm{CP}(\mathbf{X}) = \sum_x p_x^2$ where $p_x = \Pr[\mathbf{X}=x]$
- $H_\infty(\mathbf{X}) \geq t \Rightarrow \mathrm{CP}(\mathbf{X}) \leq 2^{-t}$

Collision probability of $(H , \mathbf{A} \oplus H(\mathbf{B}))$ is at most $\dfrac{1+2^{n-t-t'}}{|\mathcal{H}|\, 2^n}$

- If $\mathbf{X} \in S$ and $\mathrm{CP}(\mathbf{X}) = (1+2\varepsilon^2)/|S|$ then $X \approx_\varepsilon$ uniform

$\therefore$ $(H , \mathbf{A} \oplus H(\mathbf{B})) \approx_\varepsilon$ uniform. QED.

Using LOHL' for Encryption

LOHL': Let $\mathcal{H}=\{h_i\}$ be 2-wise : (n' bits) $\rightarrow$ (n bits)

If A,B indep., and $H_\infty(A) + H_\infty(B) \geq n + 2\log(1/\varepsilon)$ then

$(H , A \oplus H(B)) \approx_\varepsilon (H , U_n)$, when $H \leftarrow \mathcal{H}$

Schemes a) $E(m,s;h) = (h , \quad m + h(s))$

or b) $E(m,s;h) = (h , \ h(m) + s \quad)$

Here $\mathcal{H}$ contains only permutations

- Either a) set $A=M$, $B=S$

  or b) set $A=S$, $B=M$

- LOHL' $\Rightarrow$ (t,ε)-indistinguishable for $k \geq (n\text{-}t) + 2\log(1/\varepsilon)$

  $\Rightarrow$ (λ,ε)-ES for $k \geq \lambda + 2 \log(1/\varepsilon)$

[RW02]: Two constructions

1. $E(m,s) = m \oplus b(s)$, with $b : \{0,1\}^k \rightarrow \{0,1\}^n$.
   - $b(\cdot)$ is carefully chosen: range is "δ-biased set"
   - Fourier-based proof works only for uniform message
   - $k \approx 2 \log n + 3 \log (1/\varepsilon)$ (here $\lambda = 0$)
2. $E(m,s;\, i) = (\phi_i\,,\ \phi_i(m) + s)$
   - $\{\phi_i\colon \{0,1\}^n \rightarrow \{0,1\}^n\}$ are 3-wise independent permutations
   - $k \approx \lambda + 3 \log (1/\varepsilon)$ (works for all λ)
   - $3n$ bits of additional randomness, difficult proof

[RW02]: Second construction

Same scheme, new analysis:

- In particular, $\mathcal{H}=\{\phi_i\}$ is 2-wise independent permutation family
- LOHL' $\Rightarrow$ scheme secure for $\boldsymbol{k \approx \lambda + 2 \log(1/\varepsilon)}$
- Simpler schemes are possible…

2. $E(m,s;\, i) = (\phi_i\,,\, \phi_i(m) + s)$
   - $\{\phi_i\colon \{0,1\}^n \to \{0,1\}^n\}$ are 3-wise independent permutations
   - $k \approx \lambda + 3 \log(1/\varepsilon)$ (works for all λ)
   - $3n$ bits of additional randomness, difficult proof

Further simplification

- "Full" 2-wise independence unnecessary for LOHL'
- Sufficient: $\forall\, x \neq x'$: $H(x) \oplus H(x') \equiv U_n$
- Construction: $\mathcal{H} = \{x \rightarrow ax \mid a \in \mathrm{GF}(2^n)\}$
- The result: $\boldsymbol{E(m,s;a) = (\mathbf{a}\ ,\ m \oplus as)}$
  - Secure for $k \geq \lambda + 2\log(1/\varepsilon)$
  - Uses only n additional bits of randomness

Outline

- ~~Equiv. Def: Indistinguishability for high-entropy sources~~

  ~~**Intuition**: Indistinguishable schemes ≈ extractors~~

- ~~Two Simple, General Constructions:~~
  - ~~Step in an expander graph~~
  - ~~Random Hash Functions~~
- Lower bounds: $k \geq \lambda$, (special case: $k \geq \lambda + \log(1/\varepsilon)$)
- "Stronger" Equiv. Def.: all functions hard to predict (not only predicates)

Lower Bounds

- Lower Bound via Shannon Bound:

$$\boldsymbol{k \geq \lambda}$$

- Lower bound via lower bounds on extractors:

$$\boldsymbol{k \geq \lambda + \log(1/\varepsilon)}$$

  – Requires that extra randomness be public, i.e.

$$E(m,s;i) = (i\ ,\ E'(m,s;i)\ )$$

  – All the schemes discussed fit this framework

Lower Bounds

- Lower Bound via Shannon Bound:

$$\boldsymbol{k} \geq \lambda$$

- Lower bound via lower bounds on extractors:

$$\boldsymbol{k} \geq \lambda + \mathbf{log(1/\varepsilon)}$$

  – Requires that extra randomness be public, i.e.

$$E(m,s;i) = (i\ ,\ E'(m,s;i)\ )$$

  – All the schemes discussed fit this framework

Simple Lower Bound

Def: (λ,ε)-entropically secure if $\forall\, M$, $H_\infty(M) \geq n\text{-}\lambda$, $\forall\, A$ $\forall$ pred. g
$\exists\, A'$: $\left|\ \Pr[A(E(M)) = g(M)]\ -\ \Pr[A' = g(M)]\ \right| \leq \varepsilon$

Proof (reduce to bounds on regular encryption):

- $\forall\, w \in \{0,1\}^{\lambda}$, define distribution $M_w = w \,\|\, U_{n\text{-}\lambda}$
  (i.e.: $M_w = w$ followed by $n\text{-}\lambda$ random bits)
- Indistinguishability $\Rightarrow$ $\forall\, v, w$: $E(M_v) \approx_{\varepsilon} E(M_w)$
- This is regular encryption (non-entropic) of w **!**
- Need $k \geq \lambda$

Lower Bounds

- Lower bound via Shannon Bound:

$$\boldsymbol{k} \geq \lambda$$

- Lower bound via lower bounds on extractors:

$$\boldsymbol{k} \geq \lambda + \mathbf{log(1/\varepsilon)}$$

  – Requires that extra randomness be public

- These bounds are quite crude
- Probable (?) answer: $k \geq \lambda + 2\log(1/\varepsilon)$

Outline

- ~~Equiv. Def: Indistinguishability for high-entropy sources~~

  ~~**Intuition**: Indistinguishable schemes ≈ extractors~~

- ~~Two Simple, General Constructions:~~
  - ~~Step in an expander graph~~
  - ~~Hash functions~~

- ~~Lower bounds: $k \geq \lambda$, (special case: $k \geq \lambda + \log(1/\varepsilon)$)~~

- "Stronger" Equiv. Def.: all functions hard to predict (not just predicates)

Indistinguishability for High Entropy

Def: (λ,ε)-entropically secure if $\forall M$, $H_\infty(M) \geq n\text{-}\lambda$, $\forall A$ $\forall$ pred. g

$\exists A'$: $|\Pr[A(E(M)) = g(M)] - \Pr[A' = g(M)]$ | ≤

Q: Can we replace "for all predicates" with "for all functions"?

A: Yes. Resulting definition is even closer to semantic security.

Recall: (Ord

$\forall$ distribu

Definition f

$\forall$ distributions M,M' with $H_\infty(M)$, $H_\infty(M') \geq t$:

$$SD(E(M),E(M')) \leq \varepsilon$$

Proposition: **(λ,ε)-ES** equiv. to **(t , ε')-IND** for $t = n\text{-}\lambda\text{-}1$

Equivalence of Functions and Predicates

For function f, random variable $\mathbf{M}$:

$$\mathbf{pred}_f(\mathbf{M}) = \text{ most likely value } = \max_z\{ \Pr[f(\mathbf{M}) = z] \}$$

Main Lemma: Suppose

- $\mathbf{M}$ r.v. with $H_\infty(\mathbf{M}) \geq 2\log(1/\varepsilon)$
- $E()$, $A()$ randomized maps, f arbitrary function.
- $\Pr[\, A(E(\mathbf{M})) = f(\mathbf{M})\,] \;\; \geq \mathbf{pred}_f(\mathbf{M}) + \varepsilon$

Then there exist predicates B and g such that

$$\Pr[\, B(A(E(\mathbf{M}))) = g(\mathbf{M})\,] \;\; \geq \mathbf{pred}_g(\mathbf{M}) + \varepsilon / 4$$

Conclusions

- Systematic study of [RW02] notion of entropic security
  - equivalent definition
  - simple constructions, proofs, lower bounds
- "Computational issues":
  - Can these proofs preserve running time of adversaries?
  - Use computational min-entropy? (recently provided by [BSW])
- In what other contexts is ES interesting?
  - Password Hashing [CMR98]: similar definition
  - "Fuzzy fingerprints" [DRS03]