5 - EC and PA for QKD

25 października 2010
12:37

We can look at Csiszar-Körner as error correction + priv. ampl.

Error correction and privacy amplification for QKD

The difference from Csiszar-Körner is that we perform EC and PA after the communication (quantum stage).

A, B, E have sequences of n bits. (x^n, y^n, z^n)

In order to apply Csiszar-Körner theorem we need to assume that $p(x^n, y^n, z^n) = p(x,y,z)^n$ "each channel use was independent"

We assume that there is additionally an ideal public channel on which A and B can communicate, but E also hears everything.

Strictly speaking Csiszar-Körner assumes one-way communication from A to B. There is nothing that prevents us from reversing the situation and allow B to communicate to A. This could sometimes help increase the key rate

$$C_s = \max\left[I(X:Y) - I(X:Z),\ I(X:Y) - I(Y:Z)\right]$$

5.1 Error correction (see homework series 2)

mutual information of A and B is $I(A:B) = I(X^n; Y^n) = n I(X:Y)$

A and B need to exchange additionally at least

$m = n\left(H(X) - I(X:Y)\right) = n \cdot H(X|Y)$ bits to correct errors. In practice:

- Interactive error-correction protocol (1992, Bennett et al.)

Iteration:

~ A and B apply random permutation

- A and B divide their N bits in blocks of length n

$$N = k \cdot n$$

(the length of the block should be such that it is not very probable that there are more than 1 error)

— They check parities of bits in each subblock if it does not agree → bisection

— If all parity errors were corrected they repeat the iteration with larger block size.

- Maximal sensible size of the block is $k = \frac{N}{2}$ If a number of subsequent checks (e.g. 20) yield no errors it is almost certain $(1 - 2^{-20})$ that no errors are left.

🏠 - write program which would allow to compare efficiency of this protocol with Shannon limit.

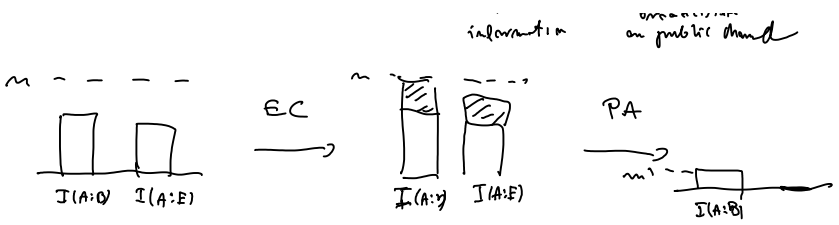
- Notice that after this procedure E has information which cannot be regarded as independent on each bit - there are some correlations from the revealed parities

5.2 Privacy amplification (first approx)

A and B have identical n-bit sequences; what is Eve's info?

$$I(A:E) \leq \underbrace{n I(X:Z)}_{\text{initial}} + m$$

└ information [illegible]



Intuitively we can shrink the sequence of A and B n bits to $n' = n - (I(A:E)) \approx n - (nI(X:Z) + m)$ so that E errors „propagate" in such a way that she knows nothing
If EC was performed at Shannon limit: $n' = n(I(X:Y) - I(X:Z))$

- a simple example that does the job
  Choose a family of $m' \times n$ random bit matrices.
  A and B choose such a matrix at random and apply

$$m' \left\{ \left[\quad \right] \right. = \left[M_{m' \times n} \right] \left\{ \left[\quad \right] \right. n$$

(example of a hashing function)

5.3 PA (privacy theory) { Maybe we return to it later

Problem: Shannon mutual information (Shannon entropy) is not really a proper quantity to describe privacy in general.

Example:
n bit random sequence X^n, let $H(X^n|E) = \frac{n}{4}$
{ This means $I(X^n:E) = n - \frac{n}{4} = \frac{3}{4}n$

How much we need to shorten the key to be sure E has no information? It seems that to $\frac{n}{4}$. This might be wrong however. Consider two cases:

a) $H(X^n|E) = \frac{n}{4}$ $\quad E = 2^n$ $\quad p(x^n, z^n) = p(x,z)^n$

$$p(x,z) = \begin{array}{c|cc} x \backslash z & 0 & 1 \\ \hline 0 & \frac{1}{2}-\varepsilon & \varepsilon \\ 1 & \varepsilon & \frac{1}{2}-\varepsilon \end{array}$$

$$(\tfrac{1}{2}-\varepsilon)\log(\tfrac{1}{2}-\varepsilon) + \varepsilon \log \varepsilon = \frac{1}{4}$$

$$\Downarrow$$

$$\varepsilon = 0{,}042 = 4{,}2\%$$

In this case reasoning is correct provided n is large after privacy amplification $I(X^m:E) = 0$.

b) $H(X^n|E) = \frac{n}{4}$
this entropy may result from a situation that E with probability $\frac{3}{4}$ knows exactly X^n and with probability $\frac{1}{4}$ knows nothing. But then no matter what PA protocol you apply always
$I(X^m:E) = \frac{3}{4} \cdot m$ It will never go to zero

If X^n are not independent you should be careful!
For asymptotic considerations $n \to \infty$, and independent variables we can use simplified approach, but in general we cannot.

5.3.1 Hashing functions

- Universal class of hash functions $h: A \to B$

• Universal class of hash functions $h: \mathcal{A} \to \mathcal{B}$

Def: $\mathcal{H}$ – universal class of hash functions

$\forall_{x_1 \neq x_2}$ size of the set $\{h \in \mathcal{H} : h(x_1) = h(x_2)\}$ is at most $\frac{|\mathcal{H}|}{|\mathcal{B}|}$.

{ intuition: different inputs yield different outputs

Przykład: $\mathcal{A} = \{0,1\}^n \quad \mathcal{B} = \{0,1\}^r$

Idea: $\mathcal{H}$ można wziąć wszystkie funkcje liniowe z $\mathcal{A} \to \mathcal{B}$

(wszystkie macierze binarne $n \times r$) – (macierze prostokątne ale dzikie)

How to prove this?

Let $a \neq a'$ $\quad Ma = b$

How many M there are such that $Ma' = b$

$M(a - a') = 0$

$a - a'$ – niezerowy wektor $\quad M = r\left[\quad\overset{n}{\quad}\quad\right]$

$\Downarrow$

r constraints

So there are at least 2^{nr-r} such matrices $= \frac{2^{nr}}{2^r}$ ok.

– dla $r = n$ tryw. chyba przykład jest identyczność

Przykład: Random binary Toeplitz matrices

T – Toeplitz matrix $T_{i+a, j+a} = T_{ij}$

(it is enough to define first row and first column)

If we put randomly 0 and 1 we get a universal class of hash functions ⌂

Advantage we need only $(n+m-1)$ instead of $n \cdot m$ bits to determine the matrix.

5.3.2 Renyi entropy

$$H_s(X) = \frac{1}{1-s} \log\left(\sum_x p(x)^s\right)$$

$$\lim_{s \to 1} H_s(X) = \frac{\frac{d}{ds} \log \sum_x p(x)^s}{\frac{d}{ds}(1-s)}\Bigg|_{s=1} = \frac{\sum_x \log p(x) \cdot p(x)^s}{\sum_x p(x)^s}\Bigg|_{s=1} = H(X)$$

Th: $H_s(X)$ – monotonically decreasing function of s

$$\frac{d}{ds} H_s(X) \leq 0$$

Proof ⌂

Fact: All Renyi entropies are equal for uniform distributions

X, $p(x) = \frac{1}{|X|}$ $\qquad H_s(X) = \log |X|$

Renyi: $s=2$ entropy

$$H_2(X) = -\log\left(\underbrace{\sum_x p(x)^2}_{\text{collision probability}}\right) = -\log\left(\sum_{x_1, x_2} p(x_1 = x_2)\right)$$

(prob. that in two

realizations you get the same outcome)

Calculate Renyi for example in 5.3

5.3.3

Theorem: Let P_{X2} probability distribution

If $H_2(X|Z=z) \geq c$ and $K = u(X)$ (hashing function treated as random variable) $u \in \mathcal{U}$ universal class of hash functions (k bit string variable)

then: $H(K|\mathcal{U}Z=z) \geq k - 2^{k-c}/\ln 2$

this means that we can distill $k = c - s$ bits (where s – security parameter) $H(K|u,z=z) \geq k - \frac{2^{-s}}{\ln 2}$

Proof:

$$\forall_{x_1 \neq x_2} \left|\{u : u(x_1) = u(x_2)\}\right| \leq \frac{|\mathcal{U}|}{2^k}$$

$$H(K|\mathcal{U}Z=z) = \frac{1}{|\mathcal{U}|}\sum_u H(K|\mathcal{U}=u\,Z=z) \geq \frac{1}{|\mathcal{U}|}\sum_u H_2(K|\mathcal{U}=u\,Z=z) =$$

$$= -\frac{1}{|\mathcal{U}|}\sum_u \log \sum_k p(k|u,z)^2$$

$$> -\log \sum_u \frac{1}{|\mathcal{U}|}\sum_k p(k|u,z)^2 = -\sum_u \log \frac{1}{|\mathcal{U}|} P(K_1 = K_2 | u,z) =$$

(can be dropped) (can be dropped)

$$= -\log \sum_u \frac{1}{|\mathcal{U}|}\left(P(K_1 = x_2 | u,z) + P(K_1 \neq x_2 | u,z) \cdot p(K_1 = k_2 | x_1 \neq x_2, u, z)\right) \geq$$

$$= -\log\left(2^{-H_2(X|Z=z)} + \left(1 - 2^{-H_2(X|Z=z)}\right)\cdot \frac{1}{2^k}\right)$$

$$\geq -\log\left(2^{-H_2(X|Z=z)} + 2^{-k}\right) = k - \log\left(1 + 2^{-H_2(X|Z)+k}\right)$$

$$\geq k - \frac{2^{-H_2(X|Z=z)+k}}{\ln 2}$$

Observation

For many independent realizations $p(x^m, z^m) = p(x,z)^m$, $m \to \infty$ provided the sequence is typical, we have uniform probability distribution which means that:

for $(x^m, z^m) \in T_\varepsilon^m$ $\quad H_2(x^m, z^m) = H(x^m, z^m) = m\,H(x,z)$

at first it might look strange, since H_2 is also additive $H_2(x^m, z^m) = m\,H_2(x,z) \neq m\,H(x,z)$, but here we have constraint that we look only on typical sequences

This is a peculiar property of H_2 that sometimes $H_2(X|Z) > H_2(X)$ (spoiling the knowledge by additional information)

The power of this theorem: unlike Csiszar-Körner it does not deal with votes, but total number of bits, hence can be applied also to finite strings.

5.3.4

The above theorem is not enough we also need to know how much $H_2(X|Z=z)$ decreases when additional bits of information are revealed (as in public error-correction)

Try to bound the difference between Shannon entropy and 2-Renyi entropies for ε-typical sequences of length n

Theorem: Eller et al. side information U.

With probability $1 - 2^{-s}$ [illegible] value u

With probability $1-2^{-s}$ U takes the value u for which:

$$H_2(X) - H_2(X|U=u) \leq 2\log|\mathcal{U}| + 2s$$

Jest lepsze ograniczenie patrz. Cachin Phd thesis: $\log|\mathcal{U}| + 2s + 2$)

* One can circumvent this problem by encrypting error correction procedure – this does not hurt in the secret key growing approach ▽

Proof:

Let $\mathcal{U}_+ = \{u : p(u) \geq p_{min}\}$ $\quad p_{min} = \frac{2^{-s}}{|\mathcal{U}|}$

$$\overline{\mathcal{U}_+} = \{u : p(u) < p_{min}\}$$

ponieważ $p_{min} \leq \frac{1}{|\mathcal{U}|}$ to $\overline{\mathcal{U}_+}$ może mieć najwyżej $|\mathcal{U}|-1$ elementów

$p(\overline{\mathcal{U}_+}) \leq \frac{2^{-s}}{|\mathcal{U}|} \cdot |\mathcal{U}|-1 \approx 2^{-s}$ $\quad p(\mathcal{U}_+) \geq 1-2^{-s}$

2 prawdop. $1-2^{-s}$ mogą więc:

$$\sum_x p(x|u)^2 = \sum_x \frac{p(x,u)^2}{p(u)^2} \leq \left(\sum_x p(x,u)^2\right) 2^{2s}|\mathcal{U}|^2 \leq \sum_x p(x)^2 2^{2s}|\mathcal{U}|^2$$

$$H_2(X|U=u) \geq H_2(X) - 2s - 2\log|\mathcal{U}|$$ ▨

Example: (showing that after conditioning Renyi entropy may increase – in analogy to example with conditioning on being typical)

X_1^m 2 - cinq random vars.

$p(x_1^m, 0) = \varepsilon\, \delta_{x,0^m}$ $\quad p(x,1) = \frac{1}{2^m}(1-\varepsilon)$

$p_{X^m}(0^m) = \varepsilon + \frac{1}{2}(1-\varepsilon)$ $\quad p_{X^m}(x^m) = \frac{1}{2^m}(1-\varepsilon)$ $\Big\}$ for $x^m \neq 0^m$

$$H_2(X) = -\log\left(\left(\varepsilon + \frac{1}{2^m}(1-\varepsilon)\right)^2 + \left(2^m-1\right)\cdot\frac{1}{2^{2m}}(1-\varepsilon)^2\right) =$$

$$= -\log\left(\varepsilon^2 + \frac{2\varepsilon(1-\varepsilon)}{2^m} + \frac{1}{2^m}(1-\varepsilon)^2\right)$$

$$H_2(X|Z) = \sum_z p(z)\, H_2(X|Z=z) =$$

$$= \varepsilon \cdot 0 + (1-\varepsilon)\cdot m = (1-\varepsilon)m$$

$$H_2(X) = -\log\varepsilon^2 - \log\left[1 + \frac{1}{2^m}(1-\varepsilon^2)\right]$$

If we fix ε and take $m \to \infty$ only first term survives $\quad H_2(X) = -\log\varepsilon^2 \leq H(X|Z)$

$-\log\varepsilon^2 \leq (1-\varepsilon)m$ for m large enough

For comparison calculate Shannon entropy

$$H(X|Z) = H_2(X|Z) = (1-\varepsilon)m$$

$$H(X) = -\left(\varepsilon + \frac{1}{2^m}(1-\varepsilon)\right)\log\left(\varepsilon + \frac{1}{2^m}(1-\varepsilon)\right)$$
$$- \left(2^m-1\right)\frac{1}{2^m}(1-\varepsilon)\log\left(\frac{1}{2^m}(1-\varepsilon)\right)$$

when $m \to \infty$.

$$H(X) \to -\varepsilon\log\varepsilon - (1-\varepsilon)\log(1-\varepsilon) + m(1-\varepsilon) \geq H(X|Z)$$ ok.

$+ \; m(1-\varepsilon) \geq H(X|Z)$ ok.