The BB84 cryptologic protocol

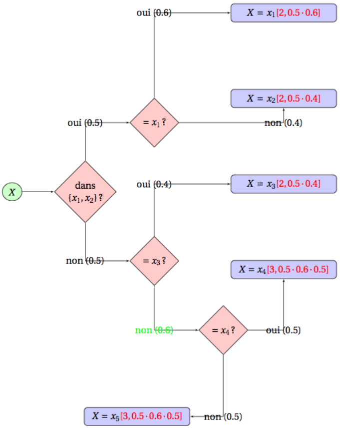
Security analysis against individual attacks

Dimitri Petritis

Institut de recherche mathématique de Rennes
Université de Rennes 1 et CNRS (UMR 6625)

Santiago, November 2013

Determing the outcome of a r.v.



$$
\begin{aligned}
X &\in \{x_1, \ldots, x_5\} \\
\mathbf{p} &= (0.3, 0.2, 0.2, 0.15, 0.15) \\
\mathbb{E}N &= 2 \cdot [0.3 + 0.2 + 0.2] + 3 \cdot [0.15 + 0.15] \\
&= 2.3.
\end{aligned}
$$

UNIVERSITÉ DE RENNES 1

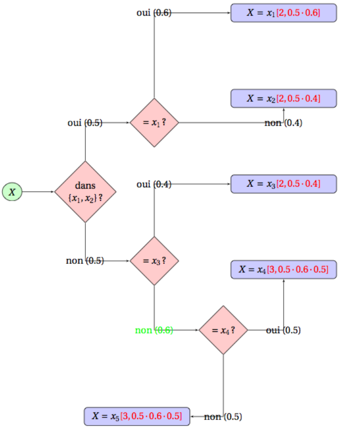
Entropy and information

- **Information:** probabilistic quantity = reduction of uncertainty when the outcome has been revealed.
- X a $\mathbb{A}$-valued r.v. with law $\mathbf{p}$ conveys information

$$H(X) = H(\mathbf{p}) := -\sum_{a \in \mathbb{A}} p(a) \log p(a) = -\mathbb{E}(\log p(X)).$$

- **Information [Shannon (1948)] = entropy [Boltzmann (1877)]**.
- First significance: **entropy is an expectation** (that makes us ageing . . .).

A second significance of entropy



$$
\begin{aligned}
X &\in \{x_1, \ldots, x_5\} \\
\mathbf{p} &= (0.3, 0.2, 0.2, 0.15, 0.15) \\
\mathbb{E}N &= 2 \cdot [0.3 + 0.2 + 0.2] + 3 \cdot [0.15 + 0.15] \\
&= 2.3 \\
H(\mathbf{p}) &= -0.3 \log 0.3 - 0.4 \log 0.2 - 0.3 \log 0.15 \\
&= 2.27
\end{aligned}
$$

Theorem

$$\mathbb{E}N \geq H(\mathbf{p}).$$

A third significance of the entropy

Definition

Let $n \geq 1$, $\mathbb{A}$ finite alphabet, $\mathbf{p} \in \mathrm{PV}_{\mathrm{card}\mathbb{A}}$, and integer $K > 0$. Sequence $\boldsymbol{\alpha} \in \mathbb{A}^n$ is **typical** ($(n, \mathbf{p}, K)$-typical) if

$$\forall a \in \mathbb{A}, \left| \frac{\nu_a(\boldsymbol{\alpha}) - np_a}{\sqrt{np_a(1-p_a)}} \right| < K.$$

Theorem

Let $\epsilon \in]0,1[$ *and* $K > \sqrt{\mathrm{card}\mathbb{A}/\epsilon}$*. For* $n \geq K$*,*

1. $\mathbb{P}(\mathbf{X}\restriction_n \notin \mathbb{T}_{n,\mathbf{p},K}) < \epsilon$*;*
2. $\mathrm{card}(\mathbb{T}_{n,\mathbf{p},K}) = 2^{n(H(\mathbf{p})+\delta_n)}$*, with* $\lim_{n\to\infty} \delta_n = 0$*;*
3. $\exists c > 0$ *s.t.* $\forall \boldsymbol{\alpha} \in \mathbb{T}_{n,\mathbf{p},K}$*,*

$$2^{-nH(\mathbf{p})-c\sqrt{n}} \leq \mathbb{P}(\mathbf{X}\restriction_n = \boldsymbol{\alpha}) \leq 2^{-nH(\mathbf{p})+c\sqrt{n}}.$$

Joint and conditional entropy

- X, Y with joint probability $\mathbf{p}$: $H(X,Y) = -\mathbb{E}(\log(p(X,Y)))$.
- **Conditional vs joint entropy:**

$$\begin{aligned} H(X|Y) &:= -\sum_{x,y\in\mathbb{A}} \mathbb{P}(X=x, Y=y)\log\mathbb{P}(X=-x|Y=y) \\ H(X,Y) &= H(Y)+H(X|Y) = H(X)+H(Y|X). \end{aligned}$$

Relative entropy and mutual information

- **Relative entropy or Kullback-Leibler contrast:** $D(\mathbf{p}\|\mathbf{q}) = \mathbb{E}_{\mathbf{p}} \log(\frac{p(X)}{q(X)}) \geq 0$. ($D$ is not a distance).
- **Mutual information:**

$$\begin{aligned} I(X:Y) &= D(\mathbb{P}_{(X,Y)}\|\mathbb{P}_X \otimes \mathbb{P}_Y) = I(Y:X) \\ &= H(X) + H(Y) - H(X,Y). \end{aligned}$$

Quantum extension

- $\rho \in \mathcal{D}(\mathbb{H})$: **entropy** $S(\rho) = -\operatorname{tr}(\rho \log \rho)$;
- $\rho_{12} \in \mathcal{D}(\mathbb{H}_1 \otimes \mathbb{H}_2)$: **joint entropy** $S(\rho_{12}) = -\operatorname{tr}(\rho_{12} \log \rho_{12})$; **conditional entropy** $S(\rho_1|\rho_2) = S(\rho_{12}) - S(\rho_2)$, where $\rho_1 = \operatorname{tr}_2 \rho_{12}$ and $\rho_2 = \operatorname{tr}_1 \rho_{12}$;
- **relative entropy** $(\operatorname{supp} X = (\ker X)^\perp)$

$$D(\rho\|\sigma) = \begin{cases} \operatorname{tr}(\rho(\log \rho - \log \sigma)) & \text{if } \operatorname{supp} \rho \subset \operatorname{supp} \sigma \\ +\infty & \text{otherwise;} \end{cases}$$

- **mutual information** $I(\rho_1 : \rho_2) = S(\rho_1) + S(\rho_2) - S(\rho_{12})$.

Theorem (Csiszár-Körner (1978) - classical and quantum)

Suppose 3 parties A, B, E posses rv having joint probability $\mathbb{P}_{ABE}$. The minimal secret key rate, parties A and B can share in presence of malevolent third party E, is given by

$$L(A, B\|E) = \max(I(A:B) - I(A:E), I(B:A) - I(B:E)).$$

The principles

Recall the main idea of BB84.

- When Bernardo's basis is the same as the one used by Alicia, their bits are perfectly correlated.
- Safety of protocol relies
  - on this perfect correlation and the fact that
  - any eavesdropping perturbs some qubits, reducing thus the correlation of bits (introducing disturbance).
- Alicia and Bernardo measure the correlation of their bits by publicly comparing subsamples of their data.
- **Question:** given a measured correlation (or equivalently a measured average disturbance), how much information Encarnación could have gained?

Intermezzo on partial traces

- Let $\mathbb{F}, \mathbb{G}$ Hilbert and consider $\mathbb{H} = \mathbb{F} \otimes \mathbb{G}$.
- f ray in $\mathbb{F}$, g ray in $\mathbb{G}$;
  $\rho_{fg} := |\, fg \,\rangle\langle\, fg \,| = |\, f \,\rangle\langle\, f \,| \otimes |\, g \,\rangle\langle\, g \,| = \rho_f \otimes \rho_g$.
- $\operatorname{tr}_{\mathbb{G}} \rho_{fg} = |\, f \,\rangle\langle\, f \,| = \rho_f \in \mathcal{D}(\mathbb{F})$;
  $\operatorname{tr}_{\mathbb{F}} \rho_{fg} = |\, g \,\rangle\langle\, g \,| = \rho_g \in \mathcal{D}(\mathbb{G})$.
- $\langle\, fg \,|\, (I_{\mathbb{F}} \otimes M) fg \,\rangle = \langle\, g \,|\, Mg \,\rangle = \operatorname{tr}(M\rho_g)$.
- Generally, for $\rho \in \mathcal{D}(\mathbb{H})$, $M \in \mathcal{B}(\mathbb{F})$, and $N \in \mathcal{B}(\mathbb{G})$,
  $\operatorname{tr}((M \otimes I_{\mathbb{G}})\rho) = \operatorname{tr}(M \operatorname{tr}_{\mathbb{F}} \rho)$ and $\operatorname{tr}((I_{\mathbb{F}} \otimes N)\rho) = \operatorname{tr}(N \operatorname{tr}_{\mathbb{G}} \rho)$.

Individual attack
Notation

- With states of every party are associated different Hilbert spaces $\mathbb{H}_A, \mathbb{H}_B$, and $\mathbb{H}_E$.
- $t \in \{0,1\}$, $\bar{t} = t - 1 \mod 2$ is the conjugate bit of t.
- $\sharp \in \{+, \times\}$, $\flat$ is the conjugate of $\sharp$, i.e. if $\sharp = +$ then $\flat = \times$ and vice versa.
- $(B^\sharp_\beta)_{\beta \in \{0,1\}}$ is the sharp resolution of the identity in $\mathbb{H}_B$ into projectors $B^\sharp_0 = |\,\epsilon^\sharp_0\,\rangle\langle\,\epsilon^\sharp_0\,|$, $B^\sharp_1 = |\,\epsilon^\sharp_1\,\rangle\langle\,\epsilon^\sharp_1\,|$, $\sum_{\beta\in\{0,1\}} B^\sharp_\beta = I_{\mathbb{H}_B}$.
- $(E_\gamma)_{\gamma\in\Gamma}$ is an unsharp resolution of $I_{\mathbb{H}_E}$ into operators $E_\gamma \geq 0$, i.e. $\sum_{\gamma\in\Gamma} E_\gamma = I_{\mathbb{H}_E}$.
- Alicia sends a qubit $\psi \in \{\epsilon^+_0, \epsilon^+_1, \epsilon^\times_0, \epsilon^\times_1\}$. Elements of this set can be decomposed into

$$|\,\epsilon^\sharp_t\,\rangle = \frac{|\,\epsilon^\flat_0\,\rangle + (-)^t|\,\epsilon^\flat_1\,\rangle}{\sqrt{2}}, t \in \{0,1\}, \sharp \in \{+,\times\}, \flat \text{ conjugate of } \sharp.$$

Individual attack

Possible actions of Encarnación

- Vector $\psi = \epsilon_t^\sharp$ produced as a pure state of $\mathbb{H}_A$; only Alicia has access on it at initial time. Once sent over the quantum channel; Alicia has no access on it any longer. When its (legal or illegal) recipient gets it, can act on it. E.g., if Bernardo receives it, he can act on it by operators of his own space $\mathbb{H}_B$, although we still write $\psi \in \mathbb{H}_A$.
- Encarnación cannot copy $\psi \in \mathbb{H}_A$ but can
  - couple $\epsilon_t^\sharp \in \mathbb{H}_A$ with a state $\phi \in \mathbb{H}_E$ of her own to produce $\Phi_t^\sharp = \epsilon_t^\sharp \otimes \phi \in \mathbb{H}_A \otimes \mathbb{H}_E$,
  - perform partial unsharp measurements $I_{\mathbb{H}_A} \otimes E_\gamma$ on $\Phi_t^\sharp$ and send first part to Bernardo.
  - Unsharp measurements can be thought as sharp measurements on some bigger Hilbert space.

Individual attack
Qualitative behaviour

- Since unitary evolution preserves pure states, suppose first that

$$
\begin{aligned}
U|\,\epsilon_t^{\sharp}\phi\,\rangle &= |\,\zeta_t^{\sharp}\phi_t^{\sharp}\,\rangle,\\
U|\,\epsilon_t^{\flat}\phi\,\rangle &= |\,\zeta_t^{\flat}\phi_t^{\flat}\,\rangle,
\end{aligned}
$$

  i.e. the transformed states remain a tensor product state. Then

$$
\frac{1}{2} = \langle\,\epsilon_t^{\flat}\,|\,\epsilon_t^{\sharp}\,\rangle = \langle\,\zeta_t^{\sharp}\,|\,\zeta_t^{\flat}\,\rangle\langle\,\phi_t^{\sharp}\,|\,\phi_t^{\flat}\,\rangle.
$$

- If $\langle\,\epsilon_t^{\flat}\,|\,\epsilon_t^{\sharp}\,\rangle = \langle\,\zeta_t^{\sharp}\,|\,\zeta_t^{\flat}\,\rangle$, i.e. the Alicia's (Bernardo's) part of the state is not altered, then $\langle\,\phi_t^{\sharp}\,|\,\phi_t^{\flat}\,\rangle = 1$ hence, states $\phi_t^{\sharp}$ and $\phi_t^{\flat}$ cannot be discriminated.
- To well discriminate these states, $|\langle\,\phi_t^{\sharp}\,|\,\phi_t^{\flat}\,\rangle|$ must be minimised, hence $|\langle\,\zeta_t^{\sharp}\,|\,\zeta_t^{\flat}\,\rangle|$ maximised, i.e. maximally disturbed.
- Idea survives even when U does not preserve tensor products.

Partial measurement

At Encarnación's side

$$
\begin{aligned}
Q^{\sharp}_{t\gamma} &= \langle \Phi^{\sharp}_t \,|\, (I_A \otimes E_\gamma)\Phi^{\sharp}_t \rangle = \langle \epsilon^{\sharp}_t \phi \,|\, (I_A \otimes E_\gamma)\epsilon^{\sharp}_t \phi \rangle \\
&= \mathbb{P}(E \text{ unsharply observes } \gamma |\ A \text{ sent } t), \\
p_t &= \mathbb{P}(A \text{ sends } t), \\
q_\gamma &= \mathbb{P}(E \text{ observes } \gamma) = \sum_{t\in\{0,1\}} p_t Q_{t\gamma}, \\
\hat{Q}_{\gamma t} &= \frac{p_t Q_{t\gamma}}{q_\gamma} = \mathbb{P}(E \text{ assigns to } t | E \text{ has observed } \gamma), \\
G_\gamma &= |\hat{Q}_{\gamma t} - \hat{Q}_{\gamma \bar{t}}| = \text{ E's gain of information}, \\
\mathbb{E}G &= \sum_{\gamma\in\Gamma} q_\gamma |\hat{Q}_{\gamma t} - \hat{Q}_{\gamma \bar{t}}|
\end{aligned}
$$

Problem reduces to estimating $q_\gamma G_\gamma = q_\gamma |\hat{Q}_{\gamma t} - \hat{Q}_{\gamma \bar{t}}|$.

Estimate of $q_\gamma G_\gamma$

Lemma

$$\begin{aligned} q_\gamma G_\gamma &= q_\gamma |\hat{Q}_{\gamma t} - \hat{Q}_{\gamma \bar{t}}| \\ &\leq \|Z_{00}^{\flat\gamma}\| \|Z_{10}^{\flat\gamma}\| + \|Z_{01}^{\flat\gamma}\| \|Z_{11}^{\flat\gamma}\|, \end{aligned}$$

where $Z_{st}^{\flat\gamma} = B_s^\flat \otimes \sqrt{E_\gamma} \Phi_s^\flat$, $s, t \in \{0, 1\}$.

Proof.

Blackboard 1: Estimate of $q_\gamma G_\gamma$ □

$$\begin{aligned} \|Z_{st}^{\flat\gamma}\|^2 &= \langle\, \Phi_s^\flat \mid B_t^\flat \otimes E_\gamma \Phi_s^\flat \,\rangle \\ &= \mathbb{P}(B \text{ measures } t, E \text{ measures } \gamma | A \text{ sends } s) \\ &= \mathbb{P}(B \text{ measures } t | E \text{ measures } \gamma, A \text{ sends } s) Q_{s\gamma}. \end{aligned}$$

UNIVERSITÉ DE RENNES 1

Distortion on conjugate basis

- $\mathbb{P}(Bs|E\gamma, As) = 1 - \mathbb{P}(B\bar{s}|E\gamma, As) = 1 - D^\flat_{s\gamma}$.
- $D^\flat_{s\gamma} = \mathbb{P}(B \text{ faults } | E \text{ measures } \gamma, A \text{ sends } s)$.

Lemma

- $q_\gamma G\gamma \leq \sqrt{Q^\flat_{0\gamma} Q^\flat_{1\gamma}} \left(\sqrt{D^\flat_{0\gamma}(1 - D^\flat_{1\gamma})} + \sqrt{D^\flat_{1\gamma}(1 - D^\flat_{0\gamma})} \right)$.

Proof.

Blackboard 2: Proof of lemma. □

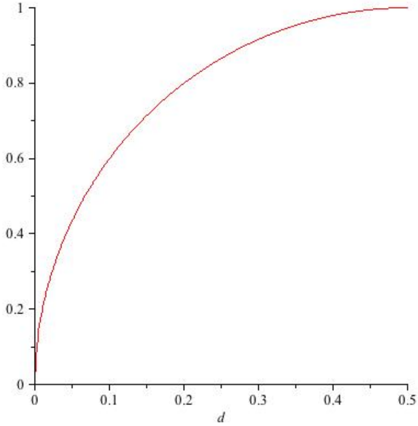
Theorem

If $D^\flat_{0\gamma} = D^\flat_{0\gamma} = d_\gamma$, *then* $\mathbb{E}G \leq 2\sqrt{\mathbb{E}d(1 - \mathbb{E}d)}$.

Proof.

Blackboard 3: Proof of theorem. □

Plot of the bound



Improvement of the bound

$$\begin{aligned}
\kappa_{t\gamma} &= p_t Q_{t\gamma} = q_\gamma \hat{Q}_{\gamma t} = \text{ joint probability on } \{0,1\} \times \Gamma. \\
H(\kappa) &= -\sum_{t,\gamma} \kappa_{t\gamma} \log \kappa_{t\gamma} \\
&= -\sum_{t,\gamma} q_\gamma \hat{Q}_{\gamma t} (\log q_\gamma + \log \hat{Q}_{\gamma t}) \\
&= H(q) - \sum_\gamma q_\gamma \sum_t \hat{Q}_{\gamma t} \log \hat{Q}_{\gamma t}.
\end{aligned}$$

Introducing $r_\gamma = \hat{Q}_{\gamma 1} - \hat{Q}_{\gamma 1} = \pm G_\gamma \in [-1, 1]$, we get

$$\hat{Q}_{\gamma 0} = \frac{1 + r_\gamma}{2}; \hat{Q}_{\gamma 1} = \frac{1 - r_\gamma}{2}.$$

Bound of relative information I

For a priori and a posteriori distributions

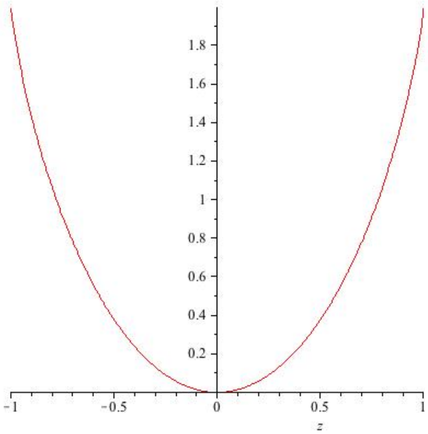
$$
\begin{aligned}
I(q:p) &= H(q) + H(p) - H(\kappa) \\
&= H(p) + H(q) - H(q) + \sum_{\gamma} q_{\gamma} \sum_{t} \hat{Q}_{\gamma t} \log \hat{Q}_{\gamma t} \\
&\overset{p=(\frac{1}{2},\frac{1}{2})}{=} \log 2 + \frac{1}{2} \sum_{\gamma} q_{\gamma} \left[(1+r_{\gamma}) \log \frac{1+r_{\gamma}}{2} + (1-r_{\gamma}) \log \frac{1-r_{\gamma}}{2} \right] \\
&= \frac{1}{2} \sum_{\gamma} q_{\gamma} g(r_{\gamma}),
\end{aligned}
$$

where $g(z) = (1+z)\log(1+z) + (1-z)\log(1-z)$. Observe that

- $g(-z) = g(z)$,
- $g'(z) = \log \frac{1+z}{1-z} > 0$ on $[0,1[$. Hence $g \uparrow$ on $[0,1[$.
- $I(q:p) = \frac{1}{2}\sum_{\gamma} q_{\gamma} g(G_g)$, because $r_{\gamma} = \pm G_{\gamma}$.

UNIVERSITÉ DE RENNES 1

Plot of the function g



Bound of relative information II

For a priori and a posteriori distributions

$$
\begin{aligned}
I(q:p) &= \frac{1}{2}\sum_{\gamma} q_\gamma g(G_g) \\
&\leq \frac{1}{2}\sum_{\gamma} q_\gamma g(2\sqrt{d_\gamma(1-d_\gamma)}); \\
\phi(t) &= g(2\sqrt{t(1-t)}) \text{ is concave on }]0,1[; \\
I(q:p) &\leq \frac{1}{2}\sum_{\gamma} q_\gamma \phi(d_\gamma) \\
&\leq \frac{1}{2}\phi(\mathbb{E}d).
\end{aligned}
$$

Information gain vs distortion

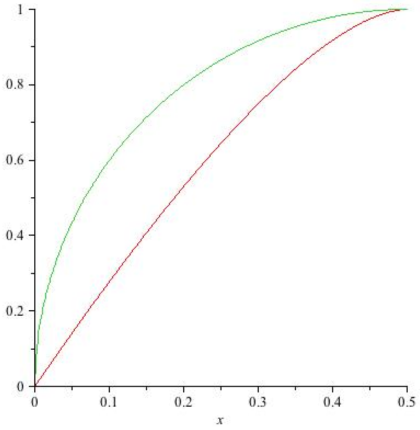


Figure: The horizontal axis represents $\mathbb{E}d$; the vertical axis for green curve represents $\mathbb{E}G$ and for the red curve $I(q:p)$.