Rate-Distortion Theory for Secrecy Systems

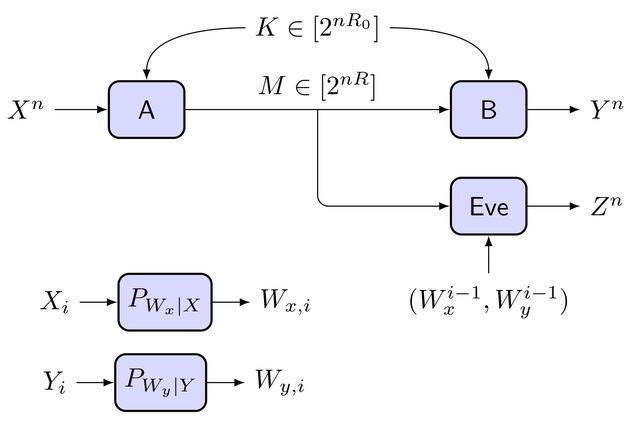
Curt Schieler and *Paul Cuff*
Department of Electrical Engineering, Princeton University

Question

Consider communication over a public noiseless channel. The transmitter and receiver share secret key, which they use to encrypt the communication. An eavesdropper observes the communication and attempts to reconstruct the source sequence.

What is the optimal tradeoff among communication rate, secret key rate, distortion at the eavesdropper, and distortion at the legitimate receiver?

Setup



Instead of two distortion functions $d_1(x,y)$ and $d_2(x,z)$, use a single payoff function $\pi(x,y,z)$. The expected payoff for a block is defined as

$$\min_{\left\{P_{Z_i|M,W_x^{i-1},W_y^{i-1}}\right\}_{i=1}^n} \mathbb{E}\frac{1}{n}\sum_{i=1}^n \pi(X_i,Y_i,Z_i)$$

Note that at each step i, the adversary is assumed to have access to a noisy version of the past behavior of the system, namely (W_x^{i-1}, W_y^{i-1}). This assumption is referred to as *causal disclosure* and plays a pivotal role.

Why causal disclosure?

Consider the following setting:

- Binary source sequence X^n.
- One bit of secret key $K \sim \text{Bern}(1/2)$.
- Encoder sends $M = Y^n$, where $Y_i = X_i \oplus K$.
- Adversary only views M (i.e., no assumption of causal disclosure)

In this scenario, any adversary incurs maximum distortion because Y_i is independent of X_i for all i. It appears as though we have achieved maximum secrecy for an n-bit source with only one bit of secret key! However, the adversary actually knows a great deal about X^n, namely that it is one of two candidate sequences. Furthermore, the adversary can determine X^n if he knows one true bit of the source sequence.

In general, if it is assumed that an adversary only observes the public message M, then an arbitrarily small rate of secret key is enough to guarantee maximum distortion. However, such secrecy is weak because the additional observation of just a few source symbols is enough for the adversary to completely identify the source sequence. A distortion-based approach to secrecy is strengthened considerably by an assumption of causal disclosure.

Main result

Theorem [1]
Fix P_X, $\pi(x,y,z)$, and causal disclosure channels $P_{W_x|X}$ and $P_{W_y|Y}$. The closure of achievable (R, R_0, Π) is equal to

$$\bigcup_{W_x - X - (U,V) - Y - W_y} \left\{ \begin{aligned} (R, R_0, \Pi):\ & R \geq I(X;U,V) \\ & R_0 \geq I(W_x W_y; V|U) \\ & \Pi \leq \min_{z(u)} \mathbb{E}\,\pi(X,Y,z(U)) \end{aligned} \right\}$$

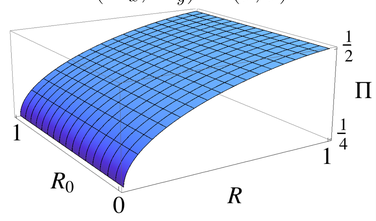
References

[1] C. Schieler and P. Cuff. "Rate-distortion theory for secrecy systems," submitted to IEEE Transactions on Information Theory, May 2013. Preprint available on arXiv.

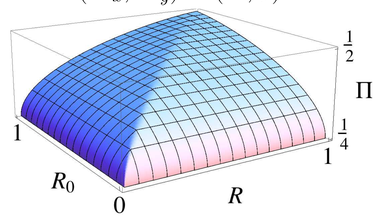
Example

Let $P_X \sim \text{Bern}(1/2)$ and $\pi(x,y,z) = \mathbf{1}\{x = y, x \neq z\}$. For this choice of payoff function, the block-average payoff is the fraction of symbols in a block that Nodes A and B are able to agree on and keep hidden from the adversary. There are four natural special cases of Theorem 1 that are obtained by setting W_x equal to $\emptyset$ or X and setting W_y equal to $\emptyset$ or Y.

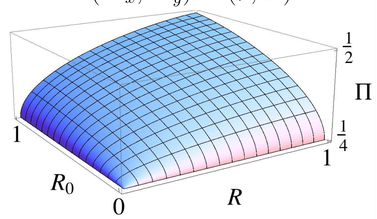
No causal disclosure: $(W_x, W_y) = (\emptyset, \emptyset)$



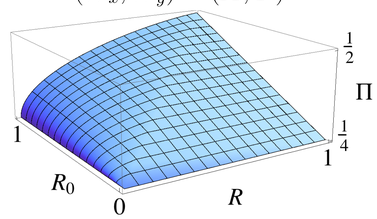
Causal disclosure of Node A: $(W_x, W_y) = (X, \emptyset)$



Causal Disclosure of Node B: $(W_x, W_y) = (\emptyset, Y)$



Causal disclosure of Nodes A and B: $(W_x, W_y) = (X, Y)$



Equivocation

Measures of secrecy based on (normalized) equivocation are a special case of the causal disclosure framework. For example, let $(W_x, W_y) = (X, \emptyset)$ and consider a payoff function $\pi : \mathcal{X} \times \mathcal{Y} \times \Delta_{\mathcal{X}} \to \mathbb{R}$ defined by

$$\pi(x,y,z) = \log \frac{1}{z(x)},$$

where z is a probability distribution on $\mathcal{X}$, and $z(x)$ denotes the probability of $x \in \mathcal{X}$ according to $z \in \Delta_{\mathcal{X}}$. With this choice of payoff function, the expected payoff is exactly the normalized equivocation $\frac{1}{n}H(X^n|M)$.

$$\begin{aligned} &\min_{\{P_{Z_i|M,X^{i-1}}\}_{i=1}^n} \mathbb{E}\frac{1}{n}\sum_{i=1}^n \pi(X_i,Y_i,Z_i) \\ &= \frac{1}{n}\sum_{i=1}^n \min_{P_{Z|M,X^{i-1}}} \mathbb{E}\,\log\frac{1}{Z(X_i)} \\ &= \frac{1}{n}\sum_{i=1}^n H(X_i|M,X^{i-1}) \\ &= \frac{1}{n}H(X^n|M). \end{aligned}$$