Shannon's Secret

Himanshu Tyagi
Indian Institute of Science

Abstract

review Shannon's notion of information theoretic secrecy

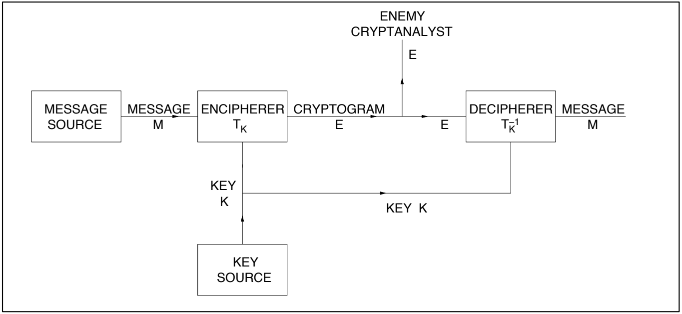
track the evolution of Shannon's ideas into modern crypto

along the way, review some major breakthroughs*

*Terms and conditions apply.

Communication Theory of Secrecy Systems

Secure Transmission Of A Message



How do we capture mathematically the notion of "secrecy"?

Shannon's "Secret"

Eavesdropper's knowledge before observing the cryptogram:

Prior distribution on the message P_M

Eavesdropper's knowledge after observing the cryptogram:

Posterior distribution on the message $\mathrm{P}_{M|E=e}$

Shannon's "Secret"

Eavesdropper's knowledge before observing the cryptogram:

Prior distribution on the message P_M

Eavesdropper's knowledge after observing the cryptogram:

Posterior distribution on the message $\mathrm{P}_{M|E=e}$

Let $f(\mathrm{P})$ denote the level of "uncertainty" in P

Secrecy of the message is defined as

$$\sigma(M;E) = f(\mathrm{P}_M) - \mathbb{E}\big[f(\mathrm{P}_{M|E})\big]$$

Shannon's "Secret"

Eavesdropper's knowledge before observing the cryptogram:

Prior distribution on the message P_M

Eavesdropper's knowledge after observing the cryptogram:

Posterior distribution on the message $\mathrm{P}_{M|E=e}$

Let $f(\mathrm{P})$ denote the level of "uncertainty" in P

Secrecy of the message is defined as

$$\sigma(M;E) = f(\mathrm{P}_M) - \mathbb{E}\big[f(\mathrm{P}_{M|E})\big]$$

Shannon chose his favorite *concave* function as f, namely

the Shannon entropy $f(\mathrm{P}) = H(\mathrm{P}) = -\sum_x \mathrm{P}(x) \log \mathrm{P}(x)$

Real World Versus Ideal World

- The *view* in the *real world*: P_{ME}
- The *view* in the *ideal world*: $\mathrm{P}_M \times \mathrm{P}_E$

$$\sigma(M;E) = H(M) - H(M|E)$$

$$= I(M \wedge E) \quad : \textit{Mutual Information} \text{ between } M \text{ and } E$$

$$= D(\mathrm{P}_{ME} \| \mathrm{P}_M \times \mathrm{P}_E)$$

$D(\mathrm{P}\|\mathrm{Q}) = \sum_x \mathrm{P}(x) \log \frac{\mathrm{P}(x)}{\mathrm{Q}(x)}$ is the Kullback-Leibler divergence

Analysis Of Theoretical Secrecy

Let M, K take values in an Abelian group $(\mathbb{G}, +)$

Consider the encryption $E = M + K$

$$
\begin{aligned}
\sigma(M;E) &= I(M \wedge E) \\
&= I(M \wedge M + K) \\
&= H(M+K) - H(M+K|M) \\
&\leq \log|\mathbb{G}| - H(M+K|M) \\
&= \log|\mathbb{G}| - H(K|M) \\
&= \log|\mathbb{G}| - H(K)
\end{aligned}
$$

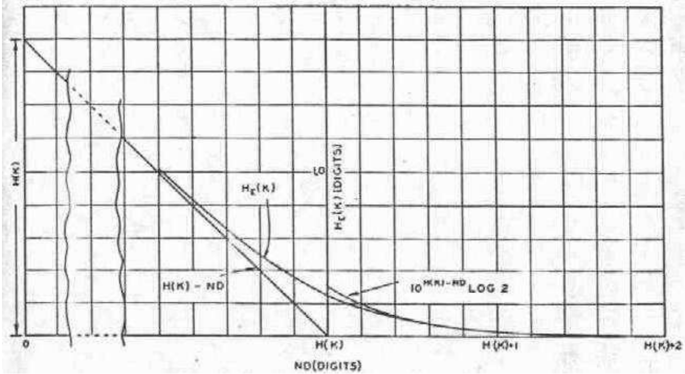
- Related the secrecy of the message to the uniformity of the key
- Used nontrivial manipulations of "uncertainty" of the cryptanalyst

Change In Secrecy Per Observed Cryptogram Bit

- Theoretical secrecy

Consider a message M that can take m^N possible values

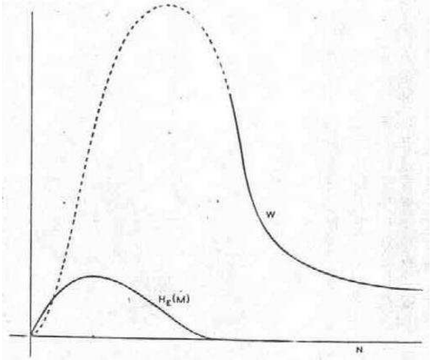
Let $ND = N \log m - H(M)$



Change In Secrecy Per Observed Cryptogram Bit

- Practical secrecy

$W(N)$: Work in "human hours" used to ascertain the posterior $\mathrm{P}_{M|E}$



Germination Of Cryptographic Thinking

Secrecy of a cipher can be established only after a thorough theoretical and practical evaluation of the power of a cryptanalyst

- Define secrecy

  keeping *the strengths and the limitations of the cryptanalyst* in mind

- Measure secrecy

  by the difference between *the real world and the ideal worlds*

- Analyze secrecy

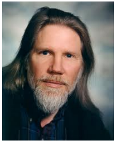
  of a message by *reducing* it to the secrecy of the corresponding key

- Quantize secrecy

  by tracking each *bit of information* leaked

Enter Diffie and Hellman

Diffie Hellman Key Exchange




"New Directions in Cryptography," 1976.

Convert a difficult number theory problem into a secure system:

A computationally limited cryptanalyst deems all answers equally likely

Diffie Hellman Key Exchange

1. Party 1 chooses a uniformly over $\mathbb{F}$ and sends g^a
2. Party 2 chooses b uniformly over $\mathbb{F}$ and sends g^b
3. Both parties compute g^{ab}

Key principle: Discrete $\exp$ is easy, discrete $\log$ is difficult

First realization of Shannon's "man hours" based practical security

Led to RSA, El Gamal's encryption scheme, ...

How Do We Quantize The Secrecy Of Such Schemes?

1. From statistical difference between the real and the ideal world to the difference in the power of a cryptanalyst in the two worlds

2. From randomness to pseudorandomness

A basic principle:

Instead of direct secrecy guarantees use *reduction arguments*

and keep a track of components with ambiguous secrecy guarantees

Building towards semantic secrecy

Step 1. An alternative definition of Information Theoretic secrecy

$$\sigma_{\texttt{var}}(M;E) = d_{\texttt{var}}(\mathrm{P}_{ME}, \mathrm{P}_M \times \mathrm{P}_E) = \mathbb{E}_{\mathrm{P}_M}\left[d_{\texttt{var}}(\mathrm{P}_{E|M}, \mathrm{P}_E)\right],$$

where $d_{\texttt{var}}(\mathrm{P}, \mathrm{Q}) = \sup_A \mathrm{P}(A) - \mathrm{Q}(A)$ is the total variation distance

Building towards semantic secrecy

Step 1. An alternative definition of Information Theoretic secrecy

$$\sigma_{\mathtt{var}}(M;E) = d_{\mathtt{var}}(\mathrm{P}_{ME}, \mathrm{P}_M \times \mathrm{P}_E) = \mathbb{E}_{\mathrm{P}_M}\left[d_{\mathtt{var}}(\mathrm{P}_{E|M}, \mathrm{P}_E)\right],$$

where $d_{\mathtt{var}}(\mathrm{P},\mathrm{Q}) = \sup_A \mathrm{P}(A) - \mathrm{Q}(A)$ is the total variation distance

The two secrecy indices are related as

$$\begin{aligned}\frac{1}{2\ln 2}\sigma_{\mathtt{var}}(M;E) &\leq \sigma(M;E)\\ &\leq \sigma_{\mathtt{var}}(M;E)\log(|\mathcal{M}|-1) + h(\min\{\sigma_{\mathtt{var}}(M;E), 2\}),\end{aligned}$$

where $\mathcal{M} \equiv$ the set of messages and $h(\cdot) \equiv$ the binary entropy function

Semantic Secrecy

Building towards semantic secrecy

Step 1. An alternative definition of Information Theoretic secrecy

$$\sigma_{\mathtt{var}}(M;E) = d_{\mathtt{var}}(\mathrm{P}_{ME}, \mathrm{P}_M \times \mathrm{P}_E) = \mathbb{E}_{\mathrm{P}_M}\left[d_{\mathtt{var}}(\mathrm{P}_{E|M}, \mathrm{P}_E)\right],$$

where $d_{\mathtt{var}}(\mathrm{P}, \mathrm{Q}) = \sup_A \mathrm{P}(A) - \mathrm{Q}(A)$ is the total variation distance

$\sigma_{\mathtt{var}}(M;E) \leq \epsilon \Rightarrow$

A randomized algorithm will attain the same performance guarantee

in the real world P_{ME} as in the secure ideal world $\mathrm{P}_M \times \mathrm{P}_E$,

up to an additional probability of error ϵ

Semantic Secrecy

Building towards semantic secrecy

Step 2. A hypothesis testing interpretation of $d_{\texttt{var}}(\mathrm{P}, \mathrm{Q})$

Let $\mathrm{P}_0 = \mathrm{P}$ and $\mathrm{P}_1 = \mathrm{Q}$.

An unbiased coin B is tossed and a sample is generated from P_B

An observer of the sample forms an estimate $\hat{B}$ of B

The least probability of error $P_e^* = \min_{\hat{B}} \Pr\left(\hat{B} \neq B\right)$ satisfies

$$\frac{1}{2} d_{\texttt{var}}(\mathrm{P}, \mathrm{Q}) = \frac{1}{2} - P_e^* = \text{ advantage over a random guess}$$

Semantic Secrecy

Building towards semantic secrecy

Step 3. Information theoretic semantic secrecy

$\sigma_{\texttt{sem}}(M;E)$ is the maximum advantage in guessing $f(M)$ from E

in the real world has over the same guess in the ideal world, namely

$$\sigma_{\texttt{sem}}(M;E) := \min_{G} \max_{f,\hat{f}} \Pr\left(\hat{f}(E) = f(M)\right) - \Pr\left(\hat{f}(G) = f(M)\right),$$

where the random variable G is independent of (M,E)

Semantic Secrecy

Building towards semantic secrecy

Step 4. Distributions free secrecy indices

- Assume the worst-case knowledge for cryptanalyst
- Encryption process is defined by $T = \mathrm{P}_{E|M}$

$$\sigma_{\texttt{var}}(\mathcal{M};T) = \sup_{\mathrm{P}_{ME}:\mathrm{P}_{E|M}=T} \sigma_{\texttt{var}}(M;E)$$

$$\sigma_{\texttt{sem}}(\mathcal{M};T) = \sup_{\mathrm{P}_{ME}:\mathrm{P}_{E|M}=T} \sigma_{\texttt{sem}}(M;E)$$

$$\sigma_{\texttt{sim}}(\mathcal{M};T) \leq \sigma_{\texttt{var}}(\mathcal{M};T) \leq 2\sigma_{\texttt{sim}}(\mathcal{M};T)$$

Semantic Secrecy

Building towards semantic secrecy

Step 4. Distributions free secrecy indices

- Assume the worst-case knowledge for cryptanalyst
- Encryption process is defined by $T = \mathrm{P}_{E|M}$

$$\sigma_{\mathtt{var}}(\mathcal{M};T) = \sup_{\mathrm{P}_{ME}:\mathrm{P}_{E|M}=T} \sigma_{\mathtt{var}}(M;E)$$

$$\sigma_{\mathtt{sem}}(\mathcal{M};T) = \sup_{\mathrm{P}_{ME}:\mathrm{P}_{E|M}=T} \sigma_{\mathtt{sem}}(M;E)$$

$$\sigma_{\mathtt{sim}}(\mathcal{M};T) \leq \sigma_{\mathtt{var}}(\mathcal{M};T) \leq 2\sigma_{\mathtt{sim}}(\mathcal{M};T)$$

Goldwasser-Micali's Semantic Secrecy

"Probabilistic Encryption," 1976.

- Restrict the power of cryptanalyst to a computational class
- *Asymptotic theory:* Parameterize secrecy index with input-size

$$n = \log|\mathcal{M}| + \log|\mathcal{K}|$$

Cryptanalyst can use only Prob. Poly. Time (PPT) in n functions $\hat{f}$

$$\sigma_{\texttt{sem}}(M;E) = \min_{G} \max_{f,\hat{f} in PPT} \Pr\left(\hat{f}(E) = f(M)\right) - \Pr\left(\hat{f}(G) = f(M)\right),$$

Tricks Of The Trade

- Formulate the problem with information theoretic secrecy
- Take a "difference in statistician's ability" view of distances
- Use reduction arguments to relate the secrecy of your system to that of a well-studied secure primitive
- Replace your information theoretic *reduction* to computational by imposing appropriate computational restrictions

Eg. 1: Distinguishing Secrecy ≡ Semantic Secrecy

$$\sigma_{\texttt{dis}}(\mathcal{M};T) = \max_{m_0,m_1\in\mathcal{M}} \left(\max_{\hat{B}\text{in PPT}} \Pr\left(\hat{B}\left(T_{m_B}\right) = B\right) - \frac{1}{2}\right)$$

Step 1. Show equivalence for IT secrecy

$$\sigma_{\texttt{dis}}(\mathcal{M};T) \le \sigma_{\texttt{sem}}(\mathcal{M};W) \le 2\sigma_{\texttt{dis}}(\mathcal{M};W)$$

Proof. For a fixed m_0, there exists m_1 such that

$$\begin{aligned}&\Pr\left(\hat{f}(T_M) = f(M)\right) - \Pr\left(\hat{f}(G) = f(M)\right)\\ &\le \Pr\left(\hat{f}(T_{m_1}) = f(m_1)\right) - \Pr\left(\hat{f}(T_{m_0}) = f(m_1)\right),\end{aligned}$$

and so, for $\hat{B}(z) = \mathbb{1}\left(\hat{f}(z) = f(m_1)\right)$

$$\Pr\left(\hat{B}(T_{m_B}) = B\right) \ge \frac{1}{2} + \frac{1}{2}\left[\Pr\left(\hat{f}(T_M) = f(M)\right) - \Pr\left(\hat{f}(G) = f(M)\right)\right]$$

Eg. 1: Distinguishing Secrecy $\equiv$ Semantic Secrecy

$$\sigma_{\texttt{dis}}(\mathcal{M};T) = \max_{m_0,m_1\in\mathcal{M}} \left(\max_{\hat{B}\text{in PPT}} \Pr\left(\hat{B}\left(T_{m_B}\right) = B\right) - \frac{1}{2}\right)$$

Step 1. Show equivalence for IT secrecy

$$\sigma_{\texttt{dis}}(\mathcal{M};T) \le \sigma_{\texttt{sem}}(\mathcal{M};W) \le 2\sigma_{\texttt{dis}}(\mathcal{M};W)$$

Proof. For a fixed m_0, there exists m_1 such that

$$\begin{aligned} &\Pr\left(\hat{f}(T_M) = f(M)\right) - \Pr\left(\hat{f}(G) = f(M)\right) \\ &\le \Pr\left(\hat{f}(T_{m_1}) = f(m_1)\right) - \Pr\left(\hat{f}(T_{m_0}) = f(m_1)\right), \end{aligned}$$

and so, for $\hat{B}(z) = \mathbb{1}\left(\hat{f}(z) = f(m_1)\right)$

$$\Pr\left(\hat{B}(T_{m_B}) = B\right) \ge \frac{1}{2} + \frac{1}{2}\left[\Pr\left(\hat{f}(T_M) = f(M)\right) - \Pr\left(\hat{f}(G) = f(M)\right)\right]$$

Step 2. Check the feasibility of steps under computational restrictions

Eg. 1: Distinguishing Secrecy ≡ Semantic Secrecy

$$\sigma_{\texttt{dis}}(\mathcal{M};T) = \max_{m_0,m_1\in\mathcal{M}} \left(\max_{\hat{B}\text{in PPT}} \Pr\left(\hat{B}\left(T_{m_B}\right) = B\right) - \frac{1}{2}\right)$$

Step 1. Show equivalence for IT secrecy

$$\sigma_{\texttt{dis}}(\mathcal{M};T) \le \sigma_{\texttt{sem}}(\mathcal{M};W) \le 2\sigma_{\texttt{dis}}(\mathcal{M};W)$$

Proof. For a fixed m_0, there exists m_1 such that

$$\begin{aligned}
&\Pr\left(\hat{f}(T_M) = f(M)\right) - \Pr\left(\hat{f}(G) = f(M)\right) \\
&\le \Pr\left(\hat{f}(T_{m_1}) = f(m_1)\right) - \Pr\left(\hat{f}(T_{m_0}) = f(m_1)\right),
\end{aligned}$$

and so, for $\hat{B}(z) = \mathbb{1}\left(\hat{f}(z) = f(m_1)\right)$ ($\hat{B}$ must be in PPT)

$$\Pr\left(\hat{B}(T_{m_B}) = B\right) \ge \frac{1}{2} + \frac{1}{2}\left[\Pr\left(\hat{f}(T_M) = f(M)\right) - \Pr\left(\hat{f}(G) = f(M)\right)\right]$$

Step 2. Check the feasibility of steps under computational restrictions

Eg. 2: Defining Pseudorandomness

Let M, K take values in an Abelian group $(\mathbb{G}, +)$

Consider the encryption $E = M + K$

Step 1. Uniform K implies IT distinguishable secrecy

Can distinguish $K + m_0$ from $K + m_1 \Rightarrow$ can distinguish K from uniform

Eg. 2: Defining Pseudorandomness

Let M, K take values in an Abelian group $(\mathbb{G}, +)$

Consider the encryption $E = M + K$

Step 1. Uniform K implies IT distinguishable secrecy

Can distinguish $K + m_0$ from $K + m_1 \Rightarrow$ can distinguish K from uniform

Step 2. "Pseudorandom" K implies IT distinguishable secrecy

Can distinguish $K + m_0$ from $K + m_1 \Rightarrow$ can distinguish K from uniform
in PPT in PPT

Eg. 2: Defining Pseudorandomness

Let M, K take values in an Abelian group $(\mathbb{G}, +)$

Consider the encryption $E = M + K$

Step 1. Uniform K implies IT distinguishable secrecy

Can distinguish $K + m_0$ from $K + m_1 \Rightarrow$ can distinguish K from uniform

Step 2. "Pseudorandom" K implies IT distinguishable secrecy

Can distinguish $K + m_0$ from $K + m_1$ in PPT $\Rightarrow$ can distinguish K from uniform in PPT

Definition of pseudorandomness

K is pseudorandom if you cannot distinguish it from uniform in PPT

Secure Public-Key Encryption Using Diffie-Hellman

Given a finite field $\mathbb{F}$ and its generator g (say):

1. Party 2 generates $b \sim \texttt{unif}(\mathbb{F})$ and publishes g^b publicly
2. Party 1 seeks to send a message $m \in \mathbb{F}$ to Party 2
   - It generates $a \sim \texttt{unif}(\mathbb{F})$ and sends $(g^a, (g^b)^a \oplus m)$
3. Party 2 observes $(g^a, (g^b)^a \oplus m)$ and computes

$$\hat{m} = (g^a)^b \oplus (g^b)^a \oplus m$$

The scheme is secure under $\sigma_{\texttt{dis}}$ if g^{ab} constitutes pseudorandomness for a "cryptanalyst with side-information" (g^a, g^b)

Active Adversaries: Chosen Plaintext Attack

Hereto, the cryptanalyst was gives access to one cryptogram

In practise, however, often a malacious cryptanalyst
can obtain cryptograms for his chosen messages $m_1, ..., m_t$

Security can ensured using a pseudorandom function, namely
a function which cannot be distinguished from a random function

Pseudorandom functions can be constructed using pseudorandomness

We need one more tool from Shannon's toolkit...

Chain Rule: The So-Called Hybrid Argument

Just like Shannon's measures of information, $d_{\mathtt{var}}$, too, "tensorizes":

$$d_{\mathtt{var}}\left(\mathrm{P}_{X_1,\ldots,X_n}, \mathrm{Q}_{X_1,\ldots,X_n}\right) \leq \sum_{i=0}^{n-1} d_{\mathtt{var}}\left(\mathrm{P}_{X^i} Q_{X_{i+1}^n|X^i}, \mathrm{P}_{X^{i+1}} Q_{X_{i+2}^n|X^{i+1}}\right)$$

Used to reduce the ϵ-secrecy of a collection of n components to ϵ/n-secrecy of one of the component

Shannon's Secret Is Secure Out In Open

An Information Theoretic approach to cryptography

- Formulate the problem requiring information theoretic secrecy
- Replace the distances with the difference in the outcome of a cryptanalyst

Shannon's Secret Is Secure Out In Open

An Information Theoretic approach to cryptography

- Formulate the problem requiring information theoretic secrecy
- Replace the distances with the difference in the outcome of a cryptanalyst
- Use chain rules, chain saws, human chains and what not to identify a basic primitive that will enable the required secure object