Bell Labs May 18, 1998

Shannon Theory and Contemporary Cryptology

James L. Massey, Prof. em. ETH Zürich
Trondhjemsgade 3, 2th, DK-2100 Copenhagen
101767.233@compuserve.com

From Shannon's **1949** *B.S.T.J.* paper: "The Communication Theory of Secrecy Systems"

Theoretical Secrecy

- Definition of a Secrecy System
- Perfect Secrecy
- Strongly Ideal Ciphers
- Unicity Distance

Practical Secrecy

- Work Function
- Confusion and Diffusion
- Use of Data Compression

Shannon's "Fig. 1", reproduced here, gives a system diagram that makes clear what the mathematical assumptions are for a ciphertext-only attack on a secrecy system.

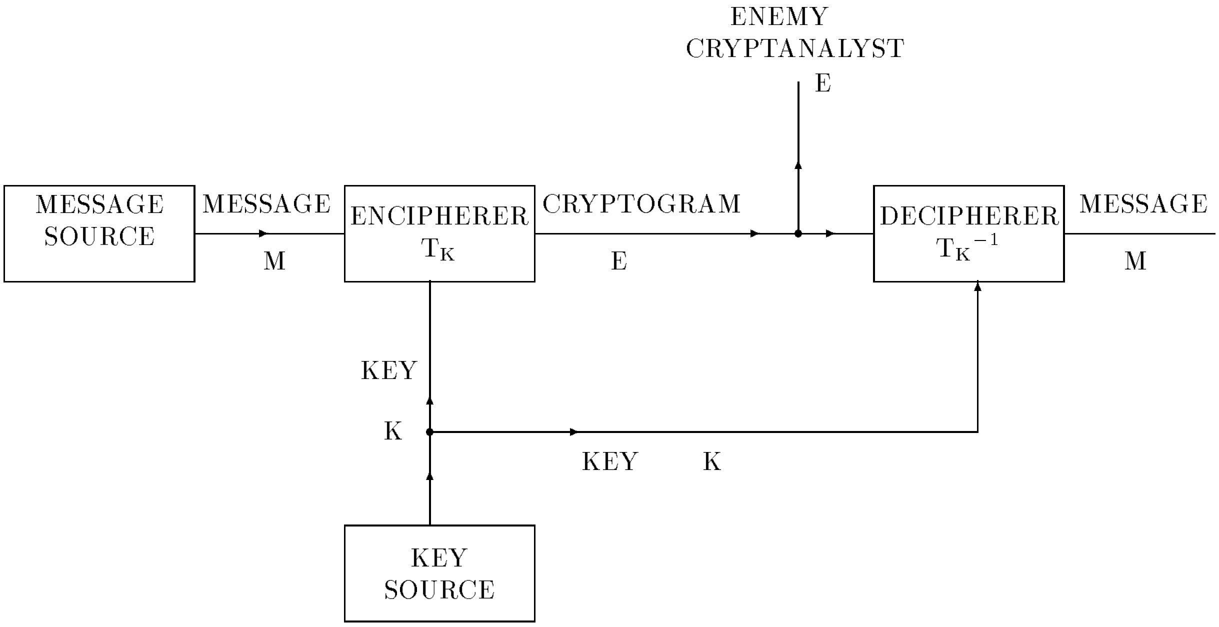


Fig. 1–Schematic of a general secrecy system.

Shannon's schematic implies

$$H(M \mid EK) = 0.$$

Shannon's bold definition:
A secrecy system provides **perfect secrecy** against a ciphertext-only attack if

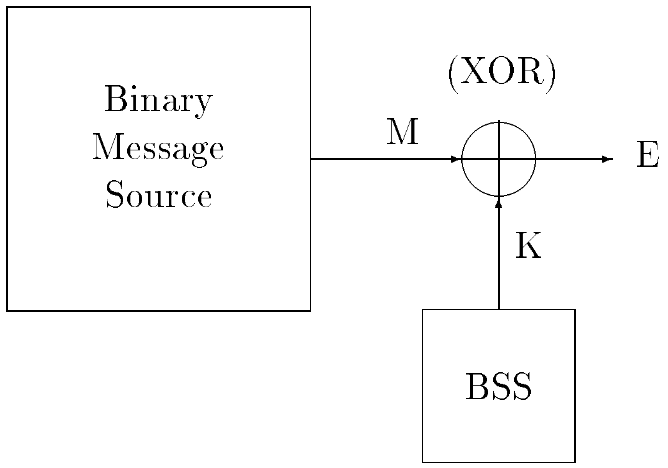
$$H(M \mid E) = H(M),$$

which is equivalent to saying that M and E are statistically independent.

This was the first precise statement of what it means to say that a secrecy system is "unbreakable"!

Shannon showed that Vernam's 1926 cipher (now usually called the "one-time pad") gave perfect secrecy.

Vernam's Cipher (One-Time Pad)



BSS = Binary Symmetric Source

BSS $\Leftrightarrow$ $P(K = k) = 2^{-N}$, all k, where N is the length in bits of the message M (and of the key K).

$\Rightarrow$ For every m,
$P(E = e \mid M = m) = 2^{-N}$, all e.

$\Rightarrow$ $H(E \mid M) = H(E)$.

$\Rightarrow$ Perfect secrecy!

Shannon's Bound on Key Size:
If a secrecy system provides perfect secrecy, then

$$H(K) \geq H(M)$$

where M is the total amount of plaintext encrypted before the key is changed.

Shannon's elegant proof:

$$\begin{aligned} H(M) &= H(M \mid E) \\ &\leq H(KM \mid E) \\ &= H(K \mid E) + H(M \mid KE) \\ &= H(K \mid E) \\ &\leq H(K). \end{aligned}$$

Shannon always measured cryptographic security in terms of **entropy** (or "**uncertainty**").

This makes a lot of sense!

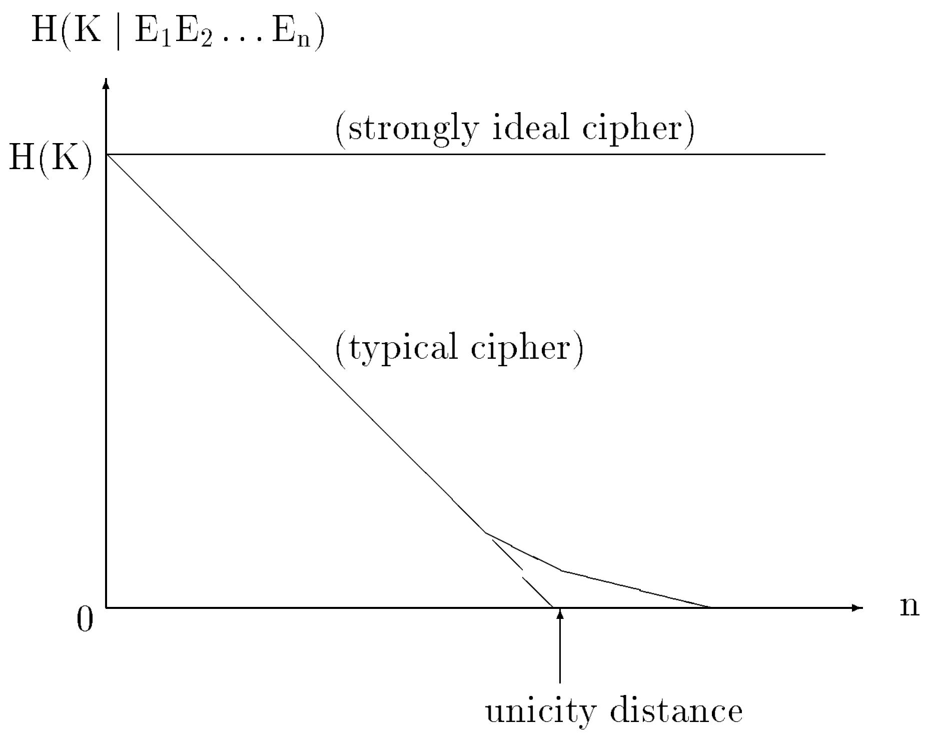
From my paper, "Guessing and Entropy," at ISIT '94:

The average number of guesses G that the enemy cryptanalyst must make to find the key K, given the cryptogram E, satisfies

$$\mathrm{E}[\mathrm{G}] \geq \frac{1}{4} 2^{\mathrm{H}(\mathrm{K}|\mathrm{E})} + 1,$$

provided that $\mathrm{H}(\mathrm{K} \mid \mathrm{E}) \geq 2$ bits.

Shannon defined the "Key Equivocation" function $H(K \mid E_1 E_2 \ldots E_n)$ to be the uncertainty (i.e., entropy) of the key given the first n digits of the ciphertext.



The "unicity distance" is the point at which there is essentially only one key consistent with the observed ciphertext.

Shannon's formula for the unicity distance,

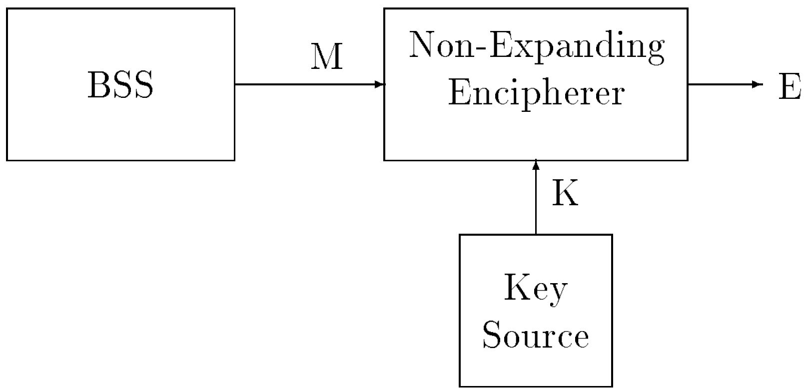
$$\text{unicity distance} = \mathrm{H(K)}/\rho,$$

where

$$\rho = \text{redundancy of enciphered message}$$

which he derived for a well-defined "random cipher", is still used today to estimate the unicity distance of virtually all ciphers.

Shannon's formula implies that any non-expanding cipher becomes strongly ideal when used to encrypt a message source with no redundancy, i.e., the BSS.



BSS $\Leftrightarrow$ $\mathrm{P}(\mathrm{M} = \mathrm{m}) = 2^{-\mathrm{N}}$, all m,
where N is the length in bits of the message (and of the cryptogram).

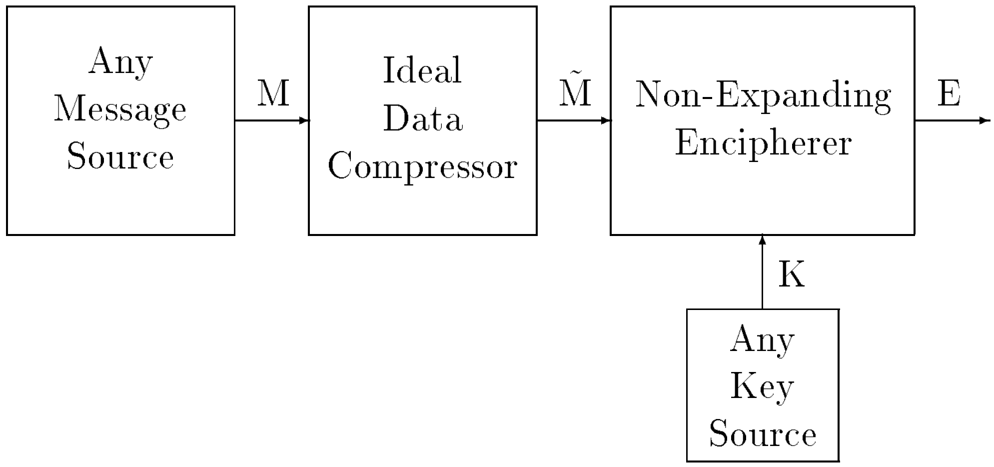
$\Rightarrow$ For every k,
$\mathrm{P}(\mathrm{E} = \mathrm{e} \mid \mathrm{K} = \mathrm{k}) = 2^{-\mathrm{N}}$, all e

$\Rightarrow$ $\mathrm{H}(\mathrm{E} \mid \mathrm{K}) = \mathrm{H}(\mathrm{E})$

$\Rightarrow$ strongly ideal!

This has implications for the exchange of session keys, where the message is well-modeled as the output of a BSS.

Shannon pointed out the usefulness of data compression in providing secrecy as several points in his 1949 paper. For instance, in the following situation:



the encrypted message $\tilde{\mathrm{M}}$ has redundancy $\rho = 0$ and hence the quite arbitrary cipher becomes strongly ideal!

No amount of effort will give the enemy cryptanalyst any information about the key in a ciphertext-only attack, no matter how much plaintext is encrypted!

Shannon realized that **practical secrecy** would have to depend on **computational security**.

"The problem of good cipher design is essentially one of finding difficult problems, subject to certain other conditions. This is a rather unusual situation, since one is ordinarily seeking the simple and easily soluble problems in a field." (p. 704)

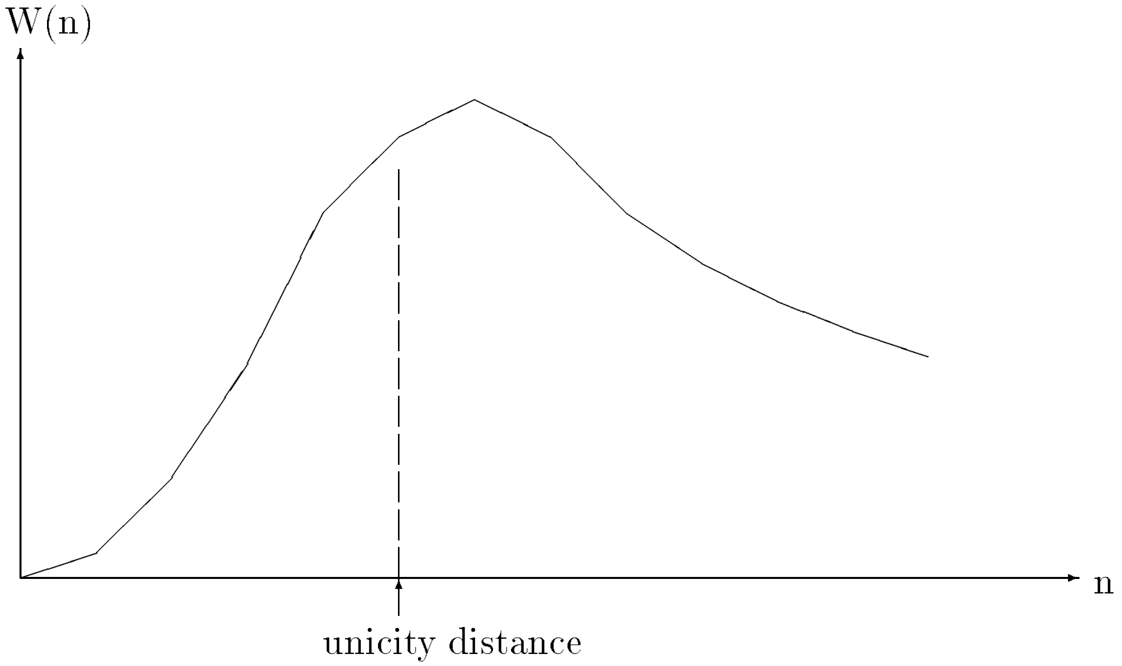
"How can we ever be sure that a system which is not ideal and therefore *has* a unique solution for sufficiently large N will require a large amount of work to break with *every* method of analysis. ... We may construct our cipher in such a way that breaking it is equivalent to (or requires at some point in the process) the solution of some problem known to be laborious." (p. 704)

Virtually all practical ciphers today, such as the Data Encryption Standard [DES], are explcitly designed on the basis of Shannon's principles of **confusion** and **diffusion**. Here is an excerpt from a consulting report that I recently wrote:

The [name suppressed] block ciphers are all iterative ciphers, i.e., ciphers in which cryptographic strength is created by repeating a "weak" enciphering operation many times. All practical block ciphers in use today such as the Data Encryption Standard [1], IDEA [2] and SAFER [3] are iterative ciphers. ... The design of iterative ciphers has mainly been guided by the principles of confusion and diffusion that were enunciated by Shannon in 1949 [4]. By "confusion" in a cipher, Shannon meant that the enciphering process should be such that the ciphertext statistics depend on the plaintext statistics in a manner too complicated to be exploited by an attacker. By "diffusion" in a cipher, Shannon meant that each bit of the input key should influence many bits of the ciphertext.

Shannon defined the **work function** W(n) of a cipher as the expected computational work to find the key, given the first n bits of the ciphertext.

Shannon postulated that the work function of a typical cipher would have the following form:



No one has yet computed, or even found a nontrivial lower bound, on the work function for any practical cipher.

Diffie and Hellman, whose 1976 paper marked the beginning of public-key cryptography, have said that they were motivated by the statements of Shannon given on transparency 11 above.

Question: **Did Shannon himself miss public-key cryptography**, i.e., did he miss the fact that a "secret key" is not necessary for secrecy?

Answer: **Maybe!**

But in fact there is no proof that **one-way functions**, much less the **trap-door one-way functions** needed for public-key cryptography exist!

Shannon often gave only intuitive arguments to support his claims but he never resorted to conjectures, which are the basis of all of public-key cryptography. Perhaps by the Centennial of Shannon's 1948 paper, if we are lucky, we will know whether he actually missed public-key cryptography or whether such cryptography is in fact impossible.